



HAL
open science

Parler aux machines, coproduire un service Intelligence artificielle et travail du client dans les services vocaux automatisés

Julia Velkovska, Valérie Beaudouin

► To cite this version:

Julia Velkovska, Valérie Beaudouin. Parler aux machines, coproduire un service Intelligence artificielle et travail du client dans les services vocaux automatisés. Emmanuel Kessous; Alexandre Mallard. La Fabrique de la vente. Le travail commercial dans les télécommunications., 2014. hal-02192400

HAL Id: hal-02192400

<https://telecom-paris.hal.science/hal-02192400>

Submitted on 23 Jul 2019

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.

Velkovska J. et Beaudouin V. (2014), « Parler aux machines, coproduire un service. Intelligence artificielle et travail du client dans les services vocaux automatisés », in Emmanuel Kessous et Alexandre Mallard (dir.), *La Fabrique de la vente. Le travail commercial dans les télécommunications.*, Paris, Presse des Mines, pp. 97-128.

Parler aux machines, coproduire un service. Intelligence artificielle et travail du client dans les services vocaux automatisés

Julia Velkovska, Valérie Beaudouin

Introduction	2
1 Les rêves de rationalisation : les automates vocaux à la croisée des utopies économiques et technologiques	3
1.1 Mutations de la relation de service et de la participation du client à la coproduction	4
1.2 L'intelligence artificielle en débat : règles, contexte, compétences	5
2 Le travail du client dans les services vocaux automatisés : gérer des simulacres de conversations	10
2.1 Terrain et méthodes.....	10
2.2 Le travail de standardisation des échanges comme condition de l'alignement avec la machine	11
2.3 Le travail d'enquête du client face aux désalignements interactionnels	15
Conclusion	19
Références	21

INTRODUCTION

Les interactions de service font sans doute partie des activités qui ont été les plus transformées ces dernières décennies par l'introduction massive de technologies de communication¹, notamment dans un souci d'automatisation (interfaces vocales, interfaces web, bornes interactives de vente et d'information, avatars). Deux éléments de justification sont avancés pour expliquer le déploiement des automates : des enjeux de rentabilité pour les entreprises, des enjeux de « modernisation de l'accueil » et « d'accessibilité des services » pour les clients (concernant notamment les moments et lieux de l'accès). L'automatisation, présentée comme une évolution économique et technologique « naturelle » et inéluctable de la société, est peut être en train de reconfigurer en profondeur les relations de service telles que nous les connaissons et d'agencer de nouveaux modes de rapport aux organisations et aux machines. Parce qu'elles constituent une des propriétés distinctives de la culture occidentale, comme le souligne Goffman (1968, [1961])², les relations de service méritent d'être étudiées au-delà de leurs fonctions strictement économiques. Dans ce cadre, il s'agit d'élucider les « présupposés » et les « théories » sur lesquelles repose le modèle de relation de service en tant que partie d'une culture qui englobe et dépasse les situations observées.

Parmi les technologies de la relation de service, les agents conversationnels intelligents (ACI)³ représentent la pointe la plus avancée de l'automatisation en étant les dispositifs les plus « anthropomorphes ». Issus des travaux en intelligence artificielle, ces technologies vocales visent à imiter au plus près la conversation humaine et sont employées pour le traitement partiel ou intégral des appels téléphoniques à des administrations ou à des entreprises. Les interactions qui s'engagent entre les appelants et les machines « parlantes » ont pour spécificité d'avoir l'apparence d'une conversation – un échange verbal séquentiellement organisé – sans en être une dans les faits. Plus précisément, comme nous allons le voir, ces interactions ne comportent pas les traits essentiels qui définissent la conversation : l'orientation commune des interlocuteurs vers sa structure temporelle et le jeu d'anticipations et de projections. Comment peut-on caractériser l'interaction d'une personne avec un automate vocal ? Quelles sont les conséquences de cette forme d'automatisation sur les modalités pratiques de la « relation » de service, notamment sur les formes de participation du client à la coproduction ?

L'article explore cette double piste à travers l'analyse de dialogues vocaux entre des clients de France Télécom et un agent conversationnel intelligent déployé dans un service client après-vente par téléphone. La nature de nos données permet d'analyser le déroulement concret de cette forme automatisée de « rencontre » et d'examiner finement les activités des clients. Ces dimensions restent inexplorées dans les travaux sur l'automatisation qui a été surtout discutée du point de vue des entreprises (rationalisation de la gestion de la relation client) et de celui des professions et des métiers (disparition ou transformation de certaines activités professionnelles). Par ailleurs, la difficulté du recueil de données naturelles sur l'interaction homme-machine comme celles de notre corpus explique sans doute la rareté des recherches basées sur ce type de données. A notre connaissance, il n'existe qu'une étude travaillant sur des données interactionnelles, celle de Wooffitt et ses collègues, mais elles proviennent de simulations de type «

¹ Plus généralement, les formes de la conversation ont été transformées en profondeur par l'introduction des technologies « conversationnelles », du téléphone à l'internet en passant par les automates vocaux (cf. Hutchby, 2001).

² « Chaque société offre diverses possibilités d'approche et de relations entre deux individus [...] Dans la société occidentale, le service dispensé ou reçu joue un rôle important dans le développement des relations entre deux individus. » (Goffman, 1968, [1961], p. 377).

³ Ces technologies sont connues sous des appellations diverses, comme automates vocaux, interfaces vocales, interfaces de dialogue naturel ou de langage naturel, serveurs vocaux interactifs (SVI), agents intelligents. Ces termes sont utilisés de façon interchangeable dans ce texte.

Magicien d'Oz », où le comportement de la machine est simulé (*cf.* Gilbert & Wooffitt 1990 ; Wooffitt & Fraser 1993 ; Wooffitt & MacDermind, 1995 ; Wooffitt & al., 1997).

Les résultats de notre enquête sont mis en perspective avec les « théories » et « les présupposés », dans le sens de Goffman, qui président à l'invention et à la mise en œuvre d'une relation de service automatisée : d'une part, les approches économiques de la rationalisation et de la rentabilité à travers la délégation de tâches aux machines, d'autre part, les avancées des travaux en intelligence artificielle dans le domaine du « dialogue naturel » visant à copier la conversation humaine. En effet, le modèle de service automatisé émerge à travers la rencontre entre une utopie économique, celle de la gestion rationnelle des contacts avec les clients, et une utopie technologique, celle de l'intelligence artificielle (partie 1). Ensuite, ces « théories » et « présupposés » qui fondent l'automatisation sont mis à l'épreuve sur des situations réelles d'accès aux services (partie 2). Ce geste fait apparaître les écarts entre le projet des concepteurs et la réalité des interactions, écarts qui forcent à remettre en question les utopies qui président à la conception des services automatisés et à appréhender leurs conséquences pratiques. L'analyse proposée articule ainsi les trois perspectives: sociologie des services, sociologie des technologies et sociologie des interactions homme-machine.

1 LES REVES DE RATIONALISATION : LES AUTOMATES VOCAUX A LA CROISEE DES UTOPIES ECONOMIQUES ET TECHNOLOGIQUES

Les interactions avec les automates vocaux constituent un objet hybride et difficile d'accès puisqu'elles défient les cadres de pensées disponibles. Simulations de l'interaction de service et plus généralement de la conversation humaine, elles ne se laissent pas saisir pleinement par les outils conceptuels et analytiques existants ni dans le domaine de la sociologie des relations de service, ni dans celui de la sociologie de l'action et de la conversation. Elles forcent à repenser les distinctions établies comme celle qui oppose conversation et rapport purement instrumental à une machine ; ou encore la distinction entre relation de service et participation opérationnelle du client à la coproduction⁴.

Dans l'univers de l'automatisation des contacts avec les clients, une des spécificités des interfaces vocales est de proposer une forme de « personnalisation » de la coproduction opérationnelle : donner l'illusion de la 'naturalité' de l'échange — en visant l'imitation de la conversation humaine — tout en répondant aux enjeux de rationalisation, de standardisation et de réduction des coûts commerciaux. Cette tension entre personnalisation et standardisation est précisément le point de rencontre entre la perspective gestionnaire et celle de l'intelligence artificielle.

Traitant des questions spécifiques venant de chaque domaine, ces deux visions se trouvent néanmoins confrontées à un problème pratique commun : comment réduire la complexité et la contingence de la conversation pour la rendre prévisible, et donc modélisable ? Pour y répondre, elles se rejoignent dans la recherche d'une modélisation et d'une conception décontextualisée de celle-ci. Nous retraçons dans cette section ces deux raisonnements et le contexte de leur rencontre qui donne naissance au modèle de l'interaction de service automatisée.

⁴ J. Gadrey insiste sur la nécessité de distinguer les notions de relation de service et coproduction opérationnelle : « De même que, à l'intérieur d'une organisation, la division technique du travail n'implique pas toujours l'existence de relations sociales directes entre les agents ainsi divisés, de même la coproduction opérationnelle (c'est-à-dire le fait que le client intervienne lui-même sur l'objet à « réparer » ou sur le problème à « traiter ») est-elle d'une autre nature que les interactions sociales entre les protagonistes du service. » (Gadrey, 1994, p. 383). Concernant la coproduction opérationnelle, l'auteur note que « [l]a participation du client peut fort bien, de son côté, revêtir un caractère purement opérationnel, parfois totalement impersonnel (dans certaines formules de « libre-service »). Par opposition, les relations de service impliquent des « relations sociales interactives dans lesquelles les acteurs sociaux agissent les uns vis-à-vis des autres, ou les uns sur les comportements des autres ».

1.1 Mutations de la relation de service et de la participation du client à la coproduction

Dès les années 1960 certains économistes observent que la production de services se distingue de la production des biens industriels entre autre par la participation nécessaire du client. Ainsi, Fuchs et Leveson en 1968, comparant le secteur des services au secteur industriel, identifient dans le premier « l'importance du consommateur comme agent coopératif dans le processus de production » (Fuchs & Leveson, 1968, p.194). La notion de *coproduction de service* trouve ici une de ses origines et souligne la contribution du client à l'accomplissement du service dans une perspective instrumentale. Cette perspective a été élargie par certains courants de la sociologie et de l'économie⁵, puisant dans l'interactionnisme, qui ont souligné que la relation de service est d'abord une relation sociale dont les aspects rituels et même civiques sont aussi importants que les composantes techniques et contractuelles⁶. A l'origine de ces courants, Goffman a identifié les propriétés interactionnelles de la coproduction qui débordent la dimension instrumentale.

