

Confrontation de données expérimentales à l'analyse d'un corpus : le cas de la récupération d'instances en production écrite

Cecilia Gunnarsson, Pierre Largy
Université de Toulouse2-Le-Mirail



Synergies Pays Scandinaves n° 5 - 2010 pp. 49-64

Résumé : *L'étude de la production écrite a fait l'objet de nombreux travaux expérimentaux depuis une vingtaine d'années. Certaines de ces recherches ont permis de modéliser la façon dont apprenants et experts produisent la morphologie flexionnelle du nombre nominal et verbal. Le présent article met ces conceptions à l'épreuve de l'examen d'un corpus de production de textes d'élèves de CE2 et de 4^{ième}. Cette étude confirme l'existence d'automatismes en production écrite et montre que la performance orthographique est sous-tendue par un phénomène de récupération d'instances en mémoire.*

Mots-clés : *développement cognitif ; production écrite ; orthographe ; accord nominal ; accord verbal ; récupération d'instance ; apprenant/expert*

Summary: *Written production has been subject to a large range of experimental studies in the last twenty years. Some of this research has enhanced models of the production of nominal and verbal flexional morphology in learners and experts. This article examines these concepts in a corpus of texts produced by pupils in the 3rd and 8th grades (CE2 and 4^{ième}). The study confirms the existence of automatisms in written production and shows that retrieval of memorized instances underlies the spelling performance.*

Keywords: *cognitive development; written production; spelling; noun agreement; verb agreement; memorized instances; learner/expert*

1. Introduction

Depuis une vingtaine d'années, de nombreux travaux expérimentaux ont tenté de comprendre puis de modéliser l'activité de production verbale écrite ; certains chercheurs ont abordé ce sujet avec l'intention d'éclairer le développement cognitif. Pour ces derniers, la question de la production de l'écrit allait permettre d'illustrer les questions plus générales de l'acquisition des connaissances et de leur mobilisation dans les activités de notre vie quotidienne. Les travaux sur l'accord en nombre ont permis de mieux comprendre la façon dont l'enfant acquiert et mobilise des connaissances dans

le domaine de l'orthographe grammaticale, au cours de son apprentissage de l'écrit. Ces travaux ont abouti à la description des processus et des étapes qui conduisent progressivement l'apprenant vers l'expertise dans la production des marques du pluriel (nominal, adjectival et verbal).

Nous entendons ici aborder les textes narratifs produits par les élèves de CE2 et de 4^{ème}, en y recherchant les erreurs qui jalonnent l'apprentissage de l'accord en nombre (nominal, adjectival et verbal) : absence de marquage, erreurs de marquage, accords par proximité. En effet, la modélisation de l'accès à l'expertise dans ce domaine repose principalement sur des travaux expérimentaux ; l'étude du présent corpus de production écrite, collecté de façon plus écologique qu'on ne peut le faire dans un contexte expérimental, nous fournit l'occasion de rechercher, à travers la nature des erreurs commises, la trace de l'accès à l'expertise.

De plus, les processus mis en évidence dans l'accès à l'expertise en production de la morphologie flexionnelle du nombre étant probablement des processus généraux d'accès à l'expertise en production verbale écrite, nous ne limiterons pas notre étude des erreurs au seul cas de l'accord en nombre. Au contraire, nous essayerons d'établir des rapprochements avec d'autres types d'erreurs en invoquant le processus général de récupération d'instances.

Ce faisant, nous observerons comment les enfants parviennent à surmonter les difficultés qui tiennent à la langue elle-même, ici le français, à travers la façon dont ils produisent d'une part une morphologie flexionnelle silencieuse, celle du nombre nominal (-s), adjectival (-s) et verbal (-nt), d'autre part des flexions verbales homophones ou quasi-homophones, celles de l'infinitif (-er), du participe passé (-é) et de l'imparfait (-ait) puis celles du participe passé (-is, -u) et du passé simple (-it, -ut).

La partie théorique de cet article débutera par une brève présentation des méthodes expérimentales d'étude de la production verbale écrite. Elle se poursuivra par la présentation des travaux ayant conduit à modéliser les étapes de l'acquisition de la morphologie flexionnelle du nombre. Le processus de récupération d'instances sera explicité à cette occasion.

2. Les méthodes d'étude de la production écrite

Les méthodes d'étude de la production écrite les plus utilisées aujourd'hui s'appuient sur les nouvelles possibilités offertes aux psycholinguistes par l'évolution des modes d'enregistrement et d'analyse de l'activité humaine. L'utilisation des tablettes graphiques permet d'accroître le nombre d'indicateurs de l'activité humaine sur lesquels faire porter nos hypothèses. En effet, en plus des paramètres spatiaux de cette activité déjà accessibles par la simple méthode « papier-crayon », nous pouvons désormais analyser un grand nombre de paramètres temporels : vitesse d'écriture, pause, accélération, etc. Des logiciels permettent ainsi d'analyser le décours temporel de l'activité manuscrite (Chesnet & Alamargot, 2005) et tapuscrite (Strömqvist & Karlsson, 2002). Ces méthodes d'investigation de la production écrite sont parfois

couplées à d'autres techniques, par exemple à celle des protocoles verbaux concomitants (p. ex. Gunnarsson, 2006) pour tenter de repérer la nature de l'activité dans laquelle le scripteur est engagé à un instant donné de sa production (planification, formulation, révision, etc.). Une autre technique consiste à parfois imposer une tâche secondaire au scripteur (Fayol & Got, 1991), cette fois pour évaluer le coût cognitif de telle ou telle dimension de l'activité rédactionnelle, ou pour étudier les conséquences sur la production écrite d'une privation de l'attention. Les travaux que nous allons présenter sur l'acquisition de la morphologie flexionnelle du nombre se sont appuyés sur ces différentes méthodes (méthode off-line d'étude de phrases et de syntagmes produits sous dictée ; méthode on-line d'analyse du décours temporel de cette activité particulière de production écrite). Ces recherches ont eu recours au paradigme de la double-tâche (tâche principale + tâche secondaire) pour discuter du caractère plus ou moins automatisé des modalités de mise en œuvre de l'accord en nombre à différentes étapes de cet apprentissage.

