

Plateforme de traitement multimodal des symptômes aphasiques

Yves Rybarczyk*, Ricardo Martins*
José Fonseca**

*Département d'Ingénierie Electronique, Université Nouvelle de Lisbonne, Portugal
y.rybarczyk@fct.unl.pt
<http://www.uninova.pt/~yr>

**Laboratoire de Recherche sur le Langage, Faculté de Médecine de Lisbonne, Portugal
jfonseca@fm.ul.pt
<http://imm.fm.ul.pt>

Résumé. Le système décrit dans cet article a été conçu pour le traitement et l'analyse des performances de patients aphasiques. Le fait que ce syndrome puisse affecter différentes modalités fonctionnelles (compréhension et expression orale et écrite) a nécessité l'élaboration d'un logiciel multimédia. Ainsi, le matériel clinique, développé par des orthophonistes, est présenté à partir de stimuli auditifs, écrits ou d'environnements virtuels, en fonction de leur pertinence pour la tâche choisie. Les résultats des exercices intègrent des contenus multimodaux et sont essentiellement enregistrés sous forme de fichiers XML, pour une meilleure flexibilité de traitement des données.

1 Introduction

1.1 Caractéristiques de l'aphasie

L'aphasie se caractérise par une perte ou déficience des fonctions langagières, telles que l'expression orale, la compréhension auditive, la lecture et l'écriture. Elle est provoquée par des lésions au cerveau. Les troubles de l'aphasie qui altèrent la programmation motrice, impliquée dans la parole et l'écriture, sont localisés dans les aires cérébrales antérieures. En revanche, les lésions postérieures produisent des troubles de la compréhension auditive et de la lecture (voir Damasio, 1991, pour une revue de synthèse). Le syndrome aphasique peut être plus ou moins sévère. Dans les jours qui suivent le traumatisme, les patients ne peuvent généralement qu'accomplir de simples tâches d'identification d'objet. Par la suite, et à force d'entraînement, la majorité d'entre eux parviennent à comprendre et construire des phrases de complexité grammaticale croissante.

1.2 Techniques de réhabilitation

1.2.1 Approche traditionnelle

De nombreuses études ont démontré l'efficacité de l'orthophonie dans la récupération des troubles aphasiques (Basso, 2003). Le traitement traditionnel se base principalement sur la

stimulation du patient à répondre à des questions ou des ordres donnés par l'orthophoniste. En effet, un facteur important dans la récupération de l'aphasie est la répétition de tâches linguistiques (Bhogal et al., 2003). Des tests avec papier-crayon et/ou utilisation d'images bidimensionnelles (qui illustrent des objets et des actions) sont le matériel le plus communément employé dans ce type de thérapie. Cependant, ces représentations rudimentaires et parfois abstraites du monde physique limitent l'évaluation des réelles capacités de compréhension et d'expression du patient. Des études ont montré, par exemple, que la discrimination visuelle des objets est plus difficile quand elle est basée sur une représentation bidimensionnelle plutôt que sur une vue tridimensionnelle plus écologique ; particulièrement chez les personnes ayant un plus faible niveau de scolarité (Reis et al., 2001).

1.2.2 Approche nouvelles technologies

A partir des années 80, des logiciels ont commencé à être utilisés dans le cadre du traitement de l'aphasie, mettant en place des programmes informatiques aussi bien pour des objectifs ciblés que pour une réhabilitation linguistique globale (Fonseca et al., 1997). Plusieurs travaux ont pu démontrer à quel point les technologies de l'information sont efficaces dans la récupération des troubles de l'aphasie (Robertson, 1990). Cependant, leur diffusion et la banalisation de leur utilisation est encore loin d'être un processus courant. En effet, les nouvelles technologies développées jusqu'ici ont l'inconvénient de présenter un coût relativement élevé et d'utiliser des équipements trop spécialisés pour être employés à grande échelle et en dehors la structure hospitalière. De plus, aucune d'entre elles ne permettent un traitement général des différents symptômes aphasiques, oubliant notamment la modalité auditive et/ou présentant des stimuli de constructions grammaticales trop basiques (Sik Lányi et al., 2006).