La spécificité du modèle goffmanien par rapport aux recherches sur la coproduction, est de placer au centre de l'analyse « les données morales et sociales, sous-jacentes aux services spécialisés » et la question de la dignité. La relation de service implique la constitution d'un lien moral entre le client et le professionnel, et la construction d'une relation de confiance. C'est aussi pour cela que Goffman exclut les services automatiques (compostage de billet, standardiste...) du périmètre de la relation de service. C'est sans doute une des pistes les plus fécondes et les moins explorées aujourd'hui pour comprendre les bouleversements récents dans les « modalité[s] de connexion entre les prestataires et les clients » (Bandt et Gadrey, 1994). Les technologies de « dialogue naturel » semblent en voie de brouiller la frontière tracée par Goffman entre le « service automatique » et « le service comme relation sociale et morale », comme en témoignent les formats interactionnels d'expression des émotions (notamment l'indignation ou la colère), de plainte et d'imputation de responsabilité à la machine en cas de dysfonctionnement. Ces formats semblent intimement liés aux modalités de l'interaction (de type conversationnel) et aux enjeux des rapports avec l'entreprise que les clients ont à traiter par ce biais. Loin des transactions simples comme le compostage de billet, les clients contemporains sont amenés à interagir avec les automates vocaux, et à leur faire confiance⁷, pour leurs factures, leurs abonnements à des services ou la résolution de divers problèmes de l'après-vente. Dans ce contexte, les transformations dans la contribution des clients à la coproduction sont liées non seulement aux propriétés des technologies vocales, mais également à la façon dont celles-ci sont intégrées dans l'écologie globale du dispositif de communication entre l'entreprise et ses clients, la « relation client », qui se caractérise par une multiplication et une spécialisation des canaux de contacts.

En effet, les transformations actuelles de la « relation client » sont liées à la gestion rationnelle des contacts (dite aussi « industrialisation ») dans un contexte d'extension des points d'adhérence entre le client et l'organisation et à la part croissante que prend la relation standardisée. A côté des automates vocaux, les formes les plus visibles de cette mutation sont d'une part l'essor des centres d'appels avec

⁵ Cf. en France les travaux de J. Gadrey (1994a), d'I. Joseph (1988) et du réseau Langage et travail (Borzeix,1992 ; M. Grosjean, 1995 ; Borzeix & Fraenkel, 2001).

⁶ Les trois registres goffmaniens de la relation de service (technique, contractuel et rituel) ont été enrichis dans les travaux en France par un quatrième dimension, « civique » ou « civile » relative aux arguments de justice (pour une synthèse, cf. Gadrey, 1994).

⁷ Prenant au sérieux la question de la confiance Justine Cassell et ses collègues cherchent à concevoir des Agents Conversationnels Incarnés (Embodied Conversational Agents), équipés non seulement de compétences techniques, mais aussi de compétences sociales (relevant par exemple du comportement non-verbal crédible ou de la capacité de conduire des échanges informels). Les auteurs font l'hypothèse que l'introduction du « papotage » (« small talk ») dans les interactions automatisées renforce la confiance et aide à la collaboration et à la résolution de problèmes. Se basant sur des études de corpus naturels, des dimensions telles que la familiarité, la solidarité et l'affection ont été explorées dans la conception d'un agent conversationnel animé pour le conseil dans l'immobilier - REA (Real Estate Agent). (Cassell & al. 2000, Cassell 2001, Cassell & Bickmore, 2004).

l'usage de « scripts » conversationnels, d'autre part le développement des applications sur internet, accessibles sur différents terminaux qui permettent au client de gérer ses services.

Ces évolutions organisationnelles et technologiques donnent lieu à de nouvelles formes de contribution du client à la production des services. Celles-ci restent à explorer et à décrire tant se multiplient ces situations nouvelles de communication entre l'entreprise et ses clients. Cet article est une contribution à cette démarche centrée sur les automates vocaux qui suscitent un intérêt vif de la part des organisations car ils sont vus comme porteurs d'une double promesse : amélioration de la rentabilité en déléguant le travail humain à des machines tout en garantissant au client une relation « de qualité » grâce au caractère « humain » de la machine, une machine dite capable de penser et de converser, portée par l'utopie de l'intelligence artificielle.

1.2 *L'intelligence artificielle en débat : règles, contexte, compétences*

La possibilité même de construire des machines qui peuvent converser au sens où les interlocuteurs humains le font est au cœur de débats pluridisciplinaires (entre l'informatique, la philosophie de l'esprit, la psychologie cognitive, la sociologie et notamment les approches de l'action située) concernant *les agents intelligents parlants*⁸. L'ethnométhodologie et ses développements dans l'analyse conversationnelle (AC) se sont naturellement intéressés aux problèmes posés par la modélisation de la conversation comme forme de l'action. Nous verrons que les positions sont divisées à l'intérieur même de ce domaine.

Ce débat s'inscrit dans une controverse plus large portant sur *les agents intelligents* qui interroge le postulat fondamental autour duquel s'est formé le domaine de recherche sur l'intelligence artificielle, à savoir la possibilité de modéliser l'action humaine. Ainsi la question de la copie – de l'action, du raisonnement, de la conversation humaine – constitue le point focal des discussions entre les tenants du courant de l'intelligence artificielle (IA) et ses critiques. La divergence des points de vue entre les deux courants porte principalement sur la nature même des actions humaines.

Pour résumer très schématiquement ce débat, notons que pour les critiques de l'IA, les propriétés de l'action et de la conversation humaine – en particulier leur caractère situé et contingent – rendent leur modélisation impossible. En revanche, pour les tenants de l'IA, il est possible d'extraire les *règles* qui gouvernent les activités humaines, de les agencer dans des séries abstraites des *contextes* concrets de leurs manifestations et de les implanter dans les machines. Ces machines seraient alors dotées de *compétences* et de capacités de raisonnement, de décision et d'action comme les humains et pourraient agir et interagir comme eux. L'atteinte de cet objectif ne serait qu'une question de temps et de progrès technologique pour surmonter des difficultés qui ne sont que d'ordre technique mais ne dépendent en aucun cas d'obstacles à la modélisation intrinsèques à la nature de l'action humaine.

Comme nous le verrons dans les deux paragraphes suivants qui présentent les arguments des débats, les concepts de *règle*, de *contexte* et de *compétences* occupent une position centrale et se situent à l'origine de certains malentendus.

1.2.1 *Copier l'intelligence ?*

Cette section donne un aperçu des débats qui ont accompagné le développement de l'intelligence artificielle dès ses origines autour des problèmes épistémologiques que pose la modélisation de l'action et de l'intelligence sous forme d'instructions désincarnées. Ce détour permet de restituer l'automatisation de la conversation (section 1.2.2), en particulier commerciale, et le phénomène de l'élargissement du travail du client que nous observons dans la partie 2 dans la continuité et dans l'épaisseur historique des recherches de règles susceptibles de gouverner le raisonnement et l'action humains.

⁸ On trouve une synthèse du débat entre l'IA et l'action située dans l'introduction de Donald Norman au numéro spécial (1993, 17 (1)) que la revue *Cognitive Science* a consacré à cette thématique.

Les années 1950 et 1960 du 20^e siècle marquent le début de l'ère de l'intelligence artificielle autour des travaux de chercheurs comme Herbert Simon, Allen Newell, Claude Shannon, Marvin Minsky, etc. Les développements de la théorie heuristique⁹ et leur mise en œuvre dans des programmes de simulation comme le *General Problem Solver* inventé par Simon et ses collègues s'accompagnent d'une série de déclarations et de promesses enthousiastes. Ainsi en 1957, H. Simon annonce « qu'il existe désormais au monde *des machines capables de penser, d'apprendre et de créer*. Qui plus est, le champ de leur possibilités est appelé à s'élargir à une cadence rapide jusqu'au jour où – dans un avenir qui n'est pas si lointain – *la gamme des problèmes qu'elles seront à même de traiter équivaudra à celle que peut appréhender l'esprit humain...* »¹⁰. Dans un article de 1958 de H. Simon et A. Newell, intitulé « Heuristic Problem Solving : The Next Advance in Operations Research », l'idée de l'effacement des frontières et des différences entre les humains et les non-humains est clairement présente : « L'intuition, l'inspiration, la perspicacité, la faculté d'apprendre *ne sont désormais plus l'apanage des humains : n'importe quel gros ordinateur puissant et rapide peut également en faire preuve lui aussi.* »¹¹.

Le projet de l'IA vise donc à concevoir des machines dotées d'intelligence. Mais qu'est-ce qu'on entend par intelligence dans ce domaine ? En analysant les travaux en IA, Hubert Dreyfus note que dès les débuts le concept d'« intelligence » retenu est ancré dans une tradition philosophique platonicienne, qui valorise le savoir propositionnel (*savoir que*) comme connaissance au détriment du savoir pratique (*savoir comment*) qui n'est qu'une « sorte de tâtonnement arbitraire » et relève de la conviction¹². Dans cette tradition, le calcul serait le modèle même du raisonnement humain puisqu'il est possible de le représenter par une série d'instructions ou de règles abstraites et reproductibles. Selon Dreyfus, c'est cette conception syntaxique (*vs.* sémantique) de la pensée qui est reprise dans le projet de l'IA des années 1970 : la poursuite de l'entreprise de l'IA s'articulera autour de la recherche de procédures permettant de traduire différentes activités humaines en une série d'instructions¹³.

Le projet du livre de Dreyfus est d'élaborer une critique de l'intelligence artificielle, en particulier d'identifier « les limites de l'intelligence chez les ordinateurs » (p. 18) pour nous éclairer sur le fonctionnement de l'intelligence humaine. Sa thèse générale pourrait être résumée ainsi : pour faire sens du monde environnant et y inscrire ses actions, les êtres humains sont capables d'une *compréhension globale d'une situation* à l'intérieur de laquelle ils peuvent discriminer des éléments périphériques ou centraux pour l'activité en cours. Cette intelligence pratique est une compétence des humains intimement liée au fait qu'ils sont dotés de corps, toujours situés, qu'ils sont socialisés dans une culture donnée et qu'ils sont capables ainsi de saisir un contexte global, de s'y situer et de lui donner sens. Dreyfus écrit : « [...] l'intelligence sous-entend la *compréhension* au sens fort, et pour doter un ordinateur de la faculté de comprendre, il faudrait aussi lui fournir ce « sens commun » dont disposent les êtres humains adultes par le simple fait qu'ils possèdent un corps, qu'ils sont en interaction avec le monde matériel, enfin, qu'ils ont été formés à une culture. » (p. 34)

⁹ Les programmes « heuristiques » de résolution de problèmes intègrent la notion de « règle pratique » et d'apprentissage et tentent de simuler les raccourcis et les cheminements de l'intelligence pratique. Ils se distinguent des programmes « algorithmiques » « au succès garanti [...] qui cheminent vers la solution selon un processus exhaustif, mais qui deviennent vite trop lourds à manier dès lors qu'ils ont trait à des problèmes pratiques. » Dreyfus, 1984, p.14.

