



**HAL**  
open science

## Rôle de la prosodie dans la perception du feedback ” oui ” dans un contexte médical

Camilla Mastriani, Caterina Petrone, Roxane Bertrand, Magalie Ochs

### ► To cite this version:

Camilla Mastriani, Caterina Petrone, Roxane Bertrand, Magalie Ochs. Rôle de la prosodie dans la perception du feedback ” oui ” dans un contexte médical. WACAI (Workshop sur les “Affects, Compagnons Artificiels et Interactions” (ACAI)), 2018, Porquerolles, France. halshs-01793222

**HAL Id: halshs-01793222**

**<https://shs.hal.science/halshs-01793222>**

Submitted on 16 May 2018

**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.

# Rôle de la prosodie dans la perception du feedback “oui” dans un contexte médical

Maria Camilla Mastriani<sup>1</sup>, Caterina Petrone<sup>2</sup>, Roxane Bertrand<sup>2</sup>, Magalie Ochs<sup>3</sup>

<sup>1</sup> Università degli Studi di Napoli «Federico II», Napoli, Italy

<sup>2</sup> Aix Marseille Univ, CNRS, LPL, Aix-en-Provence, France

<sup>3</sup> LIS UMR 7296 Aix-Marseille Université, CNRS, ENSAM, Université de Toulon, France

m.mastriani@studenti.unina.it, caterina.petrone@univ-amu.fr, roxane.bertrand@univ-amu.fr, magalie.ochs@lis-lab.fr

## ABSTRACT

Cet article porte sur l’item de feedback “oui” produit par un patient dans un environnement médical spécifique d’annonce de mauvaises nouvelles. Cette étude a pour objet de déterminer le rôle de la prosodie (intonation et délai temporel) dans la perception de ce feedback. 15 auditeurs français ont écouté de courtes interactions humain-humain impliquant deux personnes jouant le rôle du docteur et de son patient (ou d’un proche de ce dernier). Les auditeurs devaient juger du caractère plus ou moins approprié du feedback produit par le patient, en s’appuyant sur la manière dont il était produit oralement. Nous avons mesuré les effets de l’intonation (neutre/bouleversé/questionnant), du délai (court/long), et du genre des auditeurs (homme/femme) sur le score (jugé sur une échelle de 1 à 5) et le temps de réaction. Les résultats montrent le rôle de l’intonation et du délai dans une phase spécifique du dialogue (Définition du Problème) les hommes et les femmes étant sensibles à différents aspects de la prosodie. Ceci a des implications sur la modélisation de la prosodie des feedbacks d’un patient virtuel dans un contexte d’interaction homme-machine.

## Keywords

Prosodie, perception, contexte médical, feedback, fonction communicative, intonation, délai.

## 1. INTRODUCTION

L’étude en perception présentée dans ce papier s’inscrit dans un projet plus large, le projet Acorformed<sup>1</sup>, qui vise à développer un patient virtuel dans un environnement de réalité virtuelle en vue d’entraîner des médecins à l’annonce de mauvaises nouvelles. L’un des éléments-clés pour créer une interaction engageante est le comportement de feedbacks du patient virtuel. Un certain nombre de travaux de recherche ont démontré l’importance des feedbacks (appelés aussi *backchannels*) des agents conversationnels animés dans l’interaction (e.g. perception de l’agent, déroulé de la conversation, établissement d’un « rapport ») [1]. Plusieurs modèles computationnels ont été proposés pour prédire automatiquement le moment où un agent artificiel devrait produire un feedback (comme par exemple dans [2], [3]) en fonction des indices verbaux et non-verbaux exprimés par le locuteur principal.

Cependant, peu d’entre eux ont considéré le *niveau prosodique du feedback*. Un agent virtuel devrait être capable d’adapter la prosodie de son feedback en fonction de l’intention communicative qu’il souhaite exprimer. Le développement d’un tel agent requiert une meilleure connaissance du rôle de la prosodie dans la perception des feedbacks, aspect particulièrement important pour modéliser le patient dans le contexte d’annonce de mauvaises nouvelles. Afin d’atteindre cet objectif, nous avons conduit une première expérience dans le contexte médical spécifique de ce projet.

