

---

## Donner un accès aisé aux formes phoniques des mots décrits dans un dictionnaire: étude pour un dictionnaire monolingue français destiné à de jeunes utilisateurs<sup>1</sup>

Nathalie Gasiglia

UMR 8163 du CNRS – *Savoirs, Textes, Langages* – Université Lille 3, France

*Dans le cadre de cette contribution, je me propose de réfléchir à ce qui pourrait évoluer dans les dictionnaires sur support électronique concernant les descriptions des formes phoniques des unités linguistiques décrites et les modes d'accès à celles-ci. En envisageant les consultations de dictionnaires à la fois dans le cadre d'une aide à la compréhension (de ce qui est entendu ou difficile à déchiffrer) et à l'expression (énonciation ou lecture à haute voix, ou graphie des mots respectueuse de l'usage), je me propose d'examiner comment améliorer l'accès aux articles par les formes phoniques et l'utilisation des indications phonétiques fournies. Les orientations qui se dégagent de cette étude sont établies dans le cadre d'une création de dictionnaire électronique destiné à des élèves francophones de 11 à 15 ans (plus autonomes que les lecteurs débutants mais dont la maîtrise linguistique doit encore progresser) ou allophones de niveau intermédiaire ou avancé. Elles s'appuient sur ce qui est proposé dans trois dictionnaires publiés par Le Robert, l'éditeur français qui a attaché le plus de soin au traitement des prononciations dans ses produits: le Petit Robert électronique (éditions 2001 à 2008), qui est le plus élaboré des dictionnaires généraux électroniques français quant à l'accès aux descriptions des formes phoniques, le Robert junior électronique (éditions 1998 à 2006 – la dernière sous le titre Le Robert des enfants), qui dispose des mêmes fonctions de recherche que le premier, mais dont le texte, destiné aux élèves de 8 à 11 ans, est moins riche, et le Robert oral-écrit (1989), dictionnaire imprimé novateur pour apprenants (natifs ou allophones) qui permettait un accès aux graphies à partir des transcriptions de formes phoniques. Complémentairement aux modalités de traitement et de consultation inspirées de ces dictionnaires, le recours aux technologies de reconnaissance et de synthèse vocales est envisagé. Il impliquerait des partenariats de recherche et développement.*

### 0. Introduction

Dans le cadre de cette contribution, je me propose de réfléchir à ce qui pourrait évoluer dans les dictionnaires sur support électronique concernant les descriptions des formes phoniques des unités linguistiques décrites et les modes d'accès à celles-ci<sup>2</sup>. En envisageant les consultations de dictionnaires à la fois dans le cadre d'une aide à la compréhension (de ce qui est entendu ou difficile à déchiffrer<sup>3</sup>) et à l'expression (énonciation ou lecture à haute voix, ou graphie des mots respectueuse de l'usage<sup>4</sup>), je me propose d'examiner comment améliorer l'accès aux articles par les formes phoniques et l'utilisation des indications phonétiques fournies.

La recherche des traitements des formes phoniques les plus opératoires concerne ici un dictionnaire électronique général de français utilisable par de jeunes francophones (comme les collégiens de France, âgés de 11 à 15 ans) ou éventuellement par des allophones de niveau intermédiaire à avancé, les uns et les autres ayant à acquérir conjointement le code linguistique et la méthodologie de consultation des dictionnaires qui doivent les aider à résoudre leurs difficultés. La focalisation sur ce public est motivée par le fait que, bien que le marché

---

<sup>1</sup> Merci à Cyril Auran pour nos discussions liminaires et à Pierre Corbin pour ses lectures critiques.

<sup>2</sup> Landau (1989: 97), se référant à Hulbert (1968), résumait diverses difficultés posées par le traitement de la prononciation dans les dictionnaires: enregistrements d'oral malaisés à acquérir, transcriptions difficilement précises, non ambiguës et lisibles, et intérêt relativement faible du public pour ce type d'information.

<sup>3</sup> Une oralisation peut être utile pour un usager qui rencontre la graphie d'une unité linguistique qu'il connaît à l'oral mais qu'il ne reconnaît pas à l'écrit (par exemple *abbaye* [abeɪ], *aïeul* [ajœl], *août* [u(t)], etc.).

<sup>4</sup> Un accès par une forme phonique peut être utile pour un usager qui doit écrire une unité linguistique qu'il ne connaît qu'à l'oral et dont la graphie n'est pas aisément déductible de la prononciation (cf. les exemples de la note 3).

français propose des dictionnaires destinés aux jeunes lecteurs, leur éducation à la consultation de ces répertoires s’interrompt quand la complexité des articles s’accroît, ce qui les laisse confrontés à des textes qu’ils ne savent pas décoder seuls. Cette situation est d’ailleurs comparable à celle d’allophones venant de pays où il ne se vend pas de dictionnaires spécifiquement destinés au jeune public et qui peuvent n’avoir eu qu’un accès trop récent à des répertoires décrivant leur langue maternelle pour que des automatismes de consultation leur permettent de transposer leur expérience au déchiffrement d’une description en français d’items français. Ce contexte prescrit de proposer des accès aisés à des informations aussi explicites et complètes que possible.

Pour cette étude, je prendrai appui sur ce qui est proposé dans deux dictionnaires français publiés par Le Robert, l’éditeur français qui a attaché le plus de soin au traitement des prononciations dans ses produits: le *Petit Robert* électronique (éditions 2001 à 2008)<sup>5</sup>, dictionnaire général qui est destiné à des locuteurs experts et qui est le plus élaboré des dictionnaires électroniques français quant à l’accès aux descriptions des formes phoniques<sup>6</sup>, et le *Robert oral-écrit* (1989), dictionnaire imprimé novateur pour apprenants (natifs ou allophones) qui permettait un accès aux graphies à partir des transcriptions de formes phoniques et qui, faute d’un succès suffisant, ne connut qu’une commercialisation assez brève.

### 1. Consulter un dictionnaire à partir ou à la recherche d’une forme phonique

Dans les dictionnaires généraux, différents composants d’articles peuvent fournir une information sur les formes phoniques des unités linguistiques décrites en leur sein:

– la rubrique de prononciation, qui présente une transcription phonétique au moins partielle de la forme canonique (le lemme) de l’unité décrite (celle qui est mentionnée en adresse macrostructurelle – entrée principale d’article ou entrée subordonnée)<sup>7</sup> et éventuellement de certaines formes fléchies remarquables, à proximité de la première<sup>8</sup> ou au cœur des articles<sup>9</sup>, et qui peut associer une analogie phonétique aux transcriptions<sup>10</sup>;

– sur support électronique, pour une sélection d’adresses macrostructurelles, le module d’oralisation de la forme canonique et éventuellement de formes fléchies<sup>11</sup> isolées ou insérées dans des contextes minimaux<sup>12</sup>; <sup>13</sup>

<sup>5</sup> Les textes du *Petit Robert* cités ci-après seront ceux du millésime 2009 de l’édition électronique.

<sup>6</sup> Le *Robert junior* électronique, pour l’école primaire, a les mêmes fonctions de recherche que le *Petit Robert* (cf. n. 28), mais son texte moins riche motive qu’il ne serve pas de référence de base dans cette contribution.

<sup>7</sup> Le *Petit Robert* propose, s.v. **être**, la transcription ‘[ɛtʁ]’.