¹⁰ Cité dans Dreyfus, 1984, pp. 21-22, souligné par nous.

¹¹ Cité dans Dreyfus, *ibid.*, p. 16, souligné par nous.

¹² Dreyfus, *ibid.*, p. 4.

¹³ « De même que le terme « artificielle », celui d'« intelligence » peut faire naître un malentendu. Nul n'espère du robot ainsi attendu qu'il soit capable de reproduire la totalité de ce qui est tenu pour conduite intelligente chez l'être humain. Il n'aura pas besoin par exemple, de se trouver une épouse qui lui convienne, ni de traverser une avenue encombrée. Il n'est question pour lui que de rivaliser avec l'esprit humain dans les domaines les plus abstraits et les plus platoniques, comme par exemple de triompher du test de Turing. » (Dreyfus, *ibid.*, p. 17).

A partir de cette position épistémologique la poursuite de l'entreprise de l'IA – concevoir des machines intelligentes *comme* les humains – implique inévitablement une formalisation du sens commun : « Ou bien elle [l'IA] trouvera le moyen de représenter et d'ordonner ce « savoir-faire » de la vie courante dont fait preuve l'individu moyen, ou elle s'embourbera dans le fouillis de notions et de convictions qu'il faut rendre explicites pour tenter de « mettre au courant » cette entité désincarnée qu'est l'ordinateur, lequel ignore tout du vécu quotidien de l'être humain. » (p. 34).

Les idées de Dreyfus ont été reprises par la sociologie des sciences (Collins, Woolgar), puis dans le débat interne à l'ethnométhodologie et à l'analyse conversationnelle concernant plus particulièrement la modélisation de la conversation, notamment autour la *notion de règle*, et la capacité typiquement humaine de « suivre une règle »¹⁴. H. Collins (1998 et 2000) propose le concept de socialité (« socialness ») pour décrire la qualité fondamentale des êtres humains qui les différencie des autres espèces et des machines : c'est la capacité qui nous permet d'acquérir tous les savoir-faire pour lesquels nous ne pouvons pas formuler de règles et de les utiliser de façon créative¹⁵. La socialité comme capacité d'interpréter les règles permet d'anticiper le monde social et de comprendre un contexte d'action en le saisissant de façon dynamique et située « moment après moment » (Collins, 1998, p. 503). Les espèces qui possèdent la socialité se caractérisent par des *variations* entre les groupes dans leurs rapports au monde physique. Collins conclut que les animaux et les machines ne doivent pas être traités comme membres de collectivités sociales et qu'on ne doit pas confondre échange d'information et langage. La question que se pose l'intelligence artificielle sur les capacités des machines à agir comme des humains devient alors celle de la possibilité de leur socialisation.

Sans la socialité, les ordinateurs ne peuvent pas être capables d'accomplir des actions, mais seulement d'imiter certains comportements (Collins 2000). Cette distinction entre *se comporter* et *agir* permet de définir deux classes d'action en fonction de la stabilité du lien entre action et comportement. Dans la plupart des actions il n'y pas de correspondance fixe entre action et comportement, ce sont des actions « polymorphiques » : « agir de façon ouverte et créative tout en suivant des règles » (p. 35). Comprendre et mettre en forme des actions « polymorphiques » demande la capacité de comprendre la société (par exemple, écrire une lettre d'amour). Il n'est pas possible de les copier en copiant les comportements. En revanche, la deuxième classe, les actions « miméomorphiques », se caractérise par un rapport stable entre l'action et le comportement (par exemple la marche au pas, le travail à la chaîne ou certaines opérations arithmétiques). Selon Collins « [c']est seulement dans ces domaines que les ordinateurs et les autres machines peuvent directement remplacer des actions humaines sans avoir besoin de personne pour réparer les erreurs. » (p. 36).

1.2.2 Imiter la conversation ?

Comme nous venons de le voir, les débats sont vifs autour de la question de la modélisation de l'action humaine. On peut globalement distinguer deux postures : une conception maximaliste de l'IA pose qu'il est possible d'extraire des règles abstraites qui gouvernent les activités et le raisonnement humain et de les implanter dans les machines ; une conception critique de l'IA postule qu'il est impossible de modéliser

¹⁴ Cf. également à ce sujet Collins 1990 et 1993, Collins et Kush 1998, ainsi que le texte classique de C. Taylor, 1995, « Suivre une règle » qui développe la perspective wittgensteinienne sur les règles. Sans discuter explicitement les questions d'intelligence artificielle, Taylor synthétise et explicite les arguments que la philosophie et les sciences sociales opposent au fond à la modélisation de l'action. Taylor écrit : « La compréhension d'arrière-fond [...] qui sous-tend notre capacité à saisir des instructions et à suivre des règles est dans une large mesure incorporée. Ceci aide à expliquer la combinaison de traits qu'elle présente : c'est une forme de compréhension, permettant de trouver du sens aux choses et aux actions, mais en même temps entièrement informulée, tandis que, troisième point, elle peut servir de base à une formulation nouvelle » (p. 564).

¹⁵ Par exemple, nous savons produire dans le discours naturel des énoncés acceptables et inacceptables sans connaître explicitement les règles.

l'action humaine en général et donc de concevoir des agents intelligents. Les machines peuvent en revanche imiter avec succès certains comportements simples pour la résolution de problèmes circonscrits. Si l'on se limite au cas de la modélisation de la conversation comme une forme d'action spécifique, la situation est similaire : la possibilité de créer des machines qui « conversent » comme les humains est défendue par les uns, contestée par les autres.

Pourtant c'est bien par leur capacité à converser, à être un partenaire conversationnel « acceptable » qu'est éprouvée et mesurée leur « l'intelligence » depuis le test de Turing en 1950. Celui-ci permet de comprendre comment dans l'histoire de l'informatique et de l'IA la conversation humaine est devenue une référence dans la conception des technologies de parole dites de « dialogue naturel ». Pour traiter la question « *Can machines think?* »¹⁶, Turing imagine un jeu d'imitation organisé sous forme de dialogue. Un homme et une femme sont dans une pièce, l'enquêteur, dans une autre pièce, doit deviner qui est l'homme et qui est la femme. L'homme (A) cherche à tromper l'enquêteur sur son identité, tandis que la femme (B) cherche au contraire à l'aider. Le jeu se déroule sous la forme de questions-réponses retranscrites par un intermédiaire pour que la voix n'ait pas d'incidence sur l'identification du genre. Le cœur du test de Turing est le suivant : si on remplaçait A par une machine, saurait-il aussi bien que l'homme induire en erreur l'enquêteur sur son identité de genre ? Et c'est cette dernière question qui remplace la question initiale. Ce test est devenu un mythe technologique fondateur pour les recherches sur le dialogue homme-machine même si de nombreux travaux ont montré les impasses dans lequel ce test a pu conduire l'intelligence artificielle. L'évaluation de la capacité à raisonner d'une machine repose sur une conception logocentrique et dialogique de la pensée, comme le souligne François Rastier (1991). La nature du test, constitué de séquences de questions-réponses, est devenue structurante pour les travaux qui reprennent le dialogue comme forme 'naturelle' de la représentation de l'activité de pensée et de la résolution de problèmes.

Le débat qui s'est déployé dans les années 1980 et 1990 autour des machines parlantes, entre l'intelligence artificielle et les sciences sociales (en particulier dans les paradigmes issus de l'ethnométhodologie) oppose deux conceptions de la conversation. D'un côté, les ingénieurs traitent la conversation comme un échange de contenus où il s'agit de faire correspondre *ponctuellement* une requête, identifiée par des mots-clés, et une réponse, une solution, préprogrammée dans la base. Il est dès lors possible de travailler à la conception de technologies de « dialogue naturel » à partir de trois éléments de base : un module de reconnaissance vocale, un module de traitement (appariement entre requête et réponse dans une base de données) et un module de synthèse vocale. De l'autre, les ethnométhodologues, à la suite de Harold Garfinkel, opposent à cette approche une conception procédurale et située de l'action et de la conversation qui ne peuvent pas être réduites à la performance d'un plan ou d'une séquence de questions et de réponses (Suchman, 2007, Button, Coulter, Lee & Sharrock 1995, Button & Sharrock 1995). En effet, pour Garfinkel, la conversation suppose l'orientation vers un monde commun et le partage d'attentes d'arrière-fond¹⁷. La compréhension est liée au contexte, au non-dit, et possède une structure temporelle vers laquelle s'orientent les participants. Ainsi, la conversation repose sur un jeu d'anticipations et de projections et la compréhension n'est pas un événement unique et ponctuel, mais elle est produite et maintenue en permanence par un travail interactionnel : « beaucoup de choses ont été comprises en prenant les séries temporelles d'énoncés non comme un *chapelet de mots, mais comme les évidences documentaires d'une conversation se développant dans le temps.* » (p. 102).

La référence aux travaux de Garfinkel permet de situer le problème de la compréhension dans l'interaction homme-machine. En effet, dans ces échanges, contrairement à la conversation, le jeu de projections, d'anticipations et d'inférences, le « travail de compréhension » est absent de la part de

¹⁶ Turing, 1950.

¹⁷ Garfinkel, 2007, Chap. 2, "Quelques traits essentiels de la compréhension commune".

l'automate. Celui-ci traite les événements linguistiques en soi, cherche à établir des correspondances entre les requêtes et les réponses, mais fait abstraction de la dimension temporelle et contextuelle de l'échange. Ce fonctionnement n'est pas seulement contraint par les possibilités techniques actuelles mais s'enracine dans une ontologie de la conversation humaine comme simple transfert d'information. Dans cette conception l'organisation séquentielle et temporelle de la conversation apparaît comme un épiphénomène, alors qu'elle est une ressource centrale pour la compréhension mutuelle dans la perspective ethnométhodologique.