L'objectif de cet article est de présenter une nouvelle génération de plateforme multimédia pour le traitement de l'aphasie. Afin de répondre aux différences interindividuelles de chaque cas clinique, notre outil se propose d'entraîner le patient sur les quatre modalités pouvant être impliqués dans le syndrome aphasique, à savoir : la lecture, la compréhension orale, l'expression verbale et l'écriture. Par ailleurs, une préoccupation particulière a été portée au niveau de l'interface homme-machine de manière à s'adapter au plus grand nombre d'utilisateurs (sachant que ce syndrome est souvent associé à d'autres lésions cérébrales touchant, notamment, les aires motrices), tout en utilisant un matériel relativement ordinaire afin de ne pas trop élever le coût de la plateforme. Finalement, un large éventail d'exercices de degrés de complexité variables a été implémenté dans le souci d'accompagner chaque phase de récupération de la maladie. Les résultats dans les différentes modalités sont enregistrés quasi exclusivement sous un format XML, afin d'assurer une meilleure flexibilité dans le traitement ultérieur des données.

2 Le logiciel multimédia développé

2.1 Contexte d'utilisation

Le logiciel développé a pour ambition d'améliorer la réhabilitation des patients aphasiques à travers un environnement virtuel convivial et adapté aux nécessités quotidiennes de leurs utilisateurs. En outre, il assure un recueil et une systématisation d'un grand nombre

d'informations relatives aux performances obtenues dans les diverses sessions de traitement, afin de permettre une analyse postérieure des données. Bien que le malade puisse utiliser l'outil de forme indépendante, l'assistance d'une tierce personne peut-être nécessaire, du moins lors des premières sessions d'utilisation. L'outil est d'ailleurs conçu pour pouvoir être utilisé aussi bien dans le cadre domestique (sous la supervision d'un proche) que dans le contexte hospitalier (sous l'orientation d'un orthophoniste). Cette dernière situation est facilitée par la présence d'une interface permettant l'enregistrement discriminé de différents patients dans une base de données. Après que l'individu se soit enregistré, il est invité à initier une tâche dans l'une des quatre modalités symptomatiques de l'aphasie : compréhension écrite, compréhension orale, expression écrite et expression orale. Chacun de ces thèmes sont eux-mêmes subdivisés en sous-thèmes relatifs à la modalité en question. Du point de vue du matériel clinique, le logiciel charge des supports multimédias tels que des environnements 3D immersifs, des images vidéo, de la synthèse vocale, des mots, phrases et textes écrits.

2.2 La compréhension écrite (ou lecture)

Cette modalité est constituée par cinq types d'exercice. Trois d'entre eux sont présentés sous la forme d'une pièce de la maison, modélisée en 3D : « appairer un nom », « appairer une phrase » et « phrases oui/non ». Les deux autres sont simplement diffusés sous la forme de textes écrits : « responsive naming » et « mot intrus ».

2.2.1 A partir d'un environnement tridimensionnel

Pour ces exercices l'utilisateur est amené à choisir la division de la maison où il désire s'entraîner (figure 1). Une fois la pièce choisie, il va pouvoir se déplacer à l'intérieur afin d'exécuter des ordres présentés de manière écrite. Dans le cas d'un « appariement de nom », un nom d'objet apparaît à l'écran et l'individu doit pointer/cliquer sur l'objet en question (figure 2). Lorsqu'il s'agit d'un « appariement de phrase », l'objet à découvrir est indiqué sous la forme d'une phrase définissant l'objet. Par exemple, si l'ordre « endroit où l'on s'assoit pour regarder la télévision » apparaît, le patient doit pointer/cliquer sur le « fauteuil ». Enfin, pour un exercice « phrases oui/non », une phrase surgie à l'écran et l'utilisateur doit cliquer sur « oui » ou sur « non » selon que l'affirmation est vraie ou fausse par rapport à la réalité de la pièce. Un exemple de phrase possible est : « la fenêtre est ouverte... ».



FIG. 1 – Choix de la pièce.



