ALTRALANG Journal
e-ISSN: 2710-8619 p-ISSN: 2710-7922
Volume 8 Issue 2 / October 2026 pp. 583-605

Analyse des besoins et médiation linguistique : concevoir la dimension langagière d'un agent conversationnel plurilingue dans un écosystème entrepreneurial

Hanane Bouzaher*
University of Khenchela Abbes Laghrour, Algeria
bouzaher.hanane@univ-khenchela.dz

Received: 06/06/2026, **Accepted:** 08/09/2026, **Published:** 31/10/2026

Needs Analysis and Linguistic Mediation: Designing the Linguistic Dimension of a Multilingual Conversational Agent in an Entrepreneurial Ecosystem

ABSTRACT: *In digital entrepreneurial ecosystems, language is a strategic resource, yet existing support agents remain largely monolingual. In this contribution, we analyze a multilingual conversational agent's design as reverse-vector linguistic mediation: specialized language adapts to users' ordinary, hybrid, and multilingual practices. Needs analysis, which traditionally led to a course, here guides product design. This proposition is tested through NAFSIATI, a university psychological support agent. Based on a qualitative analysis of 71 responses (biology, medicine, science and technology, law), the study shows that Darija dominates in expressing distress, while French is more prominent in strongly Francophone disciplines. We derive specifications distinguishing transversal invariants (Darija, providentialist register, non-judgment) from configurable parameters (degree of mixing, primary language). This framework transforms needs analysis into a genuine decision-making operator for usage-oriented linguistic engineering.*

KEYWORDS: linguistic mediation; French for Specific Purposes (FOS); needs analysis; multilingual conversational agent; Algerian multilingualism

RÉSUMÉ : *Dans les écosystèmes entrepreneuriaux numériques, la langue est une ressource stratégique, alors que les agents de soutien existants restent largement monolingues. Nous analysons dans cette contribution la conception d'un agent conversationnel plurilingue comme une médiation linguistique à vecteur inversé : la langue de spécialité s'adapte aux pratiques hybrides et plurilingues des usagers. L'analyse des besoins, qui débouchait sur un cours, guide ici la conception d'un produit. Cette proposition est testée via NAFSIATI, agent universitaire de soutien psychologique. L'analyse qualitative de 71 réponses (biologie, médecine, sciences et technologies, droit) montre que la darija domine pour exprimer la détresse, tandis que le français progresse en milieu francophone. Nous en déduisons des spécifications distinguant invariants transversaux (darija, registre providentialiste, non-jugement) et paramètres configurables (mixage, langue-matrice). Ce cadre transforme l'analyse des besoins en un véritable opérateur de décision pour l'ingénierie linguistique orientée vers l'usage.*

MOTS-CLÉS : médiation linguistique ; Français sur Objectif Spécifique (FOS) ; analyse des besoins ; agent conversationnel plurilingue ; plurilinguisme algérien

* Corresponding author
ALTRALANG Journal / © 2026 The Author(s). Published by the University of Oran 2 Mohamed Ben Ahmed, Algeria.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

Introduction

On a longtemps pensé la langue, dans le monde de l'entreprise, comme un coût : celui de la traduction, de la formation, de la mise en conformité. Les écosystèmes entrepreneuriaux numériques en donnent une autre image. Dès qu'un produit s'adresse à ses usagers par le langage, la langue cesse d'être un coût pour devenir une condition d'adoption. Or, c'est désormais le cas d'une part croissante des services, des assistants les plus simples aux agents conversationnels les plus élaborés. Un agent qui ne rejoint pas la manière dont ses utilisateurs parlent réellement restera un objet techniquement abouti, mais socialement inerte. Il doit pouvoir reconnaître leurs hésitations, leurs alternances et les mots par lesquels ils expriment ce qui leur pèse. La question langagière n'est donc pas accessoire à l'innovation : elle en est l'une des conditions de réussite.