3. L'approche expérimentale de l'accès à l'expertise en production de l'accord en nombre

Dans le champ d'étude expérimentale de la compétence grammaticale, on s'est demandé si la maîtrise des habiletés morphologiques relevait de l'application de règles ou si elle tenait à l'existence de processus associatifs. La littérature dans ce domaine, au début des années 90, était essentiellement anglo-saxonne et se focalisait sur la gestion de la morphologie verbale « régulière » et « irrégulière » du *preterit*. Elle opposait ainsi deux conceptions, l'une soutenue par des modèles duels (p.ex. Marcus et al., 1995) selon lesquels les flexions régulières sont produites grâce à des règles (p.ex. ajouter *-ed* au verbe) et les flexions irrégulières sont listées en mémoire, l'autre défendue par des modèles connexionnistes (p. x. Plunkett & Marchman, 1993) pour lesquels la production des deux types de flexions est issue d'un même mécanisme de sélection au sein d'un réseau associatif.

Les travaux sur l'acquisition de la morphologie flexionnelle du nombre nominal et verbal du français écrit ont récemment contribué à cette réflexion. Parce que cette morphologie à la particularité d'être silencieuse et que la production ne saurait se réduire à une simple transcription phonographémique, on a pu étudier la façon dont les enfants d'école primaire puis de collège en font progressivement usage. Initialement conduites dans une perspective comparative entre l'accord nominal et l'accord verbal, ces recherches ont mis en évidence plusieurs faits. Premièrement, l'accord nominal, morphologiquement plus simple et dont la pluralité est sémantiquement motivée, est maîtrisé avant l'accord verbal (Totereau, Thevenin, & Fayol, 1997). Secondement, quel que soit l'accord, la compréhension des marques flexionnelles du nombre précède leur production. Des travaux plus spécifiquement conduits sur l'accord verbal ont permis de dégager trois étapes dans son acquisition. Selon Fayol, Hupet et Largy (1999), les enfants passent d'une phase au cours de laquelle ils n'utilisent aucune flexion du pluriel, à une phase de calcul laborieux de l'accord matérialisée sous la forme de l'application d'une règle « condition/action » (Anderson, 1992) puis, enfin, à une phase d'activation automatique de la flexion verbale appropriée (Logan, 1991).

Dans ce cas, l'expertise est mise en évidence expérimentalement par la survenue d'erreurs d'accord par proximité. Cette expertise se manifeste par le fait que la seule présence d'un item pluriel devant le verbe suffit à activer le morphème -nt (p.ex. *Le père des enfants arrivent*). Elle se manifeste aussi par l'absence de marquage du pluriel verbal, soit lorsque le sujet est inversé (p.ex. *Dans le train passe les contrôleurs*) soit lorsque ses constituants ne présentent pas la marque morphologique du pluriel (p.ex. *Le papa et son petit garçon arrive*). Nous parlons là d'automatismes de production, pas de la capacité qu'a le scripteur à gérer attentionnellement sa production orthographique.

L'analyse des erreurs collectées expérimentalement auprès d'enfants aux différentes étapes de l'accès à l'expertise a permis de mettre en évidence un phénomène initialement appelé « phénomène de surgénéralisation ». Ce phénomène avait déjà été observé par Largy, Fayol et Lemaire (1996) chez l'adulte lettré en situation de surcharge cognitive (p.ex. *Le facteur prend des enveloppes. Il les timbres*). Totereau, Barrouillet et Fayol (1998) ont rapporté chez l'apprenant une tendance à la surgénéralisation des marques flexionnelles du nombre nominal et verbal aux noms, aux adjectifs et aux verbes (p.ex. *Les montrent* ; *Ils griffes*). L'existence de ces erreurs, très liée à la structure probabiliste de la langue, plaide en faveur de l'existence d'un phénomène de récupération d'instances en mémoire. Ceci nous conduit à penser que l'application de la règle n'est pas le seul processus en jeu dans la production de la morphologie flexionnelle du nombre et que l'expertise naîtrait de la place grandissante prise par cet automatisme de production.

L'hypothèse de l'existence d'un processus de récupération d'instances a été testée dans le cas de l'accord nominal en raison de la simplicité de la morphologie impliquée (\emptyset pour le singulier vs. -s pour le pluriel). Cousin, Largy et Fayol (2002) et Largy et al. (2007) ont apporté plusieurs arguments en sa faveur. Leurs résultats montrent que la rencontre préalable par les enfants de noms fréquemment fléchis au singulier ou au pluriel a une incidence sur la façon dont ils vont accorder ces noms en nombre. Le phénomène de récupération d'instances (ici, d'instances « racine-flexion ») est mis en évidence par la présence d'une difficulté particulière à réaliser l'accord du nom dans certains cas. C'est le cas lorsque l'enfant doit accorder au singulier un nom qu'il a habituellement rencontré fléchi au pluriel (dans ce cas, l'enfant tend à produire néanmoins la flexion -s du pluriel) et, dans une moindre mesure, lorsqu'il doit accorder au pluriel un nom habituellement fléchi au singulier (dans ce cas, ce nom reçoit moins fréquemment la marque du pluriel qu'un nom habituellement rencontré au pluriel). Le fait que cet effet perdure au cours de la scolarité permet de le différencier d'un simple effet « de récence », plus transitoire. Au contraire, cet effet semble rendre compte d'un phénomène plus général qui semble participer pleinement à la façon dont s'acquiert la langue écrite et sa structure morphologique.

Dans le cas de la morphologie flexionnelle du nombre, le scripteur disposerait à la fois de règles susceptibles d'être appliquées aux items familiers ou non familiers mais disposerait aussi d'instances mémorisées. Les travaux expérimentaux sur l'accord en nombre montrent qu'il existe, en plus de la connaissance de règles, un processus présent dès le début de l'apprentissage

qui permet à l'enfant de produire par écrit des items sous la forme qu'il a le plus fréquemment rencontrée. L'intérêt de la récupération d'instances est son efficacité ; il s'agit d'une procédure rapide, non consciente, donc peu coûteuse cognitivement, qui, fondée sur la structure probabiliste de la langue écrite, conduit généralement à un accord exact.