Les feedbacks linguistiques sont des mécanismes qui permettent aux participants à une conversation d’échanger de l’information relative à quatre fonctions communicatives de base : contact, perception, compréhension et réaction attitudinale [4]. Le feedback verbal est principalement produit via des énoncés courts tels que *ouais, oui, mh, mhm, ok*. Des études ont exploré les différentes caractéristiques des feedbacks en vue d’améliorer la désambiguïsation des fonctions discursive-pragmatiques, plus particulièrement à un niveau prosodique (pour une revue voir [5]). Plus spécifiquement, les *contours intonatifs* sont considérés comme un indice saillant pour exprimer différents types de fonctions discursives ([6; 7] parmi d’autres). Parmi les études montrant la collaboration active de l’interlocuteur, plusieurs travaux ont montré que le rôle de ce dernier dans le storytelling particulièrement était de fournir des *réponses feedback* appropriées, ceci dépendant de leur fonction *générique* ou *spécifique* (compréhension versus évaluation [8], ou encore les « continuers/assessments » chez [9]), mais aussi de leur localisation et du timing avec lequel ils étaient produits [10; 11]. Des travaux en Analyse Conversationnelle ont également montré l’importance du *délai* avec lequel un interlocuteur produit une réponse spécifique. En ligne avec le principe de l’organisation préférentielle (*preference organization*) [12], l’interlocuteur tendrait à produire une réponse préférée (attendue) plutôt qu’une réponse non préférée (inattendue). Ainsi, à une question, les participants attendent une réponse, à une offre, plutôt une acceptation (réponse préférée) qu’un refus (réponse non préférée). Différentes études ont montré que les réponses préférées seraient produites avec un délai plus

<sup>1</sup> <http://www.lpl-aix.fr/~acorformed/>

court que les réponses non préférées [13; 14]. Pourtant, très peu de travaux ont porté sur les caractéristiques prosodiques et le caractère approprié ou non des feedbacks. L'objectif à long-terme est de tester si la prise en compte des caractéristiques prosodiques des feedbacks permet d'améliorer l'interaction avec les ACA en contexte médical à travers l'intégration des résultats de l'étude au sein d'un patient virtuel. Préalablement, ce papier présente une analyse perceptive du feedback « oui » tel qu'il a été produit par des humains en interaction jouant le rôle d'un docteur et d'un patient. La question principale est d'évaluer perceptivement à quel point les paramètres prosodiques (intonation et délai temporel) du feedback permettent de le juger comme une réponse appropriée (ou non) à l'énoncé qui le précède. Nous nous attendons à ce que les « oui » moins marqués prosodiquement (produits avec une intonation neutre/questionnante ou après un court délai) soient plus compatibles avec des contextes favorisant un « oui » fonctionnant comme un simple continuer, tandis que ceux produits avec une intonation plus marquée (intonation bouleversée ou après un long délai) soient jugés comme plus appropriés dans un contexte impliquant un engagement émotionnel plus important. Les temps de réaction plus longs, supposés refléter une difficulté dans la tâche, sont attendus avec des items inappropriés. Enfin, étant donné que le genre des auditeurs peut moduler la perception de l'émotion [15], l'analyse préliminaire sera reportée en distinguant les résultats des hommes et des femmes.

## 2. ANALYSE DE CORPUS D'INTERACTION HUMAIN- MACHINE

Préalablement à l'expérience perceptive qui constitue l'aspect central de ce travail, nous avons exploré un corpus d'interactions entre des médecins annonçant des mauvaises nouvelles à des acteurs jouant le rôle de patients (appelés « patients standardisés ») lors de véritables formations dans des établissements médicaux français (il n'est pas possible pour des raisons éthiques d'enregistrer de vraies situations d'annonce de mauvaises nouvelles). L'utilisation de « patients standardisés » dans la formation médicale est une pratique courante. Les acteurs sont soigneusement formés (dans notre projet, les acteurs sont aussi des infirmiers) et suivent des scénarios préétablis définis par des experts pour jouer les réactions des patients les plus fréquemment observés. En moyenne, une consultation simulée dure 9 minutes. Le corpus collecté est composé de 13 vidéos d'interaction patient-médecin (chaque vidéo implique des médecins et/ou un couple acteur-patient différent) avec différents scénarios. Le corpus initial a été annoté de manière semi-automatique sur une durée totale de 119 minutes.