<sup>8</sup> Le *Petit Robert* propose, s.v. **œil, yeux**, ‘**œil** [œj], plur. **yeux** [jø] nom masculin’.

<sup>9</sup> Le *Petit Robert* dissocie, s.v. **ail**, la prononciation de l’emploi usuel au singulier (‘**ail** [aj]’) et celle du pluriel ancien *aulx*, qui clôt la première subdivision de description (‘*Des aulx* [o] (vieilli), mod. *des ails*.’).

<sup>10</sup> Le *Petit Robert des enfants* proposait, par exemple, s.v. **abbaye**: ‘*Abbaye* [abei] rime avec *obéi*.’.

<sup>11</sup> Pour des noms comme ceux de genre masculin qui ont une finale en *-al* au singulier et *-aux* au pluriel, le *Petit Robert* fournit soit l’oralisation de la seule forme du singulier, comme s.v. **canal, aux**, soit celle des deux nombres, comme s.v. **géosynclinal, aux**, les formes oralisées étant les mêmes, que l’on clique sur le haut-parleur associé aux items dans la liste des ‘entrées’ de la ‘nomenclature complète’ ou sur celui qui figure en haut à droite de l’article. Par contre, pour *cheval*, un des noms de l’ensemble précédent, le *Petit Robert* donne accès à deux fenêtres d’oralisations différentes: (i) à partir de **cheval, aux** dans la liste des entrées, celle de l’adresse microstructurelle

- l’adresse macrostructurelle graphique, qui peut se voir associer un diacritique pour les mots dont la lettre *h* initiale est aspirée<sup>14</sup>;
- l’adresse graphique de nouveau, ou sa reprise, pour l’affichage d’un découpage syllabique<sup>15</sup> (l’accentuation des syllabes pourrait y figurer, si le français était concerné par cette propriété);
- des remarques, qui peuvent attirer l’attention des lecteurs sur une difficulté particulière<sup>16</sup>;
- des mentions d’homonymes, qui peuvent permettre d’éviter des confusions entre unités linguistiques partageant des formes phoniques (celles de leur forme canonique ou d’autres formes)<sup>17</sup>.

C’est avec l’aide de ces informations susceptibles de figurer dans les dictionnaires généraux que leurs utilisateurs peuvent construire ou conforter la composante phonique de leur

---

à *cheval*, décrite et transcrite dans la subdivision I.2. (‘**Loc. adj.** et **loc. adv.** (fin XI<sup>e</sup>) **À CHEVAL** [a]val] : sur un cheval.’); (ii) à partir de l’article **cheval, aux**, celles de à *cheval* et de trois autres adresses décrites dans la subdivision II.3.: *cheval vapeur* et son abréviation *Ch*, transcrites l’une et l’autre (‘CHEVAL-VAPEUR [ʃ(ə)valvapœʁ] **n. m.** ou **CHEVAL** (abrév. *Ch* [ʃ(ə)valvapœʁ]) : ancienne unité de puissance équivalant à 736 watts.’), et *deux-chevaux*, mentionnée en italiques mais non transcrite dans la description de *cheval fiscal* qui est postposée à la précédente (‘*Une deux(-)chevaux Citroën, une 2 CV*’). Dans la liste des entrées du dictionnaire, les items complexes construits avec *cheval* (*cheval-vapeur*, son abréviation et *deux-chevaux*) figurent accompagnés d’un haut-parleur qui donne accès à leur fenêtre d’oralisation respective, où ils sont tous transcrits et énoncés.

La mise à disposition de l’oralisation de à *cheval* à partir de **cheval, aux** dans la liste des entrées pourrait être analysée comme résultant d’une erreur locale de rattachement de la transcription. Mais ce fait, corrélé d’une part à la non-systématicité de l’oralisation de la forme du pluriel (absente pour *canaux* s.v. **canal, aux**) et d’autre part, dans la liste des entrées, à la régularité des appariements des composés de *cheval* évoqués et de leur oralisation et à l’absence de à *cheval*, semble plutôt témoigner d’une volonté d’exploitation de toutes les oralisations disponibles, même si elles n’ont pas toutes le même degré de pertinence là où elles sont présentées.

<sup>12</sup> Le *Petit Robert* propose par exemple, s.v. **être**, les oralisations de 22 formes: *être* [ɛʁ], *est-ce que ?* [ɛskə], *je suis* [ʒəsɥi], *tu es, il est, nous sommes, vous êtes, ils sont, j’étais, nous étions, je fus, je serai, nous serons, que je sois, que nous soyons, je serais, nous serions, sois* [swa], *soyons* [swajɔ̃], *soyez* [swaje], *étant* [etã], *été* [ete]. 8 formes d’appel combinent forme graphique et transcription phonétique, alors que 14 ne comportent que la graphie. Seule la transcription de l’infinitif est indiquée dans le texte de l’article, et aucune ne l’est dans le tableau de conjugaison lié.

<sup>13</sup> Landau (1989: 95) évoque la présence de descriptions de prononciations isolées et en contexte dans le *Webster’s Third New International Dictionary*, critiquées pour leur complexité mais d’une richesse indéniable.

<sup>14</sup> Le *Petit Robert* ne propose pas ce type d’indication mais le *Robert junior* le fait, s.v. **\*hache** par exemple.

<sup>15</sup> Le découpage syllabique figure rarement dans les dictionnaires de français (dans la collection *Encarta*, le dictionnaire français ne le propose pas, contrairement à celui pour l’anglais – version US), mis à part quelquesuns de ceux à visée pédagogique à l’intention de jeunes enfants comme *Mon dictionnaire pour la grande école*, dont les différentes syllabes de chaque adresse macrostructurelle sont imprimées dans une couleur distincte, ou le *Dictionnaire des mots difficiles à lire*, qui propose systématiquement une reprise de l’adresse graphique incluant une segmentation syllabique comme ‘ab/baye’ s.v. **abbaye** ou ‘ai/eul’ s.v. **aioul**.

<sup>16</sup> Le *Petit Robert* propose, s.v. **dix**, la remarque suivante: ‘Se prononce [diz] devant un mot commençant par une voyelle ou un *h* muet, [di] devant un mot commençant par une consonne, [dis] dans les autres cas.’. Le *Robert oral-écrit* fournit des remarques analogiques à la manière du *Petit Robert des enfants*, comme s.v. **soil** (cf. n. 24).

<sup>17</sup> Le *Petit Robert* propose, s.v. **être**, les homonymes de trois formes phoniques: ‘■ **HOMONYMES : Êtres, hêtre; suis : suïs (suivre); es, est : ai (1. avoir), hais (hair).**’.

compétence linguistique et qu'ils doivent, le cas échéant, résoudre les problèmes qui motivent la consultation de l'un de ces répertoires:

– S'il s'agit d'une difficulté de transcription graphique ou de compréhension d'une séquence entendue, l'utilisateur doit d'abord segmenter le flux sonore pour y repérer les bornes de la ou des unités linguistiques problématiques, puis, pour chacune, identifier une graphie susceptible de correspondre à la forme orale (que l'unité en cause soit ou non connue de lui à l'oral).<sup>18</sup>

– S'il s'agit d'une difficulté de prononciation (de quelque chose qui est écrit ou que l'utilisateur sait écrire), il doit écouter l'oralisation ou lire la transcription (partielle ou intégrale) de la forme canonique ou d'une autre forme remarquable, selon ce qui lui est fourni, puis identifier les éléments de variation existant entre la graphie de cette forme et celle de la forme qu'il souhaite prononcer et en inférer les changements à opérer pour produire une prononciation supposée adéquate, et enfin insérer harmonieusement la forme phonique spéculée dans son contexte, ce qui peut impliquer des liaisons, assimilations ou effacements.<sup>19</sup>

## 2. Atouts et faiblesses comparées du *Petit Robert* et du *Robert oral-écrit*

Certains dictionnaires portent une attention particulière à la description des formes phoniques. Parmi eux, comparons le *Petit Robert* et le *Robert oral-écrit*, dont les projets sont tout à fait distincts, ce qui se traduit par des organisations macro- et microstructurelles très différentes.

