



HAL
open science

Expressions référentielles et formes prosodiques : quand et comment la prise en compte de notre interlocuteur guide-t-elle nos choix linguistiques ?

Maud Champagne-Lavau, Noémie Moreau, Lola Rivoal, Amandine Michelas

► To cite this version:

Maud Champagne-Lavau, Noémie Moreau, Lola Rivoal, Amandine Michelas. Expressions référentielles et formes prosodiques : quand et comment la prise en compte de notre interlocuteur guide-t-elle nos choix linguistiques ?. Travaux Interdisciplinaires sur la Parole et le Langage, 2022, Panorama des recherches au Laboratoire Parole et Langage, 38, 10.4000/tipa.4935 . hal-03894933v2

HAL Id: hal-03894933

<https://hal.science/hal-03894933v2>

Submitted on 3 Feb 2023

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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License



TIPA. Travaux interdisciplinaires sur la parole et le langage

38 | 2022

Numéro spécial : Panorama des recherches
au Laboratoire Parole et Langage

Expressions référentielles et formes prosodiques : quand et comment la prise en compte de notre interlocuteur guide-t-elle nos choix linguistiques ?

*Referential expressions and prosodic forms: when and how the addressee's
perspective*

Maud Champagne-Lavau, Noémie Moreau, Lola Rivoal et Amandine
Michelas



Édition électronique

URL : <https://journals.openedition.org/tipa/4935>

DOI : 10.4000/tipa.4935

ISSN : 2264-7082

Éditeur

Laboratoire Parole et Langage

Référence électronique

Maud Champagne-Lavau, Noémie Moreau, Lola Rivoal et Amandine Michelas, « Expressions référentielles et formes prosodiques : quand et comment la prise en compte de notre interlocuteur guide-t-elle nos choix linguistiques ? », *TIPA. Travaux interdisciplinaires sur la parole et le langage* [En ligne], 38 | 2022, mis en ligne le 27 janvier 2023, consulté le 29 janvier 2023. URL : <http://journals.openedition.org/tipa/4935> ; DOI : <https://doi.org/10.4000/tipa.4935>

Ce document a été généré automatiquement le 29 janvier 2023.



Creative Commons - Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International
- CC BY-NC-ND 4.0

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Expressions référentielles et formes prosodiques : quand et comment la prise en compte de notre interlocuteur guide-t-elle nos choix linguistiques ?

Referential expressions and prosodic forms: when and how the addressee's perspective

Maud Champagne-Lavau, Noémie Moreau, Lola Rivoal et Amandine Michelas

Introduction

- 1 S'il est clair que locuteurs et auditeurs adaptent leur comportement linguistique à leur interlocuteur pour éviter et résoudre les situations d'incompréhension, la question de savoir jusqu'à quel point et dans quelle situation fait actuellement débat. Autrement dit, quand et comment prenons-nous en compte la perspective de notre interlocuteur lorsque nous choisissons des formes linguistiques particulières (e.g., expressions référentielles, formes prosodiques)? Comment l'auditeur exploite-t-il nos choix linguistiques? L'objectif de cet article est de situer les travaux que nous réalisons au LPL dans le cadre de ce questionnement. Il s'agira, d'une part, de rappeler les résultats de la littérature et, d'autre part, de montrer l'originalité de notre contribution à ce débat, à savoir : a) l'étude des choix prosodiques du locuteur en conversation en fonction de la présence d'un interlocuteur et des connaissances de cet interlocuteur, sachant que la plupart des travaux ayant contribué à ce débat ont porté sur l'utilisation de la référence ; b) l'étude de ces phénomènes via une approche clinique. L'utilisation d'une telle approche chez des individus atteints de schizophrénie, susceptibles de présenter des difficultés à attribuer des connaissances, croyances aux autres, permet

d'identifier quand et comment les choix linguistiques réalisés en conversation sont faits en tenant compte de la perspective de l'interlocuteur.

Concepts d'*audience design* et de savoir partagé

- 2 Le processus visant à prendre en compte ou à s'adapter aux besoins et connaissances de notre interlocuteur dans une conversation est appelé *audience design* (Clark & Murphy, 1982 ; Clark, 1996 ; Galati & Brennan, 2010). On parle aussi de *prise de perspective* dans la mesure où l'interlocuteur est amené à prendre en compte la perspective, le point de vue de son interlocuteur quand il formule ses énoncés (Brown-Schmidt & Heller, 2018). Comprendre quand et comment les interlocuteurs utilisent l'*audience design* revient à comprendre comment les locuteurs prennent en compte des informations qu'ils croient partager avec leur interlocuteur, autrement dit leur *common ground* ou savoir partagé. L'*audience design* implique, en effet, la distinction entre ce qui fait partie du *common ground*, c'est-à-dire les informations partagées avec un interlocuteur particulier et mutuellement reconnues comme telles (Clark & Marshall 1981 ; Clark, 1996 ; Stalnaker, 1978) et ce qui fait partie du savoir privé (i.e., *privileged ground*) propre à chaque locuteur et non connu de son interlocuteur. Le savoir partagé est mis à jour au fur et à mesure que la conversation progresse. Il comprend les connaissances partagées par le locuteur et l'auditeur et reconnues par les deux interlocuteurs pour être connues par chacun d'eux (Clark & Marshall, 1981). Ces connaissances incluent des informations sur le contexte physique actuel, des informations liées aux précédents échanges conversationnels, des connaissances encyclopédiques telles que des connaissances partagées par une communauté spécifique (e.g., les habitants d'Aix-en-Provence). Ainsi, les informations partagées ou non partagées par les interlocuteurs à un moment de la conversation détermineraient le type d'expressions utilisées par le locuteur pour son auditeur. Plusieurs études ont ainsi montré qu'après avoir développé des expressions référentielles pour désigner une figure (e.g., *Le Monsieur avec une canne*) avec un destinataire donné, le locuteur tend à ne plus utiliser ces expressions référentielles dans les conversations suivantes avec un nouveau destinataire (Wilkes-Gibbs & Clark, 1992).

Paradigmes expérimentaux classiquement utilisés pour étudier l'*audience design*

Le paradigme de communication référentielle

- 3 La plupart des travaux menés sur l'*audience design* ont porté sur les formes référentielles utilisées par les locuteurs dans des conditions où ils réfèrent au même objet à différents moments ou avec des destinataires différents, mais aussi à l'interprétation en temps réel des formes référentielles par les auditeurs telle que mesurée par le suivi des mouvements oculaires (e.g., Brown-Schmidt, 2009 ; Keysar *et al.*, 2000). Le *paradigme de communication référentielle* a été utilisé dans la plupart des études ayant contribué à ce débat (Krauss & Weinheimer, 1966 ; Clark & Wilkes-Gibbs, 1986). Ce paradigme permet de reproduire une situation de communication afin d'évaluer comment les interlocuteurs prennent en compte ou échouent à prendre en compte le savoir partagé, qu'ils marquent linguistiquement par une utilisation

différenciée des marqueurs de référence (indéfinis, définis, démonstratifs, pronoms). Ce paradigme repose sur la description de séries d'items présentés visuellement (e.g., tangrams issus d'un jeu chinois). Il implique des interactions sociales basées sur la collaboration puisqu'un participant jouant le rôle de directeur doit aider un autre participant ou un confédéré jouant le rôle de destinataire à ordonner les tangrams selon un ordre prédéterminé. Le directeur doit donc produire des informations verbales discriminantes permettant au destinataire d'identifier chacun des tangrams et de les placer un à un dans le bon ordre. Lorsque la tâche est répétée plusieurs fois (toujours avec les mêmes stimuli présentés dans des ordres différents), au fur et à mesure de la répétition des essais, une évolution du processus de négociation référentielle doit être observée, allant dans le sens d'une amélioration significative de l'efficacité de la communication. La tâche est réalisée avec un écran opaque entre le directeur et le destinataire pour éviter toute communication par les gestes. Dans ce type de tâche, le savoir partagé est constitué non seulement par le contexte physique mais aussi par la suite des échanges qui ont lieu entre les deux interlocuteurs. A la fin de chaque essai, le directeur et le destinataire vérifient que l'ordre trouvé par le destinataire est correct. En accord avec le modèle de référence collaboratif, que nous allons décrire ci-dessous (Clark & Wilkes-Gibbs, 1986), quand des locuteurs essaient d'identifier un référent comme dans le cas d'une conversation par exemple, ils commencent avec plusieurs descriptions (indéfinies) et proposent ce qu'ils pensent être une expression adéquate permettant de reconnaître ce référent. Le locuteur assume que les connaissances qu'il utilise – l'arrière-plan conversationnel – sont partagées par l'auditeur et vont permettre une formulation adéquate. L'avantage des tangrams est que ce sont des figures non nommables a priori. Ainsi au début de la tâche, directeur et destinataire n'ont aucun savoir partagé à propos des figures. La répétition du jeu permet d'évaluer la construction du savoir partagé entre les deux interlocuteurs. Plusieurs mesures permettant d'évaluer la présence ou pas d'*audience design* peuvent être prises en compte au cours de cette tâche (Clark & Wilkes-Gibbs, 1992) : nombre de mots utilisés, nombre de tours de parole entre directeur et destinataire, nombre de marqueurs indéfinis, nombre de marqueurs définis, nombre de demandes de clarification du destinataire, nombre de réponses du directeur à ces demandes de clarification, temps de réalisation de chaque essai. Ce paradigme reposant sur la description de séries d'images a été adapté en fonction du type de problématique.

