La compréhension comme médiateur entre contextualisation des assistants vocaux et expérience utilisateur : une étude exploratoire

Jérémie Muselet
Université de Lille, laboratoire Lumen
Annabel Martin
Université de Lille, laboratoire Lumen

Résumé

À l'heure où plus de 8 milliards d'assistants vocaux circulent, leur rôle domestique croît. Bien qu'ils gagnent en capacités conversationnelles, leur compréhension du langage reste limitée. Cette étude examine comment la compréhension (lexico-sémantique, contextuelle, situationnelle et pragmatique) agit comme médiateur entre technologie et expérience utilisateur. L'analyse de 278 avis sur le Google Nest Mini, via une modélisation LDA, identifie 15 clusters réunis en 6 méta-clusters. Les résultats montrent que la perception dépend de la contextualisation des réponses, de la cohérence conversationnelle et du respect des attentes implicites. La puissance et la qualité audio influencent fortement l'évaluation. L'étude révèle aussi des tensions entre attentes et stratégies commerciales, affectant confiance et engagement, et souligne l'importance d'intégrer une compréhension plus poussée dans la conception de ces dispositifs.

Mots-clés: assistants vocaux, expérience utilisateur, compréhension, intelligence artificielle

**Abstract** 

At a time when more than 8 billion voice assistants are in circulation, their domestic role continues to grow. Although their conversational capabilities are improving, their understanding of language remains limited. This study examines how understanding — in its lexico-semantic, contextual, situational, and pragmatic dimensions — acts as a mediator between technology and user experience. The analysis of 278 reviews of the Google Nest Mini, using LDA topic modelling, identifies 15 clusters grouped into 6 meta-clusters. The findings show that user perception depends on the contextualisation of responses, conversational coherence, and the fulfilment of implicit expectations. Perceived sound power and audio quality strongly influence evaluations. The study also highlights tensions between user expectations and commercial strategies, which affect trust and engagement, and emphasises the importance of integrating deeper understanding into the design of these devices.

Keywords: voice assistants, user experience, understanding, artificial intelligence

## **INTRODUCTION**

En 2024, plus de 8,4 milliards d'assistants vocaux étaient en circulation dans le monde (Statista, 2024). Ces dispositifs facilitent les tâches quotidiennes et l'automatisation domestique via la commande vocale (Richert, 2022). Des acteurs comme Amazon et Meta innovent: Alexa+ adapte ses réponses au ton de la voix (AP News, 2025) et LLaMA 4 ambitionne des dialogues plus naturels (Financial Times, 2025). Cependant, la compréhension du langage reste limitée (Richert, 2022) : elle dépend du contexte interactionnel et situationnel (Vachaudez et al., 2024; Dey et al., 2000; Kissine et al., 2016) et suppose une coconstruction du sens (Clark, 1996; Dionne et al., 2011) et une adaptation constante (Suchman, 2007; Suchelle-Ferrière et al., 2021). L'expérience utilisateur est donc centrale (Hassenzahl et al., 2006). Les utilisateurs projettent des attentes anthropomorphiques (Nass et al., 2000; Purington et al., 2017). Les paradigmes d'intelligence ambiante (Aaarts & Marzano, 2003) et d'informatique contextualisée (Dourish, 2004) soulignent l'importance de l'adaptabilité (Sciuto et al., 2018; Coutaz et al., 2005). La technologie devient ainsi un médiateur de notre rapport au monde (Verbeek, 2005). Ce travail adopte une approche exploratoire pour enrichir les recherches sur la contextualisation, en considérant la compréhension machine comme un lien entre contexte et expérience (Klein et al., 2023).