Le *Petit Robert* est un dictionnaire dont les articles ont des adresses graphiques et sont triés selon l'ordre alphabétique de ces dernières. Dans sa version imprimée, les formes phoniques ne sont donc pas prises en compte dans la structure d'accès aux articles. La version électronique propose, elle, une fonction de recherche à partir d'une forme phonique, qui doit pouvoir dispenser l'utilisateur de la recherche mentale d'une forme graphique susceptible de correspondre à la forme phonique à chercher, et qui, à condition de maîtriser l'alphabet phonétique, permet d'accéder à des adresses graphiques d'articles<sup>20</sup>, également accessibles à partir des recherches portant sur les formes graphiques en utilisant l'option de 'correction phonétique'<sup>21, 22</sup>. La

<sup>18</sup> Il faut par exemple segmenter [saʃɑ̃brestsɥsu] pour y repérer les items *sa*, *chambre*, *est* et *sens dessus dessous*.

<sup>19</sup> Pour prononcer la forme fléchie du verbe *sentir* dans la phrase *Je sentais une grande fatigue m'envahir* à l'aide du *Petit Robert*, il faut lire la transcription de l'infinitif ([sɑ̃tir]) ou écouter l'oralisation de l'une de ses formes conjuguées (le dictionnaire propose celles de 7 formes: *je sens*, *nous sentons*, *ils sentent*, *nous sentions*, *nous sentirions*, *que nous sentions*, *nous sentirions*), y retrancher et/ou ajouter ce qui convient (à partir de la transcription de l'infinitif, oraliser le pronom puis substituer au [ir] final [ɛ], ou à partir de la première forme oralisée – *je sens* – concaténer [tɛ]), puis éventuellement repérer la liaison possible avec le déterminant suivant pour produire [ʒəsɑ̃tɛ(z)ɥn...] voire [ʃsɑ̃tɛ(z)ɥn...].

<sup>20</sup> Dans le *Petit Robert*, la saisie de [sɑ̃], par exemple, dans l'interface de recherche par critère phonétique permet d'accéder à 5 adresses graphiques: **C** (adj. numér. card., 'Cent, en chiffres romains.', s.v. **2. C** 1.), **1. cent** (adj. numér. et n.), **3. cent** (n. m., 'Centième partie de l'euro.'), **sang** (n. m.) et **sans** (prép.).

<sup>21</sup> Dans le *Petit Robert*, la recherche des entrées graphiques proches de 'san' en sélectionnant l'option 'en utilisant la correction phonétique' fournit 5 adresses: les 4 dernières du résultat précédent (cf. n. 20) et **2. cent** (n. m. et f., '1. **N. m.** Centième partie de l'unité monétaire de divers pays' et '2. **N. f.** (1851) **Région.** (Canada) [...] cette unité, cette pièce.').

<sup>22</sup> Les autres dictionnaires monolingues français n'offrent pas d'accès comparable, à l'exception du *Trésor de la langue française informatisé (TLFi)*, qui permet soit de faire une 'saisie phonétique' au moyen de boutons qui correspondent à des signes phonétiques mais qui ne portent que des mentions de lettres (par exemple 'CH' pour

recherche par critère phonétique mise en œuvre présente la particularité de donner accès à toutes les adresses macrostructurelles, mais aussi à des adresses microstructurelles transcrites<sup>23</sup>.

Le *Robert oral-écrit*, qui se démarque des ‘autres dictionnaires, [où] on ne trouve un mot que si l’on connaît son orthographe ou si l’on réussit à la deviner’ (‘Préface’, p. XI), propose des articles dont les adresses principales sont des transcriptions phonétiques qui dominent une ou plusieurs sous-adresses graphiques, lesquelles introduisent les descriptions de chaque forme lexicale correspondant à la prononciation.<sup>24</sup> Ce dispositif témoigne en outre d’un souci d’accessibilité poussé pour les formes telles qu’elles s’observent dans les contextes où l’usager les trouve (cf., par exemple, la mention de *je sens*, forme du verbe *sentir*, s.v. **sã**<sup>25</sup>). Il rencontre néanmoins des limites liées au support de consultation: par exemple, s.v. **sãs** ① un **sens**, une remarque indique ‘On prononce parfois *sã* (*sens dessus dessous sãdĈesydĈesu* ou *sãtsytsu*; *sens devant derrière sãdĈEvãdĈERjĈER* ou *sãsdĈEvãdĈERjĈER*)’, ce qui n’est pas indiqué s.v. **sã** et n’est donc accessible qu’à partir de la forme hors contexte de l’item.<sup>26</sup> Une version électronique de cet ouvrage pourrait proposer un module de recherche comparable à celui du *Petit Robert*, capable d’explorer toutes les transcriptions plutôt que celles des seules adresses. Bien que susceptible de valoriser un produit original, cette édition électronique est toutefois hautement improbable à titre commercial du fait d’une part de la petite taille de la

---

[G], ‘É’ pour [e], ‘AN’ pour [ã], etc.), soit de fournir une graphie qui peut être fautive (dans les limites – assez larges en mode dit ‘forcé’ – de ce que sait gérer le ‘correcteur d’erreurs’ intégré).

<sup>23</sup> Dans le *Petit Robert*, la recherche des adresses graphiques correspondant à la forme phonétique [aze] fait accéder à A. G. s.v. **assemblée** 2. (‘ASSEMBLEE GENERALE (abrév. fam. A. G. [aze])’), mais toutes les adresses microstructurelles transcrites ne sont pas accessibles à partir de la recherche phonétique (par exemple *faux jeton* mentionnée s.v. **jeton** 2. («Par ext. UN, UNE FAUX JETON [foʒtõ]: un, une hypocrite») ou à *cheval* et *chevalvapeur* mentionnées s.v. **cheval**, **aux** (cf. n. 11) ne le sont pas).

<sup>24</sup> Les transcriptions qui figurent en adresses principales sont imprimées en bleu dans le volume et reproduites ici en gris. Toutes les transcriptions sont écrites dans un code qui respecte assez largement l’alphabet phonétique international, mais qui utilise une codification particulière pour quelques neutralisations: ‘E’ neutralise l’opposition observable entre ‘e’ et ‘ɛ’, ‘O’ celle entre ‘o’ et ‘ɔ’, ‘Ĉ’ celle entre ‘ø’, ‘œ’ et ‘ə’, etc. Ces neutralisations sont compensées par la spécification de certains traits pertinents quand ceux-ci peuvent servir à différencier des mots proches (à condition que ces mots soient prononcés par des personnes qui font la distinction). Par exemple, s.v. **sOI**, les sous-adresses – un **saule**, ① un **sol**, ② un **sol**, une **sole** – dominent chacune une remarque où le degré de fermeture ou d’ouverture du [O] est précisé: ‘O fermé (comme dans *pot*).’ sous la première, ‘O ouvert (comme dans *port*).’ sous chacune des trois suivantes.