Pour la penser, le Français sur Objectif Spécifique offre un point de départ naturel. Depuis Mangiante et Parpette (2004), on sait que toute formation langagière sérieuse commence par une analyse des besoins. Celle-ci consiste à repérer les situations de communication visées, les discours qui s'y échangent et les compétences qu'ils supposent, afin d'en déduire un programme. Le titre même de leur ouvrage trace le chemin, de l'analyse des besoins à l'élaboration d'un cours. Ce mouvement a une orientation précise : il vise à amener l'apprenant vers une langue de spécialité dont il ne dispose pas encore. L'analyse des besoins, dans cette tradition, est au service de l'enseignement.

Mais que devient l'analyse des besoins quand elle ne conduit plus à un cours ? Concevoir un agent conversationnel renverse l'orientation. Le système, lui, maîtrise déjà cette langue de spécialité : celle de la psychologie clinique, des thérapies cognitivo-comportementales et de l'accompagnement émotionnel. Ce n'est pas l'usager qu'il faut faire monter vers ce registre ; c'est ce registre qu'il faut faire redescendre vers les pratiques ordinaires, hybrides et plurilingues de l'usager.

Les notions de traduction, de localisation ou d'adaptation culturelle permettent de décrire certaines dimensions de ce travail, mais elles ne rendent compte que partiellement du problème rencontré ici. En effet, la traduction suppose le passage d'un contenu d'une langue à une autre. La localisation adapte un produit à un marché. L'adaptation culturelle ajuste des références et des normes d'interaction. Or, la difficulté rencontrée dans la conception d'un agent conversationnel plurilingue tient moins au transfert entre langues qu'à la nécessité de faire circuler un savoir spécialisé vers des pratiques langagières ordinaires, caractérisées par leur hybridité et leur instabilité. C'est précisément cette opération spécifique, qui engage simultanément registre, répertoire linguistique et finalité interactionnelle, que les catégories disponibles décrivent imparfaitement.

Dans le cadre de cette étude, nous proposons de décrire cette opération comme une forme de médiation linguistique à vecteur inversé. Cette notion, avancée à titre heuristique, rend compte du déplacement observé : entre la langue spécialisée mobilisée par le système, d'une part, et les pratiques langagières ordinaires des usagers, d'autre part. Elle prolonge la médiation au sens du Cadre européen, entendue comme la transposition d'un sens d'un code vers un autre. Toutefois, elle n'en confie plus la charge à un enseignant ou à un dispositif pédagogique, mais à l'architecture conversationnelle du produit lui-même. Cette opération entre également en résonance avec les travaux sur l'entrepreneuriat linguistique (De Costa, Park & Wee, 2016). Ceux-ci envisagent les ressources langagières comme un capital mobilisable dans les processus d'innovation et de création de valeur.

Ce renversement prend tout son relief en contexte plurilingue. Les agents conversationnels de soutien psychologique se multiplient, mais restent très majoritairement anglophones et monolingues (Sabour et al., 2023). Des synthèses récentes relèvent que l'adaptation linguistique et culturelle y demeure insuffisamment prise en compte, en particulier pour les langues dites à faibles ressources (Algumaei et al., 2025). Plusieurs études suggèrent par ailleurs que les choix de langue ne sont pas indifférents lorsqu'il s'agit d'exprimer des émotions ou des situations de détresse. Les usagers basculent, alternent et trouvent parfois dans leur langue première des nuances que l'autre ne leur offre pas (Elwahsh et al., 2025). En Algérie, l'étudiant compose en permanence avec la darija, le français et l'arabe standard. Leur mobilisation dans des

usages complémentaires et souvent alternés fait partie du quotidien. À notre connaissance, peu de travaux ont articulé explicitement l'analyse des besoins langagiers et la conception de produits conversationnels plurilingues dans un contexte entrepreneurial et sociolinguistique comparable.

