

# Étude de la lisibilité des documents de santé avec des méthodes d'oculométrie

Natalia Grabar<sup>1</sup> Emmanuel Farce<sup>3</sup> Laurent Sparrow<sup>3</sup>

(1) CNRS, UMR 8163 - STL - Savoirs Textes Langage, Univ. Lille, F-59000 Lille, France

(2) Univ. Lille, CNRS, UMR 9193 - SCALab, F-59000 Lille, France

natalia.grabar@univ-lille3.fr, emmanuel.farce@univ-lille3.fr,  
laurent.sparrow@univ-lille3.fr

## RÉSUMÉ

---

Le domaine médical fait partie de la vie quotidienne pour des raisons de santé, mais la disponibilité des informations médicales ne garantit pas leur compréhension correcte par les patients. Plusieurs études ont démontré qu'il existe une difficulté réelle dans la compréhension de contenus médicaux par les patients. Nous proposons d'exploiter les méthodes d'oculométrie pour étudier ces questions et pour détecter quelles unités linguistiques posent des difficultés de compréhension. Pour cela, des textes médicaux en version originale et simplifiée sont exploités. L'oculométrie permet de suivre le regard des participants de l'étude et de révéler les indicateurs de lecture, comme la durée des fixations, les régressions et les saccades. Les résultats indiquent qu'il existe une différence statistiquement significative lors de la lecture des versions originales et simplifiées des documents de santé testés.

## ABSTRACT

---

### **Study of readability of health documents with eye-tracking methods.**

Medical area is integral part of our lives due to health concerns, but the availability of medical information does not guarantee its correct understanding by patients. Several studies addressed this issue and pointed out real difficulty in understanding of health contents by patients. We propose to use eye-tracking methods for studying further the issue and for detecting which linguistic units in health documents are problematic. For this, original and simplified versions of medical documents are exploited. Eye-tracking permits to follow the gaze of participants and to reveal reading indicators such as duration of fixations, regressions and saccades. The results indicate that there is statistically significant difference in reading of original and simplified versions of the health documents tested.

**MOTS-CLÉS :** Lisibilité des documents, compréhension, alphabétisation médicale, oculométrie.

**KEYWORDS:** Readability of documents, understanding, health literacy, eye-tracking.

---

## 1 Introduction

Le domaine médical est de plus en plus présent dans la vie quotidienne des citoyens, essentiellement pour des raisons de santé et de soins médicaux mais également parce que ce type d'information peut être rencontré dans les médias ou la littérature. Avec l'évolution de l'Internet, les informations médicales deviennent aussi accessibles et disponibles en ligne. Il a été par exemple noté que l'Internet est devenu la deuxième source d'information pour les patients, en se plaçant après les consultations chez les médecins (Pletneva *et al.*, 2011; Fox, 2011). Ainsi, jusqu'à 24 % de la population utilise

l'Internet au moins une fois par jour pour trouver des informations relatives à leur santé et, de manière plus générale, jusqu'à 80 % de la population recourt à l'Internet pour trouver ce type d'informations. Néanmoins, la disponibilité des informations de santé ne présume pas qu'elles soient compréhensibles et correctement utilisées par les patients. Comme tous les domaines de spécialité, le domaine médical utilise des termes qui véhiculent des notions complexes, comme par exemple *tenosynovite*, *arthralgies* ou *métatarsophalangien*. Cette situation a été observée dans plusieurs études qui démontrent une faible compréhension des informations médicales par les citoyens (Patel *et al.*, 2002; Williams *et al.*, 1995; Berland *et al.*, 2001) et une communication compliquée entre les patients et les médecins (Jucks & Bromme, 2007; Tran *et al.*, 2009).

En relation avec ces questions, la complexité et la compréhension des textes et des termes sont étudiées par les chercheurs de différentes disciplines. Par exemple, la linguistique étudie la complexité morphologique des lexèmes (Iacobini, 2003; Lüdeling *et al.*, 2002), la psychologie étudie différents facteurs internes et externes aux mots qui influencent leur reconnaissance par les locuteurs (Bertram *et al.*, 2011; Lüttmann *et al.*, 2011; Bozic *et al.*, 2007; Dohmes *et al.*, 2004; Cain *et al.*, 2009), la terminologie étudie comment différencier les bons candidats-termes, qui véhiculent le sens bien spécialisé, au sein d'une extraction effectuée automatiquement (Kageura & Umino, 1996; Frantzi *et al.*, 2000; Hamon *et al.*, 2014), et le TAL propose des méthodes automatiques pour faire la différenciation entre les textes ou termes difficiles et ceux de la langue générale (Zeng *et al.*, 2005; Chmielik & Grabar, 2011; Shardlow, 2013).

L'objectif de notre travail consiste à étudier la compréhension des informations de santé par les non-experts. Plus particulièrement, nous proposons de voir quel est l'impact de la simplification des termes. Nous travaillons avec des données en français. Nous exploitons pour ceci les méthodes et les outils fournis par la psychologie, et plus particulièrement l'oculométrie (*eye-tracking*). En effet, l'étude des mouvements oculaires pendant la lecture fournit des indications objectives et précieuses sur les processus cognitifs impliqués. Plus particulièrement, la difficulté et la lisibilité d'un texte peuvent être mesurées avec différents indicateurs (Duchowski, 2007; Rayner, 1998; Sparrow *et al.*, 2003; Miellet *et al.*, 2008). Il s'agit principalement des indicateurs suivants :

- Les *saccades oculaires* sont des mouvements rapides des yeux pour aller d'un point de lecture vers un autre. Lorsque le texte est facile à lire et à assimiler, les saccades sont longues, alors qu'elles deviennent plus courtes avec un texte compliqué ;
- Les *fixations* sont des périodes pendant lesquelles les yeux sont stables. Les fixations correspondent aux moments lorsque l'information visuelle est analysée. La durée des fixations augmente avec des textes difficiles car ils nécessitent un temps d'assimilation plus important ;
- Les *régressions* sont les retours en arrière vers les endroits du texte déjà lus par le lecteur. La lecture d'un texte difficile engendre en général plus de régressions.

Ainsi, la comparaison de paramètres oculomoteurs (durée des fixations, amplitude des saccades, régressions, etc.) enregistrés pendant la lecture des textes va permettre d'évaluer plus précisément les difficultés et les points de blocage de la part des lecteurs. Selon notre hypothèse, la lecture de textes complexes et de termes inconnus conditionne notre attention et les mouvements oculaires présentent alors des comportements typiques et observables. De tels indices peuvent donc être directement corrélés avec les difficultés de compréhension.

