

# Apprentissage incident de vocabulaire et de l'orthographe en lecture de texte

Valentine Beck, Université de Strasbourg

Eva Commissaire, Université de Strasbourg, Laboratoire de Psychologie des Cognitions (UR 4440)

**Résumé.** Des recherches ont montré qu'il est possible d'apprendre l'orthographe de nouveaux mots, et d'inférer leur sens, de manière incidente en lecture silencieuse, en langue maternelle et seconde (L2). Dans cette étude, les participants devaient lire un texte en français ou anglais (L2), dans lequel étaient insérés des pseudo-mots avec une orthographe typique de la L2 (*vight*) ou partagée entre les langues (*flace*). Les résultats relatifs à la qualité de l'apprentissage de la forme et du sens de ces mots seront discutés au regard des modèles en psycholinguistique.

**Abstract.** *Incidental learning of vocabulary and spelling through text reading.* Research has shown that it is possible to learn the spelling of new words, and to infer their meaning, incidentally through silent reading, in native and second language (L2). In this study, participants were asked to read a text in French or English (L2), in which were inserted pseudowords with an L2-marked orthography (*vight*) or shared between languages (*flace*). The results concerning the quality of learning of the form and meaning of these words will be discussed with respect to psycholinguistic models.

L'enseignement de l'orthographe lexicale fait partie des missions du corps enseignant, désormais encouragé à proposer des séquences d'apprentissage progressives et explicites. Outre ce mode d'apprentissage, il est cependant entendu que des connaissances sont également acquises de manière plus incidente, notamment lors de situations de lecture naturelle de textes auxquelles sont confrontés les élèves pendant leur scolarité. Nous nous intéresserons dans cette étude aux différentes acquisitions, de sens et formelles, lors de la lecture de mots nouveaux insérés dans un texte, et ce en langue première (L1) et seconde (L2).

Le concept d'apprentissage incident fait référence à une manière d'apprendre non intentionnelle, non planifiée, pouvant avoir lieu lors d'une activité dans laquelle est engagé l'individu, telle que la lecture (Ellis, 1994). En effet, si les élèves peuvent être amenés à lire un texte pour des raisons diverses, telles que renforcer la fluidité en lecture, mémoriser le contenu, ou par simple plaisir, cette activité fournit également des opportunités pour apprendre du vocabulaire nouveau (Laufer, 2003 ; Webb et Nation, 2017) et la forme orthographique de ces mêmes mots (Share, 1995). En effet, les lecteurs génèrent des inférences lorsqu'ils lisent des textes et extraient ainsi des informations, certes partielles, relatives au sens des mots nouveaux rencontrés dans ces textes. En outre, les travaux de Share (1995) ont mis en évidence que quelques expositions à un nouveau mot lors de la lecture permettaient déjà d'en extraire des connaissances orthographiques partielles. Les élèves confrontés à des pseudo-mots insérés dans un texte pouvaient, quelques jours plus tard, reconnaître la bonne orthographe de ces items parmi plusieurs options, dont certaines de même prononciation, et produire en situation de dictée une orthographe assez proche de l'item cible (quoique comprenant encore des erreurs, soulignant bien sûr la complexité que représente l'apprentissage de l'orthographe lexicale pour un jeune élève).

Ces mécanismes d'apprentissage incident ont lieu en situation de lecture en langue maternelle (L1) mais également en langue seconde (L2, Waring et Nation, 2004 ; Webb, 2007), ce terme pouvant renvoyer à une multitude de situations de bilinguisme et d'apprentissage de langues étrangères. En L2, des études récentes ont notamment employé l'eye-tracking permettant l'enregistrement de données oculométriques lors de la lecture, renseignant ainsi sur les processus de familiarisation puis d'intégration des mots nouveaux rencontrés pendant la lecture elle-même. Ainsi, plus l'individu est exposé à un mot nouveau, plus son temps de fixation diminue ainsi que le nombre de régressions sur ce mot, jusqu'à atteindre des niveaux comparables à des mots connus (e.g., Elgort et al., 2018 ; Godfroid et al., 2018).

Jusqu'à présent, ces travaux ont principalement considéré d'une part l'impact du nombre d'expositions aux mots et, d'autre part, celui des propriétés sémantiques et contextuelles du texte sur l'apprentissage des nouveaux mots cibles et sur les patterns oculomoteurs associés. À notre connaissance, aucune étude ne s'est centrée sur les propriétés formelles des mots à apprendre, notamment une dimension orthographique dénommée depuis peu « marquage orthographique » (Casaponsa, Carreiras et Duñabeitia, 2014 ; Commissaire, Audusseau et Casalis, 2019). Les mots « marqués » d'une langue comprennent des patterns orthographiques (lettres, bi/trigrammes) typiques d'une langue tandis que les mots « non marqués » ont des patterns orthographiques plutôt plausibles dans les deux langues du bilingue. Ainsi, le mot anglais « *night* » peut être considéré comme marqué de l'anglais (-ght) tandis que le mot « *house* » est non marqué. Si des études ont montré que les individus bilingues jugeaient plus rapidement les mots marqués que ceux non marqués dans des tâches de décision de langue ou lexicale, seule une étude à notre connaissance a porté sur l'apprentissage de ces mots en L2 et montré un avantage des mots non marqués (Bartolotti et Marian, 2017).