C'est ce manque que le présent article entend combler à partir du cas de NAFSIATI, un projet d'innovation universitaire algérien consacré à la conception d'un agent conversationnel de soutien psychologique pour les étudiants. Le cas n'est pas ici le sujet, mais le support de la démonstration. Il permet d'examiner l'hypothèse selon laquelle l'analyse des besoins peut guider non plus l'élaboration d'un cours, mais la conception d'un produit. La question de recherche est la suivante : comment une analyse des besoins langagiers conduite auprès des usagers visés fonde-t-elle les choix de conception d'un agent conversationnel plurilingue dans un écosystème entrepreneurial ? Pour y répondre, nous analysons un corpus de soixante et onze réponses recueillies par questionnaire bilingue auprès d'étudiants de quatre filières : biologie, médecine, sciences et technologies, et droit. L'analyse adopte un angle résolument sociolinguistique, centré sur les choix de langue, l'alternance codique, les lacunes lexicales et les idiomes de détresse.

Précisons d'emblée la nature de notre recherche. Nous ne présentons pas un agent implémenté, mais l'analyse des besoins langagiers qui en fonde la conception. Cette analyse aboutit à un cahier des charges linguistique issu des pratiques réelles des usagers. Cette recherche met en évidence quatre apports principaux. Le premier est un cadre de médiation linguistique pour la conception d'agents conversationnels plurilingues. Le deuxième consiste à repenser l'analyse des besoins comme outil de conception de produits linguistiques numériques. Le troisième offre une cartographie exploratoire des besoins langagiers d'étudiants algériens en contexte plurilingue et montre comment ceux-ci peuvent être traduits en spécifications conversationnelles. Le quatrième suggère enfin que, dans un écosystème d'innovation, les choix langagiers peuvent influencer l'adoption du produit et sa viabilité entrepreneuriale.

Le présent article est structuré en quatre parties. La section 1 développe le cadre théorique en articulant l'analyse des besoins en FOS, l'entrepreneuriat linguistique et la médiation linguistique à vecteur inversé. La section 2 expose le contexte algérien, le terrain NAFSIATI et le protocole méthodologique. La section 3 est consacrée à l'analyse sociolinguistique du corpus et à la déduction des spécifications conversationnelles. Enfin, la section 4 discute les implications méthodologiques, théoriques, sociolinguistiques et entrepreneuriales des résultats obtenus.

1. Cadre théorique

1.1. L'analyse des besoins en FOS : un acquis et son orientation

En didactique des langues sur objectifs spécifiques, l'analyse des besoins fait figure d'évidence partagée. Mangiante et Parpette (2004) en ont donné, pour le français, la formulation la plus nette. Avant d'élaborer le moindre contenu, on commence par décrire les situations de communication visées. On s'intéresse ensuite aux discours qui s'y échangent et aux compétences qu'ils requièrent. Le programme se déduit de cette description ; il n'existe pas en amont d'elle. La tradition anglophone a formalisé cette démarche à travers une série de modèles complémentaires : analyse de la situation cible, de la situation présente et de l'environnement d'apprentissage (Brown, 2016 ; Long, 2005). Elle a par ailleurs connu une évolution notable. On ne conçoit plus l'analyse des besoins comme un relevé strictement linguistique effectué une fois pour toutes. Elle est désormais envisagée comme un processus situé et dynamique, attentif à la pluralité des acteurs et à la singularité des contextes.

Un trait demeure toutefois inchangé et nous importe particulièrement ici : son orientation. Dans toutes ses variantes, l'analyse des besoins en FOS est orientée dans un même sens. Elle part de l'usager pour le conduire vers une langue de spécialité qu'il ne maîtrise pas encore (celle du droit, de la médecine, de l'entreprise). Le besoin y est toujours un manque à combler par l'apprentissage, et la finalité, toujours

pédagogique. Or, c'est cette orientation, et non l'outil lui-même, que la conception d'un produit numérique vient renverser.

