
Accès non visuel au Web par *Tagthunder*

Fabrice Maurel

Université Caen Normandie
14032, Caen, France
fabrice.maurel@unicaen.fr

Résumé

La mise page et la typographie sont déterminantes pour l'efficacité des stratégies de lecture de haut niveau. Comment les rendre techniquement, psychologiquement et cognitivement perceptibles aux non-voyants ? Le projet *Tagthunder* vise à rendre possible le développement autonome de stratégies non visuelles efficaces pour transposer la structure visuelle des documents vers la modalité sonore ; en particulier sur le Web qui complexifie et multiplie les contrastes visuels inaccessibles à cette population d'utilisateurs.

Mots Clés

Accessibilité ; IHM ; TAL ; Systèmes de Substitution Sensorielle

Abstract

Layout and typography are critical to the effectiveness of high-level reading strategies. How to make them technically, psychologically and cognitively perceptible to the blind? The *Tagthunder* project aims to make possible the autonomous development of effective non-visual strategies for transposing the visual structure of documents into sound; in particular, on the Web which complicates and multiplies the visual contrasts inaccessible to this population of users.

Author Keywords

Accessibility; HCI; NLP; Sensory Substitution Systems

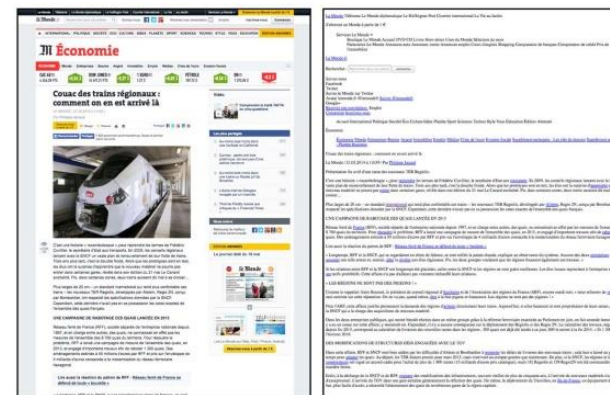
Introduction

Lors de la lecture silencieuse de documents par un voyant, nous observons souvent un enchaînement de microprocessus mentaux spécifiques en grande partie automatisés : le lecteur peut prendre de l'information d'un premier regard sur tout ou partie de la page, par un survol quasi-instantané (*skimming*) ; il peut également activer des balayages rapides du support (*scanning*), dirigés par les indices visuels et langagiers perçus pendant le *skimming*. Ces deux processus, plus ou moins conscientisés et alternant perception locale et perception globale, peuvent se répéter selon différentes combinaisons jusqu'à la satisfaction d'objectifs individuels ; puis s'ancrent en stratégies de lecture de haut niveau, en fonction des contraintes et des intentions de lecture.

La mise en page et la typographie prennent une part déterminante dans le succès et l'efficacité de ces processus, or leur restitution est quasi-absente dans les solutions existantes de lecture non visuelle. L'objectif de notre projet porte sur la possibilité de rendre accessibles rapidement aux non-voyants les différents niveaux de structure des pages Internet parcourues ; autrement dit, comment favoriser le développement de stratégies de *skimming* et de *scanning* non visuels de pages Web, en s'appuyant sur une transposition orale de leur structure visuelle, logique et thématique.

Le Concept de TAGTHUNDER

Considérons une page Web (voir Figure 1). Continuons avec sa structuration en zones cohérentes (voir Figure 2a). Extrayons pour chaque zone quelques expressions textuelles ou métatextuelles représentatives de son contenu (voir Figure 2b).



