



# Vers un dialogue oral sur le World Wide Web

José Rouillard

► **To cite this version:**

José Rouillard. Vers un dialogue oral sur le World Wide Web. 2èmes Rencontres Jeunes Chercheurs en Parole, 1997, La Rochelle, France. hal-02440290

**HAL Id: hal-02440290**

**<https://hal.inria.fr/hal-02440290>**

Submitted on 15 Jan 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# VERS UN DIALOGUE ORAL SUR LE WORLD WIDE WEB

José Rouillard

Laboratoire CLIPS/IMAG, Groupe GEOD

Campus scientifique, B.P. 53, 38041 Grenoble Cedex 09

E-mail : Jose.Rouillard@imag.fr

## 1 INTRODUCTION

Nos travaux s'insèrent dans le projet de recherche ORION, de la région Rhône-Alpes. Nous étudions de nouvelles formes d'interaction homme-machine susceptibles de faciliter la recherche d'information et la navigation sur le réseau Internet. A terme, ces travaux de thèse doivent permettre l'intégration de l'usage de la parole et la mise en œuvre d'une composante de dialogue en sur-couche à la navigation sur le World Wide Web (WWW).

## 2 PREMIERS RÉSULTATS

Les travaux réalisés jusqu'à présent dans le cadre de ma thèse sont les suivants :

- une étude des processus de navigation et de recherche d'information sur le WWW : enquête d'usage réalisée à la médiathèque de l'IMAG et étude des processus cognitifs lors de la recherche d'informations. Ces premiers résultats dévoilent quels sont les freins et/ou les besoins perçus par des usagers effectuant une recherche documentaire sur le Web grâce à une interface traditionnelle (clavier et souris) et quels pourraient être les avantages et les inconvénients d'une interface multimodale (reconnaissance et synthèse vocale, geste, etc.) dans un tel contexte.
- une étude du rôle de l'interaction multimodale dans la recherche d'informations sur le Web (travaux d'analyse statistique).
- une étude préalable de faisabilité technique à l'aide d'un logiciel "robot" (spider) naviguant automatiquement sur le Web. Cela afin d'étudier les relations entre les pages existantes et d'observer si rien ne s'oppose à la mise en place d'une architecture générique pour l'intégration d'une couche de dialogue homme-machine sur le réseau Internet. Ce travail a fait l'objet d'un article de recherche présenté à la 4ème conférence Internationale Hypertextes & Hypermédiâs, à Paris en septembre 1997 [2].

Nous présentons succinctement quelques unes de ces études ci-après.

### 2.1 ENQUÊTE D'USAGE À LA MÉDIATHÈQUE DE L'IMAG

Le but de cette enquête a été d'étudier les possibilités de mettre en place une interface multimodale pour l'accès au catalogue de la médiathèque de l'IMAG (Institut d'Informatique et de Mathématiques Appliquées de Grenoble) via le Web. Les éléments remarquables de cette enquête d'usage sont résumés ci-après :

- ♦ La recherche documentaire se fait principalement par tâtonnement et par le biais de bibliographies données par les professeurs : on n'utilise pas prioritairement l'accès au catalogue via le Web pour retrouver les documents dont on a besoin.
- ♦ 68% des usagers n'utilisent les terminaux de recherche que rarement, voire jamais. L'enquête fait ressortir deux explications possibles :
  - d'une part, des problèmes internes à la médiathèque : les terminaux ne sont pas en service, ne fonctionnent pas correctement, sont trop peu nombreux ou occupés (monopolisés) longtemps ;
  - d'autre part, des réticences des usagers : ils n'apprécient pas l'interface proposée, ils préfèrent demander à un professeur une liste d'ouvrages, ils ne voient pas l'utilité de ce catalogue, etc.
- ♦ Les utilisateurs sont favorables (68%) à une plus grande interactivité avec la machine grâce à un dialogue oral où ils pourraient dire par exemple : "non ce n'est pas cela que je recherche" ou "oui, je veux plus de détails", etc.
- ♦ Une part importante des usagers se dit prête à tester une interface vocale pour une recherche documentaire (72%), ce qui laisse supposer qu'ils ont besoin d'un meilleur dispositif permettant une communication homme-machine performante et pertinente.

Les résultats de ce travail montrent que :

(a) Même si les ouvrages sont présents et accessibles pour la plupart des usagers, ces derniers n'arrivent pas toujours à retrouver *efficacement* à partir de la page Web ce dont ils ont besoin. Ils dénoncent le moteur de recherche, pas assez précis ni pertinent, et l'interface de ce catalogue, qu'ils trouvent peu conviviale.