Les indicateurs d'oculométrie sont exploités dans différents domaines. En psychologie et en neurosciences, ils permettent de collecter les informations sur le fonctionnement du cerveau, les modalités de la lecture et de l'attention, etc. (Cooper, 1974; Molnar, 1981; Asaad & andEK Miller, 2000; Rayner & Liversedge, 2004; Clifton *et al.*, 2007). En marketing et en publicité, ils fournissent les indicateurs qui permettent d'améliorer la présentation des informations, de mettre les bonnes informations aux

bons endroits ou de comprendre à quels types d'informations les utilisateurs et acheteurs sont les plus sensibles (Andrews & Coppola, 1999; Higgins *et al.*, 2014). Dans des contextes liés à l'acquisition et au traitement de la langue, les indicateurs d'oculométrie sont typiquement exploités pour détecter les endroits dans les textes qui attirent ou bloquent le regard ou la lecture, comme par exemple : la lecture de textes en langues maternelle et secondaire (Altarriba *et al.*, 1996; Bisson *et al.*, 2014); la lecture de textes par les personnes dyslexiques (Rubino & Minden, 1973; Elterman *et al.*, 1980; Rello *et al.*, 2013; Nilsson Benfatto *et al.*, 2016) ou autistes (Yaneva *et al.*, 2015); le traitement de structures syntaxiques (Frenck-Mestre & Pynte, 1997; Clifton & Staub, 2011; Trueswell *et al.*, 1994; Singh *et al.*, 2016); la relation entre le traitement de la parole et les mouvements oculaires, où les participants montrent la tendance de fixer l'image correspondant à la phrase qu'ils entendent à l'oral (Cooper, 1974; Tanenhaus *et al.*, 1995; Wendt *et al.*, 2014); la détection et le traitement des erreurs (Keating, 2009); l'évaluation de la complexité des textes lors de la traduction (Sharmin *et al.*, 2008) ou de l'acquisition de langues (Balakrishna, 2015). Le plus souvent, c'est l'information sur les fixations qui est exploitée. Par exemple, en relation avec le marketing et la publicité, la figure 1 montre une page de résultats d'un moteur de recherche en ligne. Les couleurs chaudes indiquent les endroits qui attirent le plus le regard des utilisateurs : les premiers résultats de la recherche et, en moindre mesure, la publicité sur le côté. La zone la plus fixée forme ce qu'on appelle le *golden triangle*. Ce modèle indique aussi que l'attention des utilisateurs décroît assez rapidement.

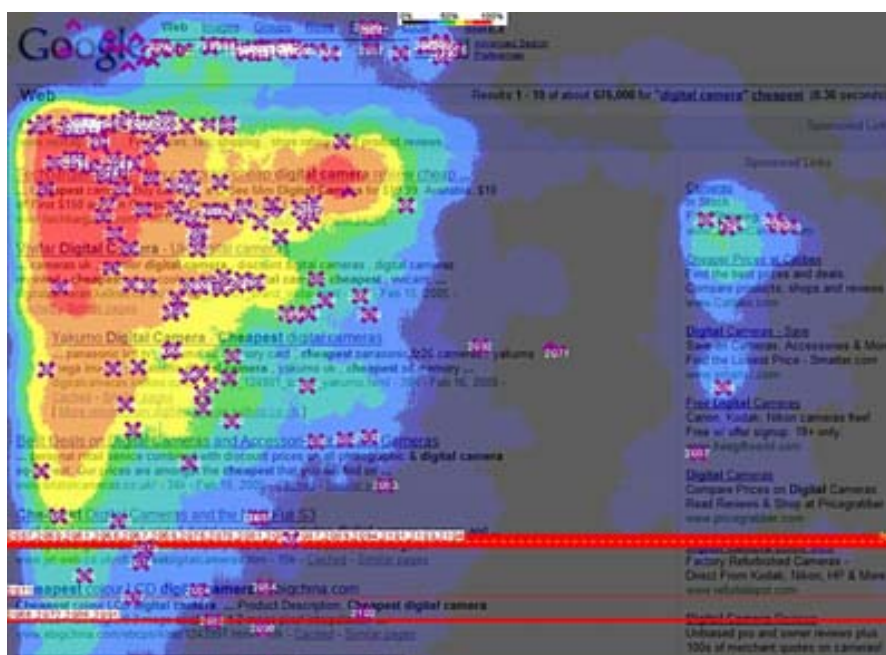


FIGURE 1 – Le modèle de fixations des pages de résultats d'un moteur de recherche.

## 2 Approche proposée

Notre approche est basée sur l'exploitation de méthodes d'oculométrie pour étudier les modalités de lecture de textes médicaux. Nous décrivons d'abord les documents utilisés et les critères d'inclusion des participants. Nous présentons ensuite le déroulement des tests et l'analyse des données obtenues.

## EXAMEN : ECHOGRAPHIE DES MAINS ET DES PIEDS

MOTIF : Bilan d'arthralgies

Mains : On ne visualise pas de ténosynovite, ou d'arthrosynovite.

Avant-pieds : On retrouve des remaniements intéressants les premières métatarsophalangiennes en rapport avec des antécédents de chirurgie d'hallux valgus.

Absence d'arthrosynovite au niveau des articulations métatarsophalangiennes.

## EXAMEN : ECHOGRAPHIE DES MAINS ET DES PIEDS

MOTIF : Bilan de douleurs articulaires

Mains : On ne visualise pas d'inflammation des tendons, ni de la membrane articulaire.

Avant-pieds : On retrouve des remaniements intéressants sur les premières articulations des pieds en rapport avec les antécédents de la chirurgie de la déformation du pied.

Absence d'inflammation de la membrane au niveau des articulations du pied.

FIGURE 2 – Texte<sub>1</sub> en version originale (en haut) et simplifiée (en bas) lu à l'étape 3 de la passation.

### 2.1 Documents cliniques et leur préparation

Deux extraits de documents cliniques sont utilisés, provenant d'un rapport de radiologie (figure 2) et d'une lettre de sortie en cardiologie (figure 3). Ces extraits sont exploités en deux versions : une version originale (technique) et une version simplifiée manuellement par un chercheur en TAL. La simplification est effectuée essentiellement au niveau lexical en utilisant les paraphrases de termes comme {arthralgie, douleur articulaire}, {ténosynovite, inflammation d'un tendon} ou {infarctus du myocarde, crise cardiaque} en partie issues de travaux antérieurs (Grabar & Hamon, 2016; Antoine & Grabar, 2016), ou bien des hyperonymes de termes {métatarsophalangien → pied}. Dans le texte<sub>1</sub>, 7 changements lexicaux sont effectués qui correspondent aux substitutions de quatre composés néoclassiques (arthralgie, ténosynovite, arthrosynovite, métatarsophalangien) et d'un emprunt latin ({hallux valgus → déformation du pied}). Dans le texte<sub>2</sub>, 10 changements sont effectués : substitutions par des équivalents ({infarctus du myocarde, crise cardiaque}, {entre 35 et 40 %, basse}), substitutions par des hyperonymes ({angioplastie → intervention chirurgicale}, {endoprothèse → un autre stent}, {IVA moyenne → artère cardiaque}, {circonflexe proximale → une autre artère}), et suppressions (antérieur, microcytaire et informations techniques sur les stents).

Les figures 2 et 3 présentent les versions originales et simplifiées des texte<sub>1</sub> et texte<sub>2</sub>, respectivement. Comme nous pouvons le voir, les versions originales et simplifiées restent équivalentes au niveau du sens. Il s'agit de textes courts (48 et 65 mots pour le texte<sub>1</sub> et 112 et 82 mots pour le texte<sub>2</sub>, version originales et simplifiées, respectivement) afin qu'ils puissent être facilement visualisés et lus sur un écran d'ordinateur. Ces textes sont exploités pour constituer deux ensembles de test :

1. texte<sub>1</sub> original et texte<sub>2</sub> simplifié,
2. texte<sub>1</sub> simplifié et texte<sub>2</sub> original.