1.2. La langue comme ressource entrepreneuriale

Si la conception d'un produit déplace l'analyse des besoins, c'est qu'elle l'inscrit dans un autre cadre : celui de l'entreprise. Or les rapports entre langue et entrepreneuriat ont fait l'objet, ces dernières années, d'une conceptualisation propre. De Costa, Park et Wee (2016) ont proposé la notion d'entrepreneuriat linguistique (*linguistic entrepreneurship*) pour décrire la manière dont les compétences langagières sont désormais envisagées : un capital à cultiver et à faire fructifier. La notion a d'abord une portée critique, puisqu'elle interroge l'injonction néolibérale faite aux individus de se rendre « employables » par la langue. Elle possède cependant aussi une valeur descriptive : dans une économie de la connaissance, savoir mobiliser plusieurs langues est devenu une ressource entrepreneuriale à part entière. Cette lecture prolonge les analyses de la commodification de la langue dans le capitalisme tardif (Heller, 2010).

L'essentiel de cette littérature envisage cependant la langue du côté de l'entrepreneur. En effet, c'est lui qui doit maîtriser l'anglais des affaires, négocier en contexte interculturel, communiquer avec des partenaires étrangers. Notre objet impose un déplacement. Dans le cas d'un produit conversationnel, la ressource langagière décisive ne se situe pas chez l'entrepreneur, mais dans la relation entre le produit et son usager. L'enjeu n'est pas que le concepteur parle la langue du marché, mais que le produit parle celle de l'utilisateur. Ce déplacement, de la langue de l'entrepreneur à la langue inscrite dans le produit, appelle un concept que la littérature existante ne fournit pas encore.

1.3. La médiation linguistique à vecteur inversé comme cadre de conception

Pour nommer l'opération par laquelle un concepteur fait parler son produit dans la langue de l'usager, nous mobilisons la notion de médiation. Le Cadre européen commun de référence en a fait, surtout depuis son volume complémentaire, une activité langagière à part entière. Médier, c'est transposer un sens d'un code, d'un registre ou d'un domaine vers un autre, pour le rendre accessible à qui ne l'atteindrait pas autrement (Conseil de l'Europe, 2001, 2018). La médiation suppose donc trois termes : une source spécialisée, un destinataire et un agent qui opère le passage, traditionnellement un enseignant, un traducteur ou un dispositif pédagogique.

Concevoir un agent conversationnel reconduit ces trois termes, mais en les redistribuant. La source spécialisée existe bel et bien : c'est le discours de la psychologie clinique, des thérapies cognitivo-comportementales et de l'accompagnement émotionnel, avec son lexique, ses protocoles, ses formulations. Le destinataire est l'étudiant, avec son répertoire plurilingue ordinaire. Mais l'agent du passage n'est plus une personne ni un cours. Il s'agit désormais de l'architecture conversationnelle du produit, qui doit reconnaître, accueillir et mobiliser les codes réellement employés par l'usager. Précisons le statut de cette expression. L'architecture conversationnelle est envisagée ici comme l'instance qui opère la médiation, non comme un objet technique. Sa description informatique excède le propos du présent article, qui porte sur les spécifications linguistiques appelées à l'alimenter. Surtout, le sens du passage s'inverse. Là où la médiation didactique élève l'apprenant vers la langue de spécialité, la médiation de conception fait redescendre la langue de spécialité vers les pratiques ordinaires, hybrides et plurilingues de l'usager. Nous proposons de désigner cette opération par l'expression médiation linguistique à vecteur inversé. Cette notion est utilisée ici comme un cadre interprétatif. Elle permet de décrire une configuration de conception dans laquelle l'adaptation s'effectue du savoir spécialisé vers les pratiques ordinaires des usagers.