Les *Map tasks* et les mesures prosodiques

- 4 Contrairement au nombre important d'études qui existent sur la référence, seul un petit nombre d'études s'est attaché à décrire comment le savoir partagé avec l'interlocuteur influence les choix prosodiques réalisés par le locuteur. Le terme prosodie fait références aux variations mélodiques et rythmiques qui accompagnent les mots et qui sont principalement repérables dans le signal de parole grâce à l'étude de trois paramètres acoustiques : la durée de formes linguistiques ainsi que les variations de fréquence fondamentale et d'intensité qui sont associées à ces formes [cf. article "La prosodie au Laboratoire Parole et Langage : histoire, recherches actuelles et perspectives"⁴ dans ce numéro]. La plupart des études qui se sont intéressées à ces variations prosodiques ont tenté de comprendre comment la durée de ces formes pouvait nous renseigner sur la mise en jeu de l'*audience design* (voir par exemple Arnold, 2012 ; Bard *et al.*, 2000 ; Bard & Aylett, 2005 ; Galati & Brennan, 2010 ; Rosa *et al.*,

2015). Ces études ont exploité le fait que lorsqu'un locuteur introduit un référent pour la première fois, le mot qui représente ce référent est prononcé clairement, avec une forme proéminente, alors que ce même mot sera produit avec une forme réduite (c'est-à-dire avec une forme dont la durée sera plus courte) lorsqu'il sera répété une seconde fois. Ces études ont cherché à déterminer si ces phénomènes de réduction de la durée des mots résulteraient de processus égocentrés ou s'ils seraient plutôt le reflet de l'*audience design*. Afin d'étudier cela, certaines de ces études ont eu recours à l'utilisation de *Map tasks* (Anderson *et al.* 1991). Les *Map tasks* sont des tâches d'élicitation de parole au cours desquelles deux locuteurs doivent interagir afin de déterminer un itinéraire défini en termes de repères qui sont étiquetés sur des cartes. Dans ce type de protocole expérimental, les deux locuteurs sont munis d'une carte assez schématique. Le directeur (désigné parmi les deux locuteurs) doit guider le destinataire d'un point à l'autre de la carte. Ce protocole comporte plusieurs avantages. Il permet tout d'abord de déclencher de la parole spontanée tout en s'assurant que certains phénomènes linguistiques seront effectivement produits par les locuteurs grâce aux éléments qui sont donnés sur la carte. Il permet également de manipuler le savoir partagé par les deux interactants au moyen de la manipulation des points de repère qui peuvent être soit identiques soit différents sur la carte du directeur et sur celle du destinataire. Par exemple dans la *Map task* utilisée par Bard *et al.* (2000), seule la moitié des points de repère étaient similaires d'une carte à l'autre et les locuteurs savaient à l'avance que leurs cartes ne correspondaient pas exactement. D'autres études ont proposé d'étudier ces phénomènes de réduction acoustique au moyen de tâches de communication référentielle au cours desquelles le participant donne des instructions à un confédéré à propos d'objets à déplacer sur une table (ex. La théière va sur le cercle rouge ; cf. Arnold *et al.* 2012 ; Rosa *et al.*, 2015). En fonction de la problématique de l'étude, le paradigme est adapté et la manipulation d'intérêt peut varier. Par exemple, dans l'étude d'Arnold *et al.* (2012), la manipulation d'intérêt concerne l'attitude du destinataire (qui était un confédéré) et qui révélait si oui ou non ce confédéré avait un savoir sur la prédictibilité de l'objet cible (ici l'objet théière). En effet, ce destinataire anticipait soit l'objet, en le ramassant avant l'instruction donnée par le locuteur, soit attendait l'instruction du locuteur. Une autre version de cette tâche consiste à étudier les formes produites par le directeur tandis que le destinataire est, soit pleinement attentif, soit distrait par une autre tâche (Rosa *et al.*, 2015). Les résultats issus de ces paradigmes expérimentaux ont contribué à l'élaboration des modèles psycholinguistiques décrits dans la section suivante.

Partie 1. Cadre théorique du débat

- 5 Plusieurs modèles d'*audience design* ont été proposés pour rendre compte du fait que les interlocuteurs ne feraient pas des inférences sur les connaissances de leur destinataire de manière continue. L'hypothèse d'un recours plus ou moins important au savoir partagé (*common ground*) dans la mise en jeu de l'*audience design* est le point de divergence entre les différents modèles existants (voir Barr & Keysar, 2006 ; Brown-Schmidt & Heller, 2018 ; Galati & Brennan, 2021 ; Knutsen & Le Bigot, 2013 pour une revue détaillée des modèles). Alors que le modèle collaboratif (Clark, 1996) considère la conversation comme une contribution constante et conjointe de l'auditeur et du locuteur au savoir partagé, d'autres modèles (Keysar *et al.*, 2000) insistent sur le comportement plus ou moins égocentrique des interlocuteurs dans la prise en compte

des besoins du destinataire et sur la mise en jeu de mécanismes automatiques peu coûteux. Nous décrivons dans cette section les principaux modèles (i.e., modèle égocentrique, modèle d'alignement interactif, modèle de traitement double, modèle probabiliste, modèle basée sur la mémoire) contribuant à ce débat sur la mise en jeu de l'*audience design*.

- 6 Selon le **modèle collaboratif** (Clark, 1996), la conversation est considérée comme une contribution constante et conjointe de l'auditeur et du locuteur au savoir partagé (*common ground*). La construction du savoir partagé est réalisée par une succession de phases de présentation et de phases d'acceptation. Ainsi, chaque interlocuteur participe à la mise à jour du savoir partagé en essayant d'établir que l'autre a suffisamment compris ses énoncés. Ceci implique un acte de participation du locuteur et de l'auditeur en termes de retour donné par l'auditeur au locuteur. Ce retour montre au locuteur que l'auditeur a bien compris. Les énoncés acceptés sont ainsi considérés comme mutuellement compris par les deux interlocuteurs et intégrés dans leur savoir partagé (processus de *grounding*). Ainsi le locuteur initie le processus mais compte aussi sur la participation active de l'auditeur. C'est un acte bilatéral nécessitant l'action à la fois du locuteur et de l'auditeur. Par conséquent, une communication réussie dépend de la construction de ce savoir partagé par les deux interlocuteurs, impliquant la prise en compte des besoins de l'interlocuteur de façon constante, ce qui requiert l'attribution d'états mentaux (e.g., croyances, connaissances) à l'interlocuteur. Le savoir partagé joue un rôle central dans le cadre de ce modèle, puisqu'il permet au locuteur des références adaptées au destinataire.
- 7 Un modèle **alternatif**, dit « **égocentrique** », issu des travaux de Keysar et collaborateurs (Barr, 2008 ; Barr & Keysar, 2002 ; Horton & Keysar, 1996 ; Keysar *et al.*, 2000) considère que la prise en compte de l'interlocuteur via le *common ground* nécessiterait une représentation distincte et coûteuse cognitivement des connaissances, croyances et intentions de l'interlocuteur. Les coûts de traitement d'un tel système seraient trop élevés pour permettre une conversation fluide. Ainsi, le locuteur ne prendrait en compte le savoir partagé que dans les étapes tardives de la production du langage afin de corriger des énoncés planifiés de manière égocentrique. Keysar et collaborateurs ont développé plusieurs travaux en psychologie cognitive portant sur la résolution de référence par un destinataire dans le cadre du paradigme de communication référentielle. Dans une étude impliquant la mesure des mouvements oculaires du destinataire, ils ont montré que dans les phases initiales de compréhension, les destinataires adoptent une perspective égocentrique en s'appuyant sur leur savoir privé (i.e., *privileged ground*) plutôt que sur le savoir partagé (Keysar *et al.*, 2000). Un tel résultat indique que la capacité des auditeurs à utiliser le *common ground* serait limitée. Ainsi, selon cette perspective égocentrique, les informations les plus saillantes ou accessibles (incluant le savoir privé) seraient utilisées de manière immédiate, la compréhension (tout comme la production) étant ensuite ajustée dans des étapes plus tardives afin de prendre en compte le savoir partagé. Par exemple, Horton & Keysar (1996) ont montré que des informations contextuelles visuellement saillantes sont intégrées de façon précoce dans la planification d'un énoncé même si ces informations ne sont pas pertinentes pour le destinataire.
- 8 Dans la lignée du modèle de Keysar et collaborateurs, le **modèle d'alignement interactif** de Pickering & Garrod (2004, 2013) propose qu'au cours de la conversation il y ait un alignement des modèles de situation des interlocuteurs. Cet alignement des