Dans ce corpus, nous avons, plus spécifiquement, analysé un ensemble de feedbacks produits par les patients en interaction avec les médecins. Fondée sur ces observations préliminaires, une expérience de perception a alors été conduite sur le feedback « oui » pour lequel ont été manipulés son intonation

et son délai temporel (Section 3). Une discussion générale et quelques perspectives sont enfin présentées (Sections 4 and 5).

### 2.1 Description du corpus

Seule une partie des sessions réelles d'entraînement collectées dans le projet Acorformed, à savoir six dialogues (sur les 13 existant jusqu'à maintenant), a été analysée (pour plus de détails concernant le corpus, voir [2]). Quel que soit le scénario (variable selon l'enregistrement et l'institution médicale), chaque dialogue a pu être divisé en 5 phases : une phase *d'ouverture*, dans laquelle le docteur accueille l'interlocuteur, une phase *d'annonce*, dans laquelle est expliquée la raison du rendez vous, une phase de *définition du problème* (désormais « PD »), dans laquelle est donnée l'information concernant la maladie du patient, une phase d'implications futures (désormais « FI ») phase, dans laquelle sont décrites les futures complications possibles, et une phase de *clôture* pour les séquences de salutations et de départ (cf. [16] pour l'organisation des phases en contexte médical).

### 2.2 Analyse et résultats

Fondés sur [17], huit types de feedback apparaissant dans les cinq phases ont été identifiés. Comme le montre la figure 1, le feedback « oui » est le seul apparaissant dans l'ensemble des phases. L'observation des occurrences de feedback montrent que les phases PD et FI concentrent le plus grand nombre de cas. Ainsi, nous avons centré notre étude sur ces deux seules phases. Un exemple d'énoncé produit par le docteur dans la phase PD est « *Il y a un obstacle dans son tube digestif* », qui décrit l'état de santé du patient. Un exemple d'énoncé dans la phase FI sur l'impact potentiel de la maladie est « *Il pourrait ne pas être conscient lorsqu'il se réveillera* ». Les deux énoncés sont suivis par un « oui » qui a été spontanément énoncé par le patient. Relativement aux délais de production des feedbacks après les énoncés produits par le médecin (à savoir le délai écoulé avant de produire un feedback), ils ont été mesurés dans les deux phases PD et FI. Ils vont d'un minimum de 0s à un maximum de 0.968s, avec une valeur médiane de 0.391s choisie comme frontière entre ce que nous avons appelé des « courts » et des « longs » délais (Figure 2). La distribution des délais courts et longs produits par le patient est significativement différente entre les deux phases considérées. Une majorité écrasante de délais courts en PD pourrait être le signe d'un large usage de « oui » fonctionnant comme un continuer dont le but est de montrer son attention au discours en cours et d'élucider plus d'informations chez le locuteur. Au contraire, pendant la phase FI, l'interlocuteur pourrait être plus impliqué émotionnellement et pourrait avoir besoin davantage d'énergie pour réaliser et intégrer la mauvaise nouvelle, ce qui pourrait expliquer le délai plus long avec lequel le « oui » est réalisé.

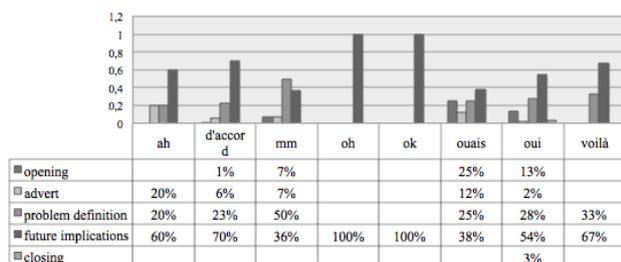


Figure 1: Distribution des Feedbacks dans les 5 phases

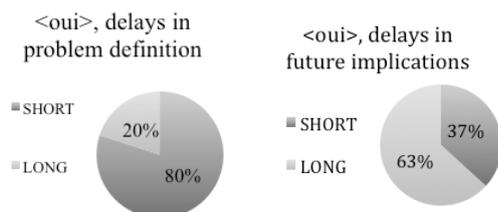


Figure 2: Long et court délais précédant les « oui » produits par le patient. Les résultats sont scindés en PD (à gauche) et FI (à droite).