Il importe de distinguer cette forme de médiation linguistique de trois opérations voisines. Elle ne relève pas de la simple traduction, qui transfère un contenu d'une langue à une autre sans en modifier le registre ni les destinataires. Elle ne se confond pas davantage avec la localisation, celle-ci adaptant un produit aux conventions d'un marché donné. En effet, la localisation présuppose en effet une langue cible relativement stable, tandis que la médiation envisagée ici s'adresse à des répertoires plurilingues

caractérisés par l'alternance codique et la coexistence de variétés non standardisées. Elle ne se réduit pas non plus à la vulgarisation, qui simplifie un contenu spécialisé sans nécessairement l'ancrer dans les pratiques langagières des usagers. Sa spécificité réside dans le mouvement qui conduit d'un savoir spécialisé vers les usages ordinaires, en prenant appui sur une analyse préalable des besoins linguistiques. On proposera donc la définition suivante, qui servira de cadre opératoire pour l'ensemble de cette étude. La médiation linguistique à vecteur inversé désigne l'opération par laquelle un dispositif conversationnel adapte un savoir spécialisé aux pratiques langagières ordinaires, hybrides et plurilingues de ses usagers, à partir d'une analyse préalable de leurs besoins. Elle inverse ainsi l'orientation classique de la médiation didactique, qui conduit l'apprenant vers la langue de spécialité.

Cette nécessité n'est pas spéculative. Les travaux ayant co-conçu des agents conversationnels de bien-être avec des locuteurs plurilingues montrent que ces derniers régulent leurs émotions de préférence dans leur première langue, qu'ils recourent au code-switching et qu'ils jugent sévèrement les réponses qui ignorent ces usages (Elwahsh et al., 2025). Les synthèses du domaine confirment quant à elles que l'adaptation linguistique et culturelle reste le point faible de ces systèmes, en particulier pour les langues à faibles ressources (Algumaei et al., 2025). Ces résultats gagnent à être interprétés à travers le cadre de la médiation linguistique, en soulignant le déplacement qui conduit à adapter une langue de spécialité aux pratiques ordinaires des usagers.

Ainsi définie, cette forme de médiation linguistique déborde potentiellement le seul cas de *NAFSIATI*. Tout dispositif conversationnel mobilisant un savoir spécialisé (médical, juridique ou administratif) est susceptible de rencontrer des enjeux comparables d'adaptation linguistique et culturelle. Le cas *NAFSIATI* ne permet pas d'établir la généralité du modèle, mais il offre un premier terrain d'observation. Le corpus ne permet d'ailleurs pas d'éprouver le concept lui-même, qui conserve un statut heuristique. Il permet plutôt d'examiner l'hypothèse de recherche qu'il encadre : la possibilité de convertir une analyse des besoins langagiers en spécifications de conception. D'autres études seront nécessaires pour en éprouver la portée dans des contextes différents.

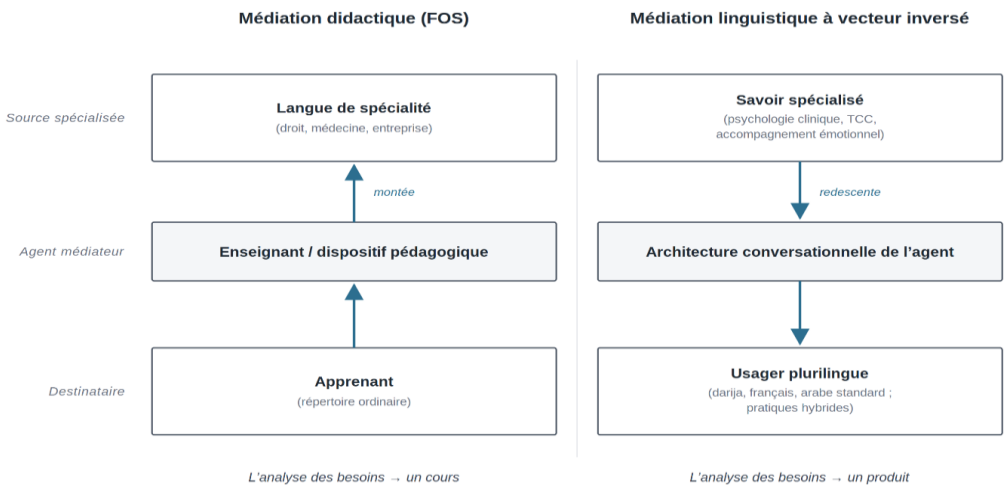


Figure 1. Inversion du vecteur d'adaptation : de l'apprenant vers la langue de spécialité (FOS) à la langue de spécialité vers les pratiques plurilingues de l'usager.