### 2.2 Critères d'inclusion

Cinquante participants sont recrutés et chaque ensemble de test est lu par 25 participants, afin que la significativité statistique des paramètres de la lecture puisse être calculée entre ces deux versions des textes, tout en prenant en compte les éventuels problèmes techniques et de calibrage.

*Cette patiente avait constitué un infarctus du myocarde antérieur en novembre 2010, pour lequel avait été réalisée une angioplastie de l'IVA moyenne avec implantation d'un stent non actif Vision de 2.75 mm x 18 mm, un complément par angioplastie au ballon seul en aval. Une endoprothèse avait également été implantée au niveau de la circonflexe proximale, avec un stent Vision 2.5 x 18 mm. La fraction d'éjection était évaluée entre 35 et 40 %.*

*Nous l'avons revue récemment, en insuffisance cardiaque, avec plusieurs autres problèmes :*

- une anémie microcytaire inexpliquée,*
- un déséquilibre important de son diabète pour lequel elle a été, entre temps, prise en charge par nos confrères diabétologues.*

*Cette patiente avait présenté une crise cardiaque en novembre 2010, pour laquelle avait été réalisée une intervention chirurgicale de l'artère cardiaque avec implantation d'un stent non actif. Un autre stent avait également été implanté au niveau d'une autre artère. La fraction d'éjection observée était basse.*

*Nous l'avons revue récemment, en insuffisance cardiaque, avec plusieurs autres problèmes :*

- une anémie inexpliquée,*
- un déséquilibre important de son diabète pour lequel elle a été, entre temps, prise en charge par nos confrères diabétologues.*

FIGURE 3 – Texte<sub>2</sub> en version originale (en haut) et simplifiée (en bas) lu à l'étape 3 de la passation.

Peuvent être inclus dans l'étude :

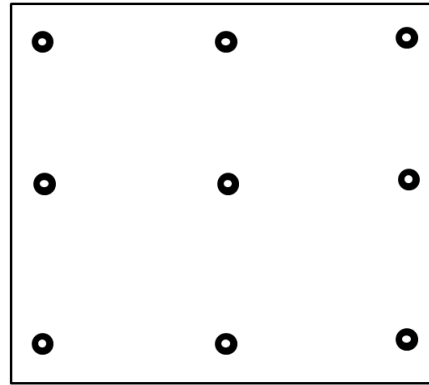
- les étudiants de L1 (niveau licence 1<sup>ère</sup> année) de différentes disciplines (psychologie, linguistique, arts plastiques...). Il s'agit d'étudiants venant de terminer leurs études secondaires et ayant suivi les premiers mois d'étude à l'université. Les étudiants poursuivant des études médicales et paramédicales sont exclus. En général, on distingue 5 niveaux d'alphabétisation (Bernèche & Perron, 2006), où les niveaux 1 et 2 correspondent aux personnes qui ont des difficultés de lecture et d'assimilation des informations, alors que les niveaux 4 et 5 fournissent la capacité de faire des déductions complexes, ce qui est souvent propre aux personnes ayant suivi des études supérieures. À notre avis, les étudiants L1 ont le plus souvent le niveau 3 d'alphabétisation : ils savent lire et écrire, ils peuvent assimiler de nouvelles informations et connaissances, mais n'ont pas de connaissances spécialisées y compris dans le domaine médical. Nous considérons ainsi que cette population peut représenter le citoyen lambda ;
- sans pathologies chroniques car cela peut impliquer une familiarité avec le domaine médical ;
- de langue maternelle française.

## 2.3 Déroulement des tests

L'approche proposée est basée sur l'exploitation d'oculométrie, qui permet d'évaluer la fluidité de la lecture avec des mesures objectives (Sparrow *et al.*, 2003; Miellet *et al.*, 2008). Comme indiqué plus haut, ces indicateurs permettent de détecter les zones de texte qui bloquent la lecture



(a) Caméra d'oculométrie



(b) Calibrage à 9 points

FIGURE 4 – Le caméra d'oculométrie et son calibrage avec 9 points de fixation (étape 1).

et la compréhension. Les textes sont présentés sur un écran d'ordinateur et une caméra spécifique enregistre les mouvements des yeux qui peuvent ensuite être superposés sur le texte.

Lors des passations de chaque participant, plusieurs étapes sont effectuées. Après la présentation des objectifs de l'étude et la collecte du formulaire de consentement, chaque participant effectue :

1. Le calibrage du caméra d'oculométrie (modèle EyeLink 1000, comme celui de la figure 4(a)). Une fois le participant installé et sa pupille repérée par le caméra, il est nécessaire d'entraîner le logiciel d'oculométrie aux mouvements des yeux de ce participant. Le calibrage à 9 points est alors effectué (figure 4(b)). Chaque point apparaît dans un ordre aléatoire sur l'écran et le participant est demandé de ne pas anticiper la localisation des apparitions mais de les suivre. Une première passe permet au logiciel d'apprendre les mouvements et positions oculaires alors que la deuxième passe permet de les précalculer, de les comparer avec les mouvements et positions réels, et de calculer ensuite la différence prédictible entre les deux positions ;
2. La lecture d'un texte général pour l'entraînement du participant. Il s'agit d'un texte pour enfant qui parle d'une petite fille qui habite dans la forêt. L'objectif de cette étape est d'entraîner le participant à la tâche de lecture et de réponse aux questions ;
3. La lecture des deux textes médicaux (figures 2 et 3), la version originale et la version simplifiée, comme décrit dans la section 2.1 ;
4. La lecture d'un texte de contrôle (figure 5) avec le contenu médical grand public qui concerne la cardiologie. L'objectif de cette étape est d'avoir une référence de lecture d'un texte médical grand public pour vérifier d'une autre manière l'objectivité des indicateurs d'oculométrie. On s'attend ici que les indicateurs de lecture soient stables avec les participants ;
5. Après la lecture de chaque texte, le participant répond à un QCM (questionnaire à choix multiples) pour évaluer sa compréhension. Par exemple, sur le texte<sub>2</sub>, une des questions posées est *La patiente présente les problèmes : 1) cardiaques, 2) cérébraux, 3) je ne sais pas.*

À la fin, si souhaité, les résultats obtenus avec un participant lui sont présentés et interprétés. Au total, la passation d'un participant dure entre 15 et 20 minutes. Les passations sont effectuées par un chercheur en TAL ou par un psychologue, en respectant les mêmes conditions expérimentales de laboratoire et sans distracteur.

*Le cœur est irrigué par les artères coronaires, qui sont elles-mêmes alimentées par une autre grande artère : l'aorte. Quand le diamètre des coronaires se rétrécit à cause de la formation progressive de plaques de graisse, le muscle cardiaque n'est plus assez alimenté en oxygène et en nutriments : il est en souffrance. Si l'artère se bouche complètement, l'infarctus guette... Les techniques du pontage ou de la pose d'un stent ont le même but : rétablir une circulation sanguine normale.*

FIGURE 5 – Texte de contrôle (étape 4).

