



**HAL**  
open science

# Hyperdialogue Homme-Machine sur le World Wide Web : Le système HALPIN

José Rouillard

► **To cite this version:**

José Rouillard. Hyperdialogue Homme-Machine sur le World Wide Web : Le système HALPIN. ERGOIA '98, 1998, Biarritz, France. pp.343-345. hal-02440335

**HAL Id: hal-02440335**

**<https://hal.science/hal-02440335>**

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.

# **Hyperdialogue Homme-Machine sur le World Wide Web : Le système HALPIN**

**José Rouillard**

Laboratoire CLIPS-IMAG, Groupe GEOD

Campus Scientifique, BP 53, 38041 Grenoble Cedex 9 – France

Téléphone : 04.76.63.56.51 - Fax : 04.76.63.55.52 - *E-mail* : Jose.Rouillard@imag.fr

## **Résumé :**

Un problème de circularité freine la conception et le développement de systèmes informatiques coopératifs adaptés aux capacités cognitives des utilisateurs : comment disposer de données pour réaliser un système qui n'a pas encore produit ces données ?

Pour supplanter les méthodes trop contraignantes de recueil de corpus de dialogues Homme-Machine (DHM) de type Magicien d'Oz et messagerie électronique, nous proposons un système d'Hyperdialogue avec un Agent en Langage Proche de l'Interaction Naturelle (HALPIN). C'est une nouvelle méthodologie de recueil de corpus via le World Wide Web (WWW) pour obtenir rapidement des données originales, qui devrait permettre, à terme, d'accroître la pertinence des modèles de DHM.

La gestion du dialogue s'effectue grâce à une technique basée sur la reconnaissance de concepts dans le discours. Les premiers résultats obtenus grâce à cette méthodologie de recueil de corpus apportent des éléments importants quant aux connaissances du domaine de la part des usagers, des expressions linguistiques utilisées et du dialogue qui s'instaure entre l'Homme et la Machine.

**Mots-clés :** Hyperdialogue, Interaction Homme-Machine, Dialogue en langue naturelle, World Wide Web.

## **Abstract :**

The conception and development of cooperative computer systems adapted to the cognitive capabilities of users are slowed down by a circular issue: how to have data to implement a system who has not yet produce those data ?

To supersede the too restricted methods of Human-Machine dialogue corpus collection such as the Wizard of Oz or electronic messages, a new methodology is employed to get quickly original data on the World Wide Web (WWW) that will allow to increase the relevance of the Human-Machine dialogue models. It is an Interactive Natural Language system based on a Hyperdialogue with an Agent (HALPIN).

The management of the dialogue is carried out thanks to a technique based on the recognition of concepts in the speech. The first results obtained thanks to this methodology of collection of corpus bring significant elements as for knowledge of the field on behalf of the users, linguistic expressions used and dialogue which is established between the Man and the Machine.

**Keywords :** Hyperdialogue, Human-Machine Interaction, Natural language dialogue, World Wide Web.

## **1. INTRODUCTION**

Dans le domaine du dialogue Homme-Machine, les chercheurs souhaitent concevoir et développer des systèmes capables de dialoguer avec un interlocuteur humain, de manière pertinente. Cependant, ils sont toujours confrontés à un problème circulaire, signalé par Siroux [5] : comment disposer de données pour réaliser un système qui n'a pas encore produit ces données ?

On note, en effet, que l'étude de la communication homme-machine se base principalement sur des observations de communication homme-homme. Or, il existe bien des phénomènes différents qui interviennent dans une interaction homme-machine. Pour progresser dans ce domaine, nous proposons un système de recueil et de traitement de corpus de dialogue homme-machine réel via le WWW. Le problème de circularité évoqué ci-dessus peut alors se résoudre en implantant ce système initial puis en l'améliorant au fur et à mesure de l'analyse des résultats obtenus en utilisation réelle.

Nous posons comme hypothèse que cette technique permettra de recueillir des corpus de DHM réels, finalisés, et en langue naturelle, de manière plus pertinente que les expériences de type Magicien d'Oz qui sont délicates et coûteuses à élaborer. Le système HALPIN que nous présentons ici se base sur ces principes et fonctionne sur le WWW.

## **2. LE SYSTEME HALPIN<sup>1</sup>**

La notion d'hyperdialogue est utilisée pour une expérience sur le WWW où l'utilisateur participe à un jeu interactif avec la machine. L'hyperdialogue est un dialogue Homme-Machine, coopératif et finalisé, dans un environnement hypertextuel [4]. Cet hyperdialogue peut être écrit, oral, voire gestuel. Dans le système HALPIN, il est à la fois écrit et oral.