En résumé, l'expertise en production de la morphologie flexionnelle du nombre est sous-tendue par la connaissance de règles mais aussi par l'efficacité d'automatismes de production qui échappent à la conscience du scripteur. Les étapes qui jalonnent l'accès à l'expertise dans la production de la morphologie flexionnelle du nombre nominal et verbal ont été mises en évidence sur la base d'erreurs produites dans des conditions expérimentales très contrôlées (i.e., dictées de phrases avec des structures syntaxiques spécifiques, couplées à l'exécution ou non de tâches secondaires) : absence de marquage du pluriel ; marquage irrégulier en fonction de l'attention que l'enfant porte à la règle au moment où il écrit (accord tantôt réussi, tantôt échoué, chez le même enfant), erreurs typiques du phénomène de récupération d'instances (p.ex. marque du pluriel sur un nom au singulier ; verbe héritant de la marque du pluriel du nom ; infinitif produit en lieu et place d'un participe passé en [e]) ; accord par proximité sous l'effet d'automatismes de production (p.ex. *L'odeur des poubelles empestent ; Dans le train monte le contrôleur*).

Nous nous proposons, dans le présent corpus, d'essayer de retrouver ces erreurs et la trace des processus évoqués. En mettant la conception développée supra de l'accès à l'expertise en production d'accords à l'épreuve de l'examen du présent corpus, nous rapprochons deux types de données, celles obtenues de manière expérimentale avec celles collectées dans un contexte plus écologique de production, celui de la classe. L'autre intérêt d'étudier ce corpus est d'avoir accès à la production de textes, alors que les études expérimentales portaient principalement sur la production de syntagmes et de phrases isolées.

4. Méthode

Parmi l'ensemble des textes du corpus général, nous avons fait le choix de centrer notre étude des erreurs sur les seuls textes narratifs, en raison du fait qu'ils sont en moyenne plus longs que les comptes-rendus scientifiques. Nous avons là plus de chances d'observer la variété des erreurs qui nous intéressent. Nous avons aussi fait ce choix parce que la production de textes narratifs est, au début de la scolarité primaire, une activité d'écriture plus souvent utilisée que le compte-rendu scientifique. De plus, souhaitant mettre en évidence les automatismes de la production écrite des enfants, nous n'avons fait porter notre analyse que sur la première version du texte. Enfin, visant à caractériser deux profils classiquement évoqués dans les travaux expérimentaux sur l'accord en nombre, celui du novice et celui de l'expert, nous n'avons retenu que les deux groupes d'élèves situés aux deux extrémités du corpus, les élèves de troisième année primaire (CE2) et ceux en troisième année du collège (4^{ième}).

La comparaison du nombre d'erreurs observées au sein des deux populations nous paraît peu informative, déjà parce que la longueur moyenne des textes

produits varie significativement du CE2 (moyenne : 126 mots [écart-type 30]) à la 4^{ième} (moyenne : 197 mots [écart-type 57]). Mais plus que le nombre d'erreurs, c'est bien la nature des erreurs qui nous intéresse. Intéressés à décrire et comprendre la performance orthographique, nous ne souhaitons pas procéder à une comparaison directe entre ces deux groupes d'enfants (CE2 vs. 4^{ième}). Nous avons préféré d'abord recenser et discuter les différentes catégories d'erreurs qui nous intéressent chez les élèves de CE2, puis chez ceux de 4^{ième}, avant de comparer des profils orthographiques au sein de ces deux populations.

Nous nous sommes ainsi successivement intéressés à l'accord en nombre, pour observer la façon dont les enfants produisaient une morphologie silencieuse (i.e., -s du pluriel sur le nom, l'adjectif et le pronom ; -nt du pluriel sur le verbe), puis au phénomène de récupération d'instances dans les situations où un autre item linguistique était produit en lieu et place de l'item attendu. Ce phénomène étant un processus général de la production de l'écrit, nous avons tenté d'en trouver la trace dans le corpus en recherchant des substitutions de radicaux (mots pleins tels que les noms, les verbes, etc.), des substitutions de morphèmes libres (mots grammaticaux tels que les prépositions, les conjonctions, etc.) et des substitutions de morphèmes liés (cas des flexions accolées à un radical).

5. Résultats

5.1 CE2

5.1.1 Accord en nombre

En données cumulées sur les 23 textes, les CE2 produisent 79 erreurs d'accord en nombre. Cinq enfants sur 23 ne commettent aucune erreur de ce type dans leur texte. Chez les 18 autres enfants, les erreurs d'accord concernent majoritairement le nom (35 erreurs), puis le verbe (27 erreurs), l'adjectif (11 erreurs) et le pronom personnel *il* (6 erreurs).

Les erreurs d'accord du nom consistent presque exclusivement en une absence de marquage du morphème -s du pluriel (31 erreurs ; p.ex. *les musicien*). Deux erreurs consistent en la présence de -s au singulier (p.ex. *une tortues*) ; deux autres en la présence de -x au singulier (*un plateaux* ; *le monsieux*). Ces deux dernières erreurs semblent relever d'un phénomène de récupération d'instances, en ce sens que les graphèmes « eaux » et « eux » existent dans la langue.

Pour le verbe, c'est aussi l'absence de marquage du morphème du pluriel qui domine (21 erreurs ; p.ex. *deux musiciens se dispute* ; *les parents avait* ; *ils lise*). Six autres erreurs consistent en la présence de -nt sur un verbe au singulier (e.g., *le servan attrapent*), dont une qui laisse entrevoir la possible influence du niveau sémantique sur l'activation de la forme plurielle du verbe (p.ex. *une famille vont*). D'autres erreurs sont associées au non marquage du pluriel sur le pronom personnel *il* (*il arriver* ; *il commendere* ; *qu'il fur* ; *qu'il partir*). Ces dernières, observées chez deux enfants, sont particulièrement intéressantes puisque l'absence du pluriel sur le pronom est concomitante de l'absence du morphème pluriel -nt sur le verbe, alors que la forme phonologique du verbe au pluriel est produite. Un autre enfant ne réalise aucun accord en nombre tout

en respectant l'information phonologique (le serreueur a pris la grenouille par c'est patte...c'est parent était fâché).

Pour l'adjectif, comme pour le pronom *il*, l'erreur consiste systématiquement en une absence de la marque -s du pluriel.