FIG. 2 – Représentation 3D de la cuisine.

Plateforme multimédia pour l'aphasie

Également, le patient peut recourir à des aides visuelles et/ou sonores lorsqu'il rencontre des difficultés pour réaliser la tâche demandée. Ces aides sont accessibles à travers un bouton et sont représentées sous la forme de photos d'objets, lorsqu'il s'agit d'indices visuels, et de prononciation de noms d'objets, quand elles se réfèrent à des indices sonores. A la fin de la série de stimuli qui compose l'exercice, une fenêtre présente à l'écran la performance du sujet ainsi la quantité d'aides à laquelle il a eu recours.

2.2.2 A partir d'un texte écrit

Pour une tâche du type « responsive naming », une phrase incomplète et quatre mots possibles apparaissent à l'écran (figure 3). Le patient doit choisir l'un des quatre mots de manière à compléter la phrase de forme logique. Finalement, dans la recherche du « mot intrus », cinq mots sont diffusés. A peine quatre mots ont une relation sémantique entre eux (figure 4). L'individu doit cliquer sur le mot qui n'a rien à voir avec les autres. Dans ces tâches comme dans les précédentes, on peut recourir à une aide qui fournit des indices visuels et/ou sonores pour résoudre l'exercice. A la fin d'une session, l'utilisateur reçoit un feedback sur ces résultats.

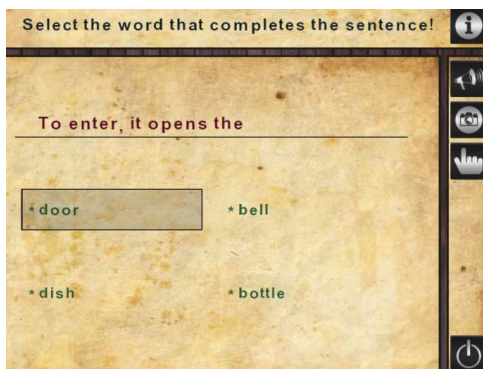


FIG. 3 – Responsive naming (aides à droite)

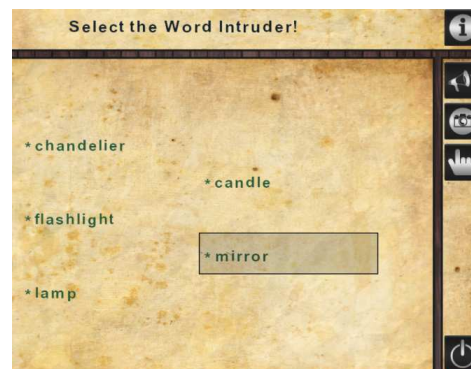


FIG. 4 – Mot intrus (aides sur la droite)

2.3 La compréhension orale

L'aphasie peut également affecter la modalité auditive. Pour cela, une seconde thématique s'attache à travailler la compréhension verbale du sujet. Ce thème est divisé en trois sous-thèmes qui sont : « identification d'objets », « phrases oui/non » et « ordres simples ». La série d'exercices diffusés dans ce cadre est relativement similaire à celle du chapitre précédent. La principale différence est que les tâches ne sont plus basées sur des stimuli écrits mais sonores (synthèse vocale). Ici, tous les exercices sont effectués dans un environnement virtuel représentant l'une des pièces de la maison choisie par l'utilisateur. Lors d'une « identification d'objet », le patient entend le nom d'un objet et doit pointer/cliquer à l'endroit où se trouve l'objet. Dans les tâches du type « phrases oui/non », il est demandé de répondre par « oui » ou « non » à une phrase émise sous la forme d'une voix de synthèse, selon que l'affirmation est correcte ou incorrecte. Enfin, les « ordres simples » sont des phrases sonores qui font référence, de manière indirecte, à un objet que le patient doit localiser dans la pièce. Un exemple d'ordre simple est : « quand on l'ouvre, il en sort de l'eau ». Une aide est éga-

lement disponible afin de fournir, si nécessaire, des indices visuels et/ou écrits pour assister le malade dans sa tâche.