modèles de situation forme la base d'un dialogue réussi. Cet alignement ou coordination entre les interlocuteurs est un mécanisme automatique guidé par un amorçage à tous les niveaux de représentation linguistique (i.e., phonétique, phonologique, lexical, syntaxique et sémantique) et par un phénomène de « percolation » entre ces niveaux. Autrement dit, l'alignement à un niveau entraîne l'alignement à d'autres niveaux. Un tel mécanisme serait automatique et aurait un faible coût cognitif. Ce processus d'alignement ne nécessiterait pas de prendre en compte les croyances et connaissances de l'interlocuteur en intégrant de façon continue les représentations du savoir partagé parce que chaque individu peut supposer implicitement que le modèle de sa situation sera suffisamment proche de la conversation de son interlocuteur (Pickering & Garrod, 2006). Les interlocuteurs seraient coordonnés via ce processus d'alignement. Ce processus d'alignement est notamment visible au niveau syntaxique, quand des interlocuteurs deviennent plus susceptibles de réutiliser des structures précédemment produites ou précédemment entendues. Les effets liés aux pactes conceptuels (i.e., accord tacite entre deux interlocuteurs concernant la conceptualisation d'un objet, d'une image, reflétée dans l'utilisation d'une expression référentielle déjà utilisée avec un interlocuteur donné ; Brennan & Clark, 1996) au cours de la communication référentielle peuvent aussi être expliqués par ce processus d'alignement. Ainsi, lorsqu'un locuteur produit une expression référentielle, cette forme linguistique crée une activation à tous les niveaux de représentation pour le locuteur et pour le destinataire. Cela peut expliquer pourquoi les locuteurs réutilisent des expressions référentielles. Cela conduit à un effet facilitateur, résultat de l'activation de la représentation linguistique dans l'esprit du destinataire. Dans le cadre de ce modèle, l'*audience design* n'apparaîtrait que dans les étapes plus tardives, en situation d'incompréhension résultant d'un alignement incomplet. Ainsi, des stratégies potentiellement plus coûteuses dépendant de la modélisation des états mentaux (e.g., connaissances, croyances, intentions) du locuteur seraient requises seulement quand les mécanismes initiaux échoueraient à produire l'alignement. Selon Pickering & Garrod (2004), prendre en compte de façon continue le point de vue de l'interlocuteur via le savoir partagé serait un processus cognitif trop coûteux qui imposerait une charge importante sur la mémoire et l'attention des interlocuteurs.

- 9 Dans la lignée des modèles dits « égocentriques », Bard et collègues (Bard *et al.*, 2000 ; Bard & Aylett, 2005) ont proposé un **modèle de traitement double** (i.e., *dual process model*) selon lequel deux types de processus cognitifs seraient impliqués dans la conversation, l'un automatique et rapide, sans coût cognitif (mise en jeu d'un amorçage qui renverrait à la propre expérience récente du locuteur), et l'autre lent et plus coûteux qui nécessiterait de faire des inférences sur les connaissances et croyances du destinataire, autrement dit qui mettrait en jeu une construction du modèle mental du destinataire. Dans une tâche de type *Map task*, Bard et collaborateurs (Bard *et al.*, 2000 ; Bard & Aylett, 2005) ont mesuré la durée des mots répétés au cours de la conversation, l'intelligibilité et la précision des expressions référentielles. Ils ont mis en évidence la présence d'*audience design* uniquement au niveau de ce dernier aspect. Ces résultats ont conduit les auteurs à suggérer que l'articulation serait encapsulée à partir des connaissances spécifiques au destinataire et serait par défaut égocentrique, alors que d'autres processus tels que la planification des expressions référentielles seraient guidés par des inférences sur les besoins de l'interlocuteur (donc mettant en jeu *audience design*). Ainsi, selon ce modèle, les différents choix linguistiques opérés par le

locuteur feraient appel à l'un ou à l'autre type de processus. Tandis que les variations phonétiques, telles que la réduction de la durée des mots, résulteraient de processus automatiques sans coût cognitif, les choix lexicaux ou syntaxiques tels que l'utilisation des pronoms impliqueraient des processus plus coûteux nécessitant une représentation mentale du destinataire.

- 10 Le **modèle probabiliste** (Brown-Schmidt & Hanna, 2011 ; Hanna, Tanenhaus, & Trueswell, 2003) tente de réconcilier les approches précédemment décrites, en considérant l'interprétation comme un processus de satisfaction de contraintes dans lequel connaissances partagées et connaissances privées influenceraient de manière simultanée la production et la compréhension des énoncés dès le début de l'échange entre les interlocuteurs. Selon l'approche de « *satisfaction de contraintes* »², lors de la résolution d'une ambiguïté dans laquelle un individu doit prendre des décisions ou réaliser des jugements basés sur des informations complexes et ambiguës, de multiples sources d'information sont évaluées et intégrées, rapidement, en parallèle et de façon continue jusqu'à ce qu'un état stable soit obtenu. Certaines informations appelées « contraintes » seraient plus susceptibles de jouer un rôle de façon précoce dans le processus de compréhension tandis que d'autres joueraient un rôle de façon plus tardive. Une telle approche suppose que les différentes sources d'information (i.e. les contraintes) fournissent un support probabiliste pour des interprétations compétitives en parallèle tout au long du temps et qu'elles auraient un poids différent dans la compréhension du langage. Les connaissances privées et les connaissances partagées influenceraient ainsi de façon distincte l'interprétation du langage en fonction de leurs poids dans la situation de communication (Brennan & Hanna, 2009 ; Brown-Schmidt & Hanna, 2011 ; Brown-Schmidt, 2012), raison pour laquelle un comportement collaboratif et un comportement égocentrique ont pu être rapportés en situation de conversation. Le fait que dans certains cas la prise en compte du destinataire émerge dans les étapes plus tardives ne signifie pas que le traitement précoce est nécessairement égocentrique. L'information concernant les besoins du destinataire n'étaient peut-être pas disponible de façon précoce (Galati & Brennan, 2010). Ainsi, le savoir partagé ne serait que l'un des multiples indices intégrés de façon simultanée pendant le traitement, ses effets sur la compréhension étant proportionnels à sa saillance pour les interlocuteurs dans la conversation. Des travaux ont montré que les connaissances du destinataire sur le point de vue du locuteur pouvaient être une des contraintes qui modulent l'interprétation des expressions référentielles de façon précoce (cf. Brown-Schmidt & Heller, 2018 pour une revue).
- 11 Selon le modèle **basé sur la mémoire** (Horton & Gerrig, 2005a, 2005b, 2016), l'adaptation au destinataire au cours de la conversation serait basée sur des processus automatiques d'encodage et de récupération en mémoire. Des informations sur un interlocuteur donné sont encodées en mémoires et gardées sous forme de traces épisodiques³. Elles sont alors accessibles automatiquement via un processus de « résonance »⁴. Ainsi lorsqu'un locuteur prépare une expression référentielle, le destinataire servirait d'indice contextuel pour un épisode dans lequel les interlocuteurs (directeur et destinataire) ont déjà été co-présents (ont conversé ensemble) et ainsi favoriserait la récupération en mémoire des informations associées à ce destinataire. Ce processus conduirait le locuteur à utiliser des informations déjà partagées concernant l'expression référentielle en désignant un objet ou une figure par son nom partagé plutôt que par un nom non-partagé entre les interlocuteurs. Cette approche permet de rendre compte de l'*audience design* en termes de mécanismes généraux (la mémoire)

non spécifiques au savoir partagé et peu coûteux cognitivement. L'*audience design* serait donc soumis aux limitations de la mémoire. Ainsi, lorsque les associations spécifiques aux interlocuteurs ne seraient pas disponibles en mémoire, les locuteurs seraient moins capables de produire des énoncés adaptés au destinataire (Horton et Gerrig, 2005b).