Si l'absence de marque du pluriel est, en apparence au moins, aisément explicable soit par la non connaissance de la règle d'accord, soit par l'absence de sa mise en œuvre pour les raisons d'un déficit attentionnel, la présence de marques du pluriel sur des mots au singulier paraît plus surprenante. Le phénomène a toutefois été étudié et une des explications fournies tient à la récupération d'instances (Largy et al, 2007) : des items (radicaux, morphèmes et parfois l'association des deux) seraient directement récupérés en mémoire et produits sous la forme sous laquelle le scripteur les a le plus fréquemment rencontrés.

5.1.2 Récupération d'instances

Le phénomène mis en évidence expérimentalement dans le domaine de l'accord nominal (Cousin et al., 2002 ; Largy et al., 2007) et de l'accord verbal (Largy et al., 1996), semble être un processus général de la production écrite. En effet, nous avons relevé dans le présent corpus un ensemble de 61 erreurs qui, selon nous, relèvent du processus de récupération d'instances. La violation d'une catégorie grammaticale mise en évidence par Largy et al. (1996), où un nom pluriel est produit en lieu et place d'un verbe pluriel, est présente dans ce corpus où de nombreux items existant dans la langue surviennent en lieu et place des items attendus. Le corpus montre un ensemble de substitutions qui dépasse largement les seules catégories nominale et verbale.

En plus des substitutions nom/verbe, évidentes (*qui allée* [au lieu de *qui allait*] ; *qui cri* [au lieu de *qui crie*] ; *les musicien jouet* [au lieu de *les musiciens jouaient*]), ou moins évidentes (*un qui souris* [au lieu de *un qui sourit*]), le corpus révèle 57 autres erreurs pouvant être imputées au phénomène de récupération d'instances. Ainsi, par exemple, un adjectif, un pronom ou une conjonction peuvent survenir à la place d'un verbe (10 erreurs : *son* [au lieu de *sont*] ; 2 erreurs : *on* [au lieu de *ont*] ; 2 erreurs : *par* [au lieu de *part*] ; 2 erreurs : *et* [au lieu de *est*]).

Le phonème [s] est une source connue d'erreurs orthographiques par substitution. Dans le présent corpus, 15 erreurs de ce type sont observées sur des morphèmes libres (monèmes grammaticaux) (p.ex. *dans ça main* [au lieu de *dans sa main*] ; *cette fois si* [au lieu de *cette fois-ci*] ; en plus des substitutions *se/ce*). A ceci s'ajoute que la proximité phonologique entre le phonème [e] et le phonème [ɛ] favorise les erreurs de substitution. Dans le corpus, on observe la variété des erreurs suivantes (*mes* [au lieu de *mais*] ; *s'est* [au lieu de *ses* ou de *ces*] ; *ses* [au lieu de *c'est*] ; en plus des substitutions *et/est*). Ici encore, les morphèmes libres sont particulièrement concernés.

Certaines erreurs concernent la morphologie flexionnelle verbale (morphèmes liés). Le participe passé du verbe est produit en lieu et place du verbe au passé

simple à la troisième personne du singulier. On relève ainsi dans le corpus les 7 erreurs suivantes : *voulu* [3 fois au lieu de *voulut*] ; *choisi* [au lieu de *choisit*] ; *sorti* [au lieu de *sortit*] ; *parti* [au lieu de *partit*] ; *mis* [au lieu de *mit*].

S'il est clair que l'information phonologique guide le passage à l'écrit, comme le montrent certaines erreurs du corpus (p.ex. *il ça soit* [au lieu de *il s'assoit*] ; *ne veut pas con emmene* [au lieu de *ne veut pas qu'on emmène*]), cette information est une source potentielle d'erreurs dans une langue telle que le français. Nous avons vu la difficulté qu'ont des enfants du CE2 à gérer certains phonèmes ou des proximités de phonèmes. Nous avons également remarqué que les erreurs produites consistent souvent en des items existants de la langue. En plus des mots grammaticaux peu épargnés par ce type d'erreurs (morphèmes libres tels que les pronoms personnels, les pronoms possessifs, les conjonctions de coordination ou de subordination, etc.), les terminaisons verbales (morphèmes liés) sont elles aussi souvent erronées. Si la terminaison verbale se prête à l'oubli chez l'apprenant, comme nous l'avons vu dans le cas du morphème flexionnel *-nt* d'un pluriel inaudible, elle se prête aussi à la substitution quand des morphèmes flexionnels verbaux se prononcent de la même façon (*-é* vs. *-er*) ou de façon très proche (*-é*, *-er* vs. *-ait*, *-aient*). Nous nous proposons de consacrer la partie suivante au recensement des erreurs de substitution de ce type.

5.1.3 Les terminaisons verbales en [e]/[ɛ]

Dans la narration produite par les élèves du CE2, nous avons dénombré 71 erreurs de terminaisons verbales impliquant les phonèmes [e] et [ɛ]. Elles sont le fait de 21 des 23 enfants de l'étude. Treize enfants les commettent occasionnellement (une à deux fois dans leur texte) et 8 les commettent plus régulièrement (cinq à dix fois). La plupart de ces erreurs renvoient à un cas typique de la récupération d'instances impliquant les terminaisons verbales de l'infinitif (en *-er*), du participe passé (en *-é*) et de l'imparfait (en *-ait*).

Trente neuf des 71 erreurs de substitution consistent en la production de l'infinitif, en lieu et place du participe passé (dans 22 cas : p.ex. *la grenouille a sauter* [au lieu de *la grenouille a sauté*]) et de l'imparfait (dans 17 cas : p.ex., *il arriver* [au lieu de *il arrivait*]). Neuf erreurs consistent en la production du morphème *-er* sur un adjectif en *-é* (p.ex. *une dame afole* [au lieu de *une dame affolée*] ; *le garçon est étonner* [au lieu de *le garçon est étonné*] ; *très fachers* [au lieu de *très fâchés*]). Ainsi, chez les CE2, l'erreur dominante (48 des 71 erreurs) consiste en la production du morphème *-er*, ce qui confirme les résultats de Brissaud (2008).

Quatorze erreurs consistent à produire la flexion *-é*, en lieu et place de l'imparfait (6 erreurs : p.ex. *qui allé* [au lieu de *qui allait*]), de l'infinitif (6 erreurs : p.ex. *pour allé* [au lieu de *pour aller*]) et de la terminaison de la deuxième personne du pluriel *-ez* (2 erreurs : *que désiré vous* [au lieu de *que désirez-vous*]).