### 3. ETUDE PERCEPTIVE

Fondée sur l'analyse préliminaire des feedbacks en corpus, les stimuli pour l'expérience de perception ont consisté en des énoncés naturels produits par deux locuteurs natifs du français. Une tâche de jugement a été conduite préalablement à l'expérience pour sélectionner les stimuli.

#### 3.1 Evaluation du matériel sélectionné

Les stimuli sont des paires d'énoncés reproduisant un énoncé « contexte » produit par le docteur suivi d'une réplique (feedback) par le patient. Les contextes consistaient en 20 énoncés naturels prononcés par un homme natif du français possédant un background en linguistique. Ce locuteur jouait le rôle du médecin. Le feedback « oui » était produit par une femme native du français avec un background en prosodie. Cette locutrice jouait le rôle de la patiente (ou d'un proche). Des détails relatifs au corpus sont présentés dans la section 3.1.1.

##### 3.1.1. Corpus

Les stimuli (énoncé contexte-feedback) ont été enregistrés dans la chambre sourde du *Laboratoire Parole et Langage* (LPL) au cours de deux sessions distinctes. Le locuteur homme a produit 20 énoncés relatifs au contexte dans le rôle du médecin donnant des mauvaises nouvelles. En particulier, huit de ces énoncés, les plus représentatifs sémantiquement et facilement isolables des dialogues, ont été extraits du corpus d'entraînement réel ; douze autres énoncés ont alors été créés en s'inspirant des huit premiers extraits du corpus. Sur les vingt énoncés contexte, la moitié présentait plus de probabilité d'apparaître dans la phase PD et l'autre moitié dans la phase FI.

Une phonéticienne française dotée d'un long entraînement en prosodie a produit le feedback « oui » en utilisant trois intonations différentes: *neutre*, *interrogative* et *bouleversée*. On lui a demandé d'imaginer qu'elle était un parent d'un patient hospitalisé recevant la mauvaise nouvelle de la part du médecin. Afin d'éliciter les différentes prosodies du feedback, elle a reçu les instructions suivantes : « montrez que vous accordez de l'attention à la parole produite » pour la production du « oui neutre » ; « vous ne montrez pas seulement de l'attention à la parole produite mais vous avez besoin de clarification ou de plus d'informations pour comprendre » pour le « oui interrogatif » ; « vous montrez non seulement que vous accordez de l'attention à la parole produite mais vous êtes bouleversée », pour le « oui bouleversé ». Des contextes visant à éliciter les différentes intonations ont été proposés (par exemple: « le docteur vous dit qu'il vous a lui-même contacté par téléphone » avec un « oui neutre » attendu ; « le docteur vous dit qu'il a rencontré des difficultés lors de l'intervention, pour un « oui interrogatif » ; « l'obstruction de l'estomac... c'est un cancer » pour un « oui bouleversé »). Au total, trois versions pour chacun des 3 « oui » différents ont été produites (au total: 9 stimuli de *oui*), classés « N1 », « N2 », « N3 » (*neutre*), « Q1 », « Q2 », « Q3 » (*questionnant*), « S1 », « S2 », « S3 » (*bouleversé*). Nous avons seulement besoin d'un « oui » de chaque (au total: 3 stimuli *oui*, le plus identifiable en termes d'intonation) pour construire notre corpus, ce qui nous a conduit à effectuer une tâche de jugement.