Notre domaine d'étude est la recherche d'information documentaire. Nous voulons analyser des paramètres qui n'ont pas encore été étudiés en réel DHM comme le type de langage employé face à une machine (que s'autorisent à dire les usagers en langue naturelle ?), les stratégies de dialogue employées, la logique d'enchaînement des tours de parole et la représentation mentale que se font les usagers du système.

Pour cela, l'ordinateur tire un livre au hasard dans une bibliothèque universitaire grenobloise, et en stocke les principales caractéristiques (auteur, éditeur, mots-clés, etc.). Le but de l'utilisateur est de retrouver cet ouvrage en posant des questions en langue naturelle. La machine tente de comprendre les énoncés et d'y répondre correctement. L'objectif est atteint lorsque l'utilisateur donne à la machine le numéro ISBN (International Standard Book Number) correspondant au livre en question.

## **3. RÉSULTATS**

Les résultats obtenus sont de plusieurs ordres :

- Nous avons montré qu'il est techniquement possible de recueillir des DHM riches et variés sur le WWW. Le corpus HALPIN compte environ 950 fichiers provenant de toute la France et de l'étranger<sup>2</sup>. Les informations ainsi recueillies seront une source importante de données pour la modélisation du DHM (richesse et complexité linguistique des énoncés, agencement des tours de parole, incidences, etc.).

---

<sup>1</sup> <http://hermes.imag.fr/~rouillar/dialogue/vocal>

<sup>2</sup> Québec, Belgique, Suisse par exemple.

- Notre gestion du DHM par reconnaissance de concept dans le discours est relativement efficace, et autorise une communication H/M proche d'un dialogue en langue naturelle.
- L'utilisation d'un système incrémental est efficace et permet son amélioration périodique. On analyse les énoncés non compris par la machine et l'on augmente son vocabulaire en conséquence, pour ainsi accroître ses capacités de dialogue.
- Le système HALPIN est capable de formuler une synthèse vocale [2] de la réponse que l'utilisateur peut entendre si l'ordinateur est équipé de matériel audio.

#### **4. CONCLUSIONS ET PERSPECTIVES**

Nous avons développé une nouvelle méthodologie de recueil de corpus via le WWW pour obtenir rapidement des données originales, en vue de modéliser les connaissances du domaine, les expressions linguistiques et le dialogue : le système HALPIN. L'intérêt de l'apprentissage incrémental du dialogue en termes de stratégies et de connaissances réside dans le fait que l'on atténue les inconvénients des techniques " Magicien d'Oz " (c'est bien la machine qui dialogue et non pas un compère) et les lacunes des méthodes " messageries électroniques " (l'interaction se fait en temps réel, sauf lenteur du réseau).

L'utilisation du WWW comme support pour le recueil de corpus Homme-Machine est également intéressante dans la mesure où il peut toucher un grand nombre et une grande variété d'utilisateurs. La notion d'hyperdialogue couplée à celle d'apprentissage par concepts devraient permettre d'apporter de meilleurs systèmes de gestion du dialogue. La prochaine étape dans cette perspective sera l'intégration d'un thésaurus pour une meilleure reconnaissance des concepts énoncés par l'utilisateur d'une part, et l'utilisation d'un plus large vocabulaire lors de la génération de réplique, d'autre part.

#### **5. REFERENCES**

[1] CAELEN, J., *Compte-rendu du " workshop " interfaces homme-machine multimodales*. GDR-PRC " Communication homme-machine ", Dourdan, Avril 1992.

[2] DUTOIT, T., PAGEL, V., PIERRET, N., BATAILLE, F., VAN DER VRECKEN, O., *The MBROLA Project: Towards a Set of High-Quality Speech Synthesizers Free of Use for Non-Commercial Purposes*, ICSLP'96, Philadelphia, vol. 3, pp. 1393-1396.

[3] GAUSSIER, E., GREFENSTETTE, G., SCHULZE, M., *Traitement du langage naturel et recherche d'informations : quelques expériences sur le français*. Premières Journées Scientifiques et Techniques du Réseau Francophone de l'Ingénierie de la Langue de l'AUFELF-UREF, Avignon, Avril 1997.

[4] ROUILLARD, J. *Les enjeux d'un dialogue Homme-Machine sur Internet - L'Hyperdialogue*. Bulletin d'informatique approfondie et applications, revue de l'université de Provence, (à paraître n° 49, 1998).

[5] SIROUX, J., GILLOUX, M., GUYOMARD, M., SORIN, C., *Le dialogue homme-machine en langue naturelle : un défi ?* Annales Télécommunication, 44, n° 1-2, 1989.

#### **6. REMERCIEMENTS**

Nos remerciements vont à la Région Rhône-Alpes, et à l'équipe de Rank-Xerox de Meylan (Eric Gaussier, Grégory Grefenstette, et Maximilian Schulze) qui nous a autorisé à utiliser son outil de lemmatisation disponible sur le Web.