## 2.4 Analyse des données

Les données collectées pendant les tests d'oculométrie sont analysées avec le *test t* pour calculer la significativité que nous observons au niveau des indicateurs de lecture des textes originaux et simplifiés, des textes d'entraînement (étape 2) et de contrôle (étape 4), et des réponses aux QCM.

Les données d'oculométrie sont aussi comparées avec les résultats d'annotation manuelle effectuée sur des données similaires. Il a été ainsi demandé aux annotateurs de marquer les segments de textes qu'ils ne peuvent pas comprendre. Nous supposons que les deux approches peuvent indiquer des unités linguistiques similaires, qui sont difficiles à comprendre, et que, contrairement aux annotations, dont le résultat correspond à des choix conscients, l'enregistrement des mouvements oculaires lors de la lecture fournit des indices non conscients de la part des participants. Il s'agit en effet des habitudes individuelles de lecture acquises lors de l'apprentissage scolaire et familial.

## 3 Résultats et discussion

	<i>Texte<sub>1</sub></i>						<i>Texte<sub>2</sub></i>					
	<i>O</i>	<i>S</i>	écart	<i>p</i>	<i>ddl</i>	<i>t-test</i>	<i>O</i>	<i>S</i>	écart	<i>p</i>	<i>ddl</i>	<i>t-test</i>
<i>TRN</i>	60,55	63,63	-3,08	0,23	45,00	1,22	62,73	59,67	3,06	0,22	45,00	1,24
<i>CRL</i>	58,88	62,06	-3,19	0,22	45,00	1,25	61,04	57,84	3,20	0,21	45,00	1,29
<i>DPF</i>	227,41	215,75	11,66	0,11	45,00	1,65	214,73	214,69	0,04	0,50	45,00	0,68
<i>NTF</i>	587,61	370,48	217,14	<b>0,00</b>	45,00	7,38	395,71	372,22	23,49	0,16	45,00	1,43
<i>AMP</i>	3,50	3,80	-0,30	<b>0,02</b>	45,00	2,44	3,33	3,82	-0,49	<b>0,00</b>	45,00	5,38
<i>REG</i>	27,26	21,21	6,06	<b>0,05</b>	45,00	2,05	21,47	19,30	2,18	0,24	45,00	1,18
<i>QCM</i>	1304,35	869,57	434,78	<b>0,02</b>	21,00	2,08	602,77	538,95	63,82	<b>0,00</b>	21,00	2,08

TABLE 1 – Résultats des passations pour les deux textes en versions originale *O* et simplifiée *S* et leur analyse statistique. Les indicateurs traités sont : les indicateurs pour les textes d'entraînement *TRN* et de contrôle *CRL* ; la durée de la première fixation *DPF*, le nombre total de fixations *NTF*, l'amplitude des saccades *AMP*, le nombre de régressions *REG* ; les réponses aux questions *QCM*. Le *p* statistiquement significatif est marqué en gras.

Le tableau 1 présente les résultats des indicateurs obtenus et leur analyse statistique pour le *texte<sub>1</sub>* et le *texte<sub>2</sub>*. L'analyse indique les valeurs suivantes : les moyennes pour les versions originales et simplifiées, l'écart-type, le *p*, le degré de liberté *ddl* et le *test t*. Les figures 6 et 7 montrent les exemples de lecture des versions originales (en haut) et simplifiées (en bas) du *texte<sub>1</sub>* et du *texte<sub>2</sub>*,

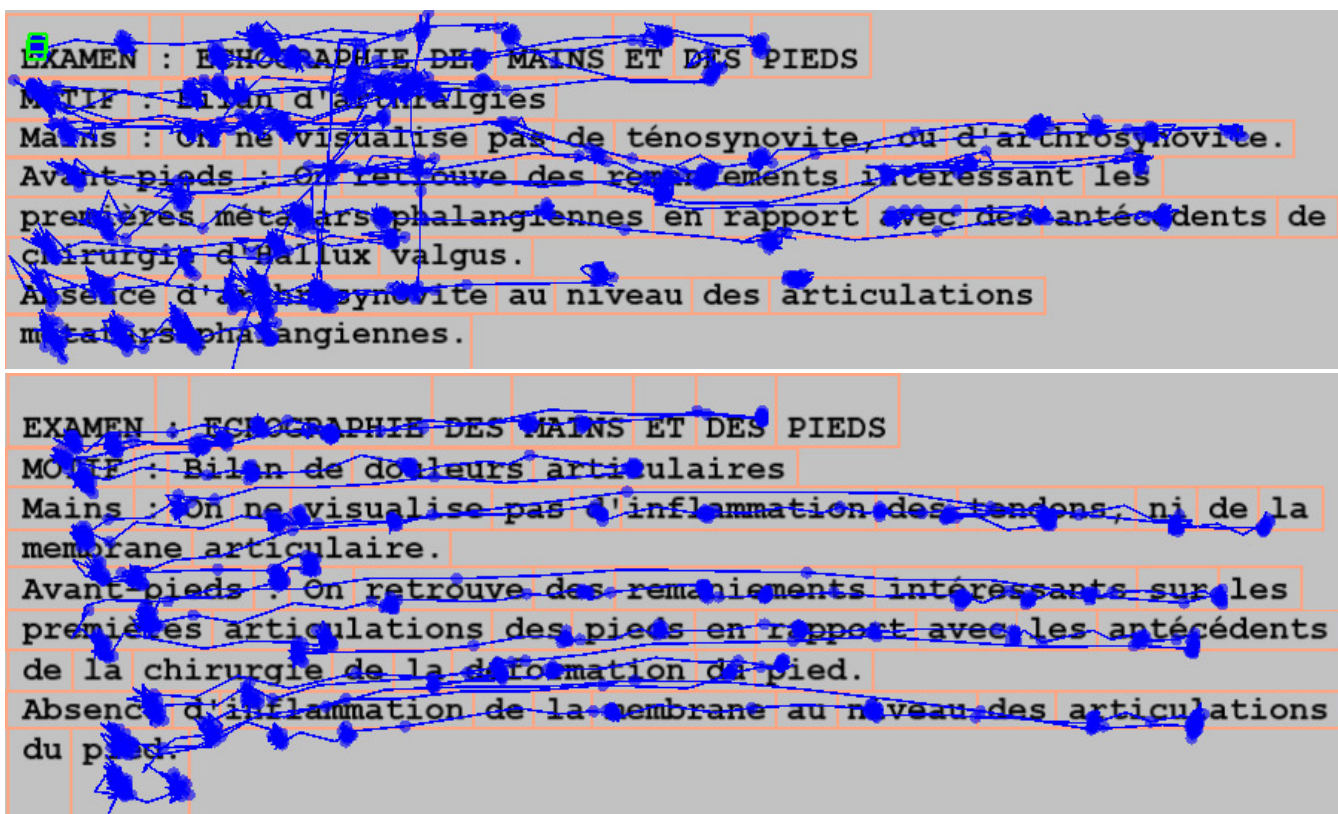


FIGURE 6 – Exemple de lecture du texte<sub>1</sub> en version originale (en haut) et simplifiée (en bas).