<sup>25</sup> **sã** domine 4 adresses graphiques: **cent**, le **sang**, **sans**, je **sens**.

<sup>26</sup> D’autres variantes signalées en remarque posent un problème d’accessibilité comparable. Elles concernent notamment:

– des liaisons: par exemple, s.v. **sã cent**, une remarque indique que ‘En liaison, on prononce parfois **sãt** (*cent ans sãtã*).’ ou, s.v. **sã sans**, une autre stipule que ‘En liaison, on prononce parfois **sãz** (*sans arrêt sãzARE*).’, ou encore s.v. **bõ** ② **bon**, il est précisé que ‘En liaison, on prononce parfois **bOn** comme au féminin, ou **bõn** (*un bon ami: ĈbOnAmi* ou *ĈbõnAmi*).’;

– des disparitions de consonnes: par exemple, s.v. **sãtimEtr** un **centimètre**, la remarque fournit la seconde forme audible (‘On prononce parfois **sãtimEt** dans une phrase dite rapidement.’), comme s.v. **kElkĈfwA** **quelquefois**, elle fournit la forme [kĈkfwA] (‘Certains prononcent **kĈkfwA**.’).

Mais certaines variantes donnent aussi lieu à des réadressages: pour *petit*, dont le *e* caduc peut disparaître, il y a un renvoi depuis **pti** vers **pĈEti** comme de **ptifis** vers **pĈEtifis**. Dans les articles de description, une variation de traitement est observable: s.v. **pĈEti** ① **petit**, la forme phonique [pti] est fournie en remarque (comme dans les cas précédents) alors que pour *petit-fils* c’est l’adresse qui mentionne les deux transcriptions (‘**pĈEtifis** ou **ptifis** un **petit-fils**’).

nomenclature du dictionnaire<sup>27</sup> et d'autre part de la transférabilité de sa richesse et de son originalité dans un autre produit électronique, ce qui invite plutôt à concevoir un ouvrage riche en transcriptions et dont l'un des modes d'affichage adopterait son principe d'adressage phonétique principal et graphique subordonné.

### 3. Vers un nouveau dictionnaire électronique valorisant les formes phoniques

Les deux dictionnaires observés offrent des modes d'accès aux formes phoniques et des informations les concernant dont j'ai montré l'utilité, qui ont été établis selon des principes distincts mais complémentaires et dont je voudrais évaluer s'ils peuvent constituer une source d'inspiration pour un nouveau dictionnaire électronique destiné à des élèves francophones de 11 à 15 ans (plus autonomes que les lecteurs débutants mais dont la maîtrise linguistique doit encore progresser) ou allophones de niveau intermédiaire ou avancé<sup>28</sup>.

Le *Petit Robert* présente deux points forts: (i) du point de vue de l'accessibilité, la puissance de son module de recherche à partir d'une transcription ou d'une forme graphique oralisable de manière approchante (cf. nn. 20 et 21), et (ii) du point de vue des descriptions fournies, la mise à disposition de transcriptions et/ou d'oralisations pour les items en adresse macrostructurelle et une sélection de ceux en adresse microstructurelle. Bien que nous ayons observé qu'il pourrait étendre les recherches à partir de requêtes phonétiques à un plus grand nombre de transcriptions de formes des paradigmes flexionnels et d'adresses microstructurelles (cf. n. 23) et augmenter le nombre des transcriptions et oralisations fournies, et qu'il serait opportun de systématiser

<sup>27</sup> 'Ce dictionnaire contient environ 17 000 mots. Ce sont avant tout les mots qui sont dans tous les dictionnaires: les 5 000 mots français les plus fréquents puis de nombreux mots du vocabulaire fondamental du *Petit Robert des Enfants*; de nombreux homonymes et quelques mots du vocabulaire technique de base utilisés dans les Lycées d'Enseignement Professionnel.' (Préface: XII).

<sup>28</sup> Pour l'heure il n'existe pas de dictionnaire électronique spécifiquement destiné à ce public en France. Il en existe en revanche deux pour les élèves du cycle 3 de l'école primaire (de 8 à 11 ans): le *Dictionnaire Auzou junior*, dont le volume de 2008 inclut un disque, et le *Robert junior*, qui a connu quatre éditions sur disques correspondant à l'édition imprimée de 1993 (rééditée jusqu'en 1998) enrichie de médias (cf. Corbin 2009): en 1998, 1999 et 2003 sous le même titre, puis 2006 sous le nom de *Robert des enfants* et dont les licences d'une version qui n'a intégré les médias qu'en janvier 2008 (cf. <http://junior.bvdep.com/demo/infos.htm>) sont achetables par les établissements scolaires via le Canal Numérique des Savoirs (<http://www.cns-edu.com/ressource-16---le-robert-junior.html>).

Le *Dictionnaire Auzou junior*, électronique comme imprimé, n'intègre aucune composante phonique.

Le *Robert junior* sur disque propose, lui, les mêmes modalités de recherche par critère phonétique et par graphie 'en utilisant la correction phonétique' que le *Petit Robert*. La saisie de [sã] dans l'interface de recherche phonétique comme celle de 'san' pour la recherche de graphies d'homophones permet d'accéder à 3 adresses graphiques: **cent** (adj.), **sang** (n. m.) et **sans** (prép.). Les résultats correspondants pour le *Petit Robert* (cf. nn. 20 et 21) sont identiques, au prorata de la différence de nomenclature des deux répertoires. Du point de vue de la description des formes phoniques, l'offre du dictionnaire junior est également comparable à celle du *Petit Robert*. S.v. **être**, des oralisations sont accessibles via le haut-parleur présent, antéposé à l'item dans la liste des adresses, en haut à gauche de l'article et dans le tableau de conjugaison. Elles concernent 21 formes: *être* [ɛtʁ], *je suis* [ʒəsɥi], *tu es*, *il est*, *nous sommes*, *vous êtes*, *ils sont*, *j'étais*, *nous étions*, *je fus*, *je serai* [ʒəsəʁe], *nous serons* [nusrɔ̃], *que je sois*, *que nous soyons*, *je serais*, *nous serions*, *sois* [swa], *soyons*, *soyez*, *étant*, *été*. L'infinitif verbal n'est pas oralisé dans le tableau des formes conjuguées et aucune transcription n'y est mentionnée. En plus des 21 formes partagées, le *Petit Robert* proposait l'oralisation de *est-ce que ?* (cf. n. 12), adverbe que le *Robert junior* mentionne dans la liste des 'Mot(s) de la même famille' mais ne traite pas s.v. **être**, préférant lui consacrer un article autonome où son oralisation est fournie. Par ailleurs, ce ne sont pas strictement les mêmes formes d'appel dans la fenêtre des prononciations qui, dans les deux dictionnaires, combinent la forme graphique et la transcription phonétique: ils partagent les transcriptions de *être*, *je suis* et *sois*, et le *Robert junior* ajoute celles de *je serai* et *nous serons*, mais ne mentionne pas celles de *soyons*, *soyez*, *étant* et *été*. Enfin, le texte de l'article **être** du *Robert junior* ne propose pas la transcription de l'infinitif, ce qui est cohérent avec le principe consistant, dans ce dictionnaire, à n'en mentionner que pour une sélection d'items jugés remarquables sur le plan phonographique.

leur principe de distribution (cf. n. 11), ces atouts me semblent très utiles pour les jeunes usagers qui seraient initiés à leur maniement<sup>29</sup> et qui pourraient ainsi en tirer un plein profit.