Malgré les incompatibilités théoriques entre la position des ingénieurs et celles des ethnométhodologues, des essais ont été faits pour utiliser, dans la conception technologique, les connaissances produites par l'analyse conversationnelle (AC), issue de l'ethnométhodologie. L'AC offre un formalisme pour l'analyse et la représentation de la conversation, séduisant pour les concepteurs. Certains de ses représentants¹⁸, ont utilisé les résultats de l'AC sur la gestion des tours de parole et les paires adjacentes pour en tirer des principes de design du dialogue homme-machine. D'autres auteurs¹⁹ contestent cette position en soutenant que le système de tours de parole (ainsi que la linguistique formelle) offre un formalisme mais n'informe pas sur la production des énoncés qui est à chaque fois localement émergente. En ce sens l'AC ne peut pas informer directement la conception, à part sensibiliser les concepteurs au caractère contingent et situé de l'activité conversationnelle.

Deux autres points posent problème pour la transposition de l'AC dans le design technologique. Le premier concerne la compétence des acteurs à interpréter et à s'ajuster à un contexte dynamique : cette compétence centrale dans la conversation n'est pas programmable, pas plus que ne l'est le « sens commun de la vie courante » pour reprendre Dreyfus. Le deuxième point concerne la confusion entre règles de conversation et règles de programmation. Alors que ces dernières sont une série d'instructions formant un algorithme, les règles de la conversation ne déterminent pas une séquence d'actions précise, mais elles *orientent* le comportement de ceux qui s'engagent dans une conversation (Button et Sharrock, 1995, p. 113)²⁰.

En conclusion, pour Button et Sharrock la machine peut accomplir certaines tâches précises et circonscrites, dans l'échange, elle peut simuler certains comportements pour des séquences standardisées et prévisibles. Mais l'ordinateur ne peut pas « comprendre » une conversation. La bonne simulation n'est pas une conversation. Cette thèse précise dans le domaine de la conversation les critiques formulées à l'égard de la modélisation de l'action.

Ce parcours des débats sur la modélisation de l'action et de la conversation montre qu'elle est systématiquement adossée à une vision abstraite de ces phénomènes, affranchie des questions de contexte. Qu'il s'agisse de compétences, de règles ou de mécanismes de la conversation, le prix à payer pour l'automatisation est une conception mécanique et décontextualisée de l'interaction. C'est également la conception sous-jacente à la rationalisation gestionnaire des contacts avec les clients : la demande du client peut être traitée par n'importe quel conseiller depuis n'importe quelle partie du monde suivant un script standardisé ou dans certains cas, par un automate vocal. Le traitement de plus en plus standardisé de ces contacts est présenté pourtant sous les apparences de la personnalisation « d'échelle » produite à partir de différentes technologies - des bases de données sur les clients, des traces d'usages, des technologies conversationnelles. Les relations de services sont ainsi traversées par la tension entre standardisation et personnalisation. L'ajustement entre ces deux tendances est souvent pris en charge par le client (par exemple, rechercher le bon interlocuteur et refaire plusieurs fois le récit de son problème) et relève de

¹⁸ Wooffitt et MacDermond, 1995 ; Wooffitt et al., 1997 ; Frolich et Luff, 1990 ; Gilbert, Wooffitt & Fraser, 1990, Fordham et Gilbert, 1995.

¹⁹ Button, 1990 ; Button, 1993, Button, Coulter et al., 1995,, Button & Sharrock, 1995.

²⁰ Cet argument est proche du développement plus général proposé par H. Collins du concept de « socialness », cette capacité typiquement humaine à interpréter des règles sans les violer.

nouvelles formes de sa contribution à la coproduction des services. Le travail d'alignement avec les interfaces vocales que nous analysons dans la suite en constitue un exemple. Nous désignons par là l'aspect unilatéral du travail de compréhension (et de réparation des malentendus) à l'œuvre dans l'interaction avec les services vocaux interactifs qui est entièrement à charge de la personne, l'interface ne faisant que réagir aux mots-clés reconnus. Nous parlerons d'alignement (et de désalignement) plutôt que de compréhension (et de malentendu) qui caractérisent la conversation humaine parce que l'interaction homme-machine ne comporte pas les traits caractéristiques et les régulations de la conversation, comme nous venons de le voir.

Si l'ambition des technologies vocales est d'imiter au plus près la conversation humaine, de s'adapter aux différents contextes et de proposer ainsi un service automatique, mais à « visage humain », dans l'état actuel de ces technologies du moins, c'est le contraire qui se produit. Le succès du dialogue (et donc l'accès au service) est corrélé à la capacité du client à s'ajuster de façon dynamique, à apprendre à interagir avec la machine et à devenir un utilisateur compétent de ces technologies. Nous mettons en évidence dans la suite quelques modalités concrètes de ces nouvelles composantes de la coproduction en contexte d'automatisation.

2 LE TRAVAIL DU CLIENT DANS LES SERVICES VOCAUX AUTOMATISÉS : GERER DES SIMULACRES DE CONVERSATIONS

Notre propos est à présent d'examiner concrètement la manière dont se déroulent les interactions entre des clients et un automate vocal intelligent. Par delà les débats organisationnels et théoriques, comment, dans la pratique, le service est-il ou non co-produit quand le représentant de l'organisation est remplacé par un automate vocal ? Y a-t-il ou non alignement entre le client et l'automate ? Quelle est la nature de ces interactions : s'agit-il de dialogue « naturel » ? Comment le client fait-il face à des situations de désalignements ? Pour ce faire, nous explorons un corpus d'interactions entre clients et automates vocaux dans le cadre de relation de service après-vente. Après avoir présenté les données sur lesquelles porte notre analyse, nous chercherons d'abord à identifier les facteurs qui favorisent la coproduction du service puis nous verrons comment les clients font face aux cas de désalignement.

2.1 Terrain et méthodes

Ce travail a été rendu possible grâce à la mise en place au sein de la R&D de France Télécom d'une infrastructure chargée de garder trace de toutes les expérimentations de relation de service avec des automates vocaux au sein de l'entreprise. L'idée était d'avoir un lieu de capitalisation pour pouvoir améliorer les services eux-mêmes, en grande partie développés en interne.

Nous avons limité notre investigation à deux services, qui correspondent chacun à un numéro d'appel (le 3000 et le 3900). Le premier, le 3000, que nous appellerons Service A, permet d'accéder par téléphone à tout un ensemble de services pour gérer la ligne de téléphone fixe (payer sa facture, faire des transferts d'appel, suivre sa consommation, accéder à la messagerie...). Le second, le 3900 ou service B, donne accès à l'assistance technique. Il permet d'identifier le client (via son numéro de téléphone) et le motif d'appel pour orienter vers les conseillers-clients adaptés. Le service B est un aiguilleur d'appels.

Les interactions entre clients et automates vocaux étudiées dans cette enquête sont produites dans des *situations naturelles* (et non expérimentales). D'une part, la situation est définie comme naturelle du point de vue des clients puisque ces derniers appellent l'entreprise avec une demande de service spécifique (information, solution à un problème...) et se trouvent confrontés à l'automate sans savoir qu'il s'agit d'une expérimentation. D'autre part, la situation est naturelle au sens où il s'agit d'échanges avec des agents intelligents et non d'expérimentation de type « Magicien d'Oz » où la machine est simulée (cf. les travaux cités plus haut de Wooffitt et ses collègues).

Dans la base de données, pour chaque interaction sont consignés le fichier audio de l'interaction, des données de description technique de l'échange (date, durée, nombre d'interventions, transcription

automatique...) et éventuellement des indications introduites par des personnes recrutées par les équipes de conception pour identifier le motif d'appel et qualifier son aboutissement : succès ou échec.

Notre démarche s'appuie sur deux approches différentes pour répondre à une même question, celle de l'alignement de l'interaction homme machine dans une relation de service. L'exploration statistique de l'ensemble des interactions a permis de produire une vue globale sur un échantillon d'interactions important (plus de 18 000) pour déterminer les facteurs qui favorisent l'accès au service. L'analyse conversationnelle a permis quant à elle de comprendre finement les mécanismes de cet alignement dans le fil de l'interaction et la nature des opérations de réparation qui se mettent en place en cas de désalignement. Pour l'exploration de ce type de données complexes, il nous a paru indispensable de combiner ces méthodes qui nous ont chacune apporté un regard complémentaire.

Sur le service A, ont été analysées toutes les interactions enregistrées dans la base entre janvier 2006 et mars 2007, soit 16777 appels, dont 10281 appels ont été qualifiés manuellement par des codeurs (motif de l'appel et aboutissement). L'exploitation quantitative porte sur des traits de description fermés (motif d'appel, aboutissement, durée, nombre d'interventions...). Pour constituer des collections d'appels pour l'analyse interactionnelle, pour chaque motif d'appel, nous avons choisi une dizaine de cas de succès et une dizaine de cas d'échec, en faisant varier la durée de l'interaction. Ces interactions ont été retranscrites manuellement à partir de l'enregistrement audio, en reprenant les principes de l'analyse conversationnelle.

Sur le service B, comme il n'y a pas eu de qualification manuelle des interactions, nous avons effectué un autre type de travail, qui a consisté à analyser les requêtes des clients, telles que le système de reconnaissance vocale les a retranscrites, dans un corpus de 1675 interactions clients, enregistrées entre avril 2006 et juin 2008. Certaines interactions du service B ont également été transcrites à la main et analysées en profondeur²¹.

2.2 Le travail de standardisation des échanges comme condition de l'alignement avec la machine

Dans cette section, nous explorons globalement le corpus des interactions qualifiées et nous cherchons à mettre en relation certains traits formels de ces interactions avec un indicateur de l'accomplissement du service (« échec » ou « réussite »). Contrairement aux promesses portées par les technologies vocales, le résultat central montre que moins le dialogue est « naturel », plus le service a des chances d'être rendu.

Pour le service A, 10281 interactions qui ont été qualifiées manuellement par des codeurs (identification du motif de l'appel et qualification en termes d'échec ou de réussite de l'interaction). La notion de réussite ou d'échec relève d'une interprétation du codeur et correspond à une évaluation externe de l'adéquation de la réponse de la machine à la demande du client.

En moyenne, les demandes des clients adressées au service A aboutissent trois fois sur quatre. La figure 1 présente la répartition des appels selon la durée et le taux de réussite. On note une concentration très forte des appels autour de la demi-minute et une longue traîne avec des appels pouvant durer très longtemps (courbe « nombre d'interactions »). En croisant la durée des interactions avec le taux de réussite, il apparaît que plus l'interaction s'étend dans le temps, plus le taux de réussite diminue. Au-delà de 70 secondes, le taux de réussite est de 60%. De fait, il existe une courte fenêtre de tir, entre 20 et 50 secondes pendant laquelle le taux de réussite est très élevé (plus de 80%).