Voyants

Non voyants

Figure 1 : page Web écran vs. lecteur d'écran

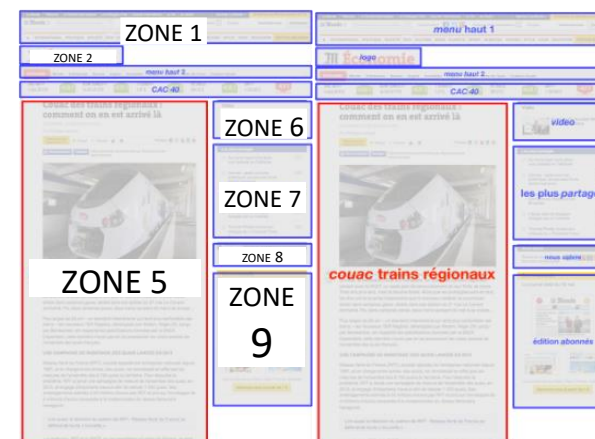


Figure 2a. et 2b. : segmentation en zone et calcul d'un représentant textuel



Figure 3 : du *Tagcloud* au *Tagthunder*

Enfin, effaçons les autres éléments de la page ; disposons les termes retenus dans la même relation spatiale que la zone qu'ils représentent et façonnons-les graphiquement pour les rendre d'autant plus saillants que la zone semble importante. Nous obtenons un stimulus visuel fréquemment utilisé dans le Web 2.0 : un *tagcloud*. L'idée que nous développons est de transposer ce concept dans le monde sonore 3D, transformant ce nuage de mots en « tonnerre » de mots où tous les stimuli sonores sont émis en parallèle (ou *tagthunder* – voir Figure 3).

Le *tagthunder* permettra d'appréhender globalement et précocement la structure visuo-logico-thématique de la page Web, tout en favorisant l'exploitation des processus de *skimming* et de *scanning* non visuels. L'analogie qui sous-tend le développement de ce concept est un prolongement de la métaphore connue par les psychologues sous le nom de « *cocktail party effect* ». En psycho-acoustique, elle dénote la possibilité de focaliser son attention auditive sur un flux verbal dans l'ambiance bruyante d'une réception ; que cela soit dirigé vers des sources sonores extérieures à sa

conversation ou vers ses interlocuteurs. En partant de travaux autour de la parole concurrente [1], nous filerons cette métaphore en considérant la relation entre le lecteur non-voyant et les zones de la page Web, comme celle entre un invité situé au centre d'une salle et les différents groupes de discussions qui s'y sont formés : les échanges sont séquentiels à l'intérieur d'un groupe mais concurrents entre les différents groupes ; l'invité, nouveau venu, doit prélever dans l'environnement sonore suffisamment d'information pour identifier rapidement la discussion à laquelle il souhaite se mêler.

Approche scientifique

Les recherches en neuro(psycho)(linguistique) et en interaction homme machine se sont intéressées à la substitution d'un sens par un ou plusieurs autres [2, 3]. Seulement quelques-unes se sont spécialisées dans la prise en compte de la structure visuelle des documents comme porteuse d'une sémantique morphe-dispositionnelle [4, 5]. Notre approche veut aller plus loin encore dans son exploitation en lui prêtant également un rôle de support de l'interaction avec le document [6]. Dans ce contexte, une critique essentielle que nous formulons à l'endroit de nombreuses propositions de présentation non visuelle de pages Web porte sur l'approche méthodologique qui est posée. Elles s'attachent, pour la plupart, à réduire la charge cognitive en recherchant dans la page Web les informations « pertinentes » et en éliminant les perturbations induites par les éléments « périphériques » [7] ; ou encore en intégrant des techniques de résumés [8]. De manière générale il s'agit de simplifier le contenu pour le rendre plus digeste aux modalités tactiles ou orales. Ce faisant, le concepteur considère que l'amélioration de l'accessibilité doit sacrifier une certaine richesse de contenu. Pourtant l'accessibilité universelle peut se définir comme « le caractère d'un produit,

procédé, service, environnement ou de l'information qui, dans un but d'équité et dans une approche inclusive, permet à toute personne de réaliser des activités de façon autonome et d'obtenir des résultats équivalents » [9]. Cette intégration de la capacité d'autodétermination de l'utilisateur dans le système interactif nous semble majeure et prioritaire à évaluer. Notre ambition est de présenter au lecteur l'ensemble de l'information dans sa complexité ; à notre charge de trouver les stimuli adéquats pour favoriser l'interaction et le déroulement d'une boucle perception/action vertueuse et éactive ; à la charge de chaque utilisateur d'apprendre à les maîtriser, à se les approprier, à pérenniser ses propres interprétations ; à la charge du temps et de la pratique de faire émerger de nouvelles stratégies.