Cinq erreurs consistent à produire la flexion *-ait*, en lieu et place de l'infinitif (3 erreurs : p.ex. *il devait rentrais* [au lieu de *il devait rentrer*]) et du participe passé (2 erreurs : p.ex. *ça a était dur* [au lieu de *ça a été dur*]).

Deux erreurs -ez sur un infinitif (p. ex. , *pour allez* [au lieu de *pour aller*]) et deux autres erreurs qui tendent à confirmer le phénomène de récupération d'instances, ici l'instance récupérée est de la taille du mot (*le musicien jouet* [au lieu de *le musicien jouait*] ; *qui allée* [au lieu de *qui allait*]).

5.2 4^{ième}

5.2.1 Accord en nombre

En données cumulées sur les 25 textes, les 4^{ième} produisent 39 erreurs d'accord en nombre. Neuf enfants sur 25 ne commettent aucune erreur de ce type dans leur texte. Chez les 16 autres enfants, les erreurs d'accord concernent majoritairement le nom (19 erreurs), puis le verbe (16 erreurs), l'adjectif (3 erreurs) et le pronom personnel *il* (1 erreur).

Les erreurs d'accord du nom consistent presque exclusivement en une absence de marquage du morphème -s du pluriel (15 erreurs ; p.ex. *ces instrument*). Quatre erreurs consistent en la présence de -s au singulier (p. ex. *l'enfants*). Pour le verbe, c'est aussi l'absence de marquage du morphème du pluriel qui domine (14 erreurs. Cependant, dans 8 des 14 erreurs, l'absence de marque renvoie typiquement au phénomène d'accord par proximité, signe de l'installation d'automatismes attribués à l'expertise (p.ex. *la salle où mange les gens ; les parents lui en donne ; les parents, la fille et le petit garçon rentre ; la mère, le père et sa sœur lui faisait*). Seules deux erreurs consistent en la présence de -nt sur un verbe au singulier (*il enlèvent ; la famille du jeune garçon étaient*), dont une, la seconde, qui laisse entrevoir la possible influence du niveau sémantique sur l'activation de la forme plurielle du verbe. Pour l'adjectif (3 erreurs : p.ex. *les musiciens mécontent*), comme pour le pronom *il* , l'erreur consiste systématiquement en une absence de la marque -s du pluriel. Si l'on assouplit le critère de maîtrise de l'accord en nombre en acceptant une erreur d'accord sur l'ensemble du texte, c'est 15 des 25 élèves de 4^{ième} qui ont atteint une forme d'expertise en production de l'accord en nombre. Le problème de gestion de ces accords simples demeure néanmoins pour 10 d'entre eux.

Parmi les erreurs d'accord qui persistent en 4^{ième}, certaines semblent tenir à l'orthographe même du radical. En effet, les noms et adjectifs qui se terminent par les lettres /ent/ apparaissent particulièrement touchés par l'absence de marquage du pluriel en -s. Ainsi, dans le corpus des 4^{ième}, les noms *parent* et *instrument* comme les adjectifs *content* et *mécontent* sont à plusieurs reprises transcrits sans -s au pluriel (p.ex. *avec ses parent ; les parent ; ces instrument ; les instrument ; ses parent n'était guerre content ; les musiciens mécontent*). On sent là tout le poids de la morphologie dans la production automatisée de l'orthographe et les compétitions qui s'exercent entre les différentes graphies possibles. Ces erreurs illustrent à leur façon le phénomène plus général de la récupération d'instances.

5.2.2 Récupération d'instances

Nous avons relevé dans le présent corpus un ensemble de 42 erreurs qui, selon nous, relèvent du processus de récupération d'instances. Ici encore, le

corpus montre un ensemble de substitutions qui dépasse largement les seules catégories nominale et verbale. On observe des substitutions nom/verbe (*le client vie* [au lieu de *le client vit*]), nom/adverbe (*guerre* [au lieu de *guère*]), nom/nom (*les vers* [au lieu de *les verres*]) mais aussi des situations où un adjectif ou un pronom peuvent survenir à la place d'un verbe (*son* [au lieu de *sont*] ; *on* [au lieu de *ont*]). La conjonction de coordination *et* est à 5 reprises écrite *est* alors que l'erreur inverse n'existe pas dans le corpus.

En 4^{ième}, le phonème [s] demeure une source d'erreurs orthographiques par substitution. Huit erreurs de ce type sont observées sur des morphèmes libres (monèmes grammaticaux) (3 erreurs : *se* [au lieu de *ce*] ; 3 erreurs : *ce* [au lieu de *se*] ; 2 erreurs : *sa* [au lieu de *ça*]). A ceci s'ajoute que la proximité phonologique entre le phonème [e] et le phonème [ɛ] favorise certaines erreurs de substitution. Dans le corpus, on relève les 3 autres erreurs suivantes (2 : *c'est* [au lieu de *ses*] ; *c'était* [au lieu de *s'était*]).

Certaines erreurs concernent la morphologie flexionnelle verbale (morphèmes liés). Le participe passé du verbe est produit en lieu et place du verbe au passé simple conjugué à la troisième personne du singulier. On relève ainsi dans le corpus les 13 erreurs suivantes : *pris* [6 fois au lieu de *prit*] ; *eu* [2 fois au lieu de *eut*] ; *choisi* [au lieu de *choisit*] ; *sorti* [au lieu de *sortit*] ; *exclu* [au lieu de *exclut*] ; *mis* [au lieu de *mit*]. Ces 13 erreurs se caractérisent par le fait que les 2 items substitués se prononcent de la même façon. Une autre erreur consistant à produire un participe passé en lieu et place du verbe au passé simple, est caractérisée par l'absence de ressemblance phonologique entre les 2 items substitués (*vu* [3 fois au lieu de *vit*]).