- 12 En résumé, la prise en compte et la mise à jour du *common ground* est centrale dans le modèle collaboratif suggérant une mise en jeu continue de l'*audience design* dans la production et la compréhension des expressions référentielles, autrement dit une modélisation de la représentation des interlocuteurs tout au long de la conversation. Ceci nécessite un traitement cognitivement coûteux remis en question par les autres modèles présentés ci-dessus. Ainsi, le modèle d'alignement interactif propose qu'un processus d'amorçage automatique à tous les niveaux de représentation linguistique entraînant une activation des mêmes représentations linguistiques chez le locuteur et le destinataire guide les choix linguistiques, l'adaptation au besoin du destinataire (et donc la prise en compte du *common ground*) n'ayant lieu que dans les situations d'incompréhension. Dans le modèle basé sur la mémoire, la prise en compte du destinataire est aussi basée sur des processus automatiques, non pas d'amorçage, mais d'encodage et de récupération en mémoire. Pour le modèle égocentrique, le locuteur ne prend en compte son destinataire que dans les phases plus tardives de production et de compréhension. Dans les phases initiales, les interlocuteurs ne s'appuient que sur leur savoir privé pour planifier leurs expressions référentielles. Pour le modèle probabiliste, savoir privé et savoir partagé influencent les choix linguistiques dès qu'ils sont disponibles dans la conversation et en fonction de leur poids. Ainsi, les informations concernant les besoins du destinataire (savoir partagé) sont prises en compte dès qu'elles sont disponibles en fonction de leur poids/de leur saillance dans la conversation. Enfin, le modèle de traitement double, tente de réconcilier le modèle collaboratif et les modèles égocentriques en proposant deux voies de traitement en fonction des choix linguistiques : la production des expressions référentielles impliquerait l'*audience design* tandis que la production de variations phonétiques (incluant les variations prosodiques telles que le choix de mots avec une durée réduite) résulterait de processus automatiques.
- 13 Parallèlement à la mise à jour de ces modèles, plusieurs travaux ont permis d'identifier un certain nombre de facteurs internes au locuteur (e.g., capacités de mémoire, de ressources attentionnelles, de théorie de l'esprit) et de facteurs propres à la situation de conversation (e.g., caractéristiques du destinataire : novice vs expert, présence d'un contexte visuel partagé ou non, difficulté de la tâche à réaliser) susceptibles d'influencer la mise en jeu de l'*audience design* (Brown-Schmidt & Heller, 2018 ; Wardlow, 2013). Ainsi, par exemple, des individus ayant de plus faibles capacités d'inhibition vont être incapables d'inhiber leur propre perspective dans une conversation et vont ainsi montrer des interprétations ou des productions de type égocentriques reflétant une absence d'*audience design* dans une conversation (Brown-Schmidt, 2009). Prendre en compte le point de vue du destinataire dans la conversation nécessite, en effet, d'inhiber son propre point de vue quand il y a un conflit entre ces deux points de vue. Par exemple, dans une tâche de communication référentielle, si une information est disponible pour le locuteur mais pas pour le destinataire, le locuteur doit inhiber sa propre perspective afin de donner une référence non ambiguë, adaptée aux connaissances du destinataire. D'autres auteurs ont montré un impact de la mémoire de travail et du contrôle exécutif sur la façon dont des destinataires prenaient en compte la perspective du locuteur pour anticiper le référent correct d'une phrase

(Brown-Schmidt, 2009 ; Ryskin et al 2015), ou un impact de la mémoire de travail sur la façon dont des participants jouant le rôle de directeur dans une tâche de communication référentielle étaient plus à même d'adapter de manière appropriée leurs énoncés au point de vue de leur destinataire (Warlow, 2013).

- 14 Comme nous l'avons souligné au début de cet article, excepté le modèle de traitement double de Bard et collaborateurs, les études ayant servi de précurseurs au développement de ces modèles ont surtout porté sur l'utilisation de la référence dans des paradigmes exploitant la communication référentielle. Les travaux que nous avons développés au LPL concernent les choix référentiels, les choix prosodiques et l'influence de facteurs internes tels que les capacités de théorie de l'esprit du locuteur et de facteurs situationnels tels que le contexte visuel sur l'*audience design*.

Partie 2. Notre contribution à ce débat

- 15 Comme nous avons pu le voir dans la section précédente, ce qui distingue les différents modèles existants c'est principalement le recours au *common ground* dans la mise en jeu de l'*audience design* ainsi que le coût cognitif lié à l'éventuelle prise en compte continue de l'interlocuteur dans la conversation. Dans le cadre du modèle collaboratif (Clark, 1996), la prise en compte de l'interlocuteur via la construction du *common ground* nécessite que locuteur et destinataire s'attribuent mutuellement des états mentaux tels que connaissances et croyances sur ce *common ground*. Cette capacité à attribuer des états mentaux à autrui réfère à la théorie de l'esprit en psychologie, c'est-à-dire la capacité à former des représentations des états mentaux (e.g. intention, croyance, connaissance) des autres et à utiliser ces représentations pour comprendre, prévoir et juger leurs énoncés et leurs comportements (Baron-Cohen *et al.*, 1985). Des troubles affectant les capacités de théorie de l'esprit sont donc susceptibles d'entraîner une non prise en compte du savoir partagé entre les deux interlocuteurs, et par conséquent un discours ne tenant pas compte des besoins du destinataire.
- 16 Ainsi, étudier l'*audience design* chez des individus présentant des difficultés à attribuer des connaissances, croyances aux autres (i.e., théorie de l'esprit), tels que des individus atteints de schizophrénie peut permettre d'identifier quand et comment les locuteurs prennent en compte la perspective de leur interlocuteur, s'adaptent à ce qu'il sait ou ne sait pas, lorsqu'ils choisissent des formes linguistiques particulières (e.g., expressions référentielles, formes prosodiques) en situation de conversation.
- 17 Dans cette section, nous présentons deux études portant sur le choix des expressions référentielles, réalisées chez des individus atteints de schizophrénie, et chez des individus atteints de la maladie d'Alzheimer présentant des difficultés à attribuer des états mentaux aux autres (Champagne-Lavau *et al.*, 2009 ; Moreau *et al.*, 2016). Nous présentons, ensuite une série d'études portant sur le choix de formes prosodiques chez des individus atteints de schizophrénie et chez des participants sains (Michelas *et al.*, 2014 ; Michelas *et al.*, 2019a, 2019b).

2.1. Choix des expressions référentielles

- 18 Dans une première étude (Champagne-Lavau *et al.*, 2009), nous avons ainsi mis en évidence que des difficultés à attribuer des états mentaux (e.g., connaissances, croyances, intentions) à l'interlocuteur dans une situation de jeu collaboratif avaient

des répercussions sur les choix linguistiques des locuteurs en termes de choix référentiels. Autrement dit, ils engendraient une non prise en compte du *common ground* construit avec l'interlocuteur. Plus précisément, nous avons utilisé le paradigme de communication référentielle chez des patients atteints de schizophrénie et des participants contrôles sains jouant le rôle de directeur avec un expérimentateur jouant le rôle de destinataire. L'analyse de l'utilisation différenciée des marqueurs de référence (descriptions indéfinies versus définies : *une montagne* pour faire référence à une information non connue de l'interlocuteur versus *la montagne* pour faire référence à une information connue de l'interlocuteur) au cours de la tâche a montré que, contrairement aux participants contrôles sains, les participants atteints de schizophrénie présentant des troubles de théorie de l'esprit (évalués par un test classique) ne tenaient pas compte du savoir partagé construit avec leur interlocuteur. Autrement dit, tout au long de la répétition des essais, ils continuaient à utiliser des descriptions indéfinies pour faire référence à des informations déjà partagées avec leur interlocuteur, alors que les participants contrôles sains utilisaient des références définies. En revanche, les participants atteints de schizophrénie ne présentant pas de difficulté de théorie de l'esprit avaient des performances semblables à celles des participants contrôles. Nous avons reproduit ce patron de performance chez des patients souffrant de la maladie d'Alzheimer présentant des difficultés à attribuer des états mentaux (Moreau *et al.*, 2016). Tout comme les participants atteints de schizophrénie, ils ne prenaient pas en compte les informations partagées au cours des échanges précédents avec leur interlocuteur. Ces études suggèrent ainsi que la théorie de l'esprit constitue un facteur interne au locuteur affectant la mise en jeu de l'*audience design* en termes de choix référentiels. Elles confirment, de plus, le rôle prépondérant de cette capacité à attribuer des connaissances et croyances à l'interlocuteur dans la construction du *common ground*.