Le *Robert oral-écrit* présente lui aussi deux points forts: (i) une attention particulière apportée à la description des variantes de prononciation (cf. *sens* [sã(s)] *supra* (§ 2.) et les exemples fournis en n. 26) et (ii) un principe d'accès aux formes graphiques à partir des formes phoniques possibles (cf. la mention de l'adresse graphique je *sens* s.v. *sũ* en n. 25). Bien que toutes les transcriptions ne figurent pas à la nomenclature principale, ce qui est compréhensible pour une édition imprimée qui a des contraintes de place, l'orientation de cet ouvrage consiste explicitement à conduire chaque usager d'une forme phonique qu'il identifie à l'ensemble des graphies qui peuvent lui correspondre.<sup>30</sup>

### 3.1. Valorisation des transcriptions de formes phoniques

Un nouveau dictionnaire, dérivé à partir d'une matrice textuelle de dictionnaire dont la structure d'accès principale repose sur le classement alphabétique des graphies des items utilisées comme adresses macrostructurelles, pourrait intégrer les modules de recherche phonétique et graphique (avec correction phonétique) proposés dans le *Petit Robert* et le *Robert junior* (cf. n. 28) mais présenter la mention du motif de recherche avant la liste d'items en adresse qu'ils produisent<sup>31</sup>. Ces modules pourraient également enrichir la liste des réponses par la mention des items et des combinaisons d'items phonétiquement soudés (cf. n. 30) qui débutent par le motif cherché<sup>32, 33</sup>. Ils pourraient enfin tirer profit d'un enrichissement des descriptions de formes phoniques: si chaque adresse macro- ou microstructurelle et chaque forme des paradigmes flexionnels était

<sup>29</sup> En France, les programmes scolaires de l'école primaire prévoient une initiation à l'utilisation de dictionnaires électroniques dès 6 ans (cf. les programmes du cycle 2, limités aux cours préparatoire et élémentaire première année ([http://www.education.gouv.fr/bo/2008/hs3/programme\\_CP\\_CE1.htm](http://www.education.gouv.fr/bo/2008/hs3/programme_CP_CE1.htm)): les élèves 'sont amenés à utiliser l'ordinateur : écriture au clavier, utilisation d'un dictionnaire électronique.'). Il serait donc souhaitable que le type de fonctionnalités envisagé ici soit également proposé dans les dictionnaires destinés à ce public scolaire plus jeune et que les enseignants puissent bien valoriser et systématiser son emploi.

<sup>30</sup> Comme la majorité des autres dictionnaires (mais en excluant les unités infralexicales prises en compte par certains), le *Robert oral-écrit* a vocation à décrire des unités linguistiques de niveau lexical (cf. n. 27) et à prendre en compte leur combinatoire régulière limitée aux 'groupes de mots soudés qui se prononcent comme un seul mot et qui sont la cause de confusions' (Préface: XII). Il est donc normal que ses usagers ne soient pas dispensés de la tâche de segmentation du flux sonore, mais ils sont soutenus dans la réalisation de celle-ci puisque certaines formes phoniques concernent des séquences de mots (un pronom préverbal et une forme verbale liée par exemple, comme s.v. *sã...*, où 2 graphies – *s'en...* et *s'em...* – sont introduites accompagnées d'un renvoi vers la section *ã*, et en particulier vers '*s'enrhumer* *ãrymE* ; *s'enfuir* *ãfyir* ; *s'envoler* *ãvOIE...*').

<sup>31</sup> Pour les deux recherches envisagées en notes 20, 21 et 28, l'expression du motif de recherche associé au nombre des items lui correspondant (nombre qui figure en bas de la liste des réponses dans le *Petit Robert* comme dans le *Robert junior*) pourrait prendre les formes respectives suivantes: '*n* mots décrits dans le dictionnaire se prononcent [sã] : ' et '*n* mots décrits dans le dictionnaire se prononcent comme la forme graphique 'san' : '.

<sup>32</sup> Cette extension de la liste pourrait être précédée des mentions '*n* mots décrits dans le dictionnaire débutent par [sã] : *rendre* (n. f.) [...]' et '*n* combinaisons de mots phonétiquement soudés décrites dans le dictionnaire débutent par [sã] : *s'amplifier* (v.) [...]'. L'item *rendre* mentionné ici correspond au premier listé par le *Robert junior* en réponse à la recherche d'un mot qui débute par [sã] (cf. n. 33). Ce type de recherche n'étant pas prévu pour les combinaisons de mots, l'item *s'amplifier* correspond, lui, au premier verbe à emploi pronominal qui débute par [ã].

<sup>33</sup> À défaut de fournir systématiquement les mots ou combinaisons de mots décrits qui débutent par un motif donné, il est souhaitable de permettre de les chercher, ce qui est d'ailleurs déjà faisable pour les mots dans le *Petit Robert* comme dans le *Robert junior* à partir des recherches soit de transcriptions phonétiques soit de graphies avec utilisation de la correction phonétique (puisque'il est possible d'inclure une étoile dans la forme cherchée: un caractère joker représentant une séquence de zéro, un ou plusieurs caractères indéterminés encadrée par ceux qui sont indiqués).

associée à une transcription phonétique et si chaque variante de prononciation régulière était répertoriée et transcrite (sans qu'elles soient systématiquement présentées aux usagers si leur nombre est jugé trop élevé et qu'il s'avère plus ergonomique de ne les fournir qu'à la demande), ces modules de recherche auraient les moyens de mettre en correspondance les transcriptions et graphies approchantes cherchées par les utilisateurs et l'ensemble de ces transcriptions<sup>34</sup>.

Un nouvel affichage, présenté aux usagers à leur demande, pourrait par ailleurs synthétiser les différents éléments d'information relatifs aux formes phoniques des mots décrits. Selon des modalités de présentation adaptées à l'ampleur des paradigmes flexionnel et combinatoire de chaque unité linguistique, il récapitulerait les indications distribuées précédemment dans la liste des items ou des combinaisons répondant aux critères phonétiques des recherches et pourrait se substituer aux composants d'homonymie actuels, dont il constituerait une version plus exhaustive: chaque forme orale transcrite dominerait des formes graphiques ne se limitant pas aux homonymes homocatégoriels<sup>35</sup>, prenant en compte les variantes de prononciation<sup>36</sup> et toutes les formes des mots décrits plutôt qu'une sélection de celles-ci<sup>37</sup>.

