



HAL
open science

Le travail émotionnel du client

Martin Aranguren

► **To cite this version:**

Martin Aranguren. Le travail émotionnel du client : La structure séquentielle des émotions dans les usages problématiques d'un serveur vocal. *Social Science Information*, 2014, 53 (3), 10.1177/0539018414523520 . hal-01335789

HAL Id: hal-01335789

<https://hal.science/hal-01335789>

Submitted on 22 Jun 2016

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.

Corresponding author:

Martin Aranguren

Ecole des Hautes Etudes en Sciences Sociales, Institut Marcel Mauss, 190 avenue de France,
75013 Paris, France

Email: martin_aranguren@yahoo.com

Le travail émotionnel du client

La structure séquentielle des émotions dans les usages problématiques d'un serveur vocal

Martin Aranguren

Institut Marcel Mauss, Ecole des Hautes Etudes en Sciences Sociales

Résumé

Des développements récents dans le secteur services ont amené les sociologues à soutenir que les clients sont de plus en plus encouragés à ‘travailler’ pour les prestataires. L’article développe l’hypothèse que le travail du client, dans le contexte de la production des services, peut comporter une dimension émotionnelle. À cet effet une révision de la notion de ‘travail émotionnel’ est proposée à partir d’une étude sur les usages problématiques d’une interface téléphonique appartenant à un opérateur de télécommunications. L’étude implique une collection d’appels réels qui ont été transcrits et codés; les énoncés ont été décrits objectivement concernant le ton et l’intensité; enfin, une méthode statistique pour la détection de *patterns* séquentiels a été appliquée. L’étude suggère que les clients répondent émotionnellement à des problèmes d’utilisabilité, et que leurs émotions se déroulent selon une structure séquentielle récurrente. L’ ‘échange correcteur’ de Goffman fournit le point de référence de l’interprétation, qui mène à la proposition que les *patterns* émotionnels découverts sont d’origine sociale, bien que non socialement situés.

Mots-clés

Recent developments in the service sector have led sociologists to suggest that customers are increasingly encouraged to ‘work’ for services organizations. The paper develops the hypothesis that the client’s work in the context of service production may involve an emotional dimension. To this purpose, a revision of the notion of ‘emotion work’ is proposed on the basis of a study on the problematic uses of a phone interface from a telecommunications operator. The study involves a collection of real calls that were transcribed and coded; utterances were described objectively regarding pitch and intensity; finally, a statistical method for detecting sequential patterns was applied.

The results suggest that clients respond emotionally to usability problems, and that their emotions unfold according to a recurrent sequential structure. Goffman's 'remedial interchange' provides the point of reference for the interpretation, which leads to the suggestion that the discovered emotional patterns are of social origin, though not socially situated.

Keywords : émotion, travail émotionnel, services, interaction homme-machine, analyse séquentielle, rituels d'interaction

Introduction

Depuis les années 1980, les gestionnaires appellent à traiter les clients des organisations de services comme des 'employés partiels' (Mills et Morris, 1986). Les transformations inspirées de ces recommandations ont pris une telle ampleur dans les années 2000 que l'on n'hésite plus, en sociologie, à parler de 'travail du consommateur' (Dujarier, 2008).

Dans le cadre de la relation de service, le travail du client est avant tout une coproduction, au sens d'une contribution indispensable du bénéficiaire à la réalisation de sa prestation (Gadrey, 1994). Dans la mesure où la coproduction a pour contexte des situations sociales marquées par des échanges verbaux, elle comporte aussi une dimension langagière essentielle (Borzeix et Fraenkel, 2001).

La possibilité explorée dans cet article est que, de manière parallèle à la coproduction et aux compétences langagières dont elle dépend, certaines interactions de service demandent au client un travail supplémentaire que l'on pourrait qualifier d'émotionnel. Pour développer cette hypothèse nous rapporterons les résultats d'une étude d'usage portant sur un serveur vocal téléphonique.

Le travail émotionnel du client ne peut être décrit qu'en élargissant une conception qui le définit en termes de mise en conformité d'un sentiment privé avec des règles du sentir obligatoires (Hochschild, 1979). La gestion des sentiments existe et elle revêt une importance indéniable, d'autant plus que son étude a grandement contribué à faire reconnaître l'intérêt sociologique des émotions. Mais une concentration exclusive sur le sentiment pourrait faire perdre de vue que le travail émotionnel peut aussi avoir la situation pour matériau (Quéré, 2012). Travailler son émotion peut prendre la forme d'un repli sur soi, mais aussi d'un

déploiement sur le monde. Cette deuxième possibilité, en déplaçant la focale du ressenti interne vers l'action dans l'environnement, fournit notre point de départ dans la description du travail émotionnel du client.

L'analyse d'une collection d'appels au serveur vocal étudié permet de dégager une structure séquentielle récurrente dans l'activité émotionnelle des clients. Nous conceptualiserons ce format, à la suite de Goffman (1967, 1971, trad. fr. 1974 et 1973), comme un 'échange' (une série ordonnée et distribuée de 'coups') par lequel se réalise un 'travail de face' (c'est-à-dire de sauvegarde de l'image de soi).

Rappelons que pour Goffman, qui dit travail de face dit rituel d'interaction. Mais, en réalité, nous verrons que l'on n'a pas besoin de rester poli pour rester soi-même, car il se peut que la réussite du travail de face n'exige pas l'accomplissement d'un rituel d'interaction. Ce n'est qu'en reconnaissant que face et rituel ne vont pas nécessairement ensemble que le travail émotionnel du client pourra être défini positivement.

La première partie de l'article se penche sur les contraintes techniques des interfaces vocales qui sont à l'origine de la possibilité structurelle du travail émotionnel du client. Ensuite est proposé un élargissement de la notion de travail émotionnel afin de la rendre pertinente, non seulement pour la gestion des sentiments, mais aussi pour la transformation des situations émouvantes. La deuxième partie présente les phases successives de l'étude sur les serveurs vocaux ainsi que ses principaux résultats, que nous illustrerons par des analyses de séquence détaillées. La troisième partie examine le format séquentiel dégagé à la lumière de la sociologie des rituels d'interaction de Goffman.

1 Problématique : du travail de communication au travail émotionnel

1.1 Aperçu des interfaces conversationnelles (1)

En général, une interface est ce avec quoi une personne communique quand elle utilise une machine. Pour simplifier, admettons que la machine est un ordinateur. Imaginons que la personne fait quelque chose, ensuite la machine change son propre état, enfin elle en informe la personne. Nous dirons alors que ce que la personne a fait, c'est produire un *input* ou une entrée, et que ce que la machine a fait, c'est traiter cet input et retourner un *output* ou une sortie. Une interface *vocale* est une version de ce schéma telle que l'entrée comporte des mots prononcés par la personne et la sortie des mots prononcés par la machine. Une interface *vocale conversationnelle* est la version de ce schéma qui encourage la personne à exprimer son input verbal dans un format 'spontané' apparenté à la conversation. L'interface vocale

conversationnelle qui fait l'objet de cette étude possède deux propriétés supplémentaires : sa finalité est commerciale et le téléphone est son canal de communication. La technologie à laquelle nous aurons affaire est donc une interface vocale, conversationnelle, commerciale et téléphonique.

1.2 Le problème du rejet

Le mode de fonctionnement des interfaces conversationnelles donne lieu à une forme particulière de ce que Lucy Suchman appelle ' le problème de la communication homme-machine ' (Suchman, 2006). Le contexte de ce problème est une tâche composée de plusieurs étapes qui prennent forme dans une interaction entre une personne et un dispositif technique. Dans ce contexte, l'accomplissement d'une tâche donnée exige un enchaînement précis de gestes. Cette coordination, nécessaire à l'accomplissement de la tâche, demande aux participants de référer chaque nouveau geste au dernier geste de l'autre ; autrement dit, de faire de chaque acte une réponse à l'acte précédent du partenaire.

Le problème général de la communication homme-machine s'enracine dans une asymétrie de compétence interactionnelle entre la personne et la technologie. Parce que les ressources de compréhension de la situation dont la machine dispose sont comparativement limitées, il y aura des cas où elle ne pourra pas référer de manière appropriée son acte suivant au geste précédent de l'utilisateur humain. Ce à quoi s'ajoute l'autre versant du problème, qui concerne l'opacité du mode de fonctionnement de la technologie pour l'utilisateur profane. En conséquence il y aura des cas où la personne estimera que la réponse de la machine est appropriée, alors qu'elle ne l'est pas, et vice versa.

Pour exposer la forme précise que ce problème général revêt dans les interfaces conversationnelles, imaginons le scénario suivant. Supposez que vous ayez un problème P chez vous et souhaitiez qu'un technicien de l'entreprise E vienne le résoudre. Vous appelez E et vous êtes redirigé vers une interface conversationnelle spécialisée dans la prise de rendez-vous. Au bout d'un court échange d'informations, la machine vous fait une proposition et vous demande de la confirmer. Son énoncé pourrait prendre la forme suivante : ' Je vous propose un rendez-vous pour vendredi à 15h. Est-ce que cela vous convient? '.

Si vous dites ' ouais ', vous déclencherez chez la machine un processus de traitement en plusieurs étapes successives, indiquées de manière simplifiée dans le tableau ci-dessous. Le tableau se lit d'abord de haut en bas, ensuite une colonne après l'autre de gauche à droite. Notez que l'output de la phase précédente devient l'input de la phase suivante.

	1. segmentation et reconnaissance	2. interprétation	3. gestion de dialogue
Input	bruit bruit uε bruit bruit	ouais	<i>oui</i>
règles	-Reconnaître des mots dans le signal ; ignorer ce qui précède et ce qui suit. -Transcrire les mots. -Sinon, <i>erreur</i> .	-Comptent pour <i>oui</i> : oui, ouais, oké, d'accord, bien sûr. -Comptent pour <i>non</i> : non, nen, pas vraiment, pas du tout. -Sinon, <i>erreur</i> .	-Si <i>oui</i> , dire ' Votre rendez-vous vient d'être confirmé. ' -Si <i>non</i> , dire ' Demande annulée. Je retourne à l'accueil. ' -Dans tous les autres cas, dire ' Je vous propose un rendez-vous pour vendredi à 15h. Est-ce que cela vous convient? '
Output	ouais	<i>oui</i>	' Votre rendez-vous vient d'être confirmé. '

Imaginez maintenant que, pour des raisons qui nous échappent, vous répondez à la demande de confirmation par ' Tiens, il fait beau aujourd'hui '. Voici comment la machine traitera votre énoncé (les règles sont les mêmes que pour le tableau précédent).