## COMPREHENSION, CONTEXTUALISATION ET PERCEPTIONS

## Niveaux de compréhension

La compréhension linguistique désigne la capacité à saisir significations implicites et subtilités d'un discours (Hajj et al., 2020). Elle comprend plusieurs niveaux : lexico-sémantique, contextuel, situationnel et pragmatique. Le niveau lexico-sémantique concerne l'interprétation des mots (Ouellette et al., 2014). Des approches comme TF-IDF pondèrent la pertinence des termes et révèlent des nuances (Lestandy et al., 2024), tandis que les bigrammes améliorent la précision des analyses (Ahmad et al., 2022). La compréhension contextuelle implique l'interprétation en fonction de l'environnement informationnel, des entités nommées et de l'historique conversationnel (Rossi et al., 2024). La compréhension situationnelle prend en compte le cadre spatial, temporel et social (Ouellette et al., 2014), par exemple via des marqueurs temporels ou références domestiques (Cen et al., 2024). Enfin, la compréhension pragmatique vise à décoder les intentions implicites (Baider et al., 2021),

notamment commerciales, soulevant des enjeux éthiques de transparence et de manipulation (Pochet, 2017).

## Contextualisation technologique comme vecteur d'amélioration de l'expérience utilisateur des assistants vocaux

La contextualisation technologique désigne la capacité d'un assistant vocal à interpréter une requête en tenant compte du lieu, du moment, de l'historique ou des préférences (Gouliaéva et al., 2020; Jézéquel, 2018; McLean et al., 2021). En intégrant ces paramètres, le système adapte ses réponses grâce à l'exploitation de données et à des algorithmes d'IA traitant le langage naturel. Cela permet d'inférer l'intention sous-jacente, même pour des requêtes ambiguës comme « Qu'est-ce que je dois porter aujourd'hui? », où moment, localisation et météo sont pris en compte (Gouliaéva et al., 2020; Scarano et al., 2023). Siri ou Alexa utilisent aussi l'historique pour anticiper les besoins, personnalisant ainsi l'expérience (Vachaudez et al., 2024; Scarano et al., 2023). Cette adaptation renforce la qualité de l'interaction et la fidélité des usagers (Jézéquel, 2018; McLean et al., 2021).

# Assistants vocaux et perception de l'expérience utilisateur : vers une meilleure compréhension des attentes

La perception de l'expérience utilisateur des assistants vocaux repose sur plusieurs dimensions. Sur le plan fonctionnel, efficacité et fluidité des interactions sont centrales (Vachaudez et al., 2024). Sur le plan émotionnel, le ton de la voix et la présence perçue influencent la qualité de l'échange (Yang et al., 2022). La confiance (Gorisse et al., 2018) et le sentiment de contrôle complètent cette évaluation (Vachaudez et al., 2024).

#### **METHODOLOGIE**

Cette recherche s'inscrit dans une épistémologie interprétativiste, voyant la compréhension comme médiation entre contextualisation technologique et expérience utilisateur. L'étude porte sur 278 avis nettoyés (N.L.P.) publiés sur Fnac.com entre novembre 2019 et octobre 2024 au sujet du Google Nest Mini<sup>1</sup>. Fnac.com a été choisi pour ses données riches<sup>2</sup>. Les concurrents ne sont pas inclus en raison de canaux de vente différents. La méthodologie

<sup>1</sup> L'absence d'évaluations pour Alexa ou Siri s'explique par leur mode de distribution : Alexa est surtout vendue sur Amazon.fr, au système d'évaluation moins détaillé que celui de la Fnac, tandis que Siri, intégré au HomePod, n'est disponible que sur le site d'Apple, sans avis publics.