Etant donné les délais avec lesquels « oui » étaient produits après le tour précédent, nous avons constaté qu'utiliser la frontière détectée dans le corpus réel pour les délais courts (à savoir 0 s) aurait paru un peu étrange étant donné que ce délai dans le corpus de perception aurait paru hors contexte pour un dialogue et plutôt associé à un stimulus isolé. Afin de conserver un délai crédible pour un stimulus isolé, nous avons décidé dans l'expérience principale d'augmenter le temps à 0.3 s (ce qui reste malgré tout une valeur inférieure à la valeur médiane de 0.391 s et respecte ainsi les valeurs analysées dans les dialogues réels). Le timing des délais longs a été arrondi à 1s. Nous avons enregistré avec PRAAT deux « silences » de 0.3s et de 1s: respectivement le délai court et le délai long.

##### 3.1.2. Participants et procédure

Huit locuteurs natifs du français, 3 hommes et 5 femmes, tous des étudiants âgés de 22 à 25 ans, ont écouté les différents échantillons isolés (i.e. en dehors du contexte) de « oui » ; le total des 9 stimuli a été joué dans le logiciel PERCEVAL [18]. Les sujets devaient juger la prosodie du feedback « oui » et choisir entre « 1: neutre », « 2: questionnant », « 3: bouleversé ».

### 3.1.3. Résultats

Pour le «oui neutre», N3 a été correctement jugé comme neutre par 6 sujets sur 2, tandis que N1 était reconnu par 2 seulement et N2 par la moitié d'entre eux. Le oui questionnant a été généralement le mieux reconnu : Q2 a été jugé correctement par 7 sujets sur 8, et Q1 et Q3 ont été jugés correctement par l'ensemble des participants. Nous avons choisi Q3 pour notre corpus. Similairement, l'intonation du «oui bouleversé» hors contexte a été également bien détectée : S1 et S2 ont été reconnus comme bouleversé par 6 juges sur 8 et S3 par l'ensemble des juges. N3, Q3 et S3 ont été utilisés pour construire notre corpus dans l'expérience principale.

## 3.2. Expérience principale

Finalement, notre corpus est constitué d'ensembles de 20 énoncés produits par un médecin donnant des mauvaises nouvelles, également répartis en phase de *définition de problème* et de  *futures implications*, soit issus soit créés sur le modèle du corpus; 3 stimuli «oui» correctement évalués et distincts en termes intonatifs ; 2 délais, court de 0.3s et long de 1s.

Nous avons combiné, respectivement, chaque ensemble avec les deux délais et les trois types d'intonation du «oui». Le corpus est finalement composé de  $20 \times 6 = 120$  stimuli.

### 3.2.1. Participants et procédure

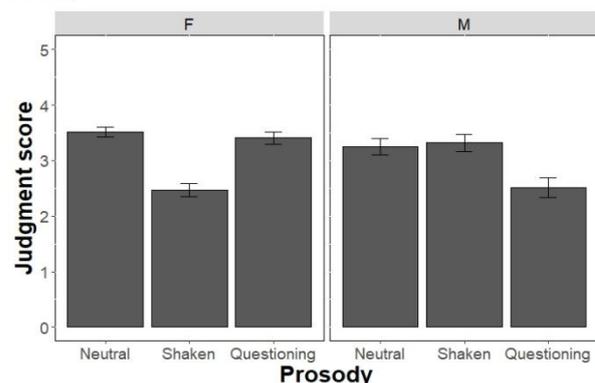
Quinze locuteurs natifs du français, 5 hommes et 10 femmes, âgés de 18 à 30 ans, ont participé à l'expérience de perception. La tâche consistait à écouter les 120 stimuli produits dans un ordinateur portable via le logiciel PERCEVAL : chaque set (20) était entendu, dans un ordre aléatoire, 6 fois, chacun étant suivi par l'un ou l'autre des deux délais et l'un des trois «oui» différents. La tâche consistait à juger de *caractère approprié* du feedback sur une échelle allant de 1 («pas du tout approprié») à 5 («complètement approprié») basé sur ce qui était dit oralement. Le jugement était exprimé sur une boîte à 5 boutons.