respectivement. Nous présentons et discutons les différents indicateurs obtenus :

- *Lecture des textes d'entraînement (TRN) et de contrôle (CRL)*. Entre les deux ensembles de tests, il n'existe pas de différence statistique dans les paramètres étudiés lors de la lecture des textes d'entraînement (TRN) et de contrôle (CRL). Cela indique que les participants ont la même capacité de lecture de manière générale : ainsi, les indicateurs collectés lors de la lecture des textes médicaux sont comparables. Cet indicateur donne d'autant plus de valeurs aux résultats obtenus avec la lecture des textes médicaux originaux et simplifiés ;
- *Durée de la première fixation (DPF)*. Pour les texte<sub>1</sub> et texte<sub>2</sub>, il n'existe pas de différence de la durée de la première fixation (DPF) entre les deux versions de ces textes. Cela indique que la lecture de ces textes commence de la même manière et sans anticipation de la part du participant. Ceci est aussi un bon indicateur qui renforce d'autres résultats car il montre que les résultats ne seront pas biaisés par l'attente ou le comportement des participants ;
- *Nombre total de fixations (NTF)*. Pour le texte<sub>1</sub>, une différence statistiquement significative est observée en relation avec le nombre total de fixations (NTF) : sur la version originale du texte<sub>1</sub>, les fixations sont plus longues et nombreuses. Cela peut être observé sur la figure 6 : sur le texte original (en haut), les points bleus des fixations sont plus grands par comparaison avec le texte simplifié (en bas). Par exemple, dans le texte original, les mots comme *arthralgie* ou *métatarsophalangiennes* montrent de longues fixations. Certains mots du texte<sub>1</sub> (*arthrosynovite*, *arthralgie*, *métatarsophalangiennes*) peuvent avoir plusieurs points de fixation. En revanche, le texte<sub>2</sub> ne montre pas de différence statistique par rapport au nombre total de fixations ;
- *Amplitude des saccades (AMP)*. Avec les texte<sub>1</sub> et texte<sub>2</sub>, la simplification mène à une augmentation statistiquement significative des amplitudes des saccades (AMP). Cela signifie que la simplification diminue effectivement la difficulté de lecture et que l'œil peut faire des sauts plus grands dans le texte pour aller d'un point de lecture vers un autre. Cela peut être observé sur les figures 6 et 7 : sur le texte original (en haut), les lignes bleues horizontales



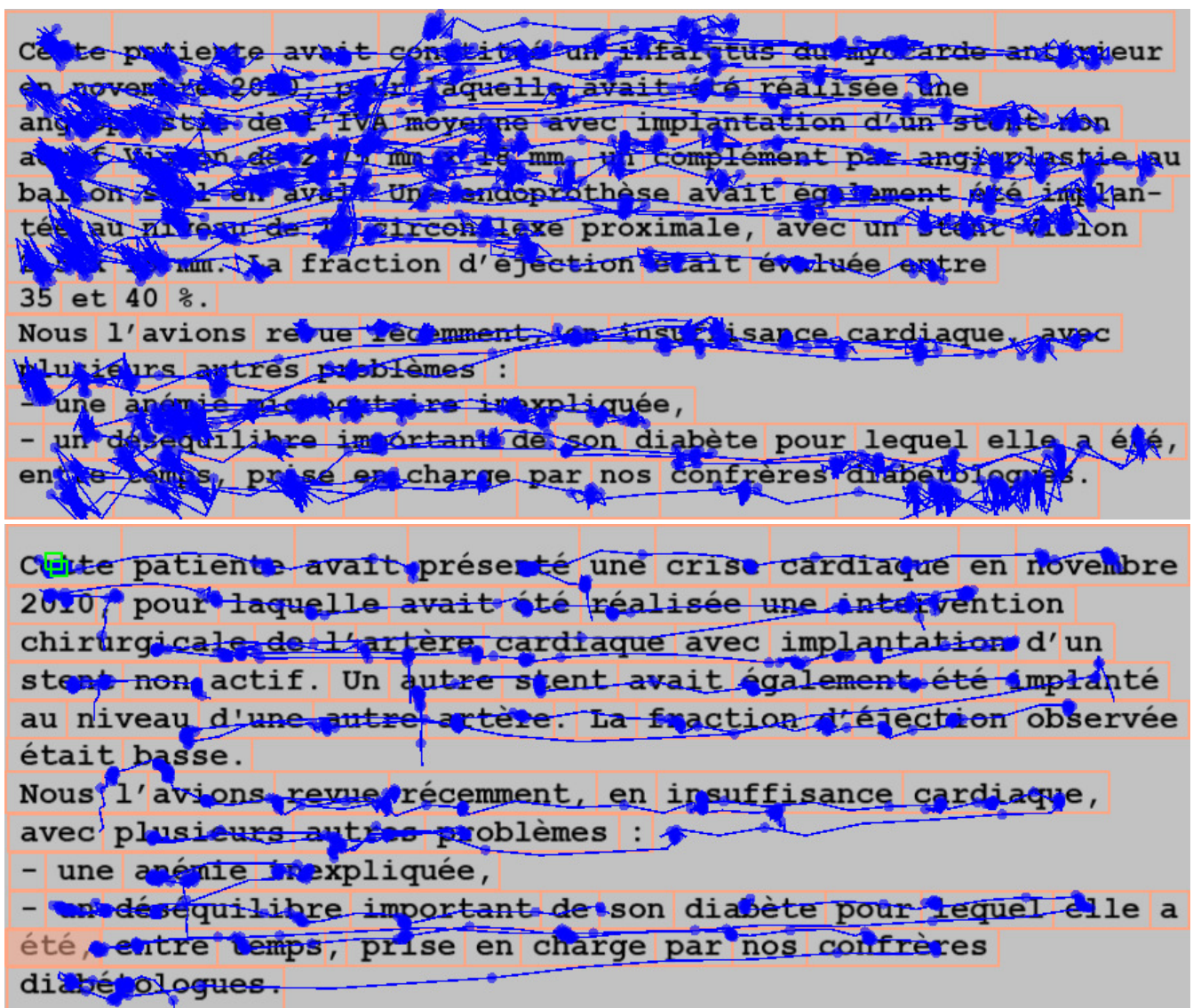


FIGURE 7 – Exemple de lecture du texte<sub>2</sub> en version originale (en haut) et simplifiée (en bas).

sont plus courtes que sur le texte simplifié ;