2.4 L'expression écrite (ou écriture)

C'est à l'heure actuelle le thème le plus développé dans notre outil. Il s'appuie sur du matériel clinique reposant sur de multiples supports d'information. De forme synthétique les stimuli sont basés sur trois types de supports : « modèles 3D », « séquences vidéos » et « information écrite ».

2.4.1 A partir d'un objet tridimensionnel

Un objet de la vie quotidienne, modélisé en 3D, est diffusé à l'écran. Le patient doit écrire au clavier de l'ordinateur le nom de l'objet en question (figure 5). Une gamme diversifiée d'aides est accessible à l'utilisateur. Ainsi, il a la possibilité d'entendre le nom de l'objet ou de voir une photographie de celui-ci. Du point de vue de la composition des mots, des indices indiquent le nombre de lettres présent dans le mot ou encore le prochain caractère qui forme le mot. Pour les individus qui nécessitent de plus d'assistance, une option permet de présenter trois possibilités dont une seule est la bonne réponse.

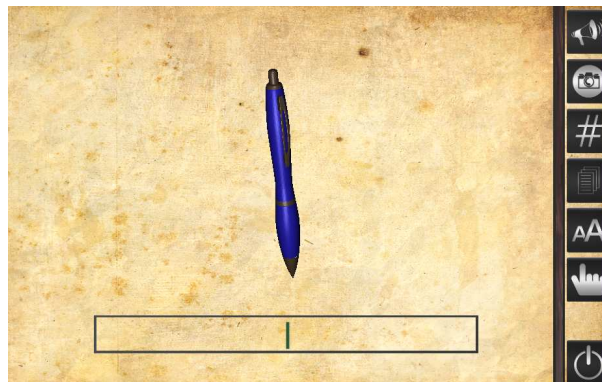


FIG. 5 – *Ecrire un substantif. Sur la droite de l'image, des icônes représentent les aides qui sont à disposition de l'utilisateur. Ainsi, de haut en bas, les fonctions disponibles sont respectivement : entendre le nom de l'objet, voir une photo de l'objet, quantité de lettres qui forme le nom, trois noms possibles, prochain caractère qui compose le nom, la réponse à l'exercice et l'option de quitter la session.*

2.4.2 A partir d'une séquence vidéo

Une séquence vidéo d'une personne réalisant une action est projetée à l'écran. La tâche du patient est d'écrire le verbe à l'infinitif qui correspond à cette action. Ainsi, si la séquence met en scène un individu en train de marquer un numéro de téléphone, l'utilisateur devra écrire le verbe « téléphoner ». Certaines scènes peuvent amener à une interprétation plus ou moins subjective de l'action. Dans ces situations, la modélisation du traitement des données permet d'accepter plusieurs réponses.

2.4.3 A partir d'une information écrite

Ici, il faut compléter/restructurer, sous la forme correcte, une phrase ou un paragraphe entier. Pour ce qui est de compléter une phrase, il peut s'agir d'insérer un nom, un verbe conjugué ou une préposition. Dans ce genre de tâche, une phrase apparaît avec un blanc qui doit être rempli par le patient. En ce qui concerne la restructuration de phrases, elle peut impliquer i) une correction d'erreurs phonologiques (certains mots sont mal orthographiés) ou ii) une réécriture de la phrase à travers un repositionnement des mots de manière à construire une phrase sémantiquement correcte. Pour finir, des tâches plus complexes demandent au sujet d'introduire des mots de nature quelconque (noms, verbes, adjectifs, ...) dans un texte de la taille d'un paragraphe.

2.5 L'expression orale

Cette modalité est la plus complexe à implémenter. En effet, pour fonctionner en mode complètement automatique, elle nécessite d'un programme de reconnaissance de la parole. Or, malgré les progrès considérables acquis dans ce domaine, la reconnaissance vocale souffre encore de difficultés de mise en œuvre et de problèmes de fiabilité (Pogue, 2010). C'est pourquoi la version de la plateforme présentée dans cet article ne permet que d'enregistrer la réponse verbale du sujet, en vue d'une analyse postérieure. Ceci étant, deux types d'exercices sont actuellement implémentés dans le software : « nommer des objets » et « nommer des actions ». Dans le premier cas, le patient doit prononcer, dans le microphone, le nom de l'objet 3D qui apparaît à l'écran. Dans l'autre cas, il doit exprimer oralement le verbe de l'action visionnée à partir d'une séquence vidéo.