<sup>&</sup>lt;sup>2</sup> Fnac.com se distingue des plateformes classiques en sollicitant de ses utilisateurs, en plus d'une note globale et d'un commentaire, l'identification d'un point fort et d'un point faible ainsi qu'une évaluation détaillée sur des critères techniques (audio, design, utilisation, puissance, connectique), offrant une base d'analyse fine de l'expérience utilisateur.

adopte une approche mixte centrée sur quatre dimensions de compréhension, utilisant la modélisation thématique LDA adaptée pour repérer les grandes préoccupations sur des échantillons de cette taille (Rajput, 2022; Maier et al., 2018; Muthusami et al., 2024). Cette analyse a structuré 15 clusters, regroupés en 6 méta-clusters. LDA est reconnue pour sa précision (Ogunwale et al., 2023) et sa lisibilité (Venkanna et al., 2023). L'innovation réside dans l'analyse des relations entre clusters, révélant la dynamique globale de l'expérience utilisateur.

#### **RESULTATS**

## Une compréhension multidimensionnelle

À partir de la modélisation thématique LDA, quatre types de compréhension ont été dégagés (Tableau 1). La compréhension lexico-sémantique reste limitée par la reconnaissance vocale et le traitement sémantique. La compréhension contextuelle rencontre des difficultés à maintenir le fil de la conversation, intégrer l'historique et gérer préférences ou entités, générant parfois de la frustration. La compréhension situationnelle varie selon la temporalité, le contexte domestique et les événements, avec un écart observé dans le délai de publication des avis selon l'émotion. Enfin, la compréhension pragmatique souligne une gêne face aux usages commerciaux et aux recommandations jugées inadéquates.

Dimensions et définitions	Manifestations	Exemples de verbatims
Compréhension lexico- sémantique : capacité à interpréter les mots et leur signification	<ul> <li>Difficultés de reconnaissance vocale</li> <li>Latence dans le traitement des commandes</li> <li>Dysfonctionnements sémantiques</li> </ul>	« problème de reconnaissance de la voix », « il ne comprend pas ce que je veux », « Lorsque que l'on dit 'Ok Google' il y a un blanc »
Compréhension contextuelle : maintien du fil conversationnel et intégration de l'historique	<ul> <li>Évaluation qualité-prix</li> <li>Reconnaissance d'entités nommées</li> <li>Intégration des préférences utilisateur</li> </ul>	« qualité-prix », « piloter des objets connectés », Utilisation fonctionnelle « météo », « radio », « Deezer », « youtube premium »
Compréhension situationnelle : adaptation au contexte spatial, temporel et social	<ul> <li>Références temporelles</li> <li>Intégration dans l'espace domestique</li> <li>Événements et occasions spéciales</li> </ul>	« semaine dernière », « il y a une semaine »,«dans la maison », « dans le salon », « toute la famille » « Noël », « offrir », « cadeau idéal »
Compréhension pragmatique : décodage des intentions implicites	<ul><li>Gêne face aux usages commerciaux</li><li>Tensions avec le modèle</li></ul>	« trop de rappels commerciaux » « il a besoin d'abonnements pour être

économique	pleinement fonctionnel »
- Attentes implicites non	[
satisfaites	

Tableau 1 - Dimensions, manifestations et verbatim de compréhension

## Médiation technologique et transformation de l'expérience utilisateur

L'analyse des avis sur le Google Nest Mini montre que la compréhension des assistants vocaux façonne l'expérience utilisateur à travers plusieurs dimensions (Tableau 2). Ces dispositifs transforment l'espace domestique en pivots domotiques et redéfinissent les zones d'interaction. Socialement et émotionnellement, ils sont perçus comme des quasi-acteurs créant un lien affectif. Leur intégration aux routines renforce le sentiment d'autonomie et de contrôle individuel. Toutefois, la confiance peut être fragilisée par des usages jugés intrusifs ou trop commerciaux, révélant une tension entre attentes et modèle économique.