1.4. Situer la médiation à vecteur inversé face aux approches de conception et de TAL

La notion proposée ici croise plusieurs cadres issus de la conception et de l'informatique, sans toutefois s'y confondre. Il convient d'en préciser les rapports. Le design centré utilisateur (Norman et Draper, 1986) et,

plus récemment, l'intelligence artificielle centrée sur l'humain (Shneiderman, 2022) partagent avec la médiation à vecteur inversé un même principe d'orientation : c'est au dispositif de s'ajuster à l'humain, non l'inverse. Ces cadres définissent toutefois l'usager par ses buts, ses tâches et ses capacités cognitives. Le répertoire linguistique n'y figure qu'à titre de paramètre d'interface, quand il n'est pas réduit à une simple préférence de langue. La médiation à vecteur inversé peut ainsi être décrite comme la spécification linguistique de ce principe général. Elle porte l'exigence d'ajustement au niveau du registre, du répertoire et de l'alternance codique, précisément ce que les approches centrées utilisateur laissent indéterminé.

Les notions issues de l'ingénierie des systèmes conversationnels s'en distinguent quant à elles par leur niveau d'intervention. Le *prompt engineering* et les techniques d'adaptation des modèles de traitement automatique des langues opèrent au moment de l'implémentation. Ils supposent donc déjà établi ce que le système doit accueillir, précisément ce que l'analyse des besoins détermine en amont. L'intelligence artificielle explicable et l'expérience utilisateur conversationnelle (Følstad et Brandtzæg, 2017) portent, pour leur part, sur la transparence des décisions du système et sur la qualité de l'interaction, non sur l'opération qui transpose un savoir spécialisé vers des pratiques ordinaires. La médiation linguistique à vecteur inversé ne concurrence donc aucune de ces notions. Elle se situe en amont, au niveau du principe de conception, là où se décident les contraintes que ces instruments auront ensuite à mettre en œuvre.

Ce niveau d'instrumentation a été profondément transformé par les architectures neuronales récentes. Il convient d'en rappeler brièvement l'évolution afin de situer le dialogue que cette étude engage avec le traitement automatique des langues. Depuis l'architecture Transformer (Vaswani et al., 2017), les modèles préentraînés de type BERT (Devlin et al., 2019), puis les grands modèles génératifs inaugurés par GPT-3 (Brown et al., 2020), ont fait de la conversation en langue naturelle une capacité générique des systèmes. Cette généralisation reste toutefois inégale : les performances de ces modèles se concentrent sur les langues à fortes ressources, et leur évaluation en contexte plurilingue demeure lacunaire (Algumaei et al., 2025). L'arabe dialectal illustre exemplairement cette asymétrie. Un modèle dédié au dialecte algérien existe certes, DziriBERT, entraîné sur un million de tweets en graphies arabe et latine (Abdaoui et al., 2021). Il atteste qu'une modélisation de cette variété peu dotée est possible. Reste que de telles architectures reposent sur une segmentation des énoncés en sous-unités apprises sur les données d'entraînement. Les formes présentes dans notre corpus, où un lexème français se conjugue selon la morphologie de la darija, ont d'autant plus de chances d'être correctement traitées qu'elles auront été identifiées et documentées en amont. C'est précisément cette documentation sociolinguistique préalable que l'analyse des besoins fournit.

Ce détour par l'état de l'art précise la place respective des deux ordres de considérations. Les grands modèles de langue et leurs adaptations dialectales constituent l'instrumentation dont disposera l'implémentation de l'agent. L'analyse des besoins fournit ce que cette instrumentation ne peut produire elle-même : la connaissance des registres, des répertoires et des formes hybrides que le système devra accueillir, ainsi que leur hiérarchisation en invariants et en paramètres. Le traitement automatique des langues est ici l'instrument de mise en œuvre ; la médiation linguistique fondée sur l'analyse des besoins en demeure le principe structurant.