### 3.2.2. Résultats

Une série de modèles à effets linéaires mixtes avec une structure aléatoire maximale a été menée séparément pour les deux phases de *définition de problème* et de  *futures implications*, dans lesquelles nous avons testé les effets des facteurs fixes DELAI (court/long), de INTONATION (neutre/bouleversé/interrogatif) et de SEXE des auditeurs (homme/femme) sur le score (jugé sur une échelle de 1 à 5) et le log du temps de réaction. Pour le facteur INTONATION, on utilise un contraste de type 'traitement' avec «neutre» en niveau de référence. AUDITEURS et ITEMS ont été inclus comme intercept aléatoire avec des pentes aléatoires pour chaque facteur fixe. Seuls les effets significatifs sont reportés ci-dessous.

Pour la phase définition du problème, nous observons une interaction significative sur les scores entre le facteur INTONATION et SEXE. Plus précisément, les femmes ont évalué le «oui bouleversé» plus bas que le «oui neutre» [ $\beta = -0.97$ ,  $SE = 0.37$ ,  $t = -2.5$ ,  $p = .01$ ] tandis qu'il n'y a pas de différence entre «neutre» et «questionnant». D'autre part, les hommes ont évalué le «oui questionnant» plus bas que le «neutre» [ $\beta = -1.004$ ,  $SE = 0.40$ ,  $t = -2.45$ ,  $p = .02$ ]; il n'y a pas de différence entre le «oui neutre» et «bouleversé». Le facteur DELAI et ses interactions ne sont pas significatifs. Figure 3 (en haut) montre le jugement moyen des scores par prosodie, séparé par le facteur sexe (les données étant réunies pour le facteur délai). Les valeurs de scores se situent entre 2.47 (pour le «oui bouleversé») et 3.51 (pour le «oui neutre») pour les femmes, et entre 2.51 (pour le «oui questionnant») à 3.31 (pour le «oui bouleversé») pour les hommes. Similairement nous observons pour les temps de réaction (RT), une triple interaction significative INTONATION x DELAI x SEXE [ $\beta = -0.54$ ,  $SE = 0.25$ ,  $t = -2.13$ ,  $p = .03$ ]. Alors qu'il n'y a pas de différence dans les RT pour les femmes sur les intonations et les délais, les hommes en revanche prennent plus de temps pour juger le «oui bouleversé» lorsqu'il est présenté après un long délai plutôt qu'après un court. Ceci est illustré en figure 3 (en bas).

Pour la phase de « futures implications », nous avons trouvé un effet significatif du délai pour le «oui neutre» [ $\beta = 0.24$ ,  $SE = 0.12$ ,  $t = 2.02$ ,  $p = .047$ ], à savoir le score moyen était légèrement plus haut après un délai court (score moyen = 3.25) qu'après un long (score moyen = 3.05). Aucun autre effet n'est significatif. De la même manière pour les temps de réaction, il n'y a pas de différence entre les facteurs étudiés.



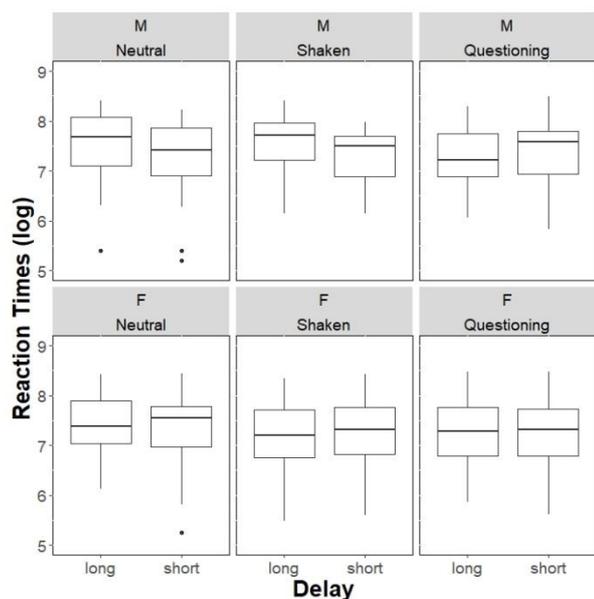


Figure 3. Scores (en haut) et temps de réaction (en bas) pour la phase « définition du problème »

La figure 3 illustre les résultats des scores moyens de jugement pour la phase PD par intonation et par sexe (quelque soit le délai de feedback).