Dans cet affichage, la diversité des usages pourrait en outre être mieux représentée si l'on parvenait à fournir des informations comparables à celles proposées dans le dictionnaire de

<sup>34</sup> Rappel d'exemples d'adresses microstructurelles, de formes fléchies et d'items sujets à des variations phoniques dont les prononciations pourraient ainsi être rendues accessibles:

– La conjonction *sans que* présentée en adresse subordonnée s.v. **sans** 2. et les locutions de valeur négative *sans arrêt*, *sans cesse*, *sans conteste*, *sans crainte*, etc. mentionnées s.v. **sans** 1. dans le *Petit Robert* figureraient dans la liste des mots phonétiquement soudés qui débutent par [sã] (cf. n. 32). Il en irait de même pour les combinaisons décrites dans le *Robert oral-écrit* s.v. **sã** ou **sãs** (cf. n. 26 et § 2.).

– La description de *je sens* incluse dans l'article **sã** du *Robert oral-écrit* (cf. n. 25) pourrait figurer dans la liste des formes verbales conjuguées: 'n formes conjuguées de verbes décrits dans le dictionnaire se prononcent [sã]: je **sens** (v. **sentir**) [...]'

– Faire leur pleine place aux variantes implique de systématiser leurs mentions pour chaque forme du paradigme flexionnel de l'item décrit et de permettre l'accès aux articles à partir de n'importe laquelle d'entre elles. Si l'exploitation des données du *Robert oral-écrit* impliquerait surtout la systématisation de cette dernière disposition, certaines variantes n'y étant accessibles qu'à partir de la prononciation standard (cf. n. 26), celle de la matière du *Petit Robert*, plus normatif, impliquerait une révision plus globale, puisque les disparitions de consonnes dans certains contextes n'y sont pas signalées (*centimètre* ou *quelquefois* y sont seulement transcrits [sãtimetr] et [kɛlkɛfwã] alors qu'ils étaient doublement transcrits dans le *Robert oral-écrit*) et que les liaisons n'y sont indiquées que si elles sont déterminées par des contraintes très spécifiques (alors que s.v. **sans** seule la transcription [sã] est fournie dans le composant de prononciation et qu'aucune remarque ne précise si d'autres formes sont possibles en contexte, une remarque spécifique s.v. **cent** (seulement transcrit [sã]) qu' 'On fait la liaison devant un nom : *cent ans* [sãtã]; [sã] dans les autres cas : *cent un* [sãœ].', et une autre s.v. **bon, bonne** (seulement transcrit [bɔ, bɔn]) fait état de la variante contextuelle dénasalisée du masculin singulier: '3. (phonét.) *bon* (adj.) se dénasalise devant un mot débutant par une voyelle ou un *h* muet : *bon anniversaire* [bɔnanivɛrsɛr].'; cf. encore n. 16 pour le cas de **dix**).

<sup>35</sup> En la matière, le *Petit Robert* semble être irrégulier, comme le montre par exemple, s.v. **être** (cf. n. 17), le fait que, alors que les homonymes mentionnés de l'infinitif ne sont pas homocatégoriels, le nom *haie* ne figure pas aux côtés des formes verbales *ai* et *hais* homonymes de *es* et *est*.

<sup>36</sup> Si les rédacteurs du *Petit Robert* ont choisi de limiter leur prise en compte des variations possibles aux cas où les mots sont 'prononcé[s] de deux ou plusieurs façons, et ceci, sans rapport avec les diverses régions de la francophonie' (Préface: XIV), ils ont simultanément décidé de 'reconsidérer certaines [...] positions sur l'**homonymie** et [d']envisager comme de possibles homonymes (notés HOM. poss. dans la rubrique finale) de nouvelles unités : 1) Des mots ne se distinguant que par les voyelles [a] et [ɑ] ou [ɛ] et [œ] prononcées de façon identique par beaucoup de Français (ex. *ta* et *tas*; *brin* et *brun*). 2) Des mots ne se distinguant que par les voyelles [e] et [ɛ] et même [o] et [ɔ] en syllabe non finale de mot (ex. *pêcheur* et *pêcheur*; *méson* et *maison*; *chauffard* et *schofar*). (Préface: XXIX).

<sup>37</sup> S.v. **être** dans le *Petit Robert*, les formes *être*, *suïs*, et *es*, *est* sont prises en compte (cf. n. 17) mais pas *sommes* homonyme de *somme* (n. f. ou n. m.) ni *été*, *étais*, etc. rapprochables d'*été* (n. m.) et *étai* (n. m.).



Martinet & Walter (1973), qui, pour chaque unité linguistique, présentait les transcriptions des prononciations observées chez 17 informateurs<sup>38</sup> dont le profil sociolinguistique était détaillé<sup>39</sup>. Outre le travail documentaire qu'impliquerait ce projet<sup>40</sup>, une question préalable me semble être l'utilisabilité de ces données par le public cible: présentées de manière brute, elles nécessiteraient trop de connaissances et de travail d'interprétation pour être plus qu'une curiosité, ce qui implique de trouver le juste degré de variation supposé perceptible, d'introduire des remarques explicitant les phénomènes à l'œuvre<sup>41</sup> et de définir des modes de présentation assurant une bonne lisibilité des informations fournies.

### 3.2. Valorisation des oralisations de formes phoniques

Quelle est la situation concernant les réalisations orales des formes phoniques et que serait-il possible d'améliorer les concernant?

– Certaines insuffisances sont inhérentes aux données orales disponibles dans les dictionnaires et aux relations qu'elles entretiennent avec ce qui peut motiver leur consultation: les oralisations sont aussi neutres et normées que possible, ce qui limite l'aide qu'elles apportent.<sup>42</sup>

Pour la compréhension, et en particulier l'identification des formes perçues, l'aide actuelle est incomplète, puisque ce qu'elle donne à entendre peut différer de ce qui est susceptible d'être entendu dans des énoncés réels du fait d'ajustements contextuels<sup>43</sup>, d'accents individuels ou régionaux des locuteurs, de variations prosodiques ou de débit, etc.

<sup>38</sup> S.v. **sens dessus dessous**, par exemple: 'sãdsydsu (cdjlmnprrtvwx) sãdãsydãsu (bgky) sãtsytsu (a)', où les lettres mentionnées entre parenthèses identifient les informateurs.

<sup>39</sup> Les 'fiches signalétiques' des informateurs sont fournies dans l'Introduction (§ II, 37-48). Elles y sont précédées d'une synthèse intitulée 'Traits généraux de la phonologie de nos informateurs.' (§I. Présentation du dictionnaire, 31-36).

<sup>40</sup> Même s'il est maintenant un peu ancien, le travail documentaire réalisé par Martinet & Walter doit prendre en compte une large partie des lemmes d'une nomenclature de langue générale actuelle: 'Pour déterminer sur quels mots porterait l'enquête, on est parti de la prononciation indiquée dans un récent dictionnaire de la langue. On a, pour chacun des quelque 50.000 mots du dictionnaire, vérifié si la prononciation indiquée était ou non conforme aux indications données dans une dizaine d'ouvrages consacrés à la prononciation du français et provenant des auteurs scientifiquement les plus sûrs. Dans le cas où aucune des autorités mises à contribution ne proposait autre chose que la prononciation de départ, cette prononciation a été considérée comme ne posant pas de problème et retenue comme définitive. Lorsqu'au contraire il y avait une divergence quelconque, le mot a été placé dans une phrase qu'on a fait prononcer aux dix-sept informateurs. Les prononciations, enregistrées et transcrites, figurent ci-dessous pour chacun des mots soumis à l'enquête.' (1973: 10).