3 Architecture software de la plateforme

L'architecture de la plateforme multimédia est organisée en quatre couches qui sont : les dispositifs d'entrées/sorties, le moteur de jeu, les scripts Python et le système de gestion des données (figure 6).

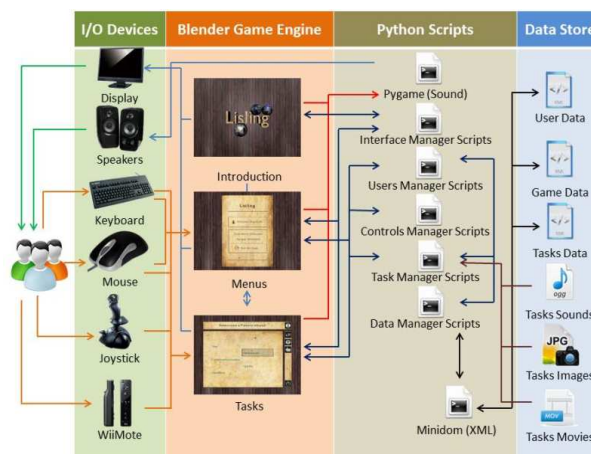


FIG. 6 – Modules d'architecture de la plateforme de traitement de l'aphasie.

3.1 Dispositifs I/O

Le fait que des déficiences motrices soient souvent associées au syndrome aphasique, a obligé à réfléchir à des interfaces de contrôles adaptées au plus grand nombre. Ces commandes doivent être les plus naturelles et intuitives possibles, car des utilisateurs non familiarisés avec l'informatique pourront également être amenés à utiliser la plateforme. Enfin, le coût matériel a aussi été un critère de choix. Ainsi, en plus des traditionnels systèmes d'inputs rencontrés sur quelconques ordinateurs (clavier et souris), un joystick et une commande Wii ont été interfacés au logiciel. Le joystick s'est avéré essentiel pour faciliter la navigation à l'intérieur des environnements 3D. La commande Wii, quand à elle, peut fournir un bénéfice en termes de liberté de mouvement, tout en conservant une bonne précision de pointage. En revanche, ce dernier dispositif nécessite davantage de configurations pour s'interfacer à l'ordinateur du patient (barre de leds pour la détection infrarouge). En termes d'outputs, l'écran et les hauts parleurs de l'ordinateur permettent de diffuser respectivement le matériel clinique visuel et sonore utilisés dans la thérapie.

3.2 Moteur de jeu Blender

La modélisation des objets qui constituent les environnements 3D des pièces de la maison a été réalisée avec le logiciel open source Blender. Ce software a l'avantage d'incorporer en un seul outil la modélisation, l'animation, l'édition vidéo, la composition de scènes, la création d'applications interactives et un moteur de jeu. Le moteur de jeu de blender permet notamment de détecter les collisions et reproduire des phénomènes physiques tels que la gravité, les forces de friction, la déformation des objets... utiles pour parfaire le réalisme des interactions avec les modèles tridimensionnels. Dans notre application, ce moteur est également impliqué dans la création/gestion du fichier des menus et gestion des utilisateurs, ainsi que des fichiers spécifiques à chaque tâche.

3.3 Scripts Python

Bien que le moteur de jeu Blender soit assez développé, des scripts programmés en Python sont nécessaires pour ajouter certaines fonctionnalités propres à la plateforme de réhabilitation. Ainsi, les scripts sont impliqués dans les opérations suivantes :

- Gestion des menus (liaison des fichiers Blender aux scènes qui les composent) ;
- Gestion des utilisateurs (création/actualisation d'utilisateurs et enregistrement de leurs données respectives) ;
- Gestion des tâches (chargement des données relatives aux exercices, validation des réponses et présentation des informations liées aux actions et aides invoquées) ;
- Gestion des commandes (réception des inputs des divers dispositifs de contrôle et traduction de ces valeurs en actions conformément à l'exercice en question).