Conséquences de la	Impact sur l'expérience	Exemples de Verbatims
compréhension		
Reconfiguration de	- Rôle central dans	« Je suis équipé en matériel
l'environnement	l'écosystème technologique	domotique un bon
domestique	- Redéfinition des zones	compagnon pour piloter
(situationnelle)	d'interaction	l'ensemble »
(Dourish, 2004)		« fonctionne que dans le
		salon »
Émergence de relations	- Perception comme quasi-	« avec cet article je me sens
affectives (pragmatique)	acteur social	moins seul à la maison »
(Naas et Moon, 2000)	- Réduction du sentiment de	
	solitude	
Renforcement de	- Intégration dans les routines	« utilisé du réveil au
l'autonomie	quotidiennes	coucher »« mon meilleur
(situationnelle)	- Sentiment de contrôle accru	allié pour organiser ma
(Dourish, 2004)		journée »
Altération de la	- Perception d'intrusion	« trop de rappels
confiance (pragmatique)	commerciale	commerciaux » « il a besoin
(Verbeek, 2005)	- Tension entre usage et modèle	d'abonnements »
	économique	

Tableau 2 - Perceptions et expérience utilisateur

L'annexe A propose un premier modèle conceptuel qui met en avant le rôle médiateur de la compréhension dans l'expérience utilisateur.

#### **DISCUSSION GENERALE**

Apport théorique. La compréhension est appréhendée comme une interface dynamique entre la capacité de contextualisation des assistants vocaux et l'expérience utilisateur. Cette recherche permet d'enrichir la théorie de la médiation technologique (Verbeek, 2005) dans le contexte des assistants vocaux. L'approche diversitaire<sup>3</sup> de Huver (2018) éclaire cette médiation qui véhicule certaines valeurs en amont de la phase d'usage proprement dite. Les dispositifs technologiques ne se contentent pas de faciliter l'usage mais participent à la circulation de nouvelles normes et pratiques domestiques. Sur le plan managérial. Les résultats soulignent qu'il ne s'agit plus uniquement d'améliorer les performances de la reconnaissance vocale, mais d'enrichir les dispositifs en tenant compte des dimensions contextuelle, situationnelle et pragmatique de la compréhension. L'analyse des avis clients met en lumière l'influence déterminante de certains attributs techniques, en particulier la puissance et la qualité sonore, dans l'évaluation globale du produit. De plus, la corrélation observée entre l'intensité émotionnelle des commentaires et leur rapidité de publication plaide en faveur de dispositifs de recueil de feedback différenciés, capables de prendre en compte à la fois les réactions immédiates et les évaluations plus réfléchies. Ces constats orientent vers trois axes stratégiques prioritaires : l'investissement massif dans l'amélioration de la qualité audio comme levier principal de satisfaction, le développement de systèmes de feedback adaptatifs distinguant les réactions émotionnelles spontanées des évaluations réfléchies, et l'intégration d'une intelligence contextuelle permettant aux assistants de s'adapter aux routines et environnements spécifiques des utilisateurs.

Cette recherche présente **plusieurs limites méthodologiques.** Le recours exclusif à des avis en ligne limite l'accès aux interactions réelles et réduit la diversité des profils utilisateurs. La modélisation de la compréhension en quatre niveaux, bien que pertinente, ne reflète que partiellement la fluidité des échanges humains. En se concentrant sur un seul assistant vocal dans un cadre domestique, l'étude reste difficilement généralisable à d'autres contextes d'usage (professionnels, mobiles). À plus long terme, l'intégration du bon sens dans les systèmes vocaux, comme le propose Choi (2022), apparaît comme une voie prometteuse pour améliorer leur compréhension. La conduite d'une collecte de données par questionnaire s'inscrirait comme une démarche méthodologique appropriée pour évaluer empiriquement la validité du modèle conceptuel élaboré dans ce contexte spécifique.

\_

<sup>&</sup>lt;sup>3</sup> La compréhension est multiple, contextuelle et évolutive.

## **Bibliographie**

Aarts, E. et Marzano, S. (2003). The new everyday: views on ambient intelligence. 010 Publishers.

AP News (2025). Amazon's new AI-powered Alexa promises to be your 'best friend in a digital world' for a monthly fee. Disponible sur:

https://apnews.com/article/017c17bddfa6742d1e78873cdda3663f [Consulté février 2025].