²¹ Pour une description plus précise de la démarche méthodologique, cf. Beaudouin & Velkovska, 2010.

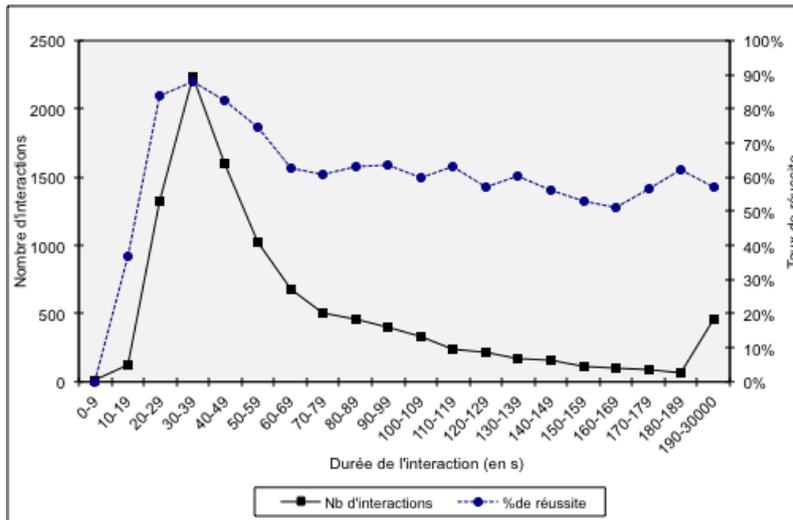


Figure 1 : Répartition des appels selon leur durée et taux de réussite associé

Quand on regarde le nombre d'interventions du client dans l'interaction, dans près de la moitié des interactions enregistrées (46%), le client n'intervient qu'une seule fois et c'est dans ces cas-là que le taux de réussite est le plus élevé (82%). L'augmentation du nombre d'interventions du client s'accompagne d'une diminution du taux de réussite. Dit simplement, plus le client parle moins il a de chances d'obtenir ce qu'il veut. Il semble y avoir un optimum d'efficacité du système quand le client n'intervient qu'une fois (cf. tableau 1). C'est une spécificité des dialogues avec les agents intelligents, les interactions ordinaires en service client se réduisant à une seule intervention du client étant plutôt inhabituelles.

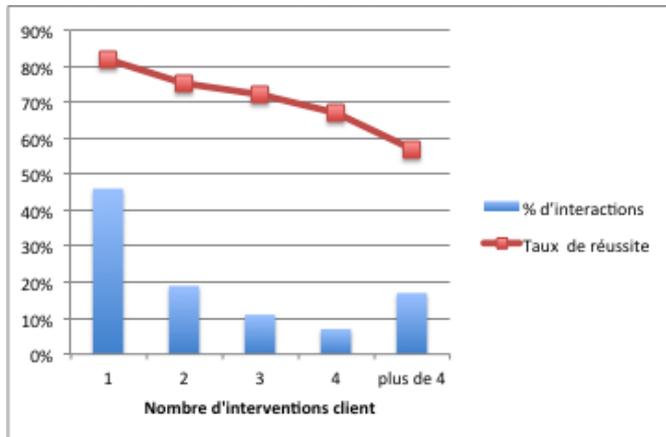


Figure 2 : Répartition des interactions en fonction du nombre d'interventions client

Si l'on essaie de différencier les interactions selon les raisons de l'appel, on note que les trois-quarts des appels correspondent à l'un des quatre motifs suivants : suivi consommation, transfert d'appel, paiement de facture et messagerie (cf. tableau 2). Le quart restant étant très dispersé.

Motif d'appel	%	nb interventions		durée moyenne médiane		Taux réussite
		moyenne	médiane	moyenne	Médiane	
Suivi consommation	33%	2,2	1	00:00:50	00:00:36	89%
Transfert d'appel	17%	4,4	3	00:01:20	00:01:04	83%

Paiement facture	15%	2,1	1	00:00:58	00:00:45	92%
Messagerie	7%	2,2	1	00:00:55	00:00:41	91%
Information	5%	6,0	5	00:02:36	00:02:04	69%
Ne sais pas	14%	3,0	2	00:01:11	00:00:59	7%
Hors périmètre	6%	3,9	3	00:01:47	00:01:33	63%

Tableau 1 : caractéristiques des interactions selon le motif de l'appel

Pour les demandes très standardisées et régulièrement utilisées (paiement de facture, messagerie, suivi consommation) le taux de réussite tourne autour des 90%. Quand il s'agit d'interaction *a priori* moins standardisées, comme les demandes d'informations ou des demandes hors périmètre du service A, le taux de réussite est beaucoup plus bas. Enfin, pour les autres demandes, non ou mal identifiées (dans 14% des cas, le codeur ne sait pas identifier le motif d'appel), l'échec est quasi systématique. Le cas du transfert d'appel est particulier puisque c'est le seul service intégralement automatisé dans le cadre du service A : cela explique le plus grand nombre d'interventions clients et la plus longue durée. En effet, pour les autres demandes, une fois que le service A reconnaît la demande, il bascule le client vers un autre service adapté.

Une régression logistique qui cherche à hiérarchiser les facteurs (motifs d'appels, nombre d'interventions clients et durée des appels) qui jouent sur le taux de réussite montre que le motif d'appel explique mieux encore le taux de réussite que la durée des appels ou le nombre d'interventions du client. Ce qui favorise l'accomplissement du service est le fait que les appels relèvent de motifs d'appels très standardisés et répétitifs.

L'analyse de la formulation de la première requête que le client adresse à l'automate, telle qu'elle est retranscrite par le système de reconnaissance vocale, confirme ces premiers résultats. On observe une très grande dispersion des requêtes avec beaucoup d'énoncés très rares et très peu d'énoncés très fréquents. Les énoncés qui n'apparaissent qu'une fois représentent 80% des énoncés, mais seulement 20% des requêtes. Ils ont un taux de succès très bas : 58% contre 75 % en moyenne. Inversement, si l'on se concentre sur les 13 énoncés les plus fréquents, ils ne représentent que 0,3 % des énoncés mais couvrent 50% des requêtes et le taux de réussite oscille entre 80 et 96%. Ainsi les énoncés les plus fréquemment adressés à la machine et qui conduisent aux taux de réussite les plus élevés sont : « Payer ma facture », « Transfert_d'appel », « Allofact », « Accéder à ma_ligne », « Messagerie vocale », « Forfaits ». Plus le client exprime sa requête avec un énoncé fréquemment attesté, plus les chances de « succès » augmentent. En cela, il apparaît que contrairement aux injonctions données par la machine qui engage à s'exprimer « naturellement », plus le langage est standardisé plus les chances d'accomplissement du service sont grandes...

Sur les données du service B nous n'avons pas pu déployer le même type d'analyse, car il n'y a pas eu de qualification manuelle des motifs d'appels, de l'échec ou de la réussite. Nous avons cherché à analyser les types de formulations adoptées par les clients pour exprimer leurs demandes. Pour ce faire, nous avons extrait les interventions des clients qui venaient juste après l'énoncé de l'automate :

1	À	D'accord avant de vous passer un téléconseiller merci d'exprimer la raison de votre appel (.) par exemple vous pouvez dire : « je n'arrive pas à installer ma livebox » c'est à vous.
---	---	---

Ce corpus a été analysé avec Alceste, un outil de statistique textuelle qui construit une typologie des énoncés en fonction du vocabulaire utilisé (Reinert, 1993).

L'analyse des motifs d'appels nous montre pour commencer que l'expression en langage naturel d'une demande ne va pas de soi. En effet, la moitié des réponses sont standardisées et suivent le modèle prescrit par l'automate (« je n'arrive pas à » ou « je n'ai plus de... »). Au cours de l'expérimentation, l'énoncé de l'automate a changé : alors que pendant un temps l'exemple donné était « je n'ai plus de tonalité sur ma ligne », il est devenu « je n'arrive pas à installer ma livebox ». Le changement de formulation de l'automate s'est répercuté immédiatement dans la manière dont une grande partie des clients ont formulé leur demande.

Un quart des réponses sont sous la forme de groupes nominaux (« problème de connexion à internet ») et empruntent le modèle de l'interrogation des moteurs de recherche. Enfin, seul un quart des réponses correspond à des énoncés « en langage naturel », exprimés librement. On peut faire l'hypothèse que ces derniers sont beaucoup plus complexes à traiter pour la machine.

La machine a donc une forte capacité de prescription et la majorité des clients s'y conforme. Elle place le client dans une situation d'injonction paradoxale : elle dit explicitement « exprimez-vous librement » et implicitement « adoptez mon langage ». Tel est le résultat principal qui ressort de l'analyse.

Quelles conclusions préliminaires pouvons-nous tirer de ces traitements à gros grain sur les interactions ?

- Il semble exister un modèle dominant de l'interaction homme-machine avec une seule intervention du client et de durée courte. C'est à la fois le modèle le plus fréquent et celui qui conduit au taux de réussite le plus élevé.
- Plus la durée de la conversation s'allonge, plus les chances d'échec augmentent.
- Plus le motif est routinier et/ou clairement circonscrit – à savoir l'accès à un service spécifique - plus le taux de succès est élevé. Les taux de réussite les plus bas concernent les demandes d'information. La standardisation des motifs d'appels favorise la réussite de l'interaction.

Enfin, l'automate joue un rôle fort de prescription sur la demande du client, en lui proposant des formulations de sa question. L'humain tend à parler comme la machine, à prendre prise sur le contenu proposé par l'automate. L'ajustement de vocabulaire qu'opère le client dans la formulation de la question suite aux « conseils » de la machine, a un effet positif sur l'aboutissement du dialogue.

Dans la suite des travaux de Wooffitt et MacDermid (1995), nous proposons de considérer l'interaction téléphonique entre le client et l'automate comme une séquence de demande d'aide/offre d'aide qui contient des séquences insérées de type question/réponse visant à avancer dans la résolution du problème. Les séquences insérées, qui sont à l'initiative de l'automate, peuvent avoir des finalités différentes comme l'identification du client, la précision du problème, la résolution par étapes. En voici un exemple :

Client : demande d'aide

{ Automate : question 1 (par exemple identification client)

{ Client : réponse

{ Automate : question 2 (par exemple précision du problème)

{ Client : réponse

...