2.2. Choix des formes prosodiques

- 19 Les principales études (Arnold, 2012 ; Bard *et al.*, 2000 ; Bard & Aylett, 2005 ; Galati & Brennan, 2010 ; Rosa *et al.*, 2015) ayant étudié la mise en jeu de l'*audience design* via les choix prosodiques ont essentiellement porté sur la durée des mots. Nos travaux apportent une contribution originale au débat sur l'*audience design* en envisageant d'autres formes prosodiques, qu'elles soient globales telles que la vitesse d'élocution, l'amplitude des mouvements mélodiques, ou plus locales telles que le découpage prosodique. Nous avons ainsi étudié si des participants atteints de schizophrénie présentant des troubles de théorie de l'esprit tenaient compte des besoins de leur interlocuteur quand ils choisissaient des formes prosodiques particulières (Michelas *et al.*, 2014 ; Michelas & Champagne-Lavau, 2019 ; Michelas, Cau & Champagne-Lavau, 2019). Pour cela, nous nous sommes intéressés au marquage prosodique de l'information nouvelle (focus) en situation d'interaction conversationnelle. Une des fonctions linguistiques de la prosodie est, en effet, de structurer l'information en accord avec les croyances du locuteur sur ce que sait son interlocuteur. Ainsi, en français, pour assurer une communication efficace, le locuteur indique à son interlocuteur ce qui est nouveau dans son message par rapport au *common ground* en utilisant le découpage prosodique (Féry 2001 ; Dohen & Loevenbruck, 2004 ; Beyssade *et al.*, 2009).

- 20 Dans une première étude (Michelas *et al.*, 2014), nous avons manipulé la saillance de l'information (nouveau/contrastif) pour induire la focalisation, dans le cadre d'une tâche interactive inspirée des *Map tasks*. Au cours de cette tâche, un directeur (le participant) devait indiquer à un destinataire (l'expérimentateur), comment rejoindre le plus rapidement et le plus précisément possible un point sur une carte. La carte était construite de manière à faire produire au directeur des contrastes prosodiques portant soit sur la forme des objets (ex. BOUGIES violettes vs. BONBONS violets) soit sur leur couleur (ex. bonbons MARRONS vs. bonbons VIOLETS). Nous avons montré que les participants atteints de schizophrénie présentant des difficultés à attribuer des états mentaux aux autres n'utilisaient pas le découpage prosodique (1 groupe prosodique [bonbons violets] versus 2 groupes prosodiques [bonbons] [violets]) pour signaler le statut contrastif des éléments du discours (éléments en focus) à leur interlocuteur. Autrement dit, alors que des participants contrôles sains redécoupaient l'élément focalisé dans un constituant prosodique à part, les participants atteints de schizophrénie ne semblaient pas le faire. Ils ne signalaient pas à leur interlocuteur la partie de l'information revêtant une importance particulière (information déjà mentionnée versus information contrastive). Au niveau acoustique, l'élément contrastif était produit avec un allongement de la durée et une augmentation de la f_0 chez les participants sains tandis qu'il n'y avait pas de différence en termes de durée et de f_0 chez les participants atteints de schizophrénie que l'élément soit contrastif ou non. L'ensemble de ces résultats suggèrent une non mise en jeu de l'*audience design* chez les participants atteints de schizophrénie en lien avec leurs troubles de théorie de l'esprit.
- 21 Suite à ce résultat, nous avons utilisé la même tâche pour étudier si la présence versus l'absence d'un interlocuteur influençait l'utilisation du découpage prosodique pour signaler le statut contrastif des éléments du discours à l'interlocuteur (Michelas & Champagne-Lavau, 2019). Ainsi, dans la condition présence d'un destinataire, quinze participants sains ont joué le rôle de directeur et devaient indiquer à un destinataire (confédéré) le chemin à prendre pour aller d'un point de départ à un point d'arrivée sur une carte. Dans la condition absence de destinataire, quinze autres participants sains devaient décrire ce chemin à haute voix. Les résultats de cette étude ont montré une influence de la présence/absence d'un interlocuteur sur l'encodage prosodique du focus en français. De manière plus spécifique, lorsqu'ils ont interagi avec un interlocuteur, les participants ont utilisé le découpage prosodique (1 groupe prosodique versus 2 groupes prosodiques) pour signaler le statut contrastif des éléments du discours. Ils ont redécoupé l'élément focalisé dans un constituant prosodique à part lorsque l'information était contrastive alors que les participants réalisant la tâche sans interlocuteur ne l'ont pas fait. Ces choix linguistiques sont en faveur du modèle de Clark (1996) impliquant une prise en compte de l'interlocuteur, et donc la mise en jeu de l'*audience design*. Cette étude nous a aussi permis de reproduire les résultats de notre précédente étude, que la tâche soit réalisée avec un expérimentateur (Michelas *et al.*, 2014) ou avec un confédéré jouant le rôle de destinataire.
- 22 Dans une nouvelle étude (Michelas *et al.*, 2019), nous avons étudié comment la manipulation des connaissances partagées entre les interlocuteurs influence la prosodie d'un locuteur jouant le rôle de directeur. Il s'agissait ainsi de déterminer comment un locuteur modifie sa prosodie lorsqu'il/elle sait que son interlocuteur ne partage pas les mêmes connaissances que lui/elle. Pour cela, nous avons manipulé le

recours au savoir partagé de la part du locuteur (i.e., les croyances du locuteur sur les connaissances de l'auditeur). Dans cette étude, nous nous sommes inspirées du paradigme expérimental de Keysar *et al.* (2000) et de notre précédente étude (Michelas *et al.*, 2014) pour développer un jeu collaboratif entre un directeur et un destinataire et manipuler le *savoir privé* et le *savoir partagé* des deux interlocuteurs concernant l'emplacement de certains objets sur une grille. Les participants jouant le rôle de directeur et un confédéré jouant le rôle de destinataire étaient chacun devant un écran d'ordinateur leur présentant une série de grilles. Chaque grille présentée était la même pour les deux interlocuteurs excepté le fait que sur celle du directeur, une croix était placée entre deux objets (les objets variant quant à leur forme et leur couleur). Le directeur devait indiquer le plus précisément possible au destinataire où se trouvait la croix afin que le destinataire puisse la placer au bon endroit (même endroit que le directeur) sur sa grille. Pour cela, le directeur indiquait entre quels objets se situait la croix en donnant pour chaque objet sa forme et sa couleur (par exemple « *tu mets la croix entre la souris violette et la maison violette* »). Nous avons manipulé deux facteurs : 1) le statut informationnel des référents (contraste sur le nom, contraste sur l'adjectif, contraste sur le fragment complet) et 2) le savoir (partagé, non-partagé) entre le directeur et le destinataire. Afin de manipuler le savoir partagé, certaines cases de la grille comportant un objet étaient noires (condition de savoir non-partagé), indiquant au directeur que l'objet présent sur ces cases pouvait être différent pour le destinataire. En revanche, les objets sur les cases blanches (condition de savoir partagé) étaient identiques pour le directeur et le destinataire. Nous avons montré, dans cette étude, que les connaissances du locuteur sur ce que sait ou ne sait pas son interlocuteur ont un impact sur ses choix prosodiques au niveau global. Ainsi, quand les participants directeurs n'avaient pas de connaissances partagées avec leur interlocuteur, ils parlaient plus lentement (diminution de la vitesse d'élocution) et augmentaient l'amplitude de leurs mouvements mélodiques. Ces résultats suggèrent, en lien avec la littérature sur le *clear speech* (Smiljanić & Bradlow, 2009), que les participants directeurs ont adapté leur prosodie à leur interlocuteur afin de les aider à comprendre leur discours dans la condition de connaissances non-partagées, anticipant ainsi une possible difficulté chez leur interlocuteur. En revanche, contrairement à ce qui était attendu au niveau local, les directeurs n'ont pas utilisé le découpage prosodique pour encoder le statut informationnel des référents quelle que soit la condition de savoir partagé. Plus spécifiquement, contrairement à nos études précédentes dans lesquelles les participants étaient en condition de savoir partagé (Michelas *et al.*, 2014 ; Michelas & Champagne-Lavau, 2019), ils ont produit majoritairement un groupe prosodique quelle que soit la condition de contraste (sur le nom, l'adjectif, le fragment). Ce résultat pourrait s'expliquer par la charge cognitive de la tâche. En effet, tenir compte des connaissances de l'interlocuteur par rapport au contexte visuel des objets cibles et des compétiteurs pourrait nécessiter une plus grande allocation de ressources attentionnelles que dans nos études précédentes, ce qui augmenterait la demande cognitive des participants directeurs. Dans ce cas, les participants directeurs choisiraient les formes linguistiques les plus économiques, ce qui consisterait dans notre tâche à ne pas utiliser le découpage prosodique pour encoder le statut informationnel des noms cibles (cf. Vogel *et al.*, 2015 ; Kantola & Van Gompel, 2016 sur la question de l'influence de la charge cognitive sur les choix linguistiques). Ces résultats seraient en accord avec le modèle double (*Dual Process Hypothesis*) de Bard *et al.* (2000) dans lequel deux processus seraient mis en jeu, l'un automatique et rapide, sans