	1. segmentation et reconnaissance	2. interprétation	3. gestion de dialogue
Input	bruit tʃɛ il fε bo oʒURdʝi bruit	tiens il fait beau aujourd'hui	<i>erreur</i>
Output	tiens il fait beau aujourd'hui	<i>erreur</i>	' Je vous propose un rendez-vous pour vendredi à 15h. Est-ce que cela vous convient? '

Supposez maintenant que, plus raisonnablement, vous répondez par ' oui ' à la proposition de rendez-vous de la machine. Mais que, à peu près au même moment où vous prononcez ce mot raisonnable, une ambulance passe, ou une porte grince, ou votre chat miaule, ou quelqu'un vous parle.

	segmentation et reconnaissance	interprétation	gestion de dialogue
Input	bruit bruit bruit bruit	<i>erreur</i>	<i>erreur</i>
Output	<i>erreur</i>	<i>erreur</i>	' Je vous propose un rendez-vous pour vendredi à 15h. Est-ce que cela vous convient? '

Les deux derniers scénarios mettent en évidence une contrainte technique de ces technologies : le système fera exactement la même chose, que le rejet tiende à un enchaînement de mots dont on ne peut pas extraire des significations (' Tiens il fait beau aujourd'hui '), ou bien à une séquence de sons dont on ne peut pas extraire des mots (le miaulement du chat). Cette

incertitude définit *le problème du rejet*, dont on ne peut jamais savoir s'il est dû à une impureté du signal ou bien à une inadéquation du vocabulaire.

1.3 Le travail de communication du client

Le problème du rejet situe le travail du client d'une interface vocale conversationnelle sur le plan de la communication. Pour le mettre en évidence, nous appliquerons au travail du client la méthode d'analyse séquentielle préconisée par le Réseau Langage et Travail pour l'étude du travail des agents. (cf. Borzeix et Fraenkel, 2001).

Afin d'illustrer le travail de communication que ces technologies demandent au client, prenons un exemple de résolution qui a été traité plus en détail ailleurs (Aranguren, 2012). ' S ' désigne le Système ; ' C ' le Client.

1	S :	Transfert d'appel. Vous pouvez demander à transférer vos appels ou bien désactiver le transfert en cours. Que voulez-vous faire ?
2	C :	Transfert d'appel.
3	S :	Transfert d'appel. Vous pouvez demander à transférer vos appels ou bien désactiver le transfert en cours. Que voulez-vous faire ?
4	C :	<i>Transférer</i> mes appels.
5	S :	Composez avec les touches de votre téléphone le numéro vers lequel vous souhaitez effectuer votre transfert.

Dans la description qui suit nous indiquons entre parenthèses le placement de la référence dans la transcription. Le système ouvre la séquence par un tour complexe composé d'une présentation du service Transfert d'appel, d'une énumération des opérations possibles et d'une invitation à en choisir une (tour 1). Le client réplique en choisissant l'option ' transférer vos appels ' par l'énoncé ' Transfert d'appel ', soit une ellipse de la phrase complète ' Je voudrais faire un transfert d'appel ' (2). Le système répond par un *reprompt*, c'est-à-dire par une répétition en boucle (3). Le client reformule alors sa demande (4). La machine accepte la demande reformulée et le dialogue passe à l'étape suivante (5). Notons que la paire (2) + (3) peut se répéter plusieurs fois avant que le client ne découvre que le problème n'est pas dans sa manière peu articulée, trop appuyée ou inaudible, de dire ' Transfert d'appel ', mais bien dans le choix des mots. Ce travail de découverte, particulièrement frustrant lorsque la recherche recycle plusieurs fois la paire (2) + (3), illustre

le travail de communication que les interfaces conversationnelles peuvent à tout moment nécessiter.

Or, notre étude d'usage suggère que le niveau d'application du client à la réalisation de ce travail est variable. Au lieu de s'efforcer de purifier le signal ou de modifier la formulation de son énoncé, il fait parfois preuve d'une certaine indiscipline communicationnelle. Face à un rejet, à l'occasion on voit le client prononcer

- les interjections primaires (2) , par ordre de complexité :

ô, â lala lala ;

- les interjections secondaires, par ordre de politesse :

punaise, putain ;

- les locutions interjectives, décrivant le degré de normalité de la situation :

c'est pas vrai, c'est pas possible, c'est du délire ;

- les locutions interjectives, décrivant l'état affectif de l'appelant :

j'en ai marre, ça m'énerve, ça me prend la tête ce truc ;

- les insultes familières :

pétasse, tu m' fais chier, je t'emmerde ;

- ou bien, plus formel :

allez vous faire enculer.

Ces expressions ne font pas partie du travail du client si celui-ci se réduit à l'effort nécessaire pour coproduire la prestation. Mais il suffit de rappeler que la relation de service se constitue d'un échange entre prestataire et bénéficiaire pour apprécier que la relation comporte non seulement la production d'un service, mais aussi l'accomplissement d'une interaction. Peut-être les expressions d'émotions que nous venons de lister remplissent-elles une fonction importante dans le processus d'interaction, sans nécessairement contribuer à la coproduction de la prestation.

1.4 Le travail émotionnel sur l'environnement

Nous allons suivre Arlie Hochschild lorsqu'elle définit le travail émotionnel (*emotion work*) comme ' l'acte d'essayer de changer en degré ou en qualité une émotion ou un sentiment ' (1979 : 561). Les expressions d'émotion des appelants peuvent être considérées comme des composantes de ce processus de gestion de l'émotion.

Dans la problématique de Hochschild cette gestion comporte la suppression d'une émotion non désirée ou bien l'évocation d'une émotion désirée. Il n'est pas inutile de rappeler

que pour fonder son modèle de l'acteur ' gestionnaire d'émotion ', elle s'est appuyée sur la tradition psychanalytique, et en particulier sur les travaux de Richard Lazarus (1966 , 1991). La suppression et, de manière secondaire, l'évocation des émotions constituent les résultats visés d'un processus que Lazarus appelle ' *coping* focalisé sur l'émotion '. Une fois qu'une émotion s'est déclenchée, le contraire de ce travail sur l'émotion n'est pas une absence d'effort émotionnel, mais plutôt une action sur l'environnement.

En d'autres mots, l'émotion peut être gérée non seulement en agissant sur son ressenti, mais aussi en agissant sur la situation. Cette forme alternative de *coping*, Lazarus la qualifie de ' focalisée sur le problème ', en entendant par ' problème ' une relation troublée entre la personne et son environnement (ou son *Lebenswelt*). Par exemple, une menace ou une insulte définissent des relations troublées entre la personne et son milieu d'action qui donnent lieu à des émotions comme la peur ou la colère. Pour les apaiser, une fois qu'elles se sont présentées, on peut les ' travailler ', par exemple, en se persuadant que la menace ou l'insulte n'existent pas réellement. Mais on peut aussi ' travailler ' l'environnement, en s'éloignant suffisamment de l'objet apeurant ou en exerçant une action vigoureuse sur l'objet agaçant – y compris un ' simple ' acte de discours. En somme, le travail émotionnel peut avoir pour matériau l'émotion ressentie, mais aussi l'environnement d'action.

Hochschild propose d'expliquer en des termes sociologiques l'organisation de cette faculté humaine à gérer ou à travailler ses émotions. Dans son schéma l'acteur cherche à se conformer à des normes, et en particulier à des règles du sentir (*feeling rules*). On peut concevoir ces règles comme une table de correspondances : dans la colonne de gauche, l'on trouve des types de situation (des ' cadrages '); dans la colonne de droite, des types d'affect ou de sentiment. Par exemple, une fois qu'un rassemblement dans un cimetière a été ' cadré ' comme un enterrement, une règle du sentir prescrit la tristesse comme sentiment obligatoire. Si vous n'en ressentez pas, en bon acteur durkheimien vous devrez faire un travail émotionnel pour conformer votre affect à la norme, que ce soit en vous forçant à être triste alors que la situation vous laisse indifférent, ou bien en supprimant un sentiment inapproprié, comme la joie de rencontrer un vieil ami parmi les présents ou de vous savoir l'héritier d'une grande fortune – ou, dans un exemple plus nuancé, de voir enfin terminées les souffrances du défunt. Dans ce modèle de l'organisation sociale du travail émotionnel, l'acteur se conforme aux règles du sentir officielles pour éviter d'éprouver ' de l'inquiétude, de la culpabilité ou de la honte ' (Hochschild, 1979: 565). Hochschild souligne enfin une particularité des règles du sentir : elles ne s'appliquent pas à l'action (' l'expression comportementale ') mais à ce qu'elle conçoit comme son antécédent, à savoir le ' sentiment interne ' ou privé (p. 566).

En résumé, un acteur capable de travailler ses sentiments se trouve toujours dans une situation d'un certain type. Des règles du sentir prescrivent des émotions obligatoires pour chaque type de situation. Si l'affect éprouvé dans une situation donnée n'est pas conforme à ces règles du sentir, l'acteur devra accomplir un travail émotionnel pour combler l'écart, sous peine de ressentir de l'inquiétude, de la culpabilité ou de la honte. Le lieu de ce travail, aussi bien que de son matériau – le sentiment –, reste le for intérieur de l'acteur. L'expression, et plus généralement le comportement observable, n'entrent en ligne de compte que plus tard.

Tout en admettant la validité de l'approche de Hochschild, il faut dire que la sociologie des émotions n'a pas besoin de formuler sa question en termes de mise en conformité, dans le for intérieur de la conscience, d'un sentiment privé de l'individu avec les normes sociales d'un groupe. Cette approche est particulièrement éclairante pour l'étude de certains objets, comme la compétence de gestion affective inhérente à certaines professions, ou encore la marchandisation de cette compétence dans l'économie capitaliste (Hochschild, 1983, 2003). Mais ni cette manière de poser la question, ni la théorie qu'elle sous-tend, nous semble-t-il, n'épuisent l'intérêt sociologique que l'on peut porter au vaste domaine du travail émotionnel.

1.5 Le format social du travail émotionnel

Notre hypothèse de travail est que la structure séquentielle du travail émotionnel sur l'environnement est un format ou un *pattern* d'origine sociale, même si l'activité qu'il organise n'a pas lieu dans le cadre d'une rencontre. Afin de la développer, nous tenterons d'identifier des séquences d'activité émotionnelle récurrentes dans notre étude d'usage. L'hypothèse trouve un antécédent dans la théorie de l'acte de Mead, qui conçoit la relation entre l'homme et son environnement avant tout comme une relation *sociale* (Mead, 1938 : 103-124 ; Joas, 2007, chapitre 7). Dans cette optique, le mécanisme de la conversation de gestes, où chaque participant à l'interaction règle sa conduite par des anticipations de la conduite de l'autre, n'organise pas seulement les échanges entre personnes, mais aussi ceux avec les 'choses physiques'. Pour Mead, l'homme peut 'adopter l'attitude' de la chose parce qu'il est capable d'anticiper la résistance à ses propres manipulations. Nous verrons que dans les usages problématiques des interfaces téléphoniques, la conversation de gestes avec la chose peut prendre plus précisément la forme d'un 'échange' (au sens de Goffman) marqué par des indices d'émotions.