Toutes les substitutions observées dans les textes des élèves montrent à quel point l'information phonologique guide la production écrite. Toutefois, certaines substitutions passent outre la ressemblance phonologique entre ce qui est produit et ce qui est attendu. C'est le cas de *vu* produit à la place de *vit* . C'est aussi le cas d'erreurs typiques des élèves de 4^{ième} (typiques, au sens où nous n'en avons pas observé dans les textes des enfants de CE2). Ces erreurs tiennent, au passé simple, à un phénomène de surgénéralisation : les verbes des 2^{ième} et 3^{ième} groupes sont conjugués comme les verbes du type le plus fréquent dans notre langue : les verbes du 1^{er} groupe (-ai ; -as ; -a ; -âmes ; -âtes ; -èrent). Onze erreurs de ce type ont été relevées dans le corpus (*le serveur prena* ; *le client prena* ; *la grenouille fesa* ; *la grenouille atterrissa* ; *il coura* ; *le garçon lui coura après* ; *le petit garçon ce renda compte* ; *qui disat* ; *je me levai et courai* ; *nous devâmes* ; *ils s'apercevèrent*). Nous n'avons pas comptabilisé ces onze erreurs parmi les erreurs de récupération d'instances puisque les 2 items substitués, l'item produit à l'écrit (p.ex. *prena*) et l'item correct attendu (p.ex. *prit*), ne présentent pas la même forme phonologique.

5.2.3 Les terminaisons verbales en [e] et [ɛ]

Nous avons dénombré 37 erreurs de terminaisons verbales impliquant les phonèmes [e] et [ɛ]. Elles sont le fait de 14 des 25 élèves de 4^{ième}. Neuf d'entre eux les commettent occasionnellement (une à deux fois dans leur texte) et 5 les commettent

plus régulièrement (trois à cinq fois). La plupart de ces erreurs renvoient à un cas typique de la récupération d'instances impliquant les terminaisons verbales de l'infinitif (en *-er*), du participe passé (en *-é*) et de l'imparfait (en *-ait*).

Seize des 37 erreurs de substitution consistent en la production de l'infinitif, en lieu et place du participe passé (dans 7 cas : p.ex. *leur avait conseiller* [au lieu de *leur avait conseillé*]) et de l'imparfait (dans 5 cas : p.ex. *un plateau qui aller* [au lieu de *un plateau qui allait*]). Trois erreurs consistent en la production du morphème *-er* sur un adjectif en *-é* (p.ex. *sa grenouille toute étoner* [au lieu de *sa grenouille toute étonnée*]). Une erreur voit survenir *-er* à la place de *-ez* (*ne la jeter pas* [au lieu de *ne la jetez pas*]).

Onze erreurs consistent à produire la flexion *-é*, en lieu et place de l'imparfait (7 erreurs : p.ex. *où été disposée* [au lieu de *où était disposée*]), de l'infinitif (4 erreurs : p.ex. *pour la récupéré* [au lieu de *pour la récupérer*]), dont une qui tutoie la forme nominale : *de la jetée* [au lieu de *de la jeter*]).

Huit erreurs consistent à produire la flexion *-ait*, en lieu et place de l'infinitif (3 erreurs : p.ex. *le petit garçon se fait disputait* [au lieu de *le petit garçon se fait disputer*]) et du participe passé (3 erreurs : p.ex., *il avait commandait* [au lieu de *il avait commandé*]). A deux reprises, *-ait* apparaît en lieu et place du verbe être (i.e., où il ait [au lieu de « où il est »]; se naît pas [au lieu de *ce n'est pas*]). L'accent circonflexe sur le *i* de *naît* plaide grandement en faveur d'une récupération d'instance.

Enfin, on observe une erreur en *-ez* sur un infinitif (celui-ci ne voulait pas écoutez [au lieu de *celui-ci ne voulait pas écouter*]) et une autre sur un imparfait (*ses parents nous amenez* [au lieu de *ses parents nous amenaient*]).

6. Discussion

Les travaux expérimentaux conduits sur l'accord en nombre ont permis de modéliser l'accès à l'expertise en production écrite des marques du pluriel (Fayol et al., 1999 ; Largy, Cousin, & Dédéyan, 2005 pour une revue). Ils ont dégagé des étapes au cours de cet apprentissage en se fondant sur l'existence de catégories d'erreurs typiques de ces étapes, de l'erreur d'absence de marquage du pluriel (typique de l'entrée dans l'écrit) à la présence d'erreurs d'accords par proximité (typiques de l'expertise), en passant par une période au cours de laquelle l'accord est parfois réussi et parfois échoué. Les auteurs ont décrit cette période intermédiaire comme une étape au cours de laquelle l'accord était géré par application d'une règle apprise. Ils ont montré que l'accord était réussi lorsque l'apprenant disposait des ressources attentionnelles nécessaires à l'application consciente de la règle. On expliquait ainsi les variations de la performance orthographique par l'attention que l'enfant avait pu porter à l'accord au moment où il avait écrit. Or, Cousin (2004) a montré que la production des marques du nombre n'était pas qu'une question de règle ou d'attention à consacrer à son application. Elle a montré que, chez le même enfant, tous les noms n'étaient pas aussi bien ou aussi mal accordés en nombre. Un contrôle rigoureux de la fréquence de rencontre par les enfants de noms au singulier ou fléchis au pluriel lui a permis de montrer expérimentalement que les enfants tendent à rappeler à l'écrit la forme sous laquelle ils ont le plus fréquemment rencontré le nom. Le phénomène a été qualifié de « récupération

d'instances en mémoire » par Cousin et al. (2002) en référence à des erreurs déjà observées dans le cas de l'accord verbal (*Le facteur prend des enveloppes. Il les timbres*) (Largy et al., 1996).

L'analyse du présent corpus a permis de retrouver les différentes erreurs caractéristiques de l'accès à l'expertise en production de l'accord en nombre. Elle confirme l'existence du phénomène de récupération d'instances, non seulement dans le cadre de la production de la morphologie flexionnelle du nombre mais aussi à travers la mise en évidence d'autres cas de substitutions opérées sur la base d'une proximité phonologique, faisant de ce phénomène de récupération mémorielle un processus général de la production écrite.