<sup>41</sup> Le manuel *Phonétique du FLE* de Léon & al. (2009), qui alerte ses lecteurs à propos des variations entre les prononciations neutres à visée orthoépique qu'il répertorie et ce qui est rencontrable dans les usages effectifs, pourrait inspirer ces remarques. Par exemple, après avoir expliqué que le [E] accentué en syllabe ouverte est généralement fermé sauf dans un certain nombre de mots pour lesquels l'ouverture est déterminée en fonction de leur nature et de leur graphie finale, il indique: 'Mais dans la pratique, et surtout dans le sud de la France, les locuteurs prononcent souvent [e], suivant ainsi la règle générale de distribution complémentaire énoncée ci-dessus.' (Léon & al. 2009 : 30).

<sup>42</sup> Les oralisations actuellement proposées ne semblent pas apporter d'amélioration par rapport à ce que Svensén (1993: 70) observait concernant la faible prise en compte des variations régionales, stylistiques et prosodiques dans les dictionnaires d'anglais.

<sup>43</sup> Certains contextes, dont les dictionnaires étudiés ne présentent qu'une sélection (cf. nn. 23, 26 et 34), imposent ou permettent des assimilations ou effacements de voyelles ou de consonnes, des liaisons, des disparitions de *e* caduc, etc.

Pour l'expression, l'aide existante contribue utilement à une énonciation non marquée<sup>44</sup> mais n'est pas suffisante puisque les items sont généralement fournis hors contexte<sup>45</sup> et sans présence systématique de remarques relatives aux ajustements que celui-ci impose ou rend possibles (cf. n. 19).

– Les enrichissements évoqués au § 3.1. associés à des oralisations de chaque forme transcrite soutiendraient mieux la réception d'énoncés ou leur production même réalisée de manière non neutre.<sup>46</sup> Mais même dans l'hypothèse où les enrichissements qui viennent d'être envisagés seraient effectifs, les données oralisées resteraient certainement présentes en trop faible quantité pour être réellement utiles puisque les énoncés oraux ne sont pas construits par concaténation de formes, alors que nous avons jusqu'ici considéré des unités linguistiques isolées ou en combinaisons régulières de taille réduite. Ce point m'amène à évaluer comment une exploitation raisonnée de technologies de reconnaissance et de synthèse de la parole serait envisageable.

Pour l'aide à la compréhension, si les dictionnaires ne semblent pas être utilisables au fil d'une conversation (sous peine de rompre celui-ci le temps de consulter un article), la reconnaissance vocale en contexte pourrait s'avérer utile lors d'écoutes d'enregistrements (instructions sonores, bandes-son de films, etc.) dont certains passages posent des problèmes d'identification de formes: un dictionnaire doté d'une interface de consultation par reconnaissance vocale permettrait de les réécouter en les lui soumettant afin qu'il y repère les unités linguistiques constituantes et qu'il donne accès à leurs descriptions<sup>47</sup>. Moins ambitieusement, il pourrait être envisagé de proposer un accès par mention orale des unités linguistiques cherchées, ce qui compléterait la recherche par transcription phonétique saisie.

Pour l'aide à l'expression, l'apport du dictionnaire pourrait consister à oraliser un mot ou une séquence de mots graphiques ou à valider l'oralisation produite par l'utilisateur. Dans la première hypothèse, si les unités linguistiques graphiées sont oralisées dans le dictionnaire, un hyperappel "intelligent" pourrait segmenter la chaîne de caractères en unités pertinentes et fournir la concaténation de leurs oralisations, mais sans les ajustements contextuels utiles le produit serait décevant. Celui d'une synthèse vocale comme celles de Jaws ou de Window-Eyes<sup>48</sup>, qui

---

<sup>44</sup> Cependant, lors de l'écriture de dialogues, dans le cadre d'un exercice scolaire par exemple, il peut être pertinent de doter les répliques d'un personnage de traits phonétiques typiques d'une émotion, d'une région, etc.

<sup>45</sup> Sauf dans le cas des citations oralisées du *Petit Robert*, mais celles-ci sont fonctionnellement anecdotiques, d'une part du fait de leur petit nombre (une centaine depuis 2001 et 200 de plus depuis le millésime 2008) et d'autre part parce que leur sélection n'est pas déterminée par des particularités de prononciation mais par l'intérêt porté à leur contenu ou à leur forme. Souvent sentencieuses, elles sont prononcées avec solennité et sont donc plutôt surnormées: les liaisons attendues y figurent ('*Les parfums, les couleurs et les sons se répondent*', s.v. **couleur** I.1.: adéquat pour un alexandrin de Baudelaire), de même que les *e* non entendus ailleurs ('*Si la connerie se mesurait il servirait de mètre étalon*', s.v. **connerie**: lecture solennisée – [silakɔ̃risɛ̃mɛzɛ̃rɛ̃vɛ̃rɛ̃dɛ̃mɛ̃trɛ̃talɔ̃] – d'une réplique du film *Le cave se rebiffe* prononcée [silakɔ̃rɪsmɛzɛ̃rɛ̃vɛ̃rɛ̃dɛ̃mɛ̃trɛ̃talɔ̃] par Jean Gabin).

<sup>46</sup> Cependant une augmentation aussi consistante des sources sonores n'irait pas sans poser des problèmes techniques, qui ne seront pas développés dans cette contribution.

<sup>47</sup> Pour ce faire, le système de reconnaissance vocale devrait composer avec la diversité des locuteurs enregistrés, des thèmes et des situations d'énonciation, et être capable de délimiter les unités linguistiques incluses dans la séquence verbale qui lui serait soumise, même si celle-ci n'est pas coupée au début de la première unité et à la fin de la dernière identifiables en son sein.

<sup>48</sup> Jaws (cf. <http://www.freedomscientific.com/products/fs/jaws-product-page.asp>) et Window-Eyes (cf.

proposent une oralisation fluide et neutre, serait meilleur, ce qui incite à envisager d'associer ce type de technologie à un dictionnaire pour la lecture de ses exemples d'emplois<sup>49</sup>, en délimitant strictement ce qui doit être oralisé<sup>50</sup>, mais ce dispositif ne documenterait pas les usagers concernant de possibles variantes de prononciation. Dans la seconde hypothèse, le dictionnaire pourrait comparer les productions des usagers avec ses prononciations de référence, à la manière de ce que proposent certaines méthodes de langues<sup>51</sup>, mais on peut se demander s'il ne serait en mesure de le faire que pour des segments de textes choisis, comme les exemples, ce qui impliquerait qu'ils soient préalablement lus par des personnes qui chercheraient à réaliser les variations les plus régulières et que les enregistrements de ces lectures soient comparés avec celles des usagers, ou s'il serait envisageable qu'il le propose pour n'importe quelles phrases soumises par ces derniers (cf. n. 49). Il s'agirait alors de les faire lire par un module de synthèse vocale et de comparer la production de celui-ci et la prononciation des usagers, ce qui demanderait idéalement une discrimination entre les variantes de prononciation licites et les fautes.

Qu'il s'agisse de synthèse ou de reconnaissance vocale, les technologies envisagées sont extérieures au champ de compétence des éditeurs et impliqueraient des partenariats de recherche et développement puis de commercialisation.