Une implication de l'hypothèse du format social est qu'elle conçoit le travail émotionnel comme pouvant se dérouler non seulement dans un for intérieur accessible aux

seuls sujets, mais aussi au niveau observable de la transaction avec l'environnement. La prise en compte de cette possibilité a pour conséquence de déplacer la focale du sentiment intérieur vers le comportement extérieur. D'un point de vue temporel, le processus central n'est plus le changement interne d'un sentiment, mais la transformation externe d'une situation. Le travail émotionnel transforme non seulement le ressenti, mais aussi la raison d'être de ce dernier, à savoir l'état du milieu d'action.

Comment approcher les émotions au niveau de l'action ? Nous ferons appel à la définition préliminaire des émotions de Frijda (1986), qui s'est inspiré, comme Hochschild, des travaux néo-freudiens de Lazarus, mais aussi de la phénoménologie des émotions de Sartre (1995) et des travaux sur l'expression faciale d'Ekman et Friesen (1975).

À la manière d'un weberien, Frijda fait remarquer que d'ordinaire nous utilisons des termes d'émotion pour décrire des conduites qui semblent s'écarter du modèle de l'action instrumentale. Parmi celles-ci, il énumère : 1) l'interruption soudaine de l'activité ; 2) l'acte instrumentalement inefficace ou gratuit ; 3) des formes d'excès ou de défaut dans la manière d'accomplir une action instrumentale.

Les expressions d'émotions que nous avons listées plus haut répondent à cette caractérisation. En quoi une exclamation (par exemple, 'C'est pas vrai !') pourrait-elle contribuer à la coproduction d'un service ? L'étude d'usage met également en évidence des manières apparemment ' excessives ' de parler à la machine. Quel pourrait être l'avantage de crier, en détachant les mots, ' Je veux des informations ! ', comparativement à une réalisation ' neutre ' du même énoncé ?

2 Méthode

2.1 Constitution de la collection

L'interrogation sur la place des émotions dans les usages problématiques d'une interface conversationnelle s'est imposée d'elle-même alors que nous examinions comment les appelants ' travaillaient ' sur le plan de la communication en résolvant le problème du rejet. Sur la base de la définition préliminaire de Frijda, nous avons constitué une collection d'appels *prima facie* émotionnels. Nous les avons prélevés sur une base d'enregistrements audio qui stocke des milliers d'appels réels de clients aux serveurs vocaux d'un opérateur téléphonique.

L'étude porte sur un serveur automatique qui se présente comme une ' agence en ligne '. Le système traite des demandes diverses allant des informations générales aux

abonnements, en passant par les consultations de son compte et des opérations téléphoniques précises (comme le transfert de ses appels de sa ligne fixe vers son mobile). Dans un ‘ dialogue ’ typique, l’interface pose une question ouverte dans le premier tour et l’appelant répond de manière relativement libre dans le deuxième. Dans le troisième tour l’interface pose une question fermée à choix multiple et l’appelant répond dans le quatrième tour dans le langage de l’interface. En voici un exemple :

1	interface :	Prononcez le nom du service qui vous intéresse ou formulez librement votre demande.
2	appelant :	Forfait.
3	interface :	Plusieurs services peuvent répondre à votre demande : votre service messagerie vocale et votre abonnement deux heures. Pour consulter l’un d’entre eux prononcez son nom.
4	appelant :	Abonnement deux heures.
5	interface :	Veillez patienter. Je vous passe votre abonnement deux heures.

La vocation primaire de la base d’enregistrements est d’alimenter la phase du développement des interfaces vocales que les concepteurs appellent *tuning*. Une fois que la technologie est déployée, le tuning est le processus par lequel les paramètres de fonctionnement de l’interface sont périodiquement ajustés aux usages observés (pour l’analyse d’un cas de tuning concernant cette interface, voir Aranguren, 2012). La première étape exige de transcrire les énoncés des appelants, c’est-à-dire de les transformer en ce que les développeurs appellent des ‘ logs ’. La base conserve non seulement l’enregistrement des appels, mais aussi, pour certains d’entre eux, une transcription en ‘ logs ’ des énoncés des utilisateurs.

Les transcriptions obtenues en phase de tuning ne précisent pas le placement séquentiel des énoncés et se bornent à leur aspect strictement lexical, mais elles fournissent un point de départ précieux permettant de sélectionner les appels qui nous intéressent à partir de mots clés. En effet, dans nos travaux précédents nous nous sommes aperçus que les appels qui contenaient des comportements émotionnels (sous la définition préliminaire plus haut) incluaient à la fois, mais pas exclusivement, des manières ‘ marquées ’ de prononcer des énoncés appartenant à la tâche, et des ‘ apartés expressifs’, c’est-à-dire des énoncés qui ne faisaient pas partie de la tâche à strictement parler mais qui semblaient entretenir un lien intelligible avec la progression de la tâche, notamment avec ses arrêts.

Afin de les entrer dans le moteur de recherche de la base de données, nous avons repéré les mots caractéristiques des apartés expressifs qui, de toute apparence, n'étaient pas instrumentalement nécessaires au dialogue (4). Par des itérations de la procédure nous avons identifié des mots qui participaient typiquement à des exclamations et qui, de ce fait, pouvaient indiquer une réponse émotionnelle en cours (5).

Cette étape de présélection nous a livré près de 200 appels. Leur écoute nous a persuadé de l'avantage de nous pencher sur les appels les plus longs, car ils permettaient de suivre le développement de la réponse *prima facie* émotionnelle d'une manière que les appels relativement plus courts ne permettaient pas. Le développement des émotions des appelants est également observable dans les séries d'appels consécutifs par un même client, même lorsque les appels individuels sont relativement courts.

Concernant le critère de la durée nous avons retenu, pour un même client, les appels uniques longs aussi bien que les séries d'appels consécutifs. Nous avons tenu compte également de l'issue de l'appel, en essayant de constituer une collection équilibrée d'échecs et de réussites.

Comme l'ont montré Baudouin et Velkovska (2010), la vaste majorité des appels à ce serveur vocal dure autour de 30 secondes. En ce sens, on peut considérer un appel comme ' long ' s'il dépasse la minute. Tous les appels ou séries d'appels que nous avons sélectionnés dépassent ce seuil, et la plupart dépassent les trois minutes. Voir le Tableau 3 en Annexe pour les statistiques de base de la collection de 30 appels sur laquelle portent nos analyses.

2.2 *Segmentation*

Dans l'étape exploratoire de l'étude nous avons utilisé une version simplifiée des conventions de Gail Jefferson (Sacks, Schegloff et Jefferson, 1974) pour transcrire les appels de la collection. Ensuite, sur le logiciel Elan (Max Planck Institute for Psycholinguistics) nous avons segmenté et codé les transcriptions à l'aide d'une grille unique. Nous avons également mesuré l'intensité et le ton des énoncés des appelants. Enfin, nous avons exporté ces données vers le logiciel Theme® (Noldus Information Technology), qui nous a permis de détecter des patterns séquentiels statistiquement significatifs. Dans ce qui suit nous explicitons brièvement ces opérations successives, en commençant par la segmentation.

Nous avons divisé le comportement de la machine en unités discrètes en prenant pour point de référence ce que les concepteurs appellent un *prompt*. Nous considérons comme un même prompt tous les énoncés et autres sons que l'interface enchaîne automatiquement sans

tenir compte du comportement de l'utilisateur. Autrement dit, font partie du même prompt tous les énoncés et tous les sons émis par l'interface qui ne réagissent pas à des signaux provenant de la ligne de l'utilisateur. Par exemple, le prompt d'accueil se compose des éléments suivants, dans cet ordre : une musique, un message de bienvenue, un bref silence, une musique, une question ouverte (par exemple, ' Que désirez-vous ? '). Tous ces éléments appartiennent à la même unité, le prompt d'accueil, parce qu'ils se suivent les uns les autres de manière automatique. En revanche, si, après la question ouverte, l'appelant ne dit pas de mot, au bout de quelques secondes l'interface reprendra la parole avec ' Je ne vous ai pas entendu '. Ce message est une réponse au silence de l'appelant. Par conséquent, nous le considérons comme un segment à part, c'est-à-dire comme un autre prompt.

Le prompt, en tant qu'unité de comportement de la machine, agit comme un équivalent fonctionnel du tour de parole, qui est l'unité de comportement ' phonatoire ' de l'appelant. Chaque fois que l'appelant émet des sons, allant des interjections primaires simples comme ' ah ! ' jusqu'à des énoncés complexes comme ' mais madame, ça ne marche pas, ça fait trois fois que vous me dites la même chose ', nous avons examiné la suite pour voir si la vocalisation n'était pas l'ouverture d'un tour de parole plus long. Ainsi, par exemple, ' ah ! ' peut préfacier le tour de parole unitaire ' ah ! (silence) c'est pas possible ! '.

Nous avons dû faire appel à cette dénomination quelque peu étrange de ' conduite phonatoire ' pour la distinguer d'une autre forme de conduite qui s'est avérée pertinente pour l'analyse des expressions d'émotions, à savoir la conduite respiratoire, en particulier les expirations marquées. Celles-ci n'impliquent pas l'appareil phonatoire mais sont perçues comme possédant un caractère expressif frappant. Lorsqu'elles se produisent de manière isolée, nous les avons considérées comme des unités à part entière.

2.3 Système de codage

La procédure de segmentation adoptée divise la conduite de l'interface en prompts et celle de l'utilisateur en tours de parole ou, à défaut, en actes respiratoires expressifs. Nous avons ensuite codé ces segments afin de soumettre les appels codés à une procédure de détection de séquences statistiquement significatives (voir la section 2.5 plus bas). À cet effet, nous avons élaboré un système de codage qui tient compte des réactions récurrentes des appelants aux différentes difficultés posées par l'interface. Les observations se fondent sur la phase exploratoire de la recherche que nous rapportons ici, ainsi que sur l'étude de l'usage d'une autre interface conversationnelle à laquelle nous avons contribué (Laroche et al, 2011).

Il est heuristiquement éclairant de supposer que les appelants, quoi qu'ils puissent être en train de faire par ailleurs, s'engagent dans le dialogue avec l'interface afin d'accomplir une tâche (6). La tâche peut avoir des buts différents : se renseigner sur l'état de sa consommation, effectuer un transfert d'appel, obtenir des informations sur un service, souscrire à un abonnement, etc. Dans ce contexte, la tâche correspond à un service que l'on peut souvent identifier à partir des énoncés de l'appelant. Si l'on suppose que l'appelant est motivé à accéder à un service particulier, on peut aussi supposer qu'il n'est pas indifférent à l'état d'avancement de la tâche. Dit positivement, l'appelant est sensible à la progression de celle-ci, à sa continuation fluide, à son achèvement espéré, mais aussi à ses interruptions imprévues.