Les résultats permettent de retrouver les profils caractéristiques de l'accès à l'expertise en production de la morphologie flexionnelle du nombre, même si 5 des 23 élèves de CE2 ne commettent déjà plus d'erreurs d'accord en nombre dans leur texte et si 16 des 25 élèves de 4^{ième} en commettent encore. Aux deux niveaux scolaires, l'erreur d'accord dominante demeure l'absence de marquage du pluriel, pour le nom et pour le verbe. Cependant, en 4^{ième}, l'erreur d'absence de -nt sur le verbe ressemble le plus souvent (8 cas sur 14 erreurs de ce type) à une erreur d'accord par proximité (e.g., la mère, le père et sa sœur lui faisait). Une autre catégorie d'erreurs d'accord par proximité aurait sans doute été plus convaincante (e.g., le père des enfants arrivent) or ce type d'erreurs n'est pas apparu dans le corpus. Nous n'en concluons pas que de telles erreurs n'existent pas dans une production libre (car nous les observons dans nos propres textes et dans les copies de nos étudiants) mais plutôt que les images retenues (i.e., l'histoire de la grenouille) étaient peu propices à induire des syntagmes sujets du type « Nom1 de Nom2 » ou « Pronom1 de Pronom2 ». De plus, l'étude off-line des textes produits ne permet pas de discuter du degré de contrôle orthographique opéré par les scripteurs et par conséquent de clairement différencier ici les erreurs qui relèvent d'automatismes de production de celles qui relèvent d'une méconnaissance de la règle. On perçoit clairement dans ce cas l'intérêt d'une étude expérimentale pour compléter l'analyse de corpus.

A ce non marquage de l'accord s'ajoutent quelques erreurs où la flexion qui apparaît n'est pas celle orthographiquement attendue (e.g., en CE2 : un plateaux ; en 4^{ième} : l'enfants). Ces erreurs d'accord sont assez typiques du phénomène de récupération d'instances mais, ici encore, le corpus textuel ne favorise pas la production d'homophones nom-verbe tels que par exemple : montre-montrer, garde-garder ou griffe-griffer. Nous ne pouvons donc pas aisément observer d'erreurs du type « il vol des montrent », « il les gardent » ou « le chat les griffes ». Par contre, le corpus offre une variété d'erreurs qui dépasse le seul cas de la morphologie silencieuse du pluriel et qui révèle, en plus des substitutions nom-verbe, des substitutions touchant à d'autres catégories grammaticales. C'est le cas de plusieurs autres situations d'homophonie, avec des erreurs concernant les phonèmes [s], [e] et [ɛ]. On observe, en CE2 mais aussi en 4^{ième}, des erreurs de substitution d'adjectifs, de pronoms ou de conjonctions à des verbes (p.ex. *son/sont, on/ont, et/est*), des substitutions de verbes à des items d'autres catégories grammaticales (p.ex. *c'est/ses, c'est/ces*) mais aussi des substitutions qui n'impliquent pas le verbe (p. ex. *mes/mais, ses/ces, sa/*

ça). Il ressort de ces observations que la production écrite fait davantage l'objet d'un guidage phonologique que syntaxique puisque, dans le cas de ces erreurs, l'information phonologique est préservée mais pas la catégorie syntaxique de l'item ; le phénomène n'est pas spécifique aux élèves les plus jeunes puisque, dans notre corpus, ces erreurs existent en CE2 mais aussi en 4^{ième}. Tout se passe comme si nous étions en présence d'un phénomène général de l'activité de production écrite.

Comme l'a montré Cousin (2004), la récupération d'instances est sensible à la fréquence de rencontre des items ; dans des situations concurrentielles ou plusieurs items présentent le même contour phonologique, l'item produit a de fortes chances d'être l'item le plus fréquemment rencontré à l'écrit. Il est généralement impossible d'évaluer la fréquence de rencontre des mots par les enfants ; on tente alors d'approcher ces fréquences en faisant usage de bases de données lexicales pour l'école primaire (p.ex. Novlex : Lambert & Chesnet, 2001 ; Manulex : Lété, Sprenger-Charolles, & Colé, 2004). A moins de faire ce qu'a fait Cousin (2004) avec une grande précision, en rendant expérimentalement fréquents ou peu fréquents des noms au départ très rares (p.ex. *bonite, drave, cupule, colobe*), on doit s'appuyer sur des phénomènes généraux de l'acquisition de l'écrit pour distinguer ce qui est fréquent de ce qui ne l'est pas. Nous pouvons le faire dans ce corpus à travers le cas de l'apprentissage du passé simple. En effet, l'enfant de CE2 ne s'est pas encore vu enseigner le passé simple, contrairement à l'élève de 4^{ième}, mais le passé simple demeure un temps peu utilisé ; les verbes au passé simple devraient donc être davantage connus des élèves de 4^{ième} que des élèves de CE2, sans toutefois que les 4^{ième} aient automatisé leur production.

Le présent corpus montre qu'en CE2 des erreurs sur le passé simple prennent la forme du participe passé du verbe (e.g., voulu/voulut ; choisi/choisit) alors que l'inverse n'est jamais observé : aucune erreur ne consiste à produire une forme du passé simple en lieu et place d'un autre item. Il ne s'agit pas non plus d'un simple oubli du -t final puisque qu'on observe également l'erreur mis/mit. On peut donc en conclure que la forme du participe passé du verbe, homophone du verbe au passé simple, apparaît en lieu et place du verbe au passé simple parce qu'elle a été rencontrée plus fréquemment par ces enfants. En CE2, pour la raison de l'ambiguïté phonologique des terminaisons en [e]/[ɛ], on retrouve un phénomène analogue avec des erreurs qui consistent à produire l'infinitif du verbe en lieu et place d'autres formes moins fréquentes à cet âge (p.ex. participe passé : la grenouille a sauter ; imparfait : il arriver). Les 48 erreurs de ce type distancent de loin les 11 erreurs consistant à produire d'autres formes à la place de l'infinitif.

En 4^{ième}, les erreurs consistant à produire l'infinitif du verbe, si leur nombre a bien diminué (16 erreurs seulement et dans des textes en moyenne plus longs qu'en CE2), sont toujours présentes. Leur proportion s'est toutefois affaiblie au regard des autres erreurs de terminaisons en [e]/[ɛ] (16 erreurs contre 7). Cette observation est compatible avec le fait que la fréquence de rencontre des formes du participe passé et de l'imparfait a probablement augmenté chez ces élèves plus âgés. En 4^{ième}, on retrouve l'erreur de substitution participe passé/

passé simple. On dénombre 13 erreurs de ce type et l'on obtient confirmation du phénomène de récupération d'instances par la présence d'erreurs pris/prit (6) et mis/mit (1). L'erreur eu/eut tend à annuler la critique selon laquelle « pris » ne serait pas le participe passé du verbe prendre mais sa conjugaison au passé simple à la première ou deuxième personne du singulier. Un phénomène nouveau apparaît en 4^{ième} : certaines erreurs au passé simple prennent la forme de régularisation des verbes des 2^{ième} et 3^{ième} groupes sur la base de l'information phonologique de la conjugaison des verbes du type le plus fréquent dans notre langue, les verbes du 1^{er} groupe (p.ex. le serveur prena ; nous devâmes ; ils s'apercevèrent). Ces erreurs de surgénéralisation surviennent classiquement au cours de l'apprentissage des langues (Bybee, 2001), langues maternelles et langues étrangères (Gunnarsson & Soum, 2009). Les nouveaux verbes qui intègrent notre langue sont aussitôt régularisés, implicitement, sur la base des verbes du 1^{er} groupe ; nous comprenons aisément la phrase « Wilander a réussi à breaker très tôt dans le premier set ».