3.4 Gestion des données

Trois types de données sont créés par l'application : les données d'utilisateurs, les données de sessions et les données de tâches. Les données d'utilisateurs se réfèrent à l'identification, aux informations personnelles et préférences de chaque patient enregistré.

Plateforme multimédia pour l'aphasie

```
<SESSION>
  <patient_id>2</patient_id>
  <session_id>5</session_id>
  <TASK>
    <task_id>3</task_id>
    <task_name>Compréhension Ecrite</task_name>
    <exercice_name>Phrase Oui/Non</exercice_name>
    <theme>Cuisine</theme>
```

FIG. 7 – Structure XML des données de sessions pour les blocs <SESSION> (identification de l'utilisateur et du numéro de session) et <TASK> (identification de la tâche).

Les données de sessions regroupent toutes les actions et tâches réalisées par le patient, qui sont enregistrées sous forme d'un fichier XML. La figure 7 présente les principaux éléments contenus sous la balise <TASK>, à savoir le nom de la tâche, le nom de la sous-tâche et la pièce dans laquelle se déroule l'exercice. Le bloc <OBJECTIVE> se réfère à chaque objectif demandé à l'utilisateur (dans la figure 8, il s'agit de localiser un 'verre'). Pour chaque objectif, un utilisateur peut donner différents types de réponses, pouvant être correctes ou fausses. Ces tentatives sont enregistrées dans la structure <ATTEMPT> (dans la figure 8, la réponse donnée est effectivement 'verre', d'où l'enregistrement du résultat comme étant 'correct'). Enfin, la structure <HELP> mémorise toutes les aides utilisées par le patient (sur l'exemple ci-dessous, la diffusion d'une photographie de l'objet à rechercher a été demandée).

```
<OBJECTIVE>
  <objective_id>verre</objective_id>
  <ATTEMPT>
    <att_response>verre</att_response>
    <att_result>correct</att_result>
    <HELP>
      <help_name>photographie</help_name>
    </HELP>
  </ATTEMPT>
</OBJECTIVE>
</TASK>
</SESSION>
```

FIG. 8 – Structure XML des données de sessions pour les blocs <OBJECTIVE> (e.g., objet à rechercher), <ATTEMPT> (réponse du patient) et <HELP> (aides demandées).

Il existe un fichier XML indépendant pour chaque tâche implémentée dans le logiciel. Ces fichiers contiennent les données, insérées au préalable, de tous les objectifs et de leurs réponses attendues. De cette manière, il est aisé d'insérer de nouveaux objectifs, puisqu'il suffit pour cela d'ajouter la structure correspondante au fichier XML de gestion des tâches. Les paragraphes suivants décrivent des exemples types de ces fichiers pour trois tâches différentes.

Phrases oui/non. Le fichier de ce type d'exercice est commun pour les tâches de compréhension écrite et orale. Chacun des objectifs est inséré dans une structure qui fait référence à une division de la maison (dans l'exemple présenté sur la figure 9, il s'agit de la salle de

séjour). Un tag spécifique (<question_id>) identifie l'objectif, ce qui permet, par la suite, de croiser les données entre les réponses fournis par l'utilisateur et les réponses attendues par le système.

```
<SALLE_SEJOUR>
  <YESNO>
    <question_id>5</question_id>
    <question_text>La porte fenêtre est-elle ouverte ?</question_text>
    <question_sound>sn5.ogg</question_sound>
    <answer>Non</answer>
  </YESNO>
</SALLE_SEJOUR>
```

FIG. 9 – Structure XML d'un fichier de tâche « Phrase oui/non ». Les tags <question_text> et <question_sound> permettent de diffuser respectivement la question sous forme textuelle ou sonore. Le tag <answer> indique quelle est la réponse exacte.