coût cognitif (mise en jeu d'un amorçage), l'autre nécessitant des inférences ou de l'attention, lent et plus coûteux en raison de la construction du modèle mental du destinataire).

- 23 Une nouvelle étude avec ce même paradigme est en cours chez des participants atteints de schizophrénie et des participants contrôles sains afin de confirmer cet ajustement du locuteur aux connaissances de son interlocuteur au niveau de la prosodie globale (c'est-à-dire de la vitesse d'élocution et de l'amplitude des mouvements mélodiques). Si ces modifications de la prosodie globale sont bien liées à une prise en compte de l'interlocuteur, nous nous attendons à ce que des participants atteints de schizophrénie présentant un déficit de théorie de l'esprit ne produisent pas ces modifications, dans la mesure où ils seront incapables de prendre en compte les connaissances de leur interlocuteur.
- 24 **En conclusion**, nos travaux, réalisés chez des populations cliniques présentant des difficultés à attribuer des connaissances, croyances aux autres et chez des participants sans pathologie, nous ont permis d'étudier jusqu'à quel point le locuteur tient compte des besoins de l'auditeur quand il choisit des formes linguistiques particulières. Ils soulignent l'influence des facteurs internes au locuteur (e.g., capacités de théorie de l'esprit, de ressources attentionnelles) et des facteurs propres à la situation de conversation (e.g., présence d'un contexte visuel partagé ou non, difficulté de la tâche à réaliser) dans la mise en jeu de l'*audience design*. Ces facteurs internes au locuteur peuvent constituer des différences individuelles chez les sujets sains affectant la façon dont les locuteurs prennent en compte et s'adaptent à la perspective de leur interlocuteur. Ces différences individuelles ainsi que les facteurs propres à la situation de conversation pourraient ainsi expliquer pourquoi l'*audience design* n'est pas mis en évidence dans certains cas et contribuer à expliquer l'absence de consensus entre les différents modèles décrits dans la section 1 sur la mise en jeu de d'*audience design*. D'autres travaux de recherche sont nécessaires afin d'identifier le rôle précis de ces facteurs afin de mieux comprendre quand et comment nous prenons en compte la perspective de notre interlocuteur lorsque nous choisissons des formes linguistiques particulières (e.g., expressions référentielles, formes prosodiques).

BIBLIOGRAPHIE

Anderson, A. H., Bader, M., Bard, E. G., Boyle, E., Doherty, G., Garrod, S., Weinert, R. (1991) The Hrc Map Task Corpus, *Language and Speech*, 34, p. 351-366.

Arnold, J. E., Kahn, J. M., & Pancani, G. C. (2012) Audience design affects acoustic reduction via production facilitation, *Psychon Bull Rev*, 19, 3, p. 505-512.

Bard, E. G., Anderson, A. H., Sotillo, C., Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000) Controlling the intelligibility of referring expressions in dialogue, *Journal of Memory and Language*, 42, 1, p. 1-22.

- Bard, E. G. & Aylett, M. P. (2005) Referential form, word duration, and modeling the listener in spoken dialogue, In J. C. Trueswell and M. K. Tanenhaus (Eds.) *Approaches to studying world-situated language use: Bridging the language-as-product and language-as action traditions*, Cambridge: MA: MIT Press, p. 173-191.
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985) Does the autistic child have a “theory of mind”?, *Cognition*, 21, 1, p. 37-46.
- Barr, D.J. & Keysar, B. (2006) Perspective taking and the coordination of meaning in language use, In Traxler, M.J. and Gernsbacher, M.A. (eds.) *Handbook of Psycholinguistics (2nd Ed.)*, Elsevier: Amsterdam, Netherlands, p. 901-938.
- Barr, D. J. (2008) Pragmatic expectations and linguistic evidence: Listeners anticipate but do not integrate common ground, *Cognition*, 109, 1, p. 18-40.
- Barr, D. J., & Keysar, B. (2002) Anchoring comprehension in linguistic precedents. *Journal of Memory and Language*, 46, 2, p. 391-418.
- Beysade, C., Hemforth, B., Marandin, J.-M. & Portes, C. (2009) Prosodic Markings of Information Focus in French, In Yoo HY, Delais-Roussarie, E. (Eds.) *Actes d’Interfaces, Discours et Prosodie*, Paris, p. 109-122.
- Brennan, S. E., & Clark, H. H. (1996) Conceptual pacts and lexical choice in conversation, *Journal of Experimental Psychology-Learning Memory and Cognition*, 22, 6, p. 1482-1493.
- Brennan, S. E., & Hanna, J. E. (2009) Partner-Specific Adaptation in Dialog, *Topics in Cognitive Science*, 1, 2, p. 274-291.
- Brown-Schmidt, S. (2009a) Partner-specific interpretation of maintained referential precedents during interactive dialog, *Journal of Memory and Language*, 61, 2, p. 171-190.
- Brown-Schmidt, S. (2009b) The role of executive function in perspective taking during online language comprehension, *Psychonomic Bulletin & Review*, 16, 5, p. 893-900.
- Brown-Schmidt, S. (2012) Beyond common and privileged: Gradient representations of common ground in real-time language use, *Language and Cognitive Processes*, 27, 1, p. 62-89.
- Brown-Schmidt, S. & Hanna, J.E. (2011) Talking in another person’s shoes: Incremental perspective-taking in language processing, *Dialogue & Discourse*, 2, 1, p.11-33.
- Brown-Schmidt, S., & Heller, D. (2018) Perspective taking during conversation, in Gaskell, G. & Rueschemeyer, S.A. (Eds) *Oxford Handbook of Psycholinguistics*, Oxford University Press, p. 551-574.
- Champagne-Lavau, M., Fossard, M., Martel, G., Chapdelaine, C., Blouin, G., Rodriguez, J. P., & Stip, E. (2009) Do patients with schizophrenia attribute mental states in a referential communication task? [Research Support, Non-U.S. Gov’t], *Cognitive Neuropsychiatry*, 14, 3, p. 217-239.
- Clark, H. H. (1996) *Using language* (Vol. 4), Cambridge: Cambridge University Press.
- Clark, H. H., & Marshall, C. R. (1981) Definite reference and mutual knowledge, in Joshi, A.K., Webber, B. & Sag, I. (Eds.) *Elements of discourse understanding*, Cambridge: Cambridge University Press, p. 10-63.
- Clark, H. H. & Murphy, G. L. (1982) Audience design in meaning and reference, *Advances in psychology*, 9, p. 287-299.
- Clark, H. H., & Wilkes-Gibbs, D. (1986) Referring as a collaborative process, *Cognition*, 22, 1, p. 1-39.