- *Régressions (REG)*. Pour le texte<sub>1</sub>, le nombre de régressions (*REG*) est significativement moins important lors de la lecture des textes simplifiés, ce qui suggère que la lecture et la compréhension de la version simplifiée sont meilleures. Cela peut aussi être observé sur la figure 6 : sur le texte original, nous pouvons voir les lignes bleues verticales. Pour le texte<sub>2</sub>, cet indicateur ne montre pas de différence statistiquement significative ;
- *Réponses des QCM*. L'analyse des réponses aux QCM indique que la compréhension des versions simplifiées est meilleure pour les texte<sub>1</sub> et texte<sub>2</sub> : nous obtenons 54 % et 50 % de réponses correctes avec les textes originaux, alors que les textes simplifiés fournissent 85 % et 78 % de réponses correctes, respectivement. Cette différence est aussi statistiquement significative pour les deux textes ;
- *Comparaison entre les fixations et les annotations manuelles*. La comparaison des fixations avec les annotations manuelles indique que les lecteurs se focalisent plus longtemps sur les types de termes et d'informations techniques qui peuvent être marqués lors de l'annotation manuelle du même type de textes. Il s'agit typiquement des composés néoclassiques (*arthralgie, métatarsophalangienne, arthrosynovite*), des termes anatomiques (*métatarsophalangienne*), des dérivés avec un sens spécialisé (*antécédent, éjection, anémie*). Dans le texte<sub>2</sub>, il y a également les valeurs numériques qui bloquent la lecture. Notons que les valeurs numériques

peuvent aussi être marquées lors de l'annotation manuelle des textes. De manière surprenante, l'emprunt latin *Hallux valgus* du texte<sub>2</sub> n'a pas bloqué le regard lors de la lecture. Pour certains participants, cet emprunt est associé avec un nom propre à cause de la majuscule au début : sa lecture n'a donc pas sollicité de fixation. Dans les exemples (1)-(3), nous montrons les résultats de l'annotation manuelle d'un même segment de texte par trois participants. Les passages en rouge sont marqués comme ne pouvant pas être compris. Nous pouvons clairement voir qu'il existe une différence importante dans les annotations. Chaque participant suit ses propres principes d'annotation : focalisation sur les mots ou sur les expressions plus complexes, annotation de la première occurrence ou de toutes les occurrences, une certaine familiarité ou une bonne connaissance des notions... De plus, l'annotation d'une unité donnée reste toujours un choix conscient avec ce type d'approche. Il s'agit donc d'une méthode plus subjective par rapport aux méthodes et indicateurs d'oculométrie. À notre avis, il s'agit d'un point très positif de l'oculométrie.

- (1) *L'échographie cardiaque montre des cavités gauches dilatées (OG 59, VG 61/40, fraction d'éjection 61%, PAPS 38, bon fonctionnement de la Saint aortique N 25 (gradient moyen à 16.5), aspect satisfaisant du montage du tube valvulaire aortique. Sur la mitrale, il s'agit d'une mitrale dystrophique avec un prolapsus de P2 et éversion systolique complète dans l'oreillette gauche par rupture de cordage, calcification mitrale postérieure, fuite mitrale cotée 4/4. L'insuffisance tricuspидienne apparaît minime.*
- (2) *L'échographie cardiaque montre des cavités gauches dilatées (OG 59, VG 61/40, fraction d'éjection 61%, PAPS 38, bon fonctionnement de la Saint aortique N 25 (gradient moyen à 16.5), aspect satisfaisant du montage du tube valvulaire aortique. Sur la mitrale, il s'agit d'une mitrale dystrophique avec un prolapsus de P2 et éversion systolique complète dans l'oreillette gauche par rupture de cordage, calcification mitrale postérieure, fuite mitrale cotée 4/4. L'insuffisance tricuspидienne apparaît minime.*
- (3) *L'échographie cardiaque montre des cavités gauches dilatées (OG 59, VG 61/40, fraction d'éjection 61%, PAPS 38, bon fonctionnement de la Saint aortique N 25 (gradient moyen à 16.5), aspect satisfaisant du montage du tube valvulaire aortique. Sur la mitrale, il s'agit d'une mitrale dystrophique avec un prolapsus de P2 et éversion systolique complète dans l'oreillette gauche par rupture de cordage, calcification mitrale postérieure, fuite mitrale cotée 4/4. L'insuffisance tricuspидienne apparaît minime.*

Si l'oculométrie offre plusieurs indicateurs objectifs sur le processus de lecture (typiquement le nombre et la durée des fixations, l'amplitude des saccades, le nombre de régressions) et la possibilité de les corrélérer avec d'autres informations comme la compréhension des textes, cette méthode présente aussi quelques limitations connues et prises en compte pendant les expériences (Duchowski, 2007) :

- Les appareils d'oculométrie permettent de capter et d'enregistrer le regard des participants. La supposition communément faite alors est que le regard coïncide avec l'attention du sujet, alors qu'en réalité l'attention peut aussi être portée sur un objet qui se trouve en périphérie du regard, par exemple. Le système de vision humaine est en effet très complexe et n'est pas complètement décodé pour le moment. Il s'agit d'une des limitations connues des méthodes d'oculométrie qui demande donc d'accepter les deux possibilités (le regard correspond à l'attention explicite ou non). Dans notre cas, avec la lecture de textes médicaux courts et le besoin de répondre aux questions, et à l'absence de distracteurs, nous pouvons supposer que

- l'attention des lecteurs coïncide avec leur regard ;
- Avec certains participants, dû aux caractéristiques physiologiques (cils longs, maquillage, paupières lourdes...), il peut être difficile d'effectuer le calibrage du caméra et de bien capter et enregistrer les mouvements oculaires. Cela peut mener à une perte des données ou bien à une mauvaise superposition des enregistrements sur le texte. En revanche, lorsque les données sont exploitables, cela n'influence pas les indicateurs de lecture ;
  - De la même manière, le port de lunettes et de lentilles de contact peut poser des difficultés lors de la capture de la pupille ;
  - Avec un même texte ou image, l'attention et donc le regard des participants varient en fonction de la tâche ou des questions qui leur sont soumises. Dans notre cas, tous les participants devaient accomplir la même tâche qui consistait en lecture des textes et en réponse aux questions de compréhension ensuite ;
  - Le matériel d'oculométrie a des limites aussi : (1) il fonctionne avec une certaine fréquence (60 Hz) et certains mouvements oculaires peuvent ne pas être captés et enregistrés ; (2) le signal enregistré est débruité, comme par exemple les clignotements ou certains mouvements périphériques ce qui peut aussi supprimer certains mouvements oculaires importants.

De manière générale, selon les résultats d'oculométrie, nous notons que la simplification des textes améliore les indicateurs de lecture : (1) la durée totale des fixations devient plus courte, (2) l'amplitude des saccades devient plus grande, (3) les régressions deviennent moins fréquentes, (4) la compréhension du contenu et les réponses aux questions s'améliorent. Tous ces indicateurs se vérifient sur le texte<sub>1</sub> et une partie de ces indicateurs se vérifie sur le texte<sub>2</sub>. Ceci est également un résultat positif du travail. Ces résultats fournissent des indicateurs cohérents obtenus suite à la lecture des textes techniques en versions originales et simplifiées. Ils indiquent qu'il existe des schémas spécifiques lors de la lecture des textes techniques où le regard est attiré par certains éléments langagiers, comme les composés néoclassiques, les termes anatomiques ou les valeurs numériques.



(a) Texte<sub>1</sub> original



(b) Texte<sub>1</sub> simplifié

FIGURE 8 – La simulation des fixations pour le texte<sub>1</sub>.