Automate : offre d'aide

On peut poser comme hypothèse, au vu de l'analyse globale des interactions faite ci-dessus, qui montre que le taux de succès de l'appel atteint son point culminant quand le client n'intervient qu'une fois, que moins le nombre de séquences insérées est élevé, plus la demande a de chances d'aboutir. L'horizon ultime de ces appels serait une séquence ne comportant que l'énoncé du problème et sa résolution.

{ C : demande d'aide

{ A : offre d'aide

Voici à titre d'exemple, une interaction de ce type :

1	A	France Télécom bonjour bienvenu sur le 3000 (.) l'appel depuis une ligne fixe est gratuit (.) (jingle) que désirez-vous ?
2	C	payer ma facture
3	A	veuillez patienter je vous passe le service telefact (jingle)

Le modèle prototypique consiste à favoriser un appariement immédiat (sans séquences insérées) entre la demande et la réponse. Plus on est proche de ce modèle — ce qui implique une seule intervention du client, l'évitement des séquences insérées, l'usage de mots-clefs ou d'expressions normalisées, et en fin de compte une connaissance plus globale de l'univers de l'offre et des requêtes auxquelles peut répondre l'automate — plus les chances d'accomplissement du service augmentent. Le bon usage de l'automate vocal consiste à interagir selon ce modèle standardisé, ce qui relève de compétences spécifiques et d'un travail d'apprentissage qui est une composante nouvelle de la contribution du client à la coproduction²².

2.3 *Le travail d'enquête du client face aux désalignements interactionnels*

Comme l'ont montré les analyses statistiques, tous les appels traités par l'automate n'aboutissent pas à l'accès au service demandé ou à une réorientation pertinente du client. A partir d'exemples tirés du corpus d'appels au service A, nous examinons maintenant la structure interactionnelle des dialogues qui s'écartent du modèle prototypique pour observer les modalités concrètes de la mise au travail du client²³ dans les environnements de parole automatisée. Pour cela, l'analyse est centrée sur les principes du désalignement interactionnel et sur les activités des clients pour y faire face, les tentatives de réparation. Un problème pratique commun caractérise le dialogue homme-machine : les clients sont confrontés à des échanges qui ont les apparences d'une conversation sans en posséder les traits essentiels (à savoir l'orientation des interlocuteurs vers un monde commun et vers des attentes d'arrière-plan qui permettent un jeu d'anticipation et de projections). Dans ce contexte, certaines routines conversationnelles s'avèrent pour partie inopérantes et même contreproductives. Les recherches sur la communication homme-machine ont déjà montré que la performance des ordinateurs dépend de notre façon d'interagir avec eux, de notre « posture intentionnelle » qui nous permet de remplir leurs manques et de négliger les erreurs (Collins, 2000). Ainsi la régulation des interactions avec les machines est entièrement à la charge de l'interlocuteur humain à travers un travail constant de réparation qui prend des formes spécifiques dans les dialogues avec les agents conversationnels. Ce travail interactionnel devient travail du client lorsqu'il conditionne l'accès au service.

Quel type d'erreurs ou de manques s'agit-il de négliger ou de réparer pour accéder à un service via une interface vocale ? Le désalignement apparaît le plus souvent comme une conséquence de la difficulté de l'agent artificiel à s'ajuster finement au contexte de l'interaction, ce qui place son partenaire humain dans des situations inhabituelles où il a du mal à identifier les liens de pertinence (ou les liens logiques) entre certains tours de parole. Le phénomène est double : il arrive fréquemment que l'intervention de l'automate soit non pertinente par rapport au tour précédent ; mais le partenaire humain, contrairement aux traitements des malentendus dans les conversations ordinaires, thématise très rarement la non pertinence de la machine, sauf dans des épisodes émotionnels au cours desquels s'expriment l'énervement ou la colère. Cette impertinence est « subie » par l'utilisateur qui cherche à la contourner (en répétant, en reformulant, en insistant) plutôt que de la réparer, de demander des explications, comme il le ferait dans une conversation ordinaire. Ce traitement témoigne d'une orientation normative vers l'interaction homme-machine qui n'est clairement pas soumise aux mêmes exigences de continuité de la compréhension mutuelle que la conversation humaine.

²² En ce sens, il est une autre forme de la mise au travail du consommateur (cf. Dujarier, M-A. 2008).

²³ Pour un développement général de ce concept dans le contexte de l'industrialisation des services, cf. Tiffon, G., 2013.

La relative impertinence des tours de la machine est liée précisément à la façon dont celle-ci traite deux aspects de l'interaction : 1) le savoir implicite ; 2) la structure séquentielle et temporelle du dialogue. Le désalignement interactionnel apparaît comme résultat de l'un ou de l'autre aspect ou d'une combinaison des deux²⁴. Les trois interactions présentées permettent d'observer trois formes typiques de travail du client pour réparer les désalignements et tenter d'accéder au service : l'ajustement, les répétitions, les reformulations.

L'automate est en effet un partenaire exigeant en termes de « savoir partagé » et la coproduction du service repose sur les capacités du client à s'aligner à cette exigence soit par l'apprentissage (comme le cas 1), soit par la recherche de solutions en situation (comme les cas 2 et 3). Prenons deux cas de personnes qui cherchent à connaître leur consommation téléphonique (exemples 1 et 2). Seule la première y parvient : par un jeu d'essai-erreur, elle retrouve le nom du service, nom qui est la clef pour une bonne réponse de l'automate.

Exemple 1²⁵. Ajustement

1	A	que désirez-vous
2	C	Bricogit
3	A	[je ne vous ai pas entendu (.) [si vous
4	C	[allofact [allofact
5	A	Veillez patienter je vous passe votre service allofact

La cliente dans l'exemple 2 adopte une procédure différente de résolution de problème : elle répète quatre fois sa demande sans la reformuler (2, 5, 7, 9) et cherche aussi à échapper au mode vocal en utilisant le clavier téléphonique (4), sans succès.

Exemple 2. Répétitions

1	A	prononcez le nom de service qui vous intéresse [ou formulez librement vot-
2	C	[FOR-FAIT]
3	A	j'ai compris que vous demandiez un for[fait sinon dites annuler=
4	C	[appuie touche
5	C	=FOR-FAIT
6	A	(5) [dans la catégorie offres tarifaires
7	C	[FOR-FAIT
8	A	j'ai compris que vous demandiez un forfait sinon [dites annuler
9	C	[OUI FORFAIT
10	A	(5) dans la catégorie offres tarifaires [vous pouvez vous abonner ou vous-
11	C	[mais non non ((raccroche))

Le désalignement émerge autour de la confusion entre les deux significations du mot « *forfait* » : pour la cliente, il s'agit d'accéder au service qui permet de consulter l'état de sa consommation ; pour le système, il s'agit de s'informer ou de s'abonner à une offre. Dans ce cadre, la stratégie de la cliente, qui consiste à répéter à l'identique sa demande, contribue à maintenir la confusion tout au long de l'interaction. Au lieu d'aboutir à une solution et de faire progresser le dialogue, ces répétitions, mènent à deux boucles interactionnelles identiques (2-6 et 7-10)²⁶ à l'intérieur desquelles le premier énoncé de la demande (2 et

²⁴ Les phénomènes de chevauchements et d'interruptions constituent une autre classe importante de problèmes interactionnels que nous ne traitons pas dans ce texte.

²⁵ Les conventions de transcription se trouvent en annexe.

²⁶ Pour mieux visualiser les deux boucles interactionnelles, la deuxième est représentée en gras.

5) active la même réponse de l'automate. A son tour sa répétition (5 et 9) est traitée comme une confirmation par le système et entraîne l'énumération des options de la rubrique « offres tarifaires ».

De façon typique, la cliente interprète le problème comme relevant de la reconnaissance vocale et cherche à le résoudre en répétant de façon très articulée et en haussant le ton (supposant que la machine entend mal ou comprend mal). Elle ne révisé pas cette définition de l'obstacle rencontré tout au long de la séquence qui se termine par le constat d'échec dans la compréhension de la machine (11). Dans les faits, elle est confrontée à un problème de nature différente qui concerne le savoir implicite supposé par la machine : pour accéder à son forfait, il faut au préalable être abonné à une offre tarifaire ce qui n'est pas le cas de cette cliente puisqu'elle se voit proposer cet abonnement (l'automate vérifie les abonnements de la ligne téléphonique dans le système d'information). Les interventions de la machine n'explicitent pas ces informations et transfèrent à l'utilisatrice la tâche de mener l'enquête, de les déduire et de trouver la solution. Les interventions de l'automate apparaissent de plus en plus impertinentes par rapport à la demande par l'accumulation des répétitions de la cliente. Les places séquentielles des différents énoncés du mot « forfait », et leur coloration émotionnelle, leur confèrent des sens différents : ainsi la deuxième répétition (7) n'a clairement pas la même signification que le premier énoncé en ligne 2 et vient même contester et corriger l'interprétation donnée par la machine en l'interrompant. Pourtant le traitement est identique quelle que soit la place séquentielle du tour de la cliente. Cette séquence permet d'observer la difficulté que rencontrent les utilisateurs pour gérer l'affaiblissement des liens de pertinence des tours de l'automate puisque aucune piste de solution n'est proposée (par exemple, une invitation à reformuler) et les épreuves émotionnelles qui peuvent surgir. En effet, ce type de désalignement qui se construit dans la progression de l'interaction est invisible pour la machine, insensible à la temporalité et à l'énerverment progressif de la cliente. Elle n'identifie aucun problème et réitère à l'infini son fonctionnement : reconnaissance d'un mot clé – réponse par un item de la base de données. Le dialogue de progresse pas.

Au fil du développement du dialogue, les tours de la machine deviennent de moins en moins pertinents : elle ne clarifie pas le savoir implicite et reste indifférente à l'ordre temporel et séquentiel, ainsi qu'à l'amplification des émotions exprimées par la cliente. La temporalité de l'interaction est traitée comme réversible par l'automate capable de réitérer autant de fois le même segment de discours que la reconnaissance du même mot-clé s'active. L'ordre temporel comme aspect essentiel de la création du sens partagé de la conversation est inaccessible aux technologies vocales, du moins dans leur état actuel. En conséquence, les personnes sont placées dans une situation paradoxale de faire sens d'un échange et d'enquêter sur les façons de s'aligner, dans le domaine de la conversation, avec un artefact qui ignore ses aspects essentiels, la temporalité et la progressivité.