(b) Installer un ordinateur à reconnaissance vocale en médiathèque implique que ce poste de travail soit insonorisé, non seulement pour éviter les bruits parasites, mais également pour ne pas gêner les autres usagers présents.

(c) Enfin, après cette enquête sur l'existant et les besoins des utilisateurs, il paraît nécessaire de disposer d'une étude sur la manière dont les usagers cherchent l'information et selon quel modèle cognitif ils évoluent.

## 2.2 ÉTUDE DU WEB EN VUE DE L'INTÉGRATION D'UNE COMPOSANTE VOCALE

Dans un dialogue en langage naturel, l'homme et la machine conversent idéalement dans une même langue. Or, nous savons que la réponse à la requête d'un usager peut se trouver n'importe où sur le réseau Internet. J'ai donc étudié les relations qui existent entre les différentes pages du Web pour vérifier que rien ne s'oppose à la mise en place d'une architecture générique pour l'intégration d'une couche de dialogue homme-machine sur Internet. Plus globalement, on peut se demander jusqu'où peut mener une page HTML si l'on suit tous les liens hypertextes qu'elle comporte et à quel moment va-t-on physiquement changer de pays (donc souvent de langue), sachant qu'actuellement rien d'explicite n'informe l'utilisateur du Web, qu'il passe d'un site à un autre, voire d'un continent à un autre.

Une étude de la propagation au sein du Web à travers les liens hypertextuels a été menée et a fait l'objet d'une publication internationale. Le robot informatique (spider) à été lancé sur le Web depuis 6 pages de base à travers le monde, de janvier à mars 1997.

Les analyses des statistiques comparatives faites sur des pages issues de différents continents montrent la manière dont ces sites se font mutuellement référence. Il apparaît que tous les pays étudiés font référence à de très nombreux autres pays ou institutions. L'étude montre aussi qu'environ 50% des images d'une page Web sont cliquables. Cela signifie qu'elles ne sont plus seulement des éléments «décoratifs» pour capter l'attention de l'utilisateur, mais qu'elles donnent, presque une fois sur deux, la possibilité d'interagir avec la machine. L'image prend de plus en plus d'importance sur le Web, et il faudra en tenir compte pour l'intégration d'une composante vocale. En effet, si le fait de passer d'un texte écrit à sa synthèse vocale ne pose pas de problèmes techniques particuliers, il n'en va de pas de même pour une image.

En bref nous avons mis en évidence que le dialogue en langage naturel homme-machine doit s'effectuer indépendamment de la requête que l'on souhaite lancer sur Internet pour éviter d'amener des ambiguïtés dans la communication. Nous avons démontré que l'on saute très facilement d'un pays à l'autre sur le Web : les données rapatriées en diverses langues ne doivent donc pas interférer avec le dialogue.

## 3 PERSPECTIVES DE RECHERCHE

Nos premiers résultats montrent qu'il est techniquement possible de piloter un feuilleton à la voix sur Internet [1]. Mais pour être efficace, ce moyen d'interaction doit être plus centré sur le dialogue entre l'homme et la machine. Nos travaux préliminaires contribuent au même objectif : l'étude des interfaces informatiques pour une meilleure communication homme-machine. Chacun des éléments étudiés (le dialogue, la multimodalité, les propriétés intrinsèques du Web, etc.) permettent de jalonner ce travail et de l'appréhender dans toutes ses dimensions.

Une étude ergonomique commencée aux archives départementales de l'Isère, en collaboration avec André Bisseret, permettra de mieux comprendre comment les différents usagers structurent leurs connaissances, et selon quel modèle cognitif ils évoluent. J'envisage d'enregistrer un corpus audio de dialogues entre lecteurs et archivistes, afin d'étudier les marqueurs (syntaxiques, prosodiques, linguistiques, etc.) qui entrent en jeu lors d'un dialogue « réel » entre novice et expert.

Je continue également à travailler dans le domaine de la recherche d'information sur Internet, et prépare une étude des dialogues en langue naturelle entre la machine et l'homme mis en situation de recherche des documents sur le World Wide Web.

## 4 BIBLIOGRAPHIE

[1] **ROUILLARD, J., CAELEN, J.** *A multimodal browser to navigate and search information on the Web.* Fourteenth International Conference on Speech Processing (ICSP97), IEEE Korea Council, IEEE Korea signal processing society, Seoul, Korea, August 1997.

[2] **ROUILLARD, J., CAELEN, J.** *Étude de la propagation au sein du Web à travers les liens hypertextes.* Quatrième conférence Internationale Hypertextes & Hypermédias, Septembre 1997, Paris. Numéro spécial de la revue Hypertextes et Hypermédias, éditions Hermès, 1997, Paris.