2. Contexte et méthodologie

2.1. Un écosystème entrepreneurial et son défi langagier

NAFSIATI est un projet d'innovation universitaire porté dans le cadre du CATI (Centre d'appui à la technologie et à l'innovation) de l'Université Abbès Laghrour de Khenchela, consacré à la conception d'un agent conversationnel de soutien psychologique pour les étudiants. Il s'inscrit dans l'écosystème entrepreneurial et d'innovation encouragé par les politiques publiques algériennes, du Label Startup Algeria (décret exécutif n° 20-254 du 15 septembre 2020) aux dispositifs universitaires de transfert et de maturation de projets. Dans cette

perspective, les besoins langagiers ne sont pas seulement des données sociolinguistiques ; ils constituent des informations stratégiques pour le développement du produit. Un agent incapable de reconnaître les pratiques linguistiques réelles de ses utilisateurs risque de voir son utilité perçue diminuer, ce qui compromet son appropriation et, à terme, sa pérennité dans l'écosystème d'innovation auquel il appartient. C'est ce constat qui a motivé, en amont de tout développement, la constitution d'un corpus sociolinguistique : non pour décrire l'état psychologique des étudiants, mais pour documenter les ressources langagières qu'un tel produit devra savoir mobiliser.

2.2. Le dispositif de conception : une démarche interdisciplinaire

La présente étude documente une étape déterminée d'une démarche de conception : l'analyse des besoins langagiers. Elle ne porte ni sur l'implémentation informatique de l'agent ni sur son évaluation technique, conformément à la délimitation méthodologique posée en 1.3. La démarche d'ensemble mérite néanmoins d'être explicitée, car elle éclaire le statut exact des résultats présentés. Après l'acceptation du projet par le CATI, *NAFSIATI* s'est doté d'un encadrement interdisciplinaire associant les sciences du langage et l'informatique. La dimension langagière du dispositif et son architecture computationnelle y sont conçues comme deux volets distincts mais articulés d'une même démarche. Le projet est conduit par deux étudiantes, qui ont assuré la collecte du corpus selon le protocole décrit en 2.4.

La conception suit un cycle en quatre phases. La première, dite de cadrage, a défini les finalités du dispositif et son inscription dans l'écosystème d'innovation universitaire. La deuxième phase, dont le présent article rend compte, constitue la méthodologie de conception linguistique. Elle suit une chaîne explicite : recueil des besoins langagiers auprès des usagers visés, analyse sociolinguistique du corpus selon six axes, dégagement de catégories de besoins, traduction de ces catégories en spécifications linguistiques hiérarchisées distinguant invariants et paramètres (section 3.4), puis formulation de leurs implications pour la conception de l'agent. Ces spécifications constituent l'interface entre les deux volets du projet. Elles forment le cahier des charges linguistique sur lequel s'appuie le volet informatique. La troisième phase, actuellement en cours, relève de ce second volet. Un prototype de l'agent est en voie de finalisation sur la base de ces spécifications. La quatrième phase portera sur les tests utilisateurs et la confrontation des spécifications aux usages réels. Elle fournira la validation expérimentale que la présente étude prépare sans la prétendre accomplir.

2.3. Le terrain plurilingue algérien

L'étudiant algérien évolue dans un répertoire pluriel. À la darija, langue première de l'oralité quotidienne, s'ajoutent le français, langue d'une part importante des études supérieures, et l'arabe standard, langue de l'écrit formel. Ces variétés ne se succèdent pas, elles s'entrelacent : l'alternance codique darija/français est un trait ordinaire de la parole estudiantine, dont la sociolinguistique algérienne a décrit aussi bien les formes que les fonctions (Amazouz et al., 2017). Les ressources computationnelles existantes pour la darija - des corpus annotés souvent en graphie latine (Arabizi) et porteurs d'annotations émotionnelles (Abainia, 2020) - confirment à la fois l'intérêt et la difficulté du traitement automatique de cette variété peu dotée. La spécialité d'études