Le comportement mécanique de l'interface a la capacité de susciter des réactions émotionnelles non en lui-même, mais en vertu de l'attente de progression qui sous-tend l'engagement de l'appelant dans le dialogue. Nous aurons l'occasion de mieux préciser la nature de cette attente générique dans la section consacrée à la discussion.

À partir de ces remarques nous avons élaboré des codes pour décrire le comportement de l'interface et de l'utilisateur. Les deux tableaux ci-dessous présentent et expliquent le système de codage utilisé.

Tableau 1 : Codes décrivant le comportement de l'interface

Code	description
accueil	prompt indiquant le début de la tâche.
suivant	prompts indiquant la continuation de la tâche. Critère : pertinence conditionnelle (Schegloff, 1968).
nonrec_rep	prompts indiquant l'interruption de la tâche (répétitions et erreurs de reconnaissance).
ret_accueil	prompt indiquant le recommencement de la tâche (message de retour à l'accueil).
Resol	prompts indiquant l'achèvement de la tâche.
misrec	prompts sans pertinence conditionnelle à l'exception des <i>nonrec_rep</i> .

Tableau 2 : Codes décrivant le comportement de l'appelant

Ling	tous les actes de parole, sauf les interjections primaires
paraling	interjections primaires
Resp	actes respiratoires perceptibles
Comp	appelant compose un numéro
Rac	appelant raccroche
Rire	appelant rit
Bruit	bruit
ling,expre	actes de parole n'ayant pas de valeur instrumentale pour la tâche (' expressions d'émotions ')
ling,expre,cntxt_dialogue	actes de parole n'ayant pas de valeur instrumentale pour la tâche à l'adresse de la machine (' insultes ').
ling,emoplus	énoncé caractérisé par 1) une augmentation relative de l'intensité (dB) et du ton (Hz) ; 2) interprétable comme le symptôme d'un état émotionnel.
ling,emomin	énoncé caractérisé par 1) une diminution relative de l'intensité (dB) et du ton (Hz) ; 2) interprétable comme le symptôme d'un état émotionnel.
Chev	chevauchement

2.4 Description acoustique des énoncés

Le caractère émotionnel d'un énoncé peut être repéré au niveau de sa réalisation acoustique, dans la mesure où les émotions s'associent à des modifications caractéristiques du signal vocal. Notamment, la dite activation émotionnelle semble être en corrélation avec la fréquence fondamentale (mesurée en Herz, Hz) et avec l'intensité (mesurée en déci-Bels, dB), comme en attestent les résultats répliqués d'études issues de traditions pourtant opposées en psychologie (cf. Scherer, 2003 , Russell et al, 2003). L'activation est classiquement conçue comme une variable bipolaire allant, au niveau du vécu, de la torpeur à l'excitation, et au

niveau physiologique, d'un minimum à un maximum d'énergie disponible dans les tissus de l'organisme (Duffy, 1962).

Étant donné qu'un changement d'activation (à la hausse ou à la baisse) est nécessaire à tout état émotionnel (Frijda, 1986, Russell, 2003), nous avons considéré une variation dans la fréquence fondamentale (le ton) ou dans l'intensité (le 'volume') des énoncés comme une condition nécessaire pour les qualifier d'émotionnellement marqués. Lorsque le niveau objectif du ton ou du volume présente des changements quantitativement significatifs, et seulement dans ce cas, nous avons complété l'attribution d'émotion par un jugement subjectif sensible à d'autres variables (7).

2.5 Détection de t-patterns

Nous avons soumis les données séquentielles issues des étapes précédentes à une procédure de détection de t-patterns. Comme la démarche est quelque peu inhabituelle en sociologie, il n'est pas inutile de donner une définition informelle du t-pattern et un bref aperçu du mode opératoire du logiciel permettant de le détecter. Pour une présentation générale et plus rigoureuse de l'un et de l'autre on peut se référer à l'article très instructif de Magnusson (2000).

Un t-pattern est avant tout une séquence de types qui se répète plusieurs fois dans une période de temps. Par exemple, la séquence (u,ling i,suivant), dans laquelle d'abord l'utilisateur (u) demande quelque chose et ensuite l'interface (i) satisfait la demande, connaît un certain nombre d'occurrences dans la totalité de notre collection.

Deuxièmement, pour les différentes occurrences d'un t-pattern donné, les composants (les types) se produisent dans le même ordre à chaque fois, et les distances temporelles séparant les composants consécutifs demeurent relativement invariables. La séquence (u,ling i,suivant), par exemple, suppose que *u,ling* a lieu avant et *i,suivant* après, et que la distance temporelle entre l'un et l'autre est relativement courte. Si le deuxième composant se produit trois tours de parole après le premier, on ne dira pas de leur distance temporelle qu'elle est courte.

Enfin, ces distances temporelles entre les composants consécutifs sont invariables en supposant que ces composants ne sont pas distribués de manière indépendante et fortuite dans le temps. Pour notre exemple, cela veut dire que la probabilité de trouver une occurrence de *i,suivant* immédiatement après une occurrence de *u,ling* est significativement plus grande que ce que l'on pourrait attribuer au hasard. Lorsque cette condition est satisfaite, on dit de deux

ou de plusieurs types qu'ils sont reliés par des relations d' 'intervalle critique '. L'Encadré 1 ci-dessous fournit une présentation concise de cette dernière notion.

Nous avons considéré les t-patterns détectés dans notre collection comme des structures séquentielles hypothétiques. Comme le travail émotionnel du client, sous notre définition, constitue une telle structure, la détection de t-patterns fournit une méthode valable pour identifier les composants du processus.

Afin de repérer les t-patterns à l'œuvre dans notre collection, nous avons utilisé le logiciel de détection automatique Theme®. Sur ce logiciel, les principaux paramètres sont le nombre minimal d'occurrences du t-pattern, dont le minimum absolu est 2, et le niveau de signification, dont le standard est 0,005. Nous avons opéré à un niveau de signification entre 0,0001 et 0,005.

Pour des bases de données faites de plusieurs périodes d'observation (dans cette étude, des appels) on peut également déterminer un nombre minimal de périodes où le t-pattern doit se produire pour être admis. Nous avons d'abord recherché les t-patterns les plus courants, en examinant les séquences ayant lieu au moins dans 15 appels. Cette première étape de détection nous a permis de retenir un t-pattern unique relativement fréquent, dont nous avons examiné les variations dans la deuxième étape, en admettant cette fois-ci un nombre plus faible d'occurrences pour chaque variante.

Encadré 1. La relation d'intervalle critique

Si A est un composant antécédent (par exemple, un rejet) et B un composant conséquent (par exemple, une exclamation) d'un même t-pattern récurrent, après l'occurrence de A au temps t , il y a un intervalle $[t+d_1, t+d_2]$ (8) qui tend à contenir au moins une occurrence de B plus souvent que ce que le hasard ne le laisserait prévoir. Cette relation s'appelle intervalle (ou relation) critique entre les distributions de A (le rejet) et de B (l'exclamation).

La figure ci-dessous illustre le genre de problèmes que la détection de t-patterns peut aider à résoudre.

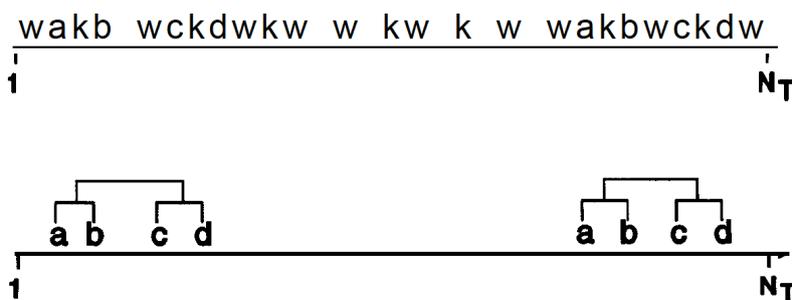


Illustration du t-pattern. Reprise de Magnusson (2000: 95)

Dans la partie supérieure de l'image apparaît une liste de comportements codés avec leur localisation dans la ligne du temps. Les comportements codés constituent des occurrences de quelques types d'événements. Par exemple A, B, C, etc. Les minuscules a, b, c, etc. indiquent leurs réalisations respectives durant la période considérée, allant de 1 à N_T . La partie inférieure est identique à la supérieure, sauf que toutes les occurrences de K et de W ont été supprimées. Un t-pattern simple (ABCD) émerge, qui reste difficile à voir tant que les autres événements sont présents. La figure montre l'arbre binaire ((AB)(CD)), mais d'autres possibilités comme ((A(BC)D)) ou (A(B(CD))) auraient été également envisageables.

3 Résultats

3.1 Énoncé, chevauchement, rejet, expression : un t-pattern récurrent

Nous avons détecté un t-pattern qui compte 28 occurrences dans toute la collection, distribuées sur 17/30 appels, impliquant 13/19 appelants. La colonne gauche du tableau ci-dessous précise l'ordre dans lequel ces quatre composants se présentent. La colonne droite indique les variations de ces composants, ainsi que la proportion de variations par rapport au total.

1	C :	énoncé	dans 9/28 cas : C hausse le volume et le ton de manière émotionnelle.
2		chevauchement	25/28 : C interrompt S.
3	S :	rejet	
4	C :	expression	08/28 : C insulte S.
5			07/28 : C raccroche.

Penchons-nous sur quelques exemplaires de cette séquence stéréotypique afin de rendre l'analyse plus concrète. Dans la transcription ci-dessous, 'i' désigne l'interface ou le système, 'u' l'utilisateur ou le client. 'TC' précise la durée d'une action ou d'un événement dans la ligne du temps. 'SD' indique la durée des silences. 'U-voem' fait référence aux vocalisations émotionnelles de l'utilisateur, le signe '+' servant à indiquer une hausse émotionnellement marquée du ton (Hz) et/ou de l'intensité (dB). La colonne tout à gauche n'indique pas nécessairement des tours de parole mais sert simplement à se repérer dans la transcription.