Dans l'exemple 3, le client adopte une autre méthode pour résoudre le problème de désalignement qu'il perçoit : la reformulation par la reprise de segments de discours de l'automate (19, 37, 48). Comme dans l'exemple précédent, l'interaction est un échec pour le client qui raccroche énervé (50), mais pas du tout pour le système, qui exécute la chaîne d'opérations sans faute : reconnaissance vocale, appariement du mot reconnu avec une information de la base de données, synthèse vocale pour la présenter. De plus, celui-ci fournit à deux reprises l'information pertinente par rapport à la demande du client (19) qui cherche à parler à un conseiller en lui indiquant qu'il faut appeler un autre numéro (cf. segments entourés). Comment expliquer alors les évaluations négatives du client (39, 44, 46) et l'échec du dialogue ?

Exemple 3 (extrait). Reformulations

17		allez-y/ c'est-à-vous\
18		(2.5)
19	C	je voudrais un conseiller\
20	A	vous êtes mis en relation avec le trois mi:l\ (.) notre
21		serveur vocal automatique qui analyse votre demande\
22		pour cette demande\ contactez le service client en appelant
23		le dix quatorze \ du lundi au samedi de huit heures à vingt heures\
24		l'appel est gratuit depuis une ligne fixe Telco\ (1)
25		vous serez alors mis en relation avec un conseiller\ (1)
26		je retourne à l'accueil\ (.) prononcez le nom du service qui
27		vous intéresse ou formulez librement votre demande sinon
28		composez le code du service si vous n'appellez pas de chez vous
29		dites d'abord \ (.) accéder à ma ligne\ c'est à vous
30		(6)
31	A	je ne vous ai pas entendu\ si vous souhaitez des exemples
		[...] [tour A lignes 32-35]
36		[allez-y prononcez le nom du serv-
37	C	[j veux le dix quatorze
38	A	à quel service souhaitez-vous vous abonner\
39	C	(3) ooo hh[hh (xxx)]
40	A	[je n'ai pas entendu le service que vous désirez]
41		(.) si vous souhaitez des exemples dites je voudrais de l'aide\ (.)
42		prononcez le nom du service qui vous intéresse ou
43		[alors formulez une description de ce service avec] une phrase simple\ (.) &
44		
44	C	[ooo ce sera toujours le même qu'il dit]
45		&je vous écoute
46	C	ALORS TU NE COMPRENDS RIEN C'EST TOUT=
47	A	=pour quel service souhaitez-vous le tarif\=
48	C	=le service client
49	A	pour cette demande contactez le service client en appelant le dix quatorze=
50	C	[ouais: ((énervé, raccroche))
51	A	=du lundi au samedi de- eu-

Cette interaction montre que la conversation ne se réduit pas à un transfert d'information et que donner une information pertinente sous une forme impertinente peut néanmoins conduire au désalignement interactionnel. La façon dont la machine participe à l'interaction et présente l'information s'avère problématique. En effet, les interventions de la machine s'insèrent mal dans le contexte interactionnel précis et n'explicitent pas les liens de pertinence avec les tours précédents. Nous retrouvons ici sous une forme plus sophistiquée les deux problèmes identifiés précédemment liés au savoir implicite et à la gestion atemporelle (ou « présentéiste ») des interactions par le système.

L'examen de l'interaction permet d'observer comment les liens de pertinence affaiblis de certains tours de parole de la machine par rapport au tour précédent du client instaurent progressivement le désalignement. La première formulation de la demande du client est suivie par une longue réponse du système qui contient la bonne information – il doit appeler un autre numéro (segment entouré, 22-26) – information qui reste inaudible pour le client jusqu'à la fin du dialogue. Si on regarde de près la structure de cette intervention du système (20-29), trois parties bien distinctes apparaissent : une auto présentation

(20-21), l'information demandée (22-25), retour au point de départ (26-29). Ainsi le renseignement est entouré par des segments de discours non-pertinents par rapport à la demande du client. D'abord l'auto-présentation du système n'a pas de lien direct avec la question posée et constitue une répétition d'un moment précédent (non-reproduit ici) sans que cela soit précisé (par exemple – « comme je l'ai déjà dit je suis un système vocal »). Suit ensuite le segment informationnel qui contient beaucoup trop de détails (comme les horaires d'ouverture), mais repose surtout sur un élément non-explicité, à savoir que le système ne peut pas procéder à une mise en relation avec le service demandé, que le client doit raccrocher et composer un autre numéro. L'ambiguïté sur ce point s'avérera cruciale pour la suite de l'échange, car jusqu'à la fin le client ne révisera pas cette croyance et cherchera à obtenir la mise en relation.

Enfin, le dernier segment du tour du système (26-29), qui annonce un retour « à l'accueil » et demande quel est le service recherché, représente une autre source de confusion car au lieu de proposer une clôture après avoir renseigné le client, il traite la temporalité de l'échange comme réversible. La succession des trois segments de l'intervention de l'automate contribue à affaiblir les liens de pertinence à l'intérieur de la paire adjacente constituée par la demande du client (19) et la réponse donnée (20-29). A partir de ce moment le désalignement installé s'approfondit progressivement puisque les deux parties sont orientées par des perspectives disjointes : le système cherche à identifier quel est le service demandé, le client cherche à être mis en relation avec le service clientèle²⁷ par des reformulations successives de sa demande initiale (37 : « j veux le dix quatorze », allant jusqu'à répondre à une question non pertinente (47) en 48 : « le service client »). De cette façon, il collabore dans les faits au maintien du désalignement en traitant les impertinences de la machine comme normales en essayant d'y remédier unilatéralement, de façon typique pour l'interaction homme-machine.

Pour résumer, le phénomène de désalignement observé présente la même forme circulaire que celui de l'exemple 2, malgré la méthode différente de résolution de problème adoptée par le client qui consiste non pas à répéter, mais à reformuler en reprenant des segments du discours de l'automate. Dans les deux cas l'action des clients est guidée par des prémisses qui s'avèrent erronées par rapport au fonctionnement de l'interface, mais ils ne trouvent aucune prise pour les réviser et mettent fin à l'interaction par des épisodes émotionnels (cf. ex. 2 – lignes 9- 11, ex. 3 - lignes 46-50)

Le désalignement contraint donc le client à s'engager dans des opérations d'enquête et de résolution de problèmes qui peuvent être pratiquement et émotionnellement éprouvantes. Si le comportement amnésique et les légers décalages interactionnels du système peuvent être surmontés à travers l'apprentissage et le passage par des séquences laborieuses, les obstacles liés au savoir implicite restent souvent infranchissables.

CONCLUSION

Une vue d'ensemble sur le corpus des appels aux deux services téléphoniques automatisés pourrait laisser croire que la promesse des technologies du dialogue naturel est sur le point d'être accomplie dans le domaine de l'après-vente : dans la majorité des cas le service est rendu. Pourtant, les résultats croisés des explorations statistiques plus poussées et de l'analyse conversationnelle mettent en lumière une réalité plus complexe. Selon la promesse portée par ces technologies les capacités conversationnelles des agents intelligents rendraient possible une prise en charge « personnalisée » de la demande du client, il suffirait que celui-ci s'exprime librement et spontanément, en « langage naturel ». L'accès aux services serait ainsi

²⁷ Une raison supplémentaire du désalignement tient au décalage entre l'horizon d'attente du client et ce que peut offrir le système : le client a une attente normative de type « guichet unique » pour l'accès aux services d'une entreprise. Cette attente entre en tension avec les nouvelles formes d'organisation de la relation client reposant sur la multiplication, la spécialisation et le cloisonnement des points de contacts. Cette tension dépasse la problématique des systèmes vocaux puisqu'elle est aussi observable dans les interactions avec les conseillers clients en face à face ou au téléphone, mais les malentendus qu'elle implique sont ingérables par des artefacts.

affranchi des contraintes imposées par l'écriture (d'un formulaire du Web par exemple) ou par des systèmes vocaux à touches de type DTMF, qui fonctionnent sur un modèle arborescent entièrement déplié face au client.

Or, les analyses statistiques et interactionnelles des dialogues montrent que l'orientation vers la conversation ordinaire est problématique. D'une part, la coproduction du service n'est jamais mieux accomplie que quand la personne se conforme à un modèle standardisé dans son échange avec l'automate, ce qui relève d'une compétence d'usage de cette technologie. Paradoxalement, plus l'interaction s'éloigne de la conversation ordinaire, plus les chances d'accéder au service augmentent. D'autre part, l'analyse conversationnelle des cas de désalignement montre que, même statistiquement minoritaires, ceux-ci ne peuvent pas être attribués au hasard des circonstances. Ils relèvent au contraire d'un problème pratique régulier consistant à créer un simulacre de conversation tout en désactivant les régulations essentielles propres à cette activité (l'orientation commune des partenaires vers un contexte dynamique, vers la séquentialité et la temporalité de l'échange, les mécanismes de maintien de la compréhension et de réparation des malentendus, l'ordre émotionnel). Dans ces situations de désalignement, le client est donc laissé à lui-même pour mener l'enquête et découvrir des modes de résolution des problèmes valables dans cet environnement technologique. Il est donc mis au travail pour enquêter et, dans certains cas, pour apprendre comment interagir avec les agents conversationnels, comment réparer les désalignements et faire sens du dialogue afin d'accéder au service.

L'automatisation de la relation de service implique ainsi de nouvelles formes de contribution du client à la coproduction en le mettant face à des simulacres de conversation : notamment un travail d'apprentissage de la standardisation et un travail de gestion des désalignements dans des échanges, affranchis non seulement de la civilité mais aussi de la temporalité comme régulations de la conversation commerciale. Ces résultats confirment sur notre terrain d'enquête les critiques à l'égard de l'IA formulées par des philosophes, des sociologues des sciences et des ethnométhodologues : les automates vocaux peuvent « rendre service » pour des tâches bien circonscrites à condition qu'ils soient débarrassés de l'illusion de naturalité et que leurs limites soient pleinement assumées et explicitées. Mais ces résultats apportent aussi un éclairage empirique nouveau à ce constat en identifiant précisément le mécanisme de désalignement à l'œuvre dans l'interaction avec les interfaces vocales. Ce mécanisme se caractérise par l'affaiblissement des liens de pertinence entre les tours de parole à travers le non-traitement des problèmes liés au savoir implicite, à la structure séquentielle de l'échange et aux expressions émotionnelles.