Notons qu'il est possible de faire des simulations d'oculométrie sur certains sites en ligne. Pour ceci, il faut soumettre une image du texte et la simulation des fixations est obtenue très rapidement. La figure 8 montre des exemples de telles simulations sur le texte<sub>1</sub> pour la version originale (figure 8(a)) et la version simplifiée (figure 8(b)). Les couleurs chaudes représentent les endroits qui attirent plus le regard. Il nous semble que cette simulation est plutôt construite selon le modèle de fixations de la figure 1, où le début du texte et les débuts des lignes attirent le plus le regard de l'utilisateur. Comme nous pouvons le voir, ce modèle de lecture est très différent de celui que nous avons obtenu avec la lecture de textes techniques (figures 6 et 7). Il s'agit d'un résultat intéressant car il montre que différents types de textes peuvent solliciter différents types d'attention et de lecture.

## 4 Conclusion

Nous avons proposé une expérience pour étudier l'effet de la simplification de textes techniques, sur l'exemple de textes médicaux, grâce aux méthodes d'oculométrie. De cette manière, nous obtenons plusieurs indicateurs de lecture objectifs, comme la durée des fixations, la durée de la première fixation, l'amplitude des saccades et les régressions. Ces mêmes indicateurs sont obtenus suite à la lecture d'un texte d'entraînement et d'un texte de contrôle. Les indicateurs collectés sont comparés entre les versions originales et simplifiées d'un texte médical donné avec des mesures statistiques (*test t*) pour analyser s'il existe une différence statistiquement significative lors de la lecture des versions originales et simplifiées de ces textes. Ensuite, nous analysons les réponses aux questionnaires QCM fournies par les participants après la lecture de chaque texte et les annotations manuelles des difficultés de compréhension.

Les résultats obtenus indiquent que la lecture des deux versions des textes, originales et simplifiées, fournit des indicateurs cohérents et stables. Par exemple, lors de la lecture des textes simplifiés, les fixations sont plus courtes, les saccades plus longues et les régressions absentes ou pas fréquentes. De plus, l'analyse des réponses aux questions indique que la compréhension des textes simplifiés est meilleure : ainsi, sur le texte<sub>1</sub>, le nombre de réponses correctes est de 54 % pour la version originale et 81 % pour la version simplifiée. Cela indique aussi que les textes techniques, et en particulier les textes médicaux, peuvent être simplifiés de manière efficace pour atteindre une meilleure compréhension de la part des non-experts. Notons aussi que la comparaison avec l'annotation manuelle, effectuée par le même type de participants, montre que les lecteurs et les annotateurs se focalisent sur les mêmes types de termes : des composés néoclassiques, des termes anatomiques, des dérivés avec un sens spécialisé, et des valeurs numériques. Nous montrons en revanche que l'annotation manuelle présente plus de subjectivité, car elle repose sur une décision consciente de la part de l'annotateur par rapport aux résultats d'oculométrie, qui eux dépendent des processus et habitudes de lecture appris lors de l'apprentissage scolaire et familial. Les méthodes et le matériel d'oculométrie ont également des limitations actuellement, que nous indiquons dans le travail.

Dans les travaux futurs, il serait intéressant d'étudier le lien entre la longueur d'un texte et sa lisibilité et compréhension. L'hypothèse serait que les textes longs, même s'ils sont simplifiés, peuvent présenter des difficultés de lecture et de compréhension. L'impact d'autres facteurs (comme les définitions, les contextes favorables, les images et illustrations) peut également être étudié. À cause des contraintes expérimentales, des extraits courts de textes sont utilisés et étudiés. Pour cette raison, il serait intéressant d'effectuer des tests supplémentaires pour augmenter la variété des types de textes et de contenus traités. De plus, comme avec la simplification manuelle, l'efficacité de la simplification automatique peut aussi être testée et évaluée en utilisant des protocoles d'oculométrie.

## Remerciements

Ce travail fait partie du projet *Termeye* financé par l'appel de l'Établissement de l'université Lille 3 et du projet *CLEAR* (*Communication, Literacy, Education, Accessibility, Readability*) financé par l'ANR sous la référence ANR-17-CE19-0016-01. Les extraits de textes médicaux proviennent du projet *RAVEL* financé par l'ANR sous la référence ANR-11-TECS-012. Nous remercions les participants de l'étude qui ont bien voulu nous consacrer leur temps. Nous remercions également les relecteurs anonymes qui ont permis d'améliorer la qualité de la version finale de ce papier.

# Références

- ALTARRIBA J., KROLL J., SHOLL A. & RAYNER K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences : Evidence from eye fixations and naming times. *Memory and Cognition*, **24**, 477–92.
- ANDREWS T. & COPPOLA D. (1999). Idiosyncratic characteristics of saccadic eye movements when viewing different visual environments. *Vision Res*, **39**(17), 2947–53.
- ANTOINE E. & GRABAR N. (2016). Exploitation de reformulations pour l'acquisition d'un vocabulaire expert/non expert. In *Traitement Automatique des Langues Naturelles (TALN)*.
- ASAAD W. & ANDEK MILLER G. R. (2000). Task-specific neural activity in the primate prefrontal cortex. *Neurophysiology*, **84**, 451–459.
- BALAKRISHNA S. (2015). *Analyzing Text Complexity and Text Simplification : Connecting Linguistics, Processing and Educational Applications*. Thèse de doctorat, Eberhard Karls Universität Tübingen, Tübingen, Germany.
- BERLAND G., ELLIOTT M., MORALES L., ALGAZY J., KRAVITZ R., BRODER M., KANOUSE D., MUNOZ J., PUYOL J. & ET AL M. L. (2001). Health information on the internet. accessibility, quality, and readability in english and spanish. *JAMA*, **285**(20), 2612–2621.
- BERNÈCHE F. & PERRON B. (2006). *Développer nos compétences en littérature : un défi porteur d'avenir. Enquête internationale sur l'alphabétisation et les compétences des adultes*. Rapport interne, Institut de la statistique du Québec, Canada.
- BERTRAM R., KUPERMAN V., BAAYEN H. R. & HYÖNÄ J. (2011). The hyphen as a segmentation cue in triconstituent compound processing : It's getting better all the time. *Scandinavian Journal of Psychology*, **52**(6), 530–544.
- BISSON M., VAN HEUVEN W., CONKLIN K. & TUNNEY R. (2014). Processing of native and foreign language subtitles in films : An eye-tracking study. *Applied Psycholinguistics*, **35**, 399–418.
- BOZIC M., MARSLÉN-WILSON W. D., STAMATAKIS E. A., DAVIS M. H. & TYLER L. K. (2007). Differentiating morphology, form, and meaning : Neural correlates of morphological complexity. *Journal of Cognitive Neuroscience*, **19**(9), 1464–1475.
- CAIN K., TOWSE A. S. & KNIGHT R. S. (2009). The development of idiom comprehension : An investigation of semantic and contextual processing skills. *Journal of Experimental Child Psychology*, **102**(3), 280–298.
- CHMIELIK J. & GRABAR N. (2011). Détection de la spécialisation scientifique et technique des documents biomédicaux grâce aux informations morphologiques. *TAL*, **51**(2), 151–179.
- CLIFTON C. & STAUB A. (2011). Syntactic influences on eye movements in reading. In S. LIVERSEDGE, I. GILCHRIST & S. EVERLING, Eds., *The Oxford handbook of eye movements*, p. 895–909. Oxford University Press.
- CLIFTON C., STAUB A. & RAYNER K. (2007). Eye movements in reading words and sentences. In R. VAN GOMPEL, M. FISCHER, W. MURRAY & R. HILL, Eds., *Eye Movements : A Window on Mind and Brain*. Oxford : Elsevier.
- COOPER R. (1974). The control of eye fixation by the meaning of spoken language : A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychol*, **6**, 84–107.
- DOHMES P., ZWITSERLOOD P. & BÖLTE J. (2004). The impact of semantic transparency of morphologically complex words on picture naming. *Brain and Language*, **90**(1-3), 203–212.