3.1.1 Exemple 1

marre5(1)

01	i	les heures france ne sont plus [commercialisées sur] ce
02		serveur,
03	TC	00:01:10.560 - 00:01:14.000
04		
05	u	[consulter.]
06	TC	00:01:12.344 - 00:01:13.253
07		
08	chev	
09	TC	00:01:12.344 - 00:01:13.253
10	SD	(1.1)
11		
12	i	actuellement vous êtes abonné aux services allofact,
13		messaging vocale de cet opérateur, et, atout plus deux
14		cent quarante. pour [consulter l'un d'en]tre eux, pr-
15	TC	00:01:14.310 - 00:01:25.562
16		
17	(1)	
18	u	[consulter.]
19	U-voem	Hz+
20	TC	00:01:23.733 - 00:01:24.637
21		
22	(2)	
23	chev	
24	TC	00:01:23.733 - 00:01:24.637
25	SD	(1.2)
26		
27	(3)	
28	i	actuellement vous êtes abonné aux
29		[services allo]fact,
30		messaging vocale de cet opérateur, et, atout plus deux
31		cent quarante. pour consulter l'un d'entre eux, prononcez
32		son nom. sinon, deman-
33	TC	00:01:25.821 - 00:01:40.265
34		
35	(4)	
36	u	[ô c'est pas vrai hein,]
37	TC	00:01:27.764 - 00:01:28.774
38	SD	(9.7)

Les chiffres entre parenthèses dans le corps de la transcription indiquent les quatre composants du t-pattern évoqué.

(1) *Enoncé.* Le premier composant de la séquence apparaît en ligne 18. Il témoigne d'une hausse du ton émotionnellement significative, comme le montre le code sur la ligne 19. En deuxième lieu, il constitue une répétition de l'énoncé de la ligne 5. Enfin, ce premier composant du t-pattern chevauche un tour de parole de l'interface qui ne semble pas donner satisfaction à la demande formulée en ligne 5. Autrement dit, l'énoncé réagit au rejet d'une demande.

(2) *Chevauchement.* Le deuxième composant du t-pattern figure sur la ligne 23. Ses coordonnées temporelles, sur la ligne 24, montrent que l'énoncé de l'appelant se superpose avec une phrase de l'interface en cours. Plutôt qu'une simple coïncidence, le chevauchement indique donc une interruption de l'interface par l'appelant.

(3) *Rejet*. Le troisième composant (lignes 28-32) est un *reprompt*, une répétition en boucle de l'interface. Tout comme la première (lignes 12-14), la deuxième occurrence de ce message ne semble pas satisfaire la demande de l'appelant.

(4) *Expression*. Le composant terminal du t-pattern, sur la ligne 36, est une expression d'émotion – une version complexe de la locution interjective ‘ c'est pas vrai ! ’ – prononcée en aparté.

3.1.2 Exemple 2

payer ma facture

01	i	je ne vous ai pas entendu. pouvez-vous répéter, en
02		parlant devant le combiné, et plus fort? prononcez le nom
03		du service qui vous intéresse ou formulez librement votre
04		demande. sinon, [prononcez le code du se]rvi-
05	TC	00:00:43.356 - 00:00:57.191
06		
07	u	[payer ma facture.]
08	U-voem	dB+ Hz+
09	TC	00:00:55.489 - 00:00:57.041
10		
11	chev	
12	TC	00:00:55.489 - 00:00:57.041
13	SD	(1.1)
14		
15	i	je ne vous ai pas entendu. [pouvez-vous répéter, en par-]
16	TC	00:00:58.095 - 00:01:01.680
17		
18	(1)	
19	u	[payer ma facture.]
20	U-voem	dB+ Hz+
21	TC	00:00:59.968 - 00:01:01.624
22		
23	(2)	
24	chev	
25	TC	00:00:59.968 - 00:01:01.624
26	SD	(1.1)
27		
28	(3)	
29	i	je ne vous ai pas enten[du. pouvez-vous répé-]
30	TC	00:01:02.720 - 00:01:05.424
31		
32	(4)	
33	u	[â mais pas possib' ça.]
34	U-voem	Hz+
35	TC	00:01:03.868 - 00:01:05.339
36	SD	(1.1)

(1) *Énoncé*. Sur la ligne 19, le premier élément du t-pattern est prononcé d'une manière émotionnellement marquée, avec des valeurs élevées d'intensité et de ton (voir la description acoustique sur la ligne 20). Par cet énoncé l'appelant répète *verbatim* sa demande de la ligne 7. La notification d'erreur de reconnaissance, que la machine effectue en ligne 15, ne constitue pas une réponse satisfaisante à cette première demande.

(2) *Chevauchement.* La notification d'erreur qui commence en ligne 15 est identique à celle qui débute en ligne 1. Il s'agit d'un tour de parole complexe, composé de plusieurs phrases. On peut raisonnablement penser que, après l'avoir écouté une fois, l'appelant sait que le message ne s'arrête pas à la première phrase. Dans cette hypothèse, le chevauchement signifie une interruption.

(3) *Rejet.* En ligne 29, le système notifie encore une erreur de reconnaissance.

(4) *Expression.* En ligne 33, en réaction à cette erreur, l'appelant s'exclame en aparté.

3.1.3 Exemple 3

petasse		
01	i	transfert d'appel. vous pouvez demander à transférer vos
02		appels, ou bien désactiver le transfert en cours.
03		[que voulez-vous faire?]
04	TC	00:00:19.819 - 00:00:29.366
05		
06	u	[transférer mes appels.
07	TC	00:00:28.290 - 00:00:29.750
08	SD	(2.1)
09		
10	chev	
11	TC	00:00:28.290 - 00:00:29.366
12		
13	(1)	
14	u	transfert [d'appel.]
15	U-voem	dB+ Hz+
16	TC	00:00:31.810 - 00:00:33.440
17		
18	(2)	
19	chev	
20	TC	00:00:32.482 - 00:00:33.440
21		
22	(3)	
23	i	[transfert d'appel].
24		vous pouvez deman[der à transférer vos] appels,
25	TC	00:00:32.482 - 00:00:37.116
26		
27	(4)	
28	u	[ô pétasse]
29	TC	00:00:35.660 - 00:00:36.530
30		
31	(5)	
32	u	((raccroche))
33	TC	00:00:36.740 - 00:00:38.530

(1) *Enoncé.* Le premier composant de la séquence, en ligne 14, est une vocalisation émotionnelle de l'appelant marquée par des valeurs élevées d'intensité et de ton. Elle reformule une demande déjà formulée en ligne 6. La redemande ne fait pas suite à la réponse de la machine, mais plutôt à son absence de réponse. La ligne 8, de fait, indique un silence supérieur à deux secondes après la première demande (9). On peut approcher cette redemande émotionnellement marquée sous le même angle que les autres : elle continue une séquence faite 1) d'une demande et 2) d'une prise en compte de celle-ci attendue mais manquante. Dans

ce contexte, on peut penser qu'un silence long est l'équivalent fonctionnel d'un rejet : les deux signifient au client que sa demande n'a pas été satisfaite.

(2) *Chevauchement*. En ligne 23 l'interface répond à la première demande retranscrite en ligne 6. La réponse a l'air d'une ' interruption ' de la redemande de l'appelant (ligne 14), mais elle tient simplement à un retard de la prise de parole de la machine dû au temps de traitement du signal de l'appelant.

(3) *Rejet*. Cette réponse apparemment interruptive répète le message qui commence en ligne 1.

(4) *Expression*. L'appelant insulte l'interface en ligne 28.

(5) *Sortie*. Cet appel exemplifie le t-pattern de cinq composants qui prolonge la séquence de base. Peu après l'expression d'émotion, ici l'insulte, l'appelant raccroche.

La figure ci-dessous cartographie la localisation temporelle, à travers les 30 appels de la collection, des 28 occurrences de ce t-pattern générique et de ses différentes variations. Pour tous les appels, chaque occurrence du t-pattern s'y trouve encadrée et ses composants énumérés.

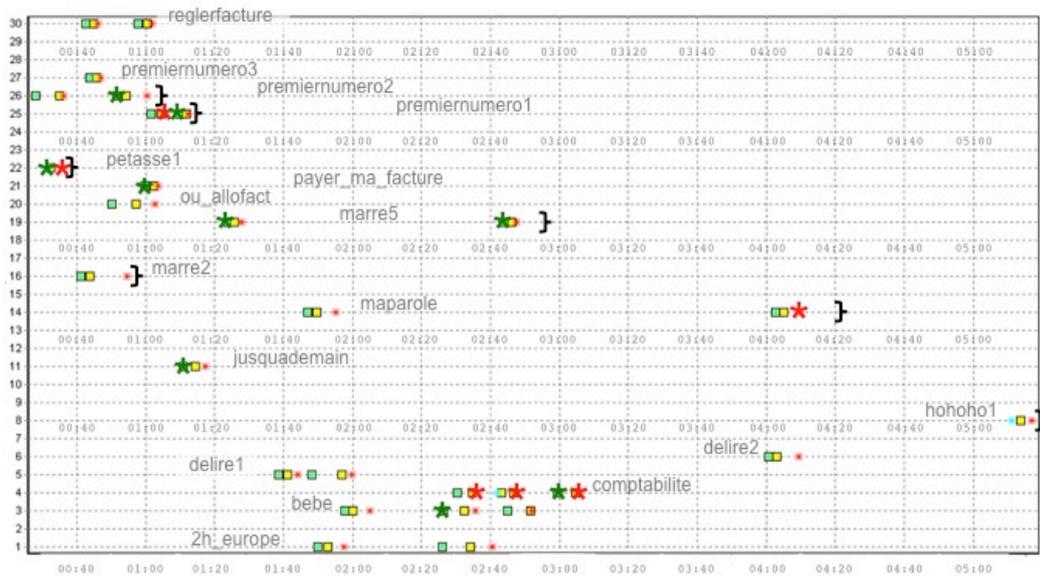
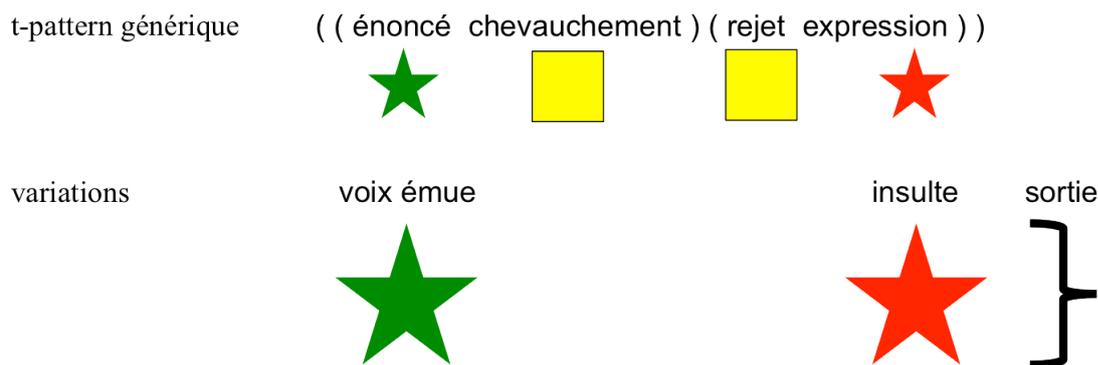


Figure 1. Cartographie du t-pattern générique et de ses variations.