L'objectif initial de notre étude était d'examiner l'influence du marquage orthographique sur l'apprentissage de mots en intégrant dans des textes des pseudo-mots marqués de l'anglais versus non marqués. Notre intérêt se centrait sur 1° l'apprentissage de leur forme orthographique et 2° la capacité à inférer leur sens, ces deux dimensions étant mesurées par des épreuves proposées après la lecture, et sur 3° les patterns oculomoteurs observés sur ces mots au cours de la lecture. L'objectif était également d'étudier l'impact de la langue du texte – et du coût cognitif associé – sur cet apprentissage en comparant les traitements sur ces items cibles insérés dans un texte en L1 (i.e., français) ou sa traduction en L2 (i.e., anglais). Compte tenu des difficultés depuis la crise sanitaire, seule l'étude pilote pourra être présentée, dont l'objectif était de vérifier le matériel de la phase de lecture et la validité des mesures d'apprentissage orthographique et sémantique.

Ainsi, si l'étude principale comprendra des adultes francophones de niveau avancé en langue anglaise, l'étude pilote ne prend en compte que le niveau de français, la L1 des participants. Ceux-ci participent à deux phases expérimentales, i.e., la lecture de texte suivie de la mesure de l'apprentissage.

**Lecture de texte.** Les textes utilisés proviennent des chapitres 1 à 4 du livre « La Guerre des Clans », traduction française de « Into the Wild » de E. Hunter. Ce texte, conseillé à partir de 9 ans, était choisi pour permettre une compréhension aisée en français et anglais, une condition *sine qua non* pour qu'un apprentissage de vocabulaire puisse avoir lieu (Nation, 2006), mais également en vue d'étendre à plus long terme cette étude à de plus jeunes lecteurs. Un total de 15 paires de pseudo-mots homophones étaient construits, dont la moitié était considérée comme marquée (e.g., *jowth*, *joath*) et l'autre comme non marquée (e.g., *blince*, *blinse*). L'un des items de la paire était présenté dans une version du texte, l'autre dans la seconde, et les versions étaient contrebalancées pour les sujets. Les items remplaçaient des mots du texte et présentaient un nombre de répétitions allant de 3 à 22, cette variable pouvant ainsi également être considérée dans les analyses.

**Mesure de l'apprentissage.** Compte tenu de la nature pilote de cette étude et de notre volonté de mesurer diverses formes d'apprentissages, dans leurs modalités perceptive et productive, un ensemble de 6 tâches était proposé aux participants. Une tâche de **rappel libre** permettait de déterminer le nombre et type de pseudo-mots spontanément rappelés par les participants. Elle était suivie d'une **dictée** de ces pseudo-mots puis d'un **choix orthographique**, épreuve dans laquelle le participant devait retrouver l'item cible parmi 4 options dont son pseudohomophone et deux distracteurs visuels. Puis, deux tâches centrées sur le sens des mots étaient proposées : une tâche de **définition** dont la consigne était de définir le pseudo-mot et proposer une traduction en langue française, suivie d'une tâche de **choix sémantique** dans laquelle le participant devait retrouver la bonne traduction en français parmi 5 options plus ou moins proches au niveau sémantique et grammatical. Une tâche de **dénomination orale** était enfin présentée et incluait les items cibles et leurs pseudohomophones. Notons que des questions de **compréhension** étaient également proposées, permettant de vérifier la bonne lecture du texte par le participant.

Seuls des résultats préliminaires pourront être présentés. Les aspects méthodologiques relatifs aux tâches utilisées et plus théoriques relatifs au concept d'apprentissage incident seront discutés. Nous pourrions ainsi nous interroger sur la place que pourrait avoir cette modalité d'apprentissage dans les séquences pédagogiques proposées en primaire et secondaire et/ou dans le cadre du travail extrascolaire, et la manière dont elle pourrait être associée à un enseignement plus explicite du vocabulaire et de l'orthographe lexicale, dont les bénéfices majeurs paraissent par ailleurs clairement avérés (e.g., Kemper et al., 2012).

## Références bibliographiques

- Bartolotti, J. et Marian, V. (2017). Orthographic knowledge and lexical form influence vocabulary learning. *Applied psycholinguistics*, 38(2), 427-456.
- Casaponsa, A., Carreiras, M. et Duñabeitia, J. A. (2014). Discriminating languages in bilingual contexts: the impact of orthographic markedness. *Frontiers in psychology*, 5, 424.
- Commissaire, E., Audusseau, J. et Casalis, S. (2019). Disentangling cross-language orthographic neighborhood from markedness effects in L2 visual word recognition. *Psychonomic bulletin & review*, 26(1), 353-359.
- Elgort, I., Brysbaert, M., Stevens, M. et Van Assche, E. (2018). Contextual word learning during reading in a second language: An eye-movement study. *Studies in Second Language Acquisition*, 40(2), 341-366.
- Ellis, N. C. (1994). Implicit and explicit processes in language acquisition: An introduction. Dans N. C. Ellis (dir.), *Implicit and explicit learning of languages* (p. 1-32). San Diego, CA: Academic Press.
- Godfroid, A., Ahn, J., Choi, I., Ballard, L., Cui, Y., Johnston, S., ... et Yoon, H. J. (2018). Incidental vocabulary learning in a natural reading context: An eye-tracking study. *Bilingualism: Language and Cognition*, 21(3), 563-584.
- Kemper, M. J., Verhoeven, L. et Bosman, A. M. (2012). Implicit and explicit instruction of spelling rules. *Learning and Individual Differences*, 22(6), 639-649.
- Laufer, B. (2003). Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence. *Canadian modern language review*, 59(4), 567-587.
- Nation, I. (2006). How large a vocabulary is needed for reading and listening? *Canadian modern language review*, 63(1), 59-82.
- Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, 55(2), 151-218.
- Waring, R. et Nation, I. S. (2004). Second language reading and incidental vocabulary learning. *Angles on the English speaking world*, 4, 97-110.
- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied linguistics*, 28(1), 46-65.
- Webb, S. et Nation, P. (2017). *How vocabulary is learned*. Oxford University Press.