Les erreurs de surgénéralisation observées dans le cas du passé simple renforcent la conception d'une production écrite sous-tendue par des automatismes. Les effets de fréquence sont bien présents car l'erreur de surgénéralisation ne touche jamais les verbes « être » et « avoir » : dans le corpus, nous n'observons aucune erreur du type « il éta » ou « il ava ». Ces erreurs sont riches d'informations et tendent à montrer que la récupération d'instances ne porte pas uniquement sur des mots entiers (e.g., guerre/guère ; qui allée/allait ; les musicien jouet/jouaient) mais aussi parfois sur des morphèmes flexionnels. En effet, dans les erreurs du type « nous devâmes » ou « ils s'apercevèrent », ces items n'ont probablement jamais été rencontrés par les élèves sous cette forme à l'écrit et pourtant la flexion verbale est correctement orthographiée.

On voit, à travers l'étude de la production de la morphologie flexionnelle et plus généralement à travers celle de la production verbale écrite, que l'approche expérimentale et l'analyse de corpus s'éclairent mutuellement. Pour continuer d'avancer vers la compréhension des processus mobilisés au cours de l'apprentissage de l'écrit, il convient désormais de croiser les regards de psychologues du développement cognitif, de linguistes et de didacticiens sur les mêmes données. C'est toute l'originalité du présent ouvrage.

Références

- Anderson, R. C. 1992. « Automaticity and the ACT*theory ». *American Journal of Psychology*, n° 105, pp. 165-180.
- Brissaud, C. 2008. *Acquisition didactique de l'orthographe*, Grenoble : Thèse HDR, Université Stendhal-Grenoble 3.
- Bybee, J. L. 2001. *Frequency and the emergence of linguistic structure*. Amsterdam : Benjamins.

- Chesnet, D., Alamargot, D. 2005. « Analyses en temps réel des activités oculaires et graphomotrices du scripteur: intérêt du dispositif 'Eye and Pen' ». *L'Année Psychologique*, n° 105(3), pp. 477-520.
- Cousin, M-P. 2004. «Apprentissage de la production écrite de l'accord en nombre : application de règles et/ou récupération d'instances ?». Thèse de doctorat : Université de Rouen.
- Cousin, M-P., Largy, P., Fayol, M. 2002. « Sometimes early learned instances interfere with the implementation of rules : The case of nominal number agreement ». *Current Psychology Letters, Behaviour, Brain & Cognition*, n° 8, pp. 51-65.
- Fayol, M., Got, C. 1991. «Automatisme et contrôle dans la production écrite ». *L'Année Psychologique*, 91, 187-205.
- Fayol, M., Hupet, M., Largy, P. 1999. « The acquisition of subject-verb agreement in written French: From novices to experts ». *Reading and Writing : An interdisciplinary Journal*, n° 11, pp. 153-174.
- Gunnarsson, C. 2006. *Fluidité, complexité et morphosyntaxe dans la production écrite en FLE*, Lund : Thèse de doctorat: Études romanes de Lund 78, Université de Lund.
- Gunnarsson, C., Soum Favaro, Ch. 2009. « Liaison et production écrite : l'écrit comme moyen d'étudier un phénomène de l'oral ». *Actes du colloque «de la France au Québec : l'écriture dans tous ces états»*, Poitiers, 12-15 novembre 2008. <http://www.poitou-charentes.iufm.fr/IMG/pdf/GunnarssonSoum.pdf>
- Lambert, E., Chesnet, D. 2001. « Novlex : une base de données lexicales pour les élèves de primaire ». *L'Année Psychologique*, 101, 561-592.
- Largy, P., Cousin, M-P., Bryant, P., Fayol, M. 2007. « When memorized instances compete with rules : The case of number-noun agreement in written French ». *Journal of Child Language*, n° 34, pp. 425-437.
- Largy, P., Cousin, M.P., Dédéyan, A. 2005. « Produire et réviser la morphologie flexionnelle du nombre : de l'accès à une expertise ». *Psychologie Française*, 50, 339-350.
- Largy, P., Fayol, M., Lemaire, P. 1996. « The homophone effect in written French: The case of verb-noun inflection errors ». *Language and Cognitive Processes*, n° 11, pp. 217-255.
- Lété, B., Sprenger-Charolles, L., Colé, P. 2004. « Manulex : a grade-level lexical database from French elementary-school readers ». *Behavior Research Methods, Instruments & Computers*, 26, 156-166.
- Logan, G. D. 1991. Automaticity and Memory. In : W. E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: Essays in Honor of Bennett Murdock*, Hillsdale : L. E. A.
- Marcus, G.F., Brinkmann, U., Clahsen, H., Wiese, R., Pinker, S. 1995. « German inflection: The exception that proves the rule ». *Cognitive Psychology*, n° 29, pp. 198-256.
- Plunkett, K., Marchman, V. 1993. « From rote learning to system building: Acquiring verb morphology in children and connectionist nets ». *Cognition*, n° 48, pp. 21-69.
- Strömquist, S. Karlsson, H. 2002. *ScriptLog - a User's Manual: Technical report*, Lund : Department of Linguistics, University of Lund, and Centre for Reading Research, University College of Stavanger.

Totereau, C., Barrouillet, P., Fayol, M. 1998. « Overgeneralizations of number inflections in the learning of written French: The case of noun and verb. *British Journal of Developmental Psychology*, n° 16, pp. 447-464.

Totereau, C., Thevenin, M-G., Fayol, M. 1997. The development of the understanding of number morphology in French. In : C. Perfetti et al. (Eds.), *Learning to spell*, Hillsdale : L.E.A., pp. 97-114.