Légende:



3.2 Demande, rejet, redemande, rejet, expression : la structure séquentielle du travail émotionnel du client

Nous avons décrit les épisodes émotionnels individuels à l'aide du t-pattern (énoncé chevauchement rejet expression), éventuellement suivi de *sortie*. Les extraits présentés ci-dessus ont été choisis parce qu'ils illustrent des tendances générales de la collection.

Lorsqu'on examine la réalisation *in situ* de ce t-pattern on s'aperçoit, comme le suggèrent les extraits, que cette séquence à quatre composants est souvent, sinon toujours, préfacée par deux autres composants : une demande initiale de l'appelant et un rejet (actif ou passif) de la machine. Dans cette séquence élargie de six composants, ce que nous avons appelé 'énoncé' fonctionne plus précisément comme la réitération d'une demande. En termes d'acte de discours, l'énoncé constitue donc une *redemande*. Typiquement, la redemande revêt une tonalité émotionnelle qui s'associe, sur le plan de la réalisation acoustique, à une hausse objective du ton ou de l'intensité, et sur le plan du séquençage des tours, à une tendance à interrompre le message en cours de l'interface. Comme le composant *chevauchement* fait référence, dans 25/28 occurrences du t-pattern, à des interruptions, il semble plus convenable de le traiter directement comme un modificateur de la redemande, au lieu de le considérer comme un composant séparé.

Nous arrivons au format séquentiel suivant pour décrire le travail émotionnel du client. La partie supérieure précise les composants génériques, la partie inférieure les modificateurs. La ligne du temps va de gauche à droite. ' C ' : client, ' S ' : système.

C,demande	S,rejet	C,redemande	S,rejet	C,expression
		voix émue ; séquençage interruptif		aparté interjectif ; insulte

Ajoutons deux remarques. En premier lieu, dans la collection, nous n'avons pas trouvé d'expressions (interjections ou insultes) qui ne seraient pas précédées par une redemande rejetée. En revanche, toutes les redemandes rejetées ne donnent pas lieu automatiquement à une expression. Il semblerait que le couple redemande-rejet puisse se répéter plusieurs fois avant que l'appelant exprime son émotion par une interjection ou par une insulte, et cela même lorsqu'une voix trop stridente ou une prise de parole trop hâtive trahissent déjà la tonalité émotionnelle de la redemande.

En deuxième lieu, l'expression en réaction à une redemande rejetée peut préfacier l'abandon de l'appel, mais le plus souvent, comme nous l'avons vu, le client persiste. Dans la collection, nous avons trouvé jusqu'à trois réalisations de la séquence type sur un même appel (voir la Figure dans l'Annexe 1).

4 Discussion

4.1 L'expression d'émotion comme réponse à une provocation

Nous allons interpréter nos résultats à partir de la notion d'échange rituel proposée par Goffman (1967, 1971). Commençons par conceptualiser à la manière goffmanienne le format découvert. L'expression comme réaction à une redemande rejetée serait une variation structurelle 'perturbatrice' d'un format d'interaction rituelle, courant dans les relations civiles, que Goffman appelle *échange réparateur*.

L'échange réparateur est le médium interactionnel par lequel se réalise le rééquilibrage d'une relation sociale troublée. Le trouble reçoit le nom d' 'offense virtuelle', terme par lequel il faut entendre, pour un acte donné, l'interprétation la plus offensive pour son destinataire et la plus diffamatoire pour son auteur. L'offense virtuelle peut guider l'interprétation d'un acte passé, mais aussi d'un acte futur. Si l'offense virtuelle se situe dans le passé, elle motive une excuse ou une explication ; si elle est dans l'avenir, une demande (10).

Si dans un repas je souhaite me servir de la salière, je peux tendre mon bras à travers la table. Si je dérange visiblement la personne qui est assise à côté de moi, je peux lui présenter mes excuses, ou bien justifier mon acte, ou bien les deux (par exemple 'Excuse-moi, je prends la salière'). Alternativement, je peux estimer moins embarrassant de demander directement à cette personne de me passer la salière. L'excuse, l'explication et la demande

visent à annuler ou à éviter l'offense virtuelle, qui consisterait à tendre mon bras, à prendre la salière en dérangeant mon voisin et à continuer le repas comme si de rien n'était.

Dans ce modèle, la motivation de l'offenseur est de démontrer sa conformité aux normes en apportant des preuves que ses actes ne constituent pas des offenses virtuelles. Corrélativement, la motivation de l'offensé est d'éviter de confirmer qu'il est victime d'une offense virtuelle, ce qui le dispose à accepter des excuses ou des explications et à satisfaire des demandes. Cela l'incline également à attirer l'attention de l'offenseur sur l'offense commise si celui-ci n'initie pas la réparation de lui-même. Dans l'exemple de la salière, si je dérange mon voisin et me montre indifférent à son inconfort, il pourra me regarder de manière 'interrogative' en attendant mon excuse ou mon explication.

L'expression d'émotion en réaction à une redemande rejetée apparaît dans le schéma de Goffman comme un échec de l'échange réparateur, d'où son caractère 'perturbateur'. Voici la séquence qui pourrait la précéder dans l'exemple de la salière :

demande	A :	Tu peux me passer la salière, s'il te plaît ?
Rejet	B :	Non.
redemande	A :	Vas-y, passe-moi la salière.
Rejet	B :	Non.
expression	A :	(irrité) Elles sont débiles tes blagues, franchement.

Dans le modèle de Goffman, les excuses, les explications et les demandes appartiennent à la classe fonctionnellement unifiée des 'remèdes'. Dans l'échange réparateur standard, le remède de l'offenseur agit comme une affirmation qui projette l'attente d'un 'soulagement' en tant que réplique de l'offensé. Accepter des excuses et des explications, ainsi que satisfaire des demandes, sont autant de manières de répondre à un remède par un soulagement. Une variation supplémentaire est le refus justifié, qui n'octroie rien d'autre que le soulagement de confirmer la pertinence de la demande et la respectabilité du demandeur. Les mots 'affirmation' et 'réplique' font penser à des actes de parole, mais ils désignent avant tout des 'coups', des gestes verbaux ou non verbaux, qui fonctionnent comme unités constitutives des interactions. Un échange se définit comme une série d'au moins deux coups accomplis chacun par un participant différent.

Lorsqu'une demande suscite un refus non justifié, l'absence de soulagement après le remède peut conduire le demandeur à voir dans ce refus la marque d'une *provocation*. Le rituel de réparation est alors mis en échec. La provocation installe un contexte alternatif, le

‘ duel de caractère ’, dans lequel son destinataire doit choisir entre le recours à la force et le retrait. Goffman fait remarquer que ce dernier s’accompagne typiquement d’expressions d’indignation, en présence ou non du provocateur (11).

4.2 Un travail émotionnel de face, pas rituel, pas socialement situé

Sous cette lumière, l’expression d’émotion de l’appelant suite au rejet de sa demande peut être interprétée comme la réponse à une provocation – l’offense virtuelle confirmée. Dans les usages de l’interface téléphonique, cette réponse expressive se réaliserait par un aparté (par exemple, ‘ c’est pas vrai ! ’) ou, moins courtoisement, par une insulte (par exemple, ‘ mais je t’emmerde ! ’). En adoptant la typologie de Goffman, du côté du duel de caractère on pourrait ranger l’insulte, et du côté du retrait indigné, l’aparté expressif et la ‘ sortie ’ par laquelle le client abandonne l’appel. Comme nous l’avons vu, des combinaisons sont possibles. Dans l’*Exemple 3* plus haut, face à une demande rejetée qui apparaît comme une provocation, l’appelant d’abord insulte la machine et ensuite raccroche.

Précisons que le terme ‘ provocation ’ ne doit pas être pris dans son sens ordinaire. Tout ce que nous voulons dire en proposant que l’expression d’émotion répond à une provocation, c’est que l’expression d’émotion est un type de coup interactionnel qui est systématiquement précédé par un autre type de coup interactionnel, auquel on peut attacher le nom ‘ provocation ’. Notre choix du terme n’a d’autre justification que le choix du terme par Goffman.

Lorsque le contexte séquentiel révèle l’offense virtuelle comme une provocation et l’offenseur virtuel comme un provocateur, le rituel interpersonnel de réparation se voit ‘ désorganisé ’. Goffman conceptualise ces échanges comme des formes de conduite obligatoire concernant des objets sacrés, ces derniers n’étant autres que les participants à l’interaction. Un rituel interpersonnel est donc une forme de conduite obligatoire par laquelle les participants s’adressent des expressions mutuelles de respect (Goffman, 1967, 19 et 57 , 1971, ch. 3). C’est ce qu’accomplit notamment l’échange réparateur, par des remèdes (excuses, demandes, explications) visant à démentir l’offense virtuelle (l’interprétation la plus diffamatoire d’un acte) (12).

Or, la notion de rituel d’interaction nous laisse avec une définition négative du travail émotionnel qui se réalise par la séquence type que nous avons dégagée : il s’agirait d’un échange réparateur malmené. Mais si l’on cherche chez Goffman une interprétation positive de cette activité, on peut la trouver dans la notion de ‘ travail de face ’ : l’effort qu’un individu

entreprend, dans des situations sociales, afin de contrecarrer les menaces à sa conception de soi (13). Goffman lui-même fait remarquer que ce travail est souvent impulsé par l'attachement émotionnel que les individus ont à l'idée qu'ils se font d'eux-mêmes. Notamment, les situations où l'on risque de perdre la face ont tendance à induire des sentiments négatifs chez l'individu, comme la honte ou l'embarras.

Goffman aligne ce travail de face, potentiellement émotionnel, avec la réalisation d'un rituel interpersonnel : afin de sauvegarder l'image positive de soi, toujours menacée par des événements imprévisibles, les individus disposent dans tout groupe social d'un répertoire limité de pratiques interactionnelles partagées, parmi lesquelles figurent les rituels d'interaction. Le couplage entre rituel et face suppose que les individus cherchent à préserver non seulement leur conception de soi, *mais aussi celle des autres* (Goffman, 1967: 11). Et que cette double motivation trouve sa voie de réalisation naturelle dans ces formes de conduite approuvées que sont les rituels interpersonnels.

Notre étude suggère, cependant, que l'échec de l'échange réparateur n'implique pas nécessairement l'échec du travail de face. La 'provocation' du serveur vocal désorganise la réparation, puisque le remède ne suscite pas un soulagement rituel, mais la sauvegarde de l'image de soi du client ne se voit pas désorganisée pour autant.

Ce qui nous amène à reconnaître que, sur le plan du travail de face, la réparation (octroi de la demande, acceptation de l'excuse ou de l'explication), la confrontation et le retrait constituent des équivalents fonctionnels. Tous trois continuent le travail de face en cours, bien que seulement la réparation le fasse en conformité avec un rituel interpersonnel, c'est-à-dire par l'expression réciproque de la valeur sacrée des participants – ou plus littéralement, dans le respect des normes de politesse.