En ce sens, la contribution de cette enquête aux débats sur l'intelligence artificielle, à travers l'analyse du phénomène particulier des désalignements et la mise en lumière des activités des clients, invite à prendre acte des limites des approches visant à copier la conversation humaine et à explorer plus en avant les spécificités des échanges avec les automates vocaux afin de mieux équiper et soutenir cette forme particulière d'interaction de service.

Annexe : Conventions de Transcription

mot/ mot\	Intonation montante Intonation descendante
mo:t	Allongement vocalique
> mot <	Débit plus rapide
mo-	Troncation d'un mot
<u>Comme, extra</u> MOT ° mot °	Prononciation accentuée, appuyée Volume de voix plus fort Volume de voix plus bas

xxx	Segment inaudible
(mot)	Segment difficilement audible, hésitations du transcripteur
[[Chevauchement de deux tours de parole
(1) (.) (. .) (...) (1.5) ital.	Pauses en secondes Pauses de moins d'une seconde (diverses longueurs) Silence correspondant à l'écoute de l'interlocuteur au téléphone
= =	Tours de parole "enchainés" de deux locuteurs (se suivant sans pause)
& &	Continuation du tour d'un même locuteur
((rire))	Phénomènes décrits de manière non détaillée
Hhh	Respiration ; la répétition de la lettre h correspond à une durée plus importante

REFERENCES

[Bandt et Gadrey, 1994] Bandt, Jacques de, Gadrey, Jean, 1994, *Relations de service, marchés de services*, Paris, CNRS Editions.

[Beaudouin et Velkovska, 2010] Beaudouin, Valérie, Velkovska, Julia, 2010, Dialogues vocaux entre clients et automates ou comment l'homme et la machine s'entendent dans la réalisation d'un service, JADT 2010 - Statistical Analysis of Textual Data, at Sapienza - Università di Roma, LED Edizioni Universitarie di Lettere Economia Diritto, 865-876.

[Borzeix 1992] Borzeix, Anni, 1992, "Qualité et bienveillance: l'épreuve de l'étranger." in Borzeix, Anni, Gardin, Bernard, Girin, Jacques, Grosjean, Michèle, Lacoste, Michèle, *EDF - GDF Bonjour : l'interaction Agent-Client à l'accueil*, rapport CRG, Paris, Ecole polytechnique.

[Borzeix et Fraenkel, 2001] Borzeix, Anni, Fraenkel, Béatrice (dir.), 2001, *Langage et travail. Communication, cognition, action*. Paris, CNRS Editions.

[Button, 1990] Button, Graham, 1990, "Going Up a Blind Alley. Conflating Conversation Analysis and Computational Modelling." In Paul Luff, Nigel Gilbert, David Frohlich (dir.), *Computers and Conversation*, London, Academic Press.

[Button, 1993] Button, Graham, 1993, *Technology in Working Order*. Edited by Routledge, London.

[Button et al., 1995] Button, Graham, Coulter, Jeff, Lee, John, Sharrock, Wes, 1995, *Computers, Minds and Conduct*, Chap. 7 "Can a Machine Talk?", Cambridge, Polity Press.

[Button et Sharrock, 1995] Button, Graham, Sharrock, Wes, 1995, "On simulacrum of conversation: Toward a clarification of relevance of conversation analysis for human-computer interaction." In Peter Thomas (dir.), *The Social and Interactional Dimensions of Human-Computer Interfaces*, 107-125, Cambridge, Cambridge University Press.

[Cassell et al., 2000] Cassell, Justine, Matthew Stone, Hao, Yan. 2000. "Coordination and context-dependence in the generation of embodied conversation." *Proceedings of the first international conference on Natural language generation*, pp. 171-178.

- [Cassell, 2001] Cassell, Justine, 2001, "Nudge Nudge Wink Wink: Elements of Face-to-Face Conversation for Embodied Conversational Agents." in Justine Cassell, Joseph Sullivan, Scott Prevost, Elizabeth Churchill (dir.) *Embodied Conversational Agents*, Cambridge: MIT Press. pp. 1-27
- [Cassell et Bickmore, 2004] Cassell, Justine, Bickmore, Timothy, 2004, "Social Dialogue with Embodied Conversational Agents." In J. van Kuppevelt, L. Dybkjaer, N. Bernsen (dir.), *Natural, Intelligent and Effective Interaction with Multimodal Dialogue Systems*, New York, Kluwer Academic.
- [Chapman, 1992] Chapman, David, 1992, "Computer Rules, Conversational Rules." *Computational Linguistics*, 18, (4).
- [1993] 1993, *Cognitive Science (A Multidisciplinary Journal of Artificial Intelligence, Linguistics, Neuroscience, Philosophy, Psychology)*, Special Issue: *Situated Action*, vol. 17, (1).
- [Collins, 1990] Collins, Harry, 1990, *Artificial Experts: Social Knowledge and Intelligent Machines*. Traduit, Cambridge, MIT Press.
- [Collins, 1992] Collins, Harry, 1992, "Hubert Dreyfus, Forms of Life, and a Simple Test for Machine Intelligence." *Social Studies of Science*, (22), 726-739.
- [Collins, 1993] Collins, Harry, 1993, "The Turing Test and Language Skills." In Graham Button (dir.), *Technology in Working Order*, 231-245, London, Routledge.
- [Collins, 1998] Collins, Harry, 1998, "Socialness and the Undersocialized Conception of Society." *Science, Technology and Human Values*, 23, (4), 494-516.
- [Collins, 2000] Collins, Harry, 2000, "Les capacités des ordinateurs et leurs limites." *Réseaux*, (100), 19-37.
- [Collins et Kush, 1998] Collins, Harry, Kush, M, 1998, *The Shape of Actions: What Humans and Machines Can do*. Traduit, Cambridge, Mass, MIT Press.
- [Denis, 2011] Denis, Jérôme, 2011, "Le travail de l'écrit en coulisses de la relation de service." *activites.org*, 8, 32-52.
- [Dreyfus, 1984] Dreyfus, Hubert, 1984, *L'intelligence artificielle. Mythes et limites*. Traduit, Paris, Flammarion.
- [Dujarier, 2008] Dujarier, Marie-Anne, 2008, *Le travail du consommateur. De McDo à eBay : comment nous coproduisons ce que nous achetons*. Traduit, Paris, La découverte.
- [Fordham et Gilbert, 1995] Fordham, Andrew , Gilbert, Nigel, 1995, "On the Nature of Rules and Conversation." *Artificial Intelligence and Society*, 9.
- [Frohlich et Luff, 1990] Frohlich, David, Luff, Paul, 1990, "Applying the technology of conversation to the technology of conversation." In Paul Luff, Nigel Gilbert, David Frohlich (dir.), *Computers and Conversation*, 187-220, London, Academic Press.
- [Fuchs, 1968] Fuchs, Victor R., 1968, *The Service Economy*. Traduit, *October*, National Bureau of Economic Research.
- [Gadrey, 1994] Gadrey, Jean, 1994, "Les relations de service et l'analyse du travail des agents." *Sociologie du travail*, 36, (3), 381-389.
- [Garfinkel, 2007, [1967]] Garfinkel, Harold, 2007, [1967], *Recherches en ethnométhodologie*. Traduit par M. Barthélemy, B. Dupret, JM. de Queiroz, Paris, PUF.
- [Gilbert et al., 1990] Gilbert, Nigel, Wooffitt, Robin, Fraser, Norman, 1990, "Organising Computer Talk." In Paul Luff, Nigel Gilbert, David Frohlich (dir.), *Computers and Conversation*, Londres, Academic Press.
- [Goffman, 1968 [1961]] Goffman, Erving, 1968 [1961], "Les hôpitaux psychiatriques et le schéma médical-type (quelques remarques sur les vicissitudes des métiers de réparateurs)." In Erving Goffman, *Asiles. Etudes sur la condition sociale des malades mentaux et autres reclus*, Paris, Les Editions de Minuit.
- [Grosjean, 1995] Grosjean, Michèle, 1995. "Sur quel ton le dites-vous?" In Isaac Joseph, Gilles Jeannot *Métiers du public. Les compétences de l'agent et l'espace de l'utilisateur*, Paris, CNRS Editions.
- [Hutchby, 2001] Hutchby, Ian, 2001, *Conversation and technology : from the telephone to the internet*, Cambridge, Polity Press.

- [Joseph, 1988] Joseph, Isaac, 1988, "La relation de service. Les interactions entre agents et voyageurs." *Les annales de la recherche urbaine*, (39).
- [Norman, 1993] Norman, Donald A., 1993, "Cognition in the Head and in the World: An Introduction to the Special Issue on Situated Action." *Cognitive Science*, 17, (1-6), 1-6.
- [Rastier, 1991] Rastier, François, 1991, *Sémantique et recherches cognitives*. Paris, PUF.
- [Reinert, 1993] Reinert, Max, 1993, "Les "mondes lexicaux" et leur logique." *Langage et société*, (n°66), 5-39.
- [Suchman, 2007] Suchman, Lucy, 2007, *Human-Machine Reconfigurations. Plans and Situated Actions 2nd Edition*. Cambridge, Cambridge University Press.
- [Taylor, 1995] Taylor, Charles, 1995, " Suivre une règle ", *Critique* (n° 579/580), 554-572.
- [Tiffon, 2013] Tiffon, Guillaume, 2014, *La mise au travail des clients*, Paris, Economica, collection Etudes sociologiques.
- [Turing, 1950] Turing, Alan, 1950, "Computing Machinery and Intelligence." *Mind*, 49, 433-460.
- [Wooffitt et Fraser, 1993] Wooffitt, Robin, Fraser, Norman, 1993, "We're off to ring the wizard, the wonderful Wizard of Oz." In Graham Button (dir.), *Technology in Working Order*, 211-230, London, Routledge.
- [Wooffitt et al., 1997] Wooffitt, Robin, Fraser, Norman, Gilbert, Nigel , McGlashan, Scott 1997, "Chapter 9 : Conversation Analysis, Simulation and System Design." In R. Wooffitt, N. Fraser, N. Gilbert, S. McGlashan (dir.), *Humans, Computers and Wizards. Analyzing human (simulated) computer interaction*, London, Routledge.
- [Wooffitt et MacDermind, 1995] Wooffitt, R., MacDermind, C., 1995, "Wizards and Social Control." In P. Thomas (dir.), *The Social and Interactional Dimensions of Human-Computer Interfaces*, 126-141, Cambridge, Cambridge University Press.
- [Woolgar, 1995] Woolgar, Steve, 1995, "Why not a sociology of machines? The case of sociology and artificial intelligence." *Sociology*, (19), 557-572.