Baider, F. et Cislaru, G. (2021). Communication et intercompréhension : regards croisés de la pragmatique interculturelle et de la pragmatique contrastive. *Langages*, 222(2), 7-24.

Calahorra-Candao, G. et Hoyos, M.J.M. (2024). The effect of anthropomorphism of virtual voice assistants on perceived safety as an antecedent to voice shopping. *Computers in Human Behavior*, 153, 108124.

Cen, Z. et Zhao, Y. (2024). Investigating the impact of ai-driven voice assistants on user productivity and satisfaction in smart homes. *Journal of Economic Theory & Business Management*, 1(6), 8-14.

Cherif, E. et Lemoine, J. (2018). Les conseillers virtuels anthropomorphes et les réactions des internautes : une expérimentation portant sur la voix du conseiller. *Recherche et Applications en Marketing*, 34(1), 29-49.

Choi, Y. (2022). The curious case of commonsense intelligence. *Daedalus*, 151(2), 139-155.

Clark, H.H. (1996). Using language. Cambridge University Press.

Coutaz, J., Crowley, J.L., Dobson, S. et Garlan, D. (2005). Context is key. *Communications of the ACM*, 48(3), 49-53.

Dey, A.K. et Abowd, G.D. (2000). Towards a better understanding of context and context-awareness. In Workshop on the What, Who, Where, When, and How of Context-Awareness, CHI 2000, 4, 1-6.

Dionne, M., McIntyre, J., Bignasca, T. et Rezzonico, S. (2021). Soutenir les habiletés pragmatiques chez les enfants ayant un trouble développemental du langage : fondements théoriques pour la pratique clinique. *Travaux Neuchâtelois de Linguistique*, 74, 111-128.

Dourish, P. (2004). Where the action is: the foundations of embodied interaction. MIT Press.

Financial Times (2025). Meta accelerates voice-powered AI push. Disponible sur: https://www.ft.com/content/a1014427-c2ce-4204-b41a-001277309cea [Consulté mars 2025].

Gorisse, G., Christmann, O. et Richir, S. (2018). De la présence à l'incarnation : proposition d'un méta-modèle pour la réalité virtuelle. *Interfaces Numériques*, 7(1).

Gouliaeva, O., Dosquet, É. et Moysan, Y. (2020). La révolution des assistants vocaux. Dunod.

Hajj, K. et Niewese, M. (2020). Compétences et difficultés pour la compréhension des écrits en contexte plurilingue : le cas du Liban. *Traduction et Langues*, 19(2), 108-127.

Hassenzahl, M. et Tractinsky, N. (2006). User experience - a research agenda. *Behaviour & Information Technology*, 25(2), 91-97.

Huver, E. (2018). Penser la médiation dans une perspective diversitaire. *Recherches en Didactique des Langues et des Cultures*, 15(2), 1-16.

Jezequel, M. (2018). Tendances, les nouvelles du management. Gestion, 43(2), 8-10.

Kissine, M., Clin, É. et de Villiers, J. (2016). La pragmatique dans les troubles du spectre autistique : développements récents. *Médecine/Sciences*, 32(10), 874-878.

Klein, A.M., Deutschlander, J., Kölln, K., Rauschenberger, M. et Escalona, M.J. (2023). Exploring the context of use for voice user interfaces: toward context-dependent user experience quality testing. *Software: Evolution and Process*, e2618.

McLean, G., Osei-Frimpong, K. et Barhorst, J. (2021). Alexa, do voice assistants influence consumer brand engagement? – examining the role of ai powered voice assistants in influencing consumer brand engagement. *Journal of Business Research*, 124, 312-328.

Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Keinert, A., Pfetsch, B., Heyer, G., Reber, U., Häussler, T., Schmid-Petri, H. et Adam, S. (2018). Applying LDA topic modeling in communication research: Toward a valid and reliable methodology. *Communication Methods and Measures*, 12(2-3), 93-118.

Muthusami, R., Kandan, N. M., Saritha, K., Prabu, P., Kumar, S. S., Priya, S. S. et Yasmin, F. (2024). Investigating topic modeling techniques through evaluation of topics discovered in short texts data across diverse domains. *Scientific Reports*, 14, 12003.

Nass, C. et Moon, Y. (2000). Machines and mindlessness: social responses to computers. *Journal of Social Issues*, 56(1), 81-103.

Ogunwale, Y. et Ajinaja, M. (2023). Application research on semantic analysis using latent dirichlet allocation and collapsed gibbs sampling for topic discovery. *Asian Journal of Research in Computer Science*, 16(4), 445-452.

Ouellette, G. et Shaw, E. (2014). Oral vocabulary and reading comprehension: an intricate affair. *L'Année Psychologique*, 114(4), 623-645.

Pochet, S. (2017). Quelles attitudes possèdent les futurs publicitaires envers la publicité et son éthique ? analyse compréhensive des représentations sociales d'un corpus d'étudiants. *Revue Communication & Professionnalisation*, 5, 94-121.

Purington, A., Taft, J.G., Sannon, S., Bazarova, N.N. et Taylor, S.H. (2017). "Alexa is my new BFF": social roles, user satisfaction, and personification of the Amazon Echo. In Proceedings of the CHI Conference Extended Abstracts on Human Factors in Computing Systems (pp. 2853-2859). ACM.

Rajput, L. (2022). Sentiment analysis using latent dirichlet allocation for aspect term extraction. *Journal of Computers, Mechanical and Management*, 1(2), 30-35.

Richert, F. (2022). Du fétichisme des sémiomarchandises « parlantes » à l'asservissement machinique : le cas des assistants personnels intelligents. *Cygne Noir*, 7, 143-169.

Sangle-Ferriere, M. et Voyer, B. (2021). La propension du client à éviter de demander assistance : conceptualisation et développement d'une échelle. *Recherche et Applications en Marketing*, 37(3), 3-27.

Scarano, M., Regany, F. et Özçağlar □ Toulouse, N. (2023). L'engagement du consommateur envers la marque : perspective socio-culturelle et voies de recherche. *Recherche et Applications en Marketing*, 38(4), 71-97.

Sciuto, A., Saini, A., Forlizzi, J. et Hong, J.I. (2018). "Hey Alexa, what's up?": a mixed-methods studies of in-home conversational agent usage. In Proceedings of the Designing Interactive Systems Conference (pp. 857-868). ACM.

Statista (2024). Number of digital voice assistants in use worldwide from 2019 to 2024. Disponible sur: https://www.statista.com/statistics/973815/worldwide-digital-voice-assistant-in-use/ [Consulté décembre 2024].

Suchman, L.A. (2007). Human-machine reconfigurations: plans and situated actions. Cambridge University Press.

Vachaudez, A. et Koubi, C. (2024). Ok google : « Pourquoi les utilisateurs d'assistants vocaux maintiennent-ils leur utilisation de la technologie dans le temps ? ». *Décisions Marketing*, 112(4), 27-50.

Venkanna, G. et Bharati, K. (2023). A hybrid approach for generative process model with topic modelling towards efficient and dynamic document clustering. *International Journal of Advanced Technology and Engineering Exploration*, 10(106), 1184-1197.

Verbeek, P.P. (2005). What things do: philosophical reflections on technology, agency, and design. Pennsylvania State University Press.

Yang, X. et Garnier, M. (2022). Repenser l'impact des technologies interactives sur l'expérience de distribution : synthèse, approche conceptuelle et agenda de recherche. *Recherche et Applications en Marketing*, 37(4), 15-62.

Annexe A - Proposition d'un modèle conceptuel