- Dohen, M. & Loevenbruck, H. (2004) Pre-focal rephrasing, focal enhancement and postfocal deaccentuation in French, in *Proceedings of Interspeech*, 4 - 8 October, 2004, Korea, p. 785-788.
- Féry, C. (2001) Intonation of focus in French, in Féry, C. & Sternefeld, W. (Eds), *Audiatur Vox Sapientes: A Festschrift for Arnim von Stechow*, Berlin: Akademi Verlag, p. 153-181.
- Galati, A., & Brennan, S. E. (2010) Attenuating information in spoken communication: For the speaker, or for the addressee?, *Journal of Memory and Language*, 62, 1, p. 35-51.
- Galati, A., & Brennan, S. E. (2021) What is retained about common ground? Distinct effects of linguistic and visual co-presence, *Cognition*, 215.
- Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003) The effects of common ground and perspective on domains of referential interpretation, *Journal of Memory and Language*, 49, 1, p. 43-61.
- Horton, W. S., & Gerrig, R. J. (2005a) Conversational common ground and memory processes in language production, *Discourse Processes*, 40, 1, p. 1-35.
- Horton, W. S., & Gerrig, R. J. (2005b) The impact of memory demands on audience design during language production, *Cognition*, 96, 2, p. 127-142.
- Horton, W. S., & Gerrig, R. J. (2016) Revisiting the Memory-Based Processing Approach to Common Ground, *Topics in Cognitive Science*, 8, 4, p. 780-795.
- Horton, W. S., & Keysar, B. (1996) When do speakers take into account common ground?, *Cognition*, 59, 1, p. 91-117.
- Kantola, L., & van Gompel, R. P. G. (2016) Is anaphoric reference cooperative?, *Quarterly Journal of Experimental Psychology*, 69, 6, p. 1109-1128.
- Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000) Taking perspective in conversation: the role of mutual knowledge in comprehension, *Psychological Science*, 11, 1, p. 32-38.
- Knutsen, D., & Le Bigot, L. (2013) La production et la compréhension de références dans les théories psychologiques actuelles du dialogue, *Psychologie française*, 58, 4, p. 277-296.
- Krauss, R. M., & Weinheimer, S. (1966) Concurrent feedback, confirmation, and the encoding of referents in verbal communication, *Journal of Personality and Social Psychology*, 4, 3, p. 343-346.
- Michelas, A., Cau, C., & Champagne-Lavau, M. (2019) The impact of shared knowledge on speakers' prosody, *PloS One*, 14, 10, p. e0223640.
- Michelas, A. & Champagne-Lavau, M. (2019) To what extent the French prosodic encoding of contrast is addressee-oriented?, in *Proceedings of the 3rd Phonetics and Phonology in Europe (PaPE) conference*, 17 - June 19, 2019, Lecce, Italy, p. 201-202.
- Michelas, A., Faget, C., Portes, C., Lienhart, A.-S., Boyer, L., Lancon, C., & Champagne-Lavau, M. (2014) Do patients with schizophrenia use prosody to encode contrastive discourse status?, *Front Psychol*, 5.
- Moreau, N., Rauzy, S., Viallet, F., & Champagne-Lavau, M. (2016) Theory of Mind in Alzheimer Disease: Evidence of Authentic Impairment During Social Interaction, *Neuropsychology*, 30, 3, p. 312-321.
- Pickering, M. J., & Garrod, S. (2004) Toward a mechanistic psychology of dialogue, *Behav Brain Sci*, 27, 2, p. 169-190; discussion 190-226.
- Pickering, M. J., & Garrod, S. (2006) Alignment as the Basis for Successful Communication, *Research on Language and Computation*, 4, p. 203-228.

- Pickering, M. J., & Garrod, S. (2013) An integrated theory of language production and comprehension, *Behavioral and brain sciences*, 36, 4, p. 329-347.
- Rosa, E. C., Finch, K. H., Bergeson, M., & Arnold, J. E. (2015) The effects of addressee attention on prosodic prominence, *Language Cognition and Neuroscience*, 30, 1-2, p. 48-56.
- Ryskin, R. A., Benjamin, A. S., Tullis, J., & Brown-Schmidt, S. (2015) Perspective-taking in comprehension, production, and memory: An individual differences approach, [Research Support, U.S. Gov't, Non-P.H.S.], *J Exp Psychol Gen*, 144, 5, p. 898-915.
- Smiljanić R. & Bradlow A. R. (2009) Speaking and hearing clearly: Talker and listener factors in speaking style changes, *Language and linguistics compass*, 3, 1, p. 236-264.
- Stalnaker, R. (1978) Assertion, in Cole, P. (Ed.) *Syntax and semantics 9: Pragmatics*, New York: Academic Press, p. 315-332.
- Thagard, P. (2000) *Coherence in thought and action*, Cambridge MIT Press.
- Vogels, J., Krahmer, E., & Maes, A. (2015) How Cognitive Load Influences Speakers' Choice of Referring Expressions, *Cogn Sci*, 39, 6, p. 1396-1418.
- Wardlow, L. (2013) Individual differences in speakers' perspective taking: The roles of executive control and working memory, *Psychon Bull Rev*, 20, 4, p. 766-772.
- Wilkes-Gibbs, D., & Clark, H. H. (1992) Coordinating Beliefs in Conversation, *Journal of Memory and Language*, 31, 2, p. 183-194.

NOTES

1. <http://journals.openedition.org/tipa/5210>
2. Cette approche a déjà été utilisée dans la résolution d'ambiguïtés syntaxiques et dans la prise de décision par exemple (cf. Thagard, 2000 pour une revue).
3. La mémoire épisodique et la capacité à se rappeler des événements passés et plus particulièrement à se remémorer les événements personnellement vécus (mémoire autobiographique) dans un contexte particulier (temps, lieu, personnes, état émotionnel).
4. L'interlocuteur servirait alors de repère résonnant avec des informations stockées en mémoire associées à cet interlocuteur.

RÉSUMÉS

Le processus visant à prendre en compte ou à s'adapter aux besoins et connaissances de notre interlocuteur dans une conversation est appelé *audience design* (Clark & Murphy, 1982 ; Clark, 1996 ; Galati & Brennan, 2010). On parle aussi de *prise de perspective* dans la mesure où l'interlocuteur est amené à prendre en compte la perspective de son interlocuteur quand il formule ses énoncés (Brown-Schmidt & Heller, 2018). Comprendre quand et comment les interlocuteurs utilisent l'*audience design* revient à comprendre comment les locuteurs prennent en compte des informations qu'ils croient partager avec leur interlocuteur, autrement dit leur *common ground*. L'*audience design* implique, en effet, la distinction entre ce qui fait partie du

common ground, c'est-à-dire les informations partagées avec un interlocuteur particulier et mutuellement reconnues comme telles (e.g., Clark, 1996) et ce qui fait partie du savoir privé (i.e., *privileged ground*) propre à chaque locuteur et non connu de l'interlocuteur. Le *common ground* comprend les connaissances - incluant des informations sur le contexte physique actuel, des informations liées aux précédents échanges conversationnels, des connaissances encyclopédiques telles que des connaissances partagées par une communauté spécifique (e.g., les habitants d'Aix-en-Provence) - partagées par le locuteur et l'auditeur et reconnues par les deux interlocuteurs pour être connues par chacun d'eux (Clark & Marshall, 1981). Ainsi, les informations partagées ou non partagées par les interlocuteurs à un moment de la conversation détermineraient le type d'expressions ou d'énoncés utilisés par le locuteur pour son interlocuteur.

S'il est clair que locuteurs et auditeurs adaptent leur comportement linguistique à leur interlocuteur pour éviter et résoudre des situations d'incompréhension, la question de savoir jusqu'à quel point et dans quelle situation l'*audience design* est mis en jeu fait actuellement débat. Autrement dit, quand et comment prenons-nous en compte la perspective de notre interlocuteur lorsque nous choisissons des formes linguistiques particulières (e.g., expressions référentielles, formes prosodiques) ? L'objectif de cet article est de situer les travaux que nous réalisons au LPL dans le cadre de ce questionnement.