Ecrire un verbe à l'infinitif. Il s'agit d'écrire le verbe à l'infinitif correspondant à l'action présentée à partir d'un clip vidéo. La structure <ANSWER> fait le listage d'une série de plusieurs réponses possibles (la structure XML de la figure 10 correspond à un film où un individu ouvre sa porte à un autre). Cette tâche peut également être réalisée sous forme de choix multiples.

```
<MOVIE>
  <id>3</id>
  <film>vi3.flv</film>
  <ANSWER>
    <response>ouvrir</response>
    <response>accueillir</response>
    <response>recevoir</response>
  </ANSWER>
  <choix_multiples>fermer,ouvrir,battre</choix_multiples>
</MOVIE>
```

FIG. 10 – Structure XML d'un fichier de tâche « Ecrire un verbe à l'infinitif ».

Ecrire un nom d'objet. Il s'agit d'écrire le nom d'un objet diffusé sous la forme d'un modèle 3D. Ici encore, plusieurs réponses sont possibles. La figure 11 représente la structure XML pour l'objet stylo (<obj_name> et <obj_sound> pour identifier respectivement le modèle et le son chargé). Différentes natures d'aides sont disponibles, telles que des propositions (<help_multiple>), le son de la première syllabe (<help_1_sound>), première syllabe du mot (<help_1_sil>), nombre de lettres (<help_n_letters>) ou photo de l'objet (<help_picture>). Le bloc <ANSWERS> liste toutes les réponses considérées comme correctes.

```
<OBJECT>
  <obj_id>2</obj_id>
  <obj_name>stylo</obj_name>
  <obj_sound>stylo.ogg</obj_sound>
  <help_multiple>stylo,brosse à dents,couteau</help_multiple>
  <help_1_sound>stylo_1.ogg</help_1_sound>
```

Plateforme multimédia pour l'aphasie

```
<help_1_sil>st</help_1_sil>  
<help_n_letters>*****</help_n_letters>  
<help_picture>stylo.jpeg</help_picture>  
<ANSWERS>  
  <ans>stylo</ans>  
  <ans>stylo-bille</ans>  
  <ans>stylo-feutre</ans>  
  <ans>feutre</ans>  
</ANSWERS>  
</OBJECT>
```

FIG. 11 – Structure XML d'un fichier de tâche « Ecrire un nom d'objet ».

Références

- Basso, A. (2003). *Aphasia and its Therapy*. New York, NY: Oxford University Press.
- Bhagal, S.K., R. Teasell, et M. Speechley (2003). Intensity of aphasia therapy, impact on recovery. *Stroke*, 34: 987-993.
- Damasio, H. (1991). Neuroanatomical correlates of the aphasias. *M.T. Sarno (Ed.), Acquired Aphasia*. San Diego, CA: Academic Press, 45-71.
- Fonseca, J., G. Leal, L. Farrajota, J. Carneiro, A. Castro-Caldas (1997). Lising – Multimedia program for aphasics training. *G. Anogianaki, C. Bühler and M. Soede (Ed.), AAATE*. Chalkidike, Greece.
- Pogue, D. (2010). Talk to the machine. *Scientific American*, 303: 18.
- Reis, A., K.M. Petersson, A. Castro-Caldas, M. Ingvar (2001). Formal schooling influences two-but not three-dimensional naming skills. *Brain and Cognition*, 47: 397-411.
- Robertson, I. (1990). Does computerized cognitive rehabilitation work? *Aphasiology*, 4: 381-405.
- Sik Lányi, C., Z. Geiszt, V. Magyar (2006). Using IT to inform and rehabilitate aphasic patients. *Informing Science Journal*, 9: 163-179.

Summary

The system described in this article is designed for the treatment and performance analysis of aphasic patients. Since this syndrome affects different functional modalities (written and spoken comprehension and expression) an implementation of a multimedia software has been necessary. Thus, the clinical material – developed by speech and language therapists – is displayed in the form of auditive, writing or virtual environments stimuli, following their relevance to the chosen task. Exercise results integrate multimodal contents and are essentially recorded into XML files, for a better flexibility of the data processing.