Le travail émotionnel du client est donc un effort séquentiellement structuré pour sauvegarder la face (l'honneur, la dignité, etc.) dans une situation qui fait apparaître le serveur vocal comme la source d'une provocation. Le fait que cette structure ne réalise pas un rituel d'interaction constitue un premier point de contraste avec l'échange réparateur de Goffman. Le deuxième est encore plus évident : l'échange auquel nous avons affaire n'a pas lieu dans une situation sociale, c'est-à-dire de coprésence mutuelle entre deux individus ou plus. Les appelants sont bien au courant qu'ils interagissent avec une machine et non avec un opérateur humain. Et pourtant leur travail émotionnel obéit à un format qui semble être d'origine sociale, comme en atteste la ressemblance avec l'échange réparateur et ses ratés. On peut dire que le travail émotionnel que nous avons analysé est un travail de face sans être rituel, d'origine sociale sans être socialement situé.

On pourrait admettre l'origine sociale de la séquence découverte tout en remettant en question la pertinence de la conceptualiser comme un travail de face. Ce qui pose une question plus fondamentale : quel est le critère du travail de face ? Sur quelle base décide-t-on si telle ou telle séquence de gestes relève de ce travail ? Si l'on pouvait expliciter un ou plusieurs critères, on serait en état de juger si la notion de travail de face correspond aux usages documentés. En ce sens, la stratégie la plus naturelle serait de chercher ces critères chez Goffman.

On peut bien les chercher, mais on ne les trouvera pas. Goffman ne découvre pas que les gens font un travail de face : il le stipule. C'est un *a priori* de son approche. Goffman suppose le *face-work* et cherche ensuite à en déduire les phénomènes observés. Bien entendu, il n'y a rien d'étrange dans cette démarche. De manière générale, elle consiste à postuler une cause inobservable, à déduire les effets que cette cause inobservable aurait si elle existait, et à comparer ces prévisions aux faits effectivement observés. Si les faits se conforment à la prévision, on peut maintenir l'affirmation de la cause inobservable, notamment si elle fournit une explication relativement simple des observations.

On peut encore objecter que Goffman aurait dû fonder sur des faits l'explication de certaines structures d'interaction observées par un *face-work*. Par exemple, il aurait pu se servir de descriptions de l'expérience des participants à des interactions sociales. Mais ce procédé l'aurait-il vraiment aidé ? Si, concernant mon comportement en public, on me demande pourquoi je fais ceci ou pourquoi je fais cela, il ne me vient pas à l'esprit de tout expliquer en termes de travail de face. Dans ce domaine d'activité, les actions que j'accomplis *de manière réfléchie* dans le but de sauvegarder une certaine image de moi-même sont minoritaires. Cependant, Goffman stipule que *toutes* nos conduites en public relèvent d'un travail de face (14). La phénoménologie profane de ma conduite en public constitue-t-elle un test pour le postulat de Goffman ?

4.3 *Le travail émotionnel dans les interactions de service*

Le mode de fonctionnement des interfaces vocales conversationnelles comporte la possibilité du ' problème du rejet ', qui délègue à l'utilisateur le travail de découvrir si le rejet tient à un défaut de signal ou à un mauvais choix de vocabulaire. Notre étude montre que parfois ce travail de communication se double d'un travail émotionnel qui porte moins sur la difficulté à transmettre des informations que sur l'image inadmissible du client (et de ses besoins) que les rejets injustifiés de la machine tendent à projeter. Nous ne connaissons pas l'ampleur de ce

travail émotionnel parmi les cas de problème du rejet, ni pour cette interface conversationnelle ni pour toutes les autres qui peuplent le marché de services aujourd'hui. Mais nous pouvons dire avec confiance que le phénomène existe et qu'il s'enracine dans une contrainte indépassable de la technologie.

Les contraintes propres aux interactions de service non automatisées peuvent elles aussi donner lieu au genre de situation qui exigera du client un travail émotionnel de face. Cela concerne notamment les situations où le client d'une entreprise (ou l'utilisateur d'un service public) estime avoir une demande légitime et urgente que le représentant de l'organisation ne peut pas prendre en charge – par exemple, parce que le client ou l'utilisateur ne s'est pas adressé au bon interlocuteur, ou parce que l'organisation ne s'occupe pas de ce type de demande. Le client ou l'utilisateur peut n'accepter ni le refus ni sa justification, et à sa place 'recycler' sa demande, parfois d'une manière émotionnellement marquée. L'agent peut alors 'recycler' son refus justifié. Le couple redemande ému --- refus justifié peut se répéter plusieurs fois. L'agent pourra mettre fin à la boucle par une explication impliquant une évaluation péjorative implicite de la demande exprimée, par exemple, en alléguant que d'autres clients attendent ou qu'il est sollicité par d'autres tâches urgentes. La péjoration implicite peut s'accompagner d'un ton clairement plus sec, ou d'une préface tranchante comme 'Écoutez, ...'. Le client peut trouver cette péjoration inadmissible, et y répondre par une expression d'indignation. Nous avons vu ce format à l'œuvre dans des observations informelles en tant que participants fortuits, mais aussi dans une collection d'appels aux pompiers dont certains mettent en scène des appelants qui donnent des alertes en dehors du périmètre du service, par exemple un chat perché sur un arbre (Chave, 2010).

4.4 Travail émotionnel et usages problématiques d'objets techniques

Nos analyses indiquent que le travail émotionnel du client possède une structure séquentielle en quatre coups par lesquels, d'abord de gré et puis de force, l'appelant cherche à 'sauver la face'. Ce format est d'origine sociale, mais les activités qu'elle organise, comme dans l'étude rapportée plus haut, ne sont pas toujours socialement situées.

À partir de ce constat, on peut se demander si l'expression après la redemande rejetée ne structure pas bien d'autres situations émotionnelles d'usage problématique d'un objet technique. Un geste non verbal, par exemple, peut fonctionner comme une demande et installer un contexte dans lequel, immédiatement après ce geste, tout ce qui proviendra de l'objet technique sera interrogé dans sa capacité à satisfaire la demande exprimée. Si dans ce

créneau (*slot*) séquentiel l'objet technique ne se conforme pas à l'attente, l'utilisateur pourra comprendre, probablement de manière non réfléchie, qu'il vient de faire l'objet d'une sorte de 'provocation', en tout cas d'un 'refus' inadmissible. Pris dans cette structure séquentielle, il pourra réagir par une expression d'émotion – une interjection, une insulte, un geste plus ou moins vigoureux, etc.

D'autre part, le problème du rejet n'est que la déclinaison du problème de la communication homme-machine pour les interfaces vocales conversationnelles. Toutes les technologies d'automatisation de services peuvent donner lieu à des problèmes de communication fondés sur une asymétrie de compétence interactionnelle avec le client. On peut penser qu'au moins dans certains cas ces derniers traiteront le problème de communication comme une offense, et qu'un travail émotionnel 'de face' s'avérera nécessaire.

Conclusion

Le travail émotionnel du client d'un serveur conversationnel est un effort pour maintenir une conception positive de soi face à des événements qui se présentent comme des offenses et qui, en tant qu'événements, sont inévitables en raison du mode de fonctionnement de la technologie. Quelle est, au juste, la portée de ce travail dans les usages des services automatisés en général ?

D'autre part, dans les services non automatisés, la division du travail d'attention au public et la définition organisationnelle des requêtes pertinentes comporte la possibilité structurelle pour les agents de devoir refuser des demandes que les clients/usagers persisteront à estimer légitimes. Des situations émotionnelles de format comparable à celui que nous avons dégagé peuvent en résulter. Quelle est l'ampleur de ce travail émotionnel dans les interactions de service en face-à-face ?

Notre étude laisse ces questions ouvertes. Mais elle montre que le problème de l'interaction homme-machine n'est pas seulement un problème de communication, mais aussi un problème de face, au moins dans le cadre de la relation de service. Et que ce n'est pas seulement dans l'intériorité de la conscience, mais aussi dans la transaction avec l'environnement que s'effectue ce processus séquentiel que nous avons appelé travail émotionnel. L'étude des formats sociaux qui structurent nos transactions émotionnelles pointe, peut-être, vers un nouveau domaine en sociologie des émotions.

Remerciements

Je remercie les participants aux journées d'études 'Émotions et sentiments moraux', organisées par l'Institut Marcel Mauss (UMR 8178) en septembre 2012, pour leurs commentaires sur la première version de cet article. Je tiens à remercier également Alexandra Bidet, Stéphane Tonnelat et Louis Quéré pour leur relecture d'une version ultérieure.

Financement

Convention Cifre 0043/2009.

Notes

1. Les descriptions techniques des interfaces vocales s'appuient sur (Morgan et Balentine, 1999, Cohen et al, 2004, Harris, 2005).

2. Nous reprenons la typologie des interjections proposée par (Ameka, 1992). Il est des mots qui sont des interjections primaires, au sens où ils ne sont pas employés autrement ; par exemple, 'ah !' et 'oh !'. En revanche, les interjections secondaires sont des mots qui appartiennent à d'autres classes selon un critère sémantique, mais qui en vertu des usages peuvent fonctionner comme des interjections primaires, c'est-à-dire comme des énoncés non elliptiques d'un seul mot utilisés 'pour exprimer une attitude ou un état mental' (Ameka, 1992: 111), par exemple, 'punaise !'. Les interjections primaires ne couvrent pas seulement des mots monomorphémiques comme 'oh !', mais aussi des mots multimorphémiques, aussi présents dans notre collection, comme 'oh là là !'. On appelle ces derniers, interjections primaires *complexes*. A côté des interjections primaires, simples ou complexes, et des interjections secondaires, il est également des énoncés entiers qui servent à exprimer un état d'esprit, affectif ou autre, et que l'on appelle de ce fait *locutions interjectives*. Par exemple, 'c'est pas vrai !'.

3. Chez Durkheim, de fait, l'émotion comme conduite obligatoire coexiste avec l'émotion comme révélateur de l'obligation reconnue, comme l'a montré Patricia Paperman (1995).

4. Par exemple, l'énoncé 'C'est pas possible !' peut être considéré comme un aparté expressif, par opposition à un énoncé demandé par la tâche, donc instrumental, comme 'Je voudrais connaître ma consommation'. Ne faisant partie d'aucun énoncé instrumental en ce sens, le mot 'possible' peut servir d'indice pour repérer des appels où des apartés de ce type se produisent, et par là pour détecter des épisodes émotionnels.