#### 4. Conclusion

Les pistes de réflexion concernant les possibilités d'amélioration de l'accès aux formes phoniques et leur utilisation par des apprenants natifs ou allophones sont multiples, mais la capacité des éditeurs à les proposer et le profit que pourraient en tirer les usagers semblent inversement proportionnels:

- les oralisations posent des problèmes de production et de gestion informatique<sup>52</sup>, mais elles sont plus aisées à exploiter par des utilisateurs de dictionnaires que les transcriptions phonétiques, ce qui doit inciter à leur faire une place plus importante et à les rendre plus aisément accessibles;
- les transcriptions ne posent pas de problème de traitement que les éditeurs n'aient déjà résolu, mais elles sont malaisées à lire par des usagers qui ne maîtrisent pas tous le décodage des séquences en alphabet phonétique<sup>53</sup>, ce qui suggère de toujours les associer aux oralisations de manière à améliorer la capacité des lecteurs à les décoder;

---

<https://www.gwmicro.com/Window-Eyes/>) sont des logiciels permettant aux non-voyants d'utiliser un ordinateur via la lecture de ce qui est à l'écran par un outil de synthèse vocale ou de conversion en braille.

<sup>49</sup> Voir des segments des textes externes à partir desquels le dictionnaire serait consulté par hyperappel.

<sup>50</sup> *Mediadico*, dictionnaire de français en ligne, propose une version oralisée de l'intégralité de ses articles, ce qui répond à un objectif différent de celui envisagé ici.

<sup>51</sup> Certains didacticiels de langues, comme *Tell me more*, proposent d'analyser les oralisations des usagers (cf. [http://fr.tellmemore.com/boutique\\_particuliers\\_france/tell\\_me\\_more\\_v10\\_france/tell\\_me\\_more\\_universal/tell\\_me\\_more\\_universal\\_francais#](http://fr.tellmemore.com/boutique_particuliers_france/tell_me_more_v10_france/tell_me_more_universal/tell_me_more_universal_francais#)).

<sup>52</sup> Les données audio, qui occupent beaucoup de place en mémoire, ralentissent les interfaces de consultation.

<sup>53</sup> Landau en infère d'ailleurs un manque d'intérêt pour la prononciation (cf. n. 2) qui n'est peut-être pas tant fondé sur l'objet que sur son mode de description.

– les modes d'accès aux formes phoniques (oralisations et transcriptions) semblent actuellement fondés sur l'exploration des transcriptions phonétiques ou sur des liens préenregistrés, ce qui implique la systématisation de leur présence, une gestion adaptée des éventuels remplacements d'enregistrements par de la parole synthétisée et une intégration de fonctions de recherche basées sur la reconnaissance vocale qui rendent l'ensemble des descriptions de formes phoniques du dictionnaire accessibles.

Les pistes d'amélioration des descriptions et de l'accès envisagées ci-dessus présentent des implications documentaires ou de développement qui paraissent constituer des charges qu'aucun éditeur ne serait susceptible de supporter seul dans le contexte économique actuel. En conséquence, si un prototype devait être élaboré, il semble que sa conception devrait mobiliser, dans le cadre d'un partenariat de recherche public-privé, des phonéticiens, des linguistes informaticiens spécialisés en synthèse et en reconnaissance vocale et des informaticiens éditoriaux concepteurs d'interfaces de consultation en plus des lexicographes.

## Références

### Dictionnaires

- Dictionnaire Auzou junior*, avec CD-ROM PC / Mac OS, Paris: Éditions Philippe Auzou, 2008.
- Dictionnaire des mots difficiles à lire* = Presse N. (2004). *Dictionnaire des mots difficiles à lire*, Paris: Rue des écoles, 2004.
- Encarta Dictionnaire français* [en ligne]. [Microsoft Corporation, 2002-2009] <http://fr.encarta.msn.com/encnet/features/dictionary/dictionaryhome.aspx> [consulté le 25 février 2010].
- Encarta U.S. English Dictionary* [en ligne]. [World English Dictionary [North American Edition], Microsoft Corporation. Developed for Microsoft by Bloomsbury Publishing Plc, 2009.] <http://g.msn.com/OENNAELF/EncartaDictionary>. [consulté le 25 février 2010].
- Martinet A. & Walter H. (1973). *Dictionnaire de la prononciation française dans son usage réel*, Paris: France Expansion.
- Mediadico* [en ligne]. [Mediadico, 1987-2010] <http://www.mediadico.com/>. [consulté le 25 février 2010]
- Mon dictionnaire pour la grande école* = Échaudemaison B.-P. (dir.) (2009). *Mon dictionnaire pour la grande école*, Saint-Cloud: Éclairs de Plume.
- Petit Robert = Le Nouveau Petit Robert. Dictionnaire alphabétique et analogique de la langue française*, millésime 2009, Paris: Dictionnaires Le Robert - SEJER, 2008.
- Petit Robert des enfants* = Rey-Debove J. (dir.) (1988). *Le Petit Robert des enfants. Dictionnaire de la langue française*, Paris: Dictionnaires Le Robert.
- Petit Robert électronique = Le Nouveau Petit Robert de la langue française. Dictionnaire alphabétique et analogique de la langue française*, PC / Mac, version 2.0, Paris: Dictionnaires Le Robert - VUEF / Bruxelles: Bureau van Dijk, 2001. Nouv. éd.: millésime 2007, version 3.0, Dictionnaires Le Robert - SEJER / Bureau van Dijk, 2006; millésime 2008, version 3.1, 2007; millésime 2009, version 3.2, 2008.
- Robert junior = Le Robert junior illustré*, Paris: Dictionnaires Le Robert, 1993.
- Robert junior électronique = Le Robert junior*, CD-ROM PC, version 1.0, Paris: Dictionnaires Le Robert - Havas Interactive / Bruxelles: Bureau van Dijk, 1998; version 1.1, PC / Mac, 1999; nouv. pressage, 2003; version 1.2, *Le Robert des enfants*, Dictionnaires Le Robert - SEJER / Bureau van Dijk, 2006.
- Robert oral-écrit* = Taulelle D. (dir.) (1989). *Le Robert oral-écrit. L'orthographe par la prononciation*, Paris: Dictionnaires Le Robert.
- Trésor de la langue française informatisé* [en ligne]. [CNRS, version 4, 2002], <http://atilf.atilf.fr/tlfv3.htm> [consulté le 25 février 2010]; CD-ROM PC / Mac OS + livre d'accompagnement, Paris: CNRS Éditions, 2004.
- Webster's Third New International Dictionary of the English Language*, Springfield, G. & C. Merriam, 1961; [en ligne] <http://www.merriam-webster.com/> [consulté le 25 février 2010].

### Autres références

- Corbin F. (2009). 'L'iconographie du *Robert junior* (1993-2004)', Dans *Lexique* 19. 69-115.
- Hulbert J. R. (1968). *Dictionaries: British and American*, éd. revue. London: André Deutsch.
- Landau S. (1989). *Dictionaries. The art and craft of lexicography*, Cambridge (U.K.): Cambridge University Press; 1<sup>e</sup> éd., 1984; nouv. éd., 2001.
- Léon P.; Léon M.; Léon F. & Thomas A. (2009). *Phonétique du FLE. Prononciation: de la lettre au son*, coll. U, Paris: Armand Colin.
- Svensén B. (1993). *Practical Lexicography. Principles and methods of dictionary-making*, traduit du suédois par J. Sykes & K. Schofield, Oxford: Oxford University Press.