- DUCHOWSKI A. (2007). *Eye Tracking Methodology. Theory and practice*. London, UK : Springer.
- ELTERMAN R., ABEL L., DAROFF R., DELL'OSSO L. & BORNSTEIN J. (1980). Eye movement patterns in dyslexic children. *J Learn Disabil*, **13**, 16–21.
- FOX S. (2011). *Health topics. 80% of internet users look for health information online*. Rapport interne, Pew Internet & American Life Project, Washington DC.
- FRANTZI K. T., ANANIADOU S. & MIMA H. (2000). Automatic recognition of multi-word terms : the C-Value/NC-Value method. *Int J on Digital Libraries*, **3**(2), 115–130.
- FRENCK-MESTRE C. & PYNTE J. (1997). Syntactic ambiguity resolution while reading in a second and native languages. *The Quarterly Journal of Experimental Psychology*, **50**(A), 119–48.
- GRABAR N. & HAMON T. (2016). Exploitation de la morphologie pour l'extraction automatique de paraphrases grand public des termes médicaux. *TAL*, **57**(1), 85–109.
- HAMON T., ENGSTRÖM C. & SILVESTROV S. (2014). Term ranking adaptation to the domain : genetic algorithm based optimisation of the C-Value. In *PolTAL 2014*, p. 71–83.
- HIGGINS E., LEINENGER M. & RAYNER K. (2014). Eye movements when viewing advertisements. *Front Psychol*, **5**, 210.
- IACOBINI C. (2003). Composizione con elementi neoclassici. In M. GROSSMANN & F. RAINER, Eds., *La formazione delle parole in italiano*, p. 69–96. Walter de Gruyter.
- JUCKS R. & BROMME R. (2007). Choice of words in doctor-patient communication : an analysis of health-related internet sites. *Health Commun*, **21**(3), 267–77.
- KAGEURA K. & UMINO B. (1996). Methods of automatic term recognition. In *National Center for Science Information Systems*, p. 1–22.
- KEATING G. (2009). Sensitivity to violations of gender agreement in native and non-native spanish : An eye-movement investigation. *Language Learning*, **59**, 503–35.
- LÜDELING A., SCHMIDT T. & KIOKPASOGLU S. (2002). Neoclassical word formation in german. *Yearbook of Morphology*, p. 253–283.
- LÜTTMANN H., ZWITSERLOOD P. & BÖLTE J. (2011). Sharing morphemes without sharing meaning : Production and comprehension of german verbs in the context of morphological relatives. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, **65**(3), 173–191.
- MIELLET S., SPARROW L. & SERENO S. (2008). The effects of frequency and predictability in french : An evaluation of the e-z reader model. *Psychonomic Bulletin & Review*, **14**, 762–769.
- MOLNAR F. (1981). About the role of visual exploration in aesthetics. In H. DAY, Ed., *Advances in Intrinsic Motivation and Aesthetics*, p. 385–414. New York : Plenum Press.
- NILSSON BENFATTO M., ÖQVIST SEIMYR G., YGGE J., PANSELL T., RYDBERG A. & JACOBSON C. (2016). Screening for dyslexia using eye tracking during reading. *PLoS ONE*, **11**(12), e0165508.
- PATEL V., BRANCH T. & AROCHA J. (2002). Errors in interpreting quantities as procedures : The case of pharmaceutical labels. *Int Journ Med Inform*, **65**(3), 193–211.
- PLETNEVA N., VARGAS A. & BOYER C. (2011). *How do general public search online health information ?* Rapport interne, Health On the Net Foundation.
- RAYNER K. (1998). Eye movements in reading and information processing : 20 years of research. *Psychological bulletin*, **124**(3), 372–373.

- RAYNER K. & LIVERSEDGE S. (2004). *Visual and linguistic processing during eye fixations in reading*, In J. HENDERSON & F. FERREIRA, Eds., *The Interface of Language, Vision, and Action : Eye movements and the visual world*.
- RELLO L., BAEZA-YATES R., BOTT S. & SAGGION H. (2013). Simplify or help ? : text simplification strategies for people with dyslexia. In ACM, Ed., *Int Cross-Disciplinary Conference on Web Accessibility*, p. 15–25.
- RUBINO C. & MINDEN H. (1973). Analysis of eye-movements in children with reading disability. *Cortex*, **9**, 217–220.
- SHARDLOW M. (2013). A comparison of techniques to automatically identify complex words. In *ACL Student Research Workshop*, p. 103–109.
- SHARMIN S., SPAKOV O., RÄIHÄ K. & JAKOBSEN A. (2008). Effects of time pressure and text complexity on translators' fixations. In *ETNA*, p. 123–126.
- SINGH A., MEHTA P., HUSAIN S. & RAJKUMAR R. (2016). Quantifying sentence complexity based on eye-tracking measures. In *Workshop on Computational Linguistics for Linguistic Complexity*, p. 202–212.
- SPARROW L., MIELLET S. & COELLO Y. (2003). The effects of frequency and predictability on eye fixations in reading : An evaluation of the E-Z reader model. *Behavioral and Brain Sciences*, **26**, 503–505.
- TANENHAUS M., SPIVEY-KNOWITON M., EBERHARDA K. & SEDIVY J. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, **268**, 1632–1634.
- TRAN T., CHEKROUD H., THIERY P. & JULIENNE A. (2009). Internet et soins : un tiers invisible dans la relation médecine/patient ? *Ethica Clinica*, **53**, 34–43.
- TRUESWELL J., TANENHAUS M. & GARNSEY S. (1994). Semantic influences on parsing : Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, **33**, 285–318.
- WENDT D., BRAND T. & KOLLMEIER B. (2014). An eye-tracking paradigm for analyzing the processing time of sentences with different linguistic complexities. *PLoS ONE*, **9**(6), e100186.
- WILLIAMS M., PARKER R., BAKER D., PARIKH N., PITKIN K., COATES W. & NURSS J. (1995). Inadequate functional health literacy among patients at two public hospitals. *JAMA*, **274**(21), 1677–1682.
- YANEVA V., TEMNIKOVA I. & MITKOV R. (2015). Accessible texts for autism : An eye-tracking study. In ACM, Ed., *Int ACM SIGACCESS Conference on Computers & Accessibility*, p. 49–57.
- ZENG Q. T., TSE T., CROWELL J., DIVITA G., ROTH L. & BROWNE A. C. (2005). Identifying consumer-friendly display (CFD) names for health concepts. In *Ann Symp Am Med Inform Assoc (AMIA)*, p. 859–63.