Plusieurs modèles d'*audience design* ont été proposés en psycholinguistique (e.g., modèle collaboratif, modèle egocentrique, modèle d'alignement interactif, modèle de traitement double, modèle probabiliste, modèle basé sur la mémoire) pour rendre compte du fait que les locuteurs ne feraient pas des inférences sur les connaissances de leur interlocuteur de manière continue. L'hypothèse d'un recours plus ou moins important au *common ground* dans la mise en jeu de l'*audience design* est le point de divergence entre les différents modèles existants. Alors que le modèle collaboratif (Clark, 1996) considère la conversation comme une contribution constante et conjointe du locuteur et de l'auditeur au *common ground*, suggérant une mise en jeu continue de l'*audience design* dans la conversation, d'autres modèles insistent sur le comportement plus ou moins egocentrique des interlocuteurs dans la prise en compte des besoins du destinataire et sur la mise en jeu de mécanismes automatiques peu coûteux. En effet, la principale critique faite au modèle collaboratif est que la prise en compte de l'interlocuteur via le *common ground* nécessiterait une représentation distincte et coûteuse cognitivement des connaissances, croyances et intentions de l'interlocuteur. Les coûts de traitement d'un tel système seraient trop élevés pour permettre une conversation fluide. Les modèles egocentrique (e.g., Keysar *et al.*, 2000), d'alignement interactif (e.g., Pickering & Garrod 2004), probabiliste (e.g., Brown-Schmidt & Hanna, 2011), et basé sur la mémoire (e.g., Horton & Gerring, 2005) visent à tenir compte de ce coût cognitif. Le modèle de traitement double (e.g., Bard *et al.*, 2000), tente de réconcilier le modèle collaboratif et les modèles egocentriques en proposant deux voies de traitement en fonction des choix linguistiques, la production des expressions référentielles impliquant des processus lents et plus coûteux qui nécessiteraient de faire des inférences sur les connaissances et croyances du destinataire, autrement dit mettant en jeu l'*audience design*, tandis que la production de variations phonétiques résulterait de processus automatiques, rapides et sans coût cognitif. Parallèlement à la mise à jour de ces modèles, plusieurs travaux ont permis d'identifier un certain nombre de facteurs internes (e.g., mémoire, contrôle exécutif, théorie de l'esprit) au locuteur et de facteurs propres à la situation de conversation (e.g., contexte visuel) susceptibles d'influencer la mise en jeu de l'*audience design*. Excepté le modèle de traitement double de Bard et collaborateurs, les études ayant servi de précurseurs au développement de l'ensemble des modèles de l'*audience design* ont surtout porté sur les formes référentielles utilisées par les locuteurs dans des conditions où ils réfèrent au même objet à différents moments ou avec des destinataires différents, mais aussi à l'interprétation en temps réel des formes référentielles par les auditeurs telle que mesurée par le suivi des mouvements oculaires (e.g., Keysar *et al.*, 2000).

Contrairement à ces études, notre approche consiste à étudier les choix prosodiques (c'est-à-dire les choix mélodiques et rythmiques) qui sont opérés par le locuteur en conversation en fonction de la présence d'un interlocuteur et des connaissances de cet interlocuteur.

Cet article vise, dans un premier temps, à mettre en perspective les principaux modèles psycholinguistiques contribuant à comprendre quand et comment l'*audience design* est mis en jeu dans la conversation. Dans un deuxième temps, nous décrivons nos travaux, développés au LPL, montrant l'intérêt d'une approche clinique pour étudier l'*audience design* chez des individus atteints de schizophrénie en évaluant leurs choix prosodiques en lien avec leurs difficultés à attribuer des états mentaux (e.g., connaissances, croyances) aux autres. Nos autres travaux réalisés chez des participants sains montrent 1) qu'un locuteur n'effectue pas les mêmes choix prosodiques afin de mettre en valeur la partie de son message qui revêt une importance particulière lorsqu'il s'exprime en présence d'un interlocuteur réel et lorsqu'il s'exprime sans cet interlocuteur; 2) qu'un locuteur adapte sa prosodie globale (i.e., sa vitesse d'élocution et l'amplitude de ses variations mélodiques) à son interlocuteur, afin de faciliter la communication, selon qu'il partage ou non le même contexte visuel avec lui.

The process of accommodating or adapting to the needs and knowledge of our interlocutor in a conversation is called *audience design* (Clark & Murphy, 1982; Clark, 1996; Galati & Brennan, 2010). We also speak of perspective taking since interlocutors have to take into account the perspective of their interlocutor when they formulate their statements (Brown-Schmidt & Heller, 2018). To understand when and how the interlocutors use audience design means to understand how speakers take into account information that they believe they share with their interlocutor, in other words their *common ground*. Audience design implies the distinction between what is part of the common ground (ie, the information shared with a particular interlocutor and mutually recognized as such; Clark, 1996), and what is part of the private knowledge (ie, privileged ground) specific to each speaker and unknown to the interlocutor. The common ground includes contextual knowledge (on the current physical context or related to previous conversational exchanges) and encyclopedic knowledge (such as knowledge shared by a specific community, e.g., the residents of Aix-en-Provence) shared by the speaker and the listener and recognized by the two interlocutors to be known by each of them (Clark & Marshall, 1981). Thus, the information shared or not shared by the interlocutors at one point in the conversation would determine the type of expressions or utterances used by the speaker for his/her interlocutor.

While it is clear that speakers and listeners adapt their linguistic behavior to their interlocutor in order to avoid and resolve situations of incomprehension, to what extent, and in which situation, audience design is involved is still a question under debate. In other words, when, and how, do we take into account the perspective of our interlocutor when we choose particular linguistic forms (e.g., referential expressions, prosodic forms)? The aim of this article is to frame the studies we conducted at the LPL as part of this debate.

Several audience design models have been proposed in psycholinguistics (e.g., collaborative model, egocentric model, interactive alignment model, dual processing model, probabilistic model, memory-based model) to account for the idea that speakers would not do inferences about the knowledge of their interlocutor continuously. These models diverge on the hypothesis of a more or less systematic use of the common ground in the implementation of audience design. While the collaborative model (Clark, 1996) considers conversation as a constant and joint contribution of the speaker and the listener to the common ground, suggesting a continuous engagement of the audience design in conversation, other models focus on the more or less speaker's egocentric behavior and on the use of less-demanding automatic mechanisms. The main criticism against the collaborative model challenges the fact that perspective taking via the common ground would require a distinct and cognitively costly representation of knowledge, beliefs and intentions of the interlocutor. The processing costs of such a system would be too

high to allow a smooth conversation. The egocentric (e.g., Keysar *et al.*, 2000), interactive alignment (e.g., Pickering & Garrod 2004), probabilistic (e.g., Brown-Schmidt & Hanna, 2011), and memory-based (e.g., Horton & Gerring, 2005) models aim to take into consideration this cognitive cost. The dual processing model (e.g., Bard *et al.*, 2000) attempts to reconcile collaborative and egocentric models by proposing two types of cognitive processes that prevail for linguistic choices: the production of referential expressions that requires making inferences about the knowledge of the interlocutor would involve slow and cognitively demanding processes, while the production of phonetic variations that refers to the speaker's own recent experience would result from more automatic and rapid processes. At the same time as updating these models, several studies have identified speaker-internal constraints (e.g., memory, executive control, theory of mind) and situational constraints (e.g., visual context) which are likely to influence the involvement of design audience. Except for the dual processing model of Bard *et al.*, models of audience design are mainly based on studies focusing either on the way speakers produce referential forms when referring to the same object at different times and/or with different addressees, or on the way listeners interpret referential forms online, as measured by the monitoring of eye movements (e.g., Keysar *et al.*, 2000). Unlike these studies, our approach consists in studying the prosodic choices (i.e., the melodic and rhythmic choices) the speaker makes during conversation according to the presence of an interlocutor and the knowledge of this interlocutor.

The aim of the current article is to describe: (1) the main psycholinguistic models contributing to understand when and how audience design takes place during conversation, and (2) to highlight the work we developed at the LPL in which we adopt a clinical approach to the study of audience design by evaluating the prosodic choices made by individuals with schizophrenia in relation to their difficulties in attributing mental states to others. More generally, our work with healthy participant shows that speakers 1) make different prosodic choices to indicate the relevant part of their message when speaking in the presence of a real interactional partner and when speaking without interlocutor 2) adapt their overall prosody relative to the visual context they share with the interlocutor to facilitate communication.

INDEX

Mots-clés : audience design, modèles psycholinguistiques, prosodie, expressions référentielles, théorie de l'esprit, ressources cognitives, schizophrénie, maladie d'Alzheimer

Keywords : audience design, psycholinguistic models, prosody, referring expressions, theory of mind, cognitive resources, Schizophrenia, Alzheimer's disease

AUTEURS

MAUD CHAMPAGNE-LAVAU

Aix-Marseille Université, CNRS, LPL, Aix-en-Provence, France
maud.champagne-lavau@univ-amu.fr

NOÉMIE MOREAU

Aix-Marseille Université, CNRS, LPL, Aix-en-Provence, France
Centre Hospitalier Intercommunal Aix-Pertuis, Service de Neurologie, Aix-en-Provence, France
noemie.moreau@univ-amu.fr

LOLA RIVOAL

Aix-Marseille Université, CNRS, LPL, Aix-en-Provence, France

AMANDINE MICHELAS

Aix-Marseille Université, CNRS, LPL, Aix-en-Provence, France

amandine.michelas@univ-amu.fr