Développements applicatifs

Nous proposons de présenter les avancées de notre projet à travers (1) les résultats obtenus dans le projet ANR ART-ADN entre 2013 et 2016, (2) les développements logiciels et matériels du projet FSN « Accessibilité numérique » entamé en septembre 2017 en collaboration avec un industriel et (3) les avancées d'une thèse ministérielle initiée en octobre 2017 [10, 11]. La confrontation aux différents spécialistes « terrain » nous semble aujourd'hui une étape importante pour mieux orienter nos choix et couvrir scientifiquement les opérations d'extraction de zones dans les pages Web, d'association de mots clés aux zones, de sélection/génération de termes à oraliser, de spatialisation des sources d'émission, d'interactions avec le *tagthunder*, et d'évaluations globales et modulaires.

Bibliographie

1. Guerreiro J. (2013), Using simultaneous audio sources to speed-up blind people's web scanning, 10th International Cross-Disciplinary Conference on Web Accessibility (W4A), p. 1-2.
2. Bach-Y-Rita P., Collins C.C., Saunders F.A., White B., Scadden L. (1969), Visual substitution by tactile image projection, *Nature*, 221, p. 963-964.
3. Lenay C., Gapenne O., Hanneton S., Marque C., Genouëlle C. (2003), Sensory Substitution, Limits and Perspectives, In *Touch for Knowing Cognitive psychology of haptic manual perception*, Amsterdam, p. 275-292.
4. Virbel J., The Contribution of linguistic knowledge to the interpretation of text Structure, J. André, V. Quint et R. Furuta (eds.), *Structured Documents*, Cambridge University Press, p. 161-181.
5. Maurel F., Mojahid M., Vigouroux N., Virbel J. (2006), Documents numériques et transmodalité, *Transposition automatique à l'oral des structures visuelles de texte*, Document numérique, Hermès, Vol. 9, N. 1, p. 25-42.
6. Maurel F., Dias G., Routoure J.-M., Vautier M., Beust P., Molina M., Sann C. (2012), *Haptic Perception of Document*
7. Giudice N., Palani H., Brenner E., Kramer K. (2012), *Learning non-visual graphical information using a touch-based vibro-audio interface*, ASSETS '12 Proceedings of the 14th international ACM SIGACCESS conference on computers and accessibility, p. 103-110, 2012.
8. Borodin Y., Bigham J.P., Stent A., Ramakrishnan I.V. (2008), *Towards one world web with HearSay3*, Proceedings of the 2008 International Cross-Disciplinary Conference on Web Accessibility (W4A), Beijing, China, p. 130-131.
9. Extrait de la plateforme Alter&Go - <http://www.alter-go.fr/>
10. Safi W., Maurel F., Routoure J.-M., Beust P., Dias G. (2015), An Empirical Study for Examining the Performance of Visually Impaired People in Recognizing Shapes through a Vibro-tactile Feedback, ASSETS2015, the 17th International ACM SIGACCESS Conference on Computers and Accessibility, October 26-28 Lisbon, Portugal.
11. Lecarpentier J.-M., Manishina E., Maurel F., Ferrari S., Busson M. (2016), Tag Thunder: Web Page Skimming in Non Visual Environment Using Concurrent Speech, Proceedings of the 7th Workshop on Speech and Language Processing for Assistive Technologies (SLPAT), INTERSPEECH 2016, San Francisco, CA, USA.