#### 4. DISCUSSION

Cette étude a exploré le rôle des caractéristiques prosodiques dans la perception du feedback linguistique « oui » dans une interaction en contexte médical impliquant un patient et un médecin. Pour le patient (et futur agent virtuel), fournir des réponses appropriées vise l'amélioration des performances des médecins s'entraînant à annoncer une mauvaise nouvelle.

Avant de conduire la principale étude perceptive, une analyse de corpus a été menée nous permettant de définir comme réponse appropriée différents types de feedback produits par le patient. Nous avons considéré l'item « oui » en raison du fait qu'il était non seulement un des plus fréquents feedbacks de notre corpus de dialogue (juste après "ouais") mais parce qu'il était aussi très fréquent dans les deux phases examinées ici (définition du problème -PD- ou futures implications -FI).

Les moyennes pour les scores de jugement se situent dans la gamme moyenne de l'échelle de jugement, ce qui suggère que les auditeurs ont interprété les « oui » comme modérément appropriés au contexte de l'énoncé précédent à la fois pour PD et FI. Cependant, les résultats pour l'intonation modulent le score de jugement selon le genre des auditeurs: pour les femmes, le « oui bouleversé » est jugé comme l'item le moins approprié tandis que c'est le « oui questionnant » pour les hommes. Si le résultat pour les femmes confirme nos attentes, à savoir l'item le plus approprié en PD étant un oui neutre ou questionnant, le résultat concernant les hommes est plus inattendu. En effet, la phase PD non seulement fournit l'opportunité de produire des

feedback à fonction de « continuer » (marque explicite d'écoute et de compréhension permettant aux locuteurs principaux de poursuivre le discours en cours) mais aussi l'item questionnant qui peut signaler un trouble ponctuel ou une incompréhension. Il y a donc plus de probabilité que cette phase PD soit ponctuée par un « oui » neutre ou questionnant qu'un « oui » bouleversé. Ceci est confirmé pour les femmes mais pas pour les hommes.

Par ailleurs, les résultats relatifs aux temps de réaction pour les hommes montrent que ces derniers, s'ils ne considèrent pas le « oui » bouleversé comme un item inapproprié, présentent un temps de réaction plus grand lorsqu'ils doivent juger le « oui » bouleversé après un long délai. Si nous rappelons que la phase PD tendrait plutôt à favoriser un délai court, ce résultat suggère que les hommes réagissent plus au délai qu'à l'intonation.

Enfin, la phase FI ne présente pas de résultats significatifs, excepté un très petit effet sur le délai. Alors que nous attendrions plutôt des items bouleversés après un long délai comme réponse préférée dans cette phase durant laquelle le patient pourrait exprimer son trouble sur les futures implications liées à l'annonce des mauvaises nouvelles, aucun des items n'est vraiment distinguable. Ceci pourrait s'expliquer en partie par l'étude de corpus elle-même dans laquelle la distribution des items « oui » en FI apparaît comme moins tranchée qu'en DP (plus grande acceptabilité pour FI) ceci étant possiblement lié au manque de données. Des investigations futures impliquant plus de données ainsi que plus de participants pourraient combler ce manque et permettre d'affiner la différence relative au genre qui a émergé dans ce travail.

#### 5. PERSPECTIVES

Dans le contexte du projet Acorformed, le feedback du patient virtuel est d'une grande importance étant donné son rôle principal de « celui qui écoute », ici une mauvaise nouvelle. Afin d'améliorer l'interaction humain-machine, et en particulier la perception du patient virtuel et l'engagement de l'utilisateur, la prochaine étape consiste à utiliser les résultats présentés dans cette étude pour modéliser les caractéristiques prosodiques des feedbacks du patient virtuel. L'implémentation des caractéristiques prosodiques dans le patient virtuel nous permettra d'explorer, à travers une étude perceptive dans le contexte d'interaction humain-machine, l'effet de l'apparence virtuelle du patient sur la perception de l'utilisateur comparé aux résultats obtenus dans cette étude perceptive humain-humain.