5. La liste inclut les mots 'vrai', 'possible' et 'énerve', que l'on trouve en tant que composants des locutions interjectives 'c'est pas vrai/possible' et 'ça m'énerve'. Elle se poursuit, moins diplomatiquement, avec 'putain' et 'merde', des mots exprimant typiquement des jurons. Elle se termine très peu diplomatiquement, avec une série de termes qui participent à des insultes : 'chier', 'pétasse', 'conasse', 'salope', 'foutre', 'enculer'. Ce processus itératif a mis également en évidence que la présélection sur la base des seuls mots 'putain' et 'merde' était trop grande pour permettre l'écoute de tous les appels retenus. Cependant, comme les appels fournis par le moteur de recherche à partir d'autres mots clés s'avéraient contenir dans leur grande majorité, à un moment ou à un autre du dialogue, ces deux termes également, nous les avons exclus de cette étape de présélection, afin de rendre possible l'écoute de la totalité des appels retenus.

6. Il faut noter que de fait tous les appelants ne s'engagent pas dans la tâche. On peut constater que dans certains appels les utilisateurs prononcent des jurons et des insultes au moment où ils s'aperçoivent qu'ils ont affaire à un serveur automatique. Des appels de ce type ne font pas partie de notre collection.

7. Les deux codes indiquant des ‘vocalisations émotionnelles’ sont ling,emoplus et ling,emonin. Voir le Tableau 2 plus haut.

8. $d1$ = limite inférieure de l’intervalle ; $d2$ = limite supérieure. Dans le cas du t-pattern, la limite inférieure est égale ou supérieure à t (le moment où le premier composant, A, se produit).

9. Rappelons que dans la conversation informelle un silence de plus d’une seconde après avoir laissé la parole peut être pris comme un ralentissement du rythme normal du séquençage des tours (Jefferson 1989).

10. Notons que le mot anglais request, que nous traduisons par ‘demande’, apparaît dans la traduction française comme prière (Goffman, 1973, 1974).

11. Le refus non justifié d’une demande est l’équivalent fonctionnel du refus de l’offenseur de s’excuser, une fois que l’offensé a attiré l’attention sur l’offense (il s’agirait donc du refus d’une ‘interrogation’). L’épisode suivant, observé fortuitement par l’auteur, en fournit un exemple : ‘Printemps 2013 à Paris. Dans la rue, sur la voie obligatoire pour les bus et pour les vélos, facultative pour les taxis. Il vient de pleuvoir copieusement ; un filet d’eau coule sur le bord droit de la voie. Afin de l’éviter, un cycliste roule légèrement plus à gauche de ce qui serait physiquement possible. Il se fait claxoner par un taxi qui l’approche par derrière. En gardant l’orientation vers le devant, sans tourner la tête, le cycliste tend le bras gauche et, en réunissant les pointes des doigts, fait un geste d’interrogation ‘à la napolitaine’ avec un mouvement court et sec vers le taxi. Celui-ci commence une manœuvre de dépassement en s’écartant amplement du cycliste et en accélérant de manière bruyante. Ils croisent les regards lorsqu’ils sont à la même hauteur. Le chauffeur du taxi bouge la tête de droite à gauche ostensiblement.’

12. Pour se démarquer de l’Analyse de Conversation (cf. Sacks, Schegloff et Jefferson 1974), Goffman (1976) distingue son analyse séquentielle de l’interaction qua rituel de l’analyse séquentielle de la conversation qua transmission de messages.

13. Cf. l’essai ‘On face-work’ dans (Goffman 1967, trad. fr. ‘Perdre la face ou faire bonne figure ?’ dans Goffman, 1974).

14. ‘Toute personne vit dans un monde qui l’amène à avoir des contacts (*encounters*) (...). Lors de ces contacts, l’individu tend à extérioriser (*to act out*) ce qu’on nomme parfois une *ligne de conduite*, c’est-à-dire, un canevas (*pattern*) d’actes verbaux et non verbaux qui lui sert à exprimer son point de vue sur la situation, et, par là, l’appréciation qu’il porte sur les participants, et en particulier d’elle-même. (...) On peut définir le terme *face* comme étant la valeur sociale positive qu’une personne revendique effectivement à travers la ligne d’action que les autres supposent qu’elle a adoptée au cours d’un contact particulier. La face est une image de moi délinéée selon certains attributs sociaux approuvés (...).’ (Goffman, 1974 : 9 , c’est nous qui ajoutons les mots de l’original en italiques)

Références

Ameka F (1992) Interjections: the universal yet neglected part of speech. *Journal of Pragmatics* 18 (2): 101–118.

Aranguren M (2012) Sociologie du tuning. Réciprocité de l’ajustement et automatisme des choix techniques dans l’explication du succès des innovations. *Réseaux* (5): 231–263.

- Baudoin V, Velkovska J (2010) Dialogues vocaux entre clients et automates ou comment l'homme et la machine s'entendent dans la réalisation d'un service. In: *10èmes journées internationales d'analyse statistique de données textuelles*. Rome.
- Borzeix A, Fraenkel B(2001) *Langage et travail : communication, cognition, action*. Paris: CNRS Editions.
- Boutet J (2005) *Paroles Au Travail*. Paris, L'Harmattan.
- Chave F (2010) Tiers en urgences. Les interactions de secours, de l'appel au 18 à l'accueil en service d'urgences pédiatriques. Contribution à une sociologie du tiers. Doctorat en sociologie. Université Paris Ouest Nanterre la Défense.
- Cohen H, Giangola JP ,Balogh J (2004) *Voice User Interface Design*. Boston: Addison-Wesley Professional.
- Duffy E(1962) *Activation and behavior*. New York: Wiley.
- Dujarier M-A (2008) *Le travail du consommateur. De mcdonalds à ebay : comment nous coproduisons ce que nous achetons*. Paris: La Découverte.
- Ekman P, Friesen W (1975) *Unmasking the face. A guide to recognizing emotions from facial expressions*. Englewood Cliffs, NJ: Prentice-Hall.
- Frijda NH (1986) *The Emotions*. Paris: Editions de la Maison des Sciences de l'Homme, New York: Cambridge University Press.
- Gadrey J (1994) Les relations de service et l'analyse du travail des agents. *Sociologie du Travail* 36 (3): 381–389.
- Girin J., Grosjean M(1996) *La transgression des règles au travail*. Paris: Editions L'Harmattan.
- Goffman E(1967) *Interaction ritual: essays on face-to-face interaction*. New York: Anchor Books.
- Goffman E (1971) *Relations in public: microstudies of the social order*. New York: Basic Books.
- Grosjean M, Lacoste M (1999) *Communication et intelligence collective: le travail à l'hôpital*. Paris: Presses universitaires de France.
- Harris R A (2005) *Voice interaction design: crafting the new conversational speech systems*. Amsterdam: Elsevier.
- Hochschild A R (1979) Emotion work, feeling rules, and social structure. *American Journal of Sociology*: 551–575.
- Hochschild AR (1983) *The managed heart: commercialization of human feelings*. Berkeley: University of California Press.

Hochschild AR (2003) *The commercialization of intimate life: notes from home and work*.

Berkeley: University of California Press.

Laroche R, Putois G, Bretier P, Aranguren M, Velkovska J, Hastie H, Keizer S et al. (2011) CLASSiC. Computational Learning in Adaptive Systems for Spoken Conversation. Deliverable 6.4. Edinburgh: Heriot-Watt University (School of Mathematical and Computer Sciences)

Lazarus RS (1991) *Emotion and adaptation*. New York: Oxford University Press.

Lazarus R S (1966) *Psychological stress and the coping process*. New York: McGraw-Hill.

Magnusson MS (2000) Discovering hidden time patterns in behavior: t-patterns and their detection. *Behavior Research Methods* 32 (1): 93–110.

Max Planck Institute for Psycholinguistics Elan - Linguistic Annotator. Nijmegen. Version 4.2. [Désolé, l'année n'est pas disponible, seulement la version]

Mills P K, Morris JH (1986) Clients as 'partial' employees of service organizations: role development in client participation. *Academy of Management Review*: 726–735.

Morgan D P, Balentine B (1999) *How to build a speech recognition application: a style guide for telephony dialogues*. Enterprise Integration Group.

Noldus Information Technology. "Theme® Version 5". Wageningen.

Pène S., Borzeix A, Fraenkel B (2001) *Le langage dans les organisations*. Paris: L'Harmattan.

Quérel L (2012) Le travail des émotions dans l'expérience publique. Marées vertes en Bretagne. *Raisons Pratiques* 22: 135–162.

Russell JA (2003) Core affect and the psychological construction of emotion. *Psychological Review* 110 (1): 145.

Russell, JA, Bachorowski, JA, Fernández-Dols JM (2003) Facial and vocal expressions of emotion. *Annual Review of Psychology* 54 (1): 329–349.

Sacks H, Schegloff EA, Jefferson G (1974) A simplest systematics for the organization of turn-taking for conversation. *Language*: 696–735.

Sartre, J-P (1995) *Esquisse d'une théorie des émotions*. Paris: Hermann.

Schegloff EA (1968) Sequencing in conversational openings. *American Anthropologist* 70 (6): 1075–1095.

Scherer K R (2003) Vocal communication of emotion: a review of research paradigms. *Speech Communication* 40 (1): 227–256.

Suchman L (2006) *Human-Machine Reconfigurations: Plans and Situated Actions*.

Cambridge (New York USA): Cambridge University Press.

Tableau 3. Statistiques de base de la collection d'appels

#	identifiant	durée		issue
		mins	secs	
1	2h_europe	4	50	-
2	atoutplus240	2	20	resol
3	bebe	3	30	resol
4	comptabilite	4	10	-
5	delire1	3	6	resol
6	delire2	8	40	-
7	funtones	5	57	resol
8	hohoho1	5	22	-
9	hohoho2	2	36	-
10	illimite	4	2	resol
11	jusqu_a_demain	2	45	-
12	lignebloquee	7	54	-
13	maismadame	3	40	-
14	maparole	4	24	-
15	marre1	0	37	-
16	marre2	0	58	-
17	marre3	0	29	-
18	marre4	0	36	-
19	marre5	2	58	-
20	ou_allofact	1	6	resol
21	payer_ma_facture	1	34	-
22	petasse1	0	38	-
23	petasse2	0	36	-
24	petasse3	1	32	resol
25	premiernumero1	1	15	-
26	premiernumero2	1	7	-
27	premiernumero3	1	12	-
28	premiernumero4	1	1	resol
29	premiernumero_mob	2	20	resol
30	reglerfacture	1	33	resol

Lorsque l'identifiant se répète, les chiffres après la partie alphabétique désignent le placement dans une série d'appels par un même appelant. Le code ' resol ' indique que l'issue du dialogue a été une résolution, une réussite. Ensemble, les 30 appels de la collection s'élèvent à une durée de 83 minutes. Les 30 appels portent sur 19 transactions, dont 10 constituent des réussites.