#### 6. REMERCIEMENTS

Ce travail de recherche est financé par l'ANR (ACORFORMED ANR-14-CE24-0034-02) et soutenu par

## 7. REFERENCES

- [1] J. Gratch, A. Okhmatovskaia, F. Lamothe, S. Marsella, M. Morales, R.J. van der Werf, & L.P. Morency, "Virtual rapport", *International Workshop on Intelligent Virtual Agents*, 14-27, Springer Berlin Heidelberg. 2006.
- [2] C. Porhet, M. Ochs, J. Saubesty, G. de Montcheuil, R. Bertrand, "Mining a Multimodal Corpus of Doctor's Training for Virtual Patient's Feedbacks", *International Conference on Multimodal Interaction (ICMI)*. 2017.
- [3] R. Poppe, K.P. Truong, D. Reidsma, D. Heylen, "Backchannel strategies for artificial listeners", *International Conference on Intelligent Virtual Agents*, 146-158, Springer Berlin Heidelberg. 2010.
- [4] Allwood, J., J. Nivre, and E. Ahlsen, "On the semantics and pragmatics of linguistic feedback", *Journal of Semantics*, 9 (1):1-30. 1992.
- [5] A. Gravano, J. Hirschberg, S. Benus, "Affirmative cue words in task-oriented dialogue". *Computational Linguistics*. Vol. 38 (1), 1-39, 2012.
- [6] Hockey, B. A., "Prosody and the role of 'okay' and 'uh-huh' in discourse", *Proceedings of the Eastern States Conference on Linguistics*, 128-136, Columbus, OH. 1993.
- [7] J. Kowtko, *The Function of Intonation in Task-Oriented Dialogue*. Ph.D. thesis, University of Edinburgh. 1996.
- [8] J. B. Bavelas, L. Coates, T. Johnson, "Listeners as co-narrators", *J. Pers. Soc. Psychol.* 79, 941--952. 2000. <http://dx.doi.org/10.1037/0022-3514.79.6.941>
- [9] E. A. Schegloff, "Discourse as an interactional achievement: some uses of "uhhuh" and other things that come between sentences", *Analyzing Discourse: Text and Talk*. Georgetown University Press, Washington, DC, pp. 71--93. 1982
- [10] R. Bertrand, R. Espesser, "Storyteller and listener's behavior in French conversational storytelling", *Journal of Pragmatics* 111, 33-53, 2017.
- [11] R. Bertrand, B. Priego-Valverde, "Listing practice in French conversation. From collaborative achievement to interactional convergence", *Discours* (online), 2017.
- [12] A. Pomerantz and J. Heritage, "Preference", In J. Sidnell & T. Stivers (eds), *The Handbook of Conversation Analysis*, Blackwell Publishing Ltd, 210-228. 2013.
- [13] Stivers, T., Enfield, N.J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoyman, G., Rossano, F., de Ruiter, J.P. Yoon, K. & S.C. Levinson, "Universals and cultural variation in turn-taking in conversation", *Proceedings of the National Academy of Sciences*, 106(26), 10587-92. 2009.
- [14] S.C. Levinson and F. Torreira, "Timing in turn-taking and its implications for processing models of language", In J. Holler, K.H. Kendrick, M. Casillas, M., S.C. Levinson eds. *Turn-Taking in Human Communicative Interaction*. Lausanne: Frontiers Media. doi: 10.3389/978-2-88919-825-2. 2016.
- [15] S. Verheul, A. Hartmann, R. Supheert, A. Chen, "Gender differences in the perception of speech prosody and visual prosody in L2". 2016.
- [16] J. D. Robinson, "An interactional structure of medical activities during acute visits and its implications for patients' participation", *Health Communication*, 15, 27-57, 2003.
- [17] L. Prevot, B. Bigi, and R. Bertrand, "A quantitative view of feedback lexical markers in conversational French", *14th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, 1-4, 2013.
- [18] C. André, A. Ghio, C. Cavé, B. Teston, "Perceval: a Computer-Driven System for Experimentation on Auditory and Visual Perception", *Proceedings of XVth ICPhS*, Barcelone, 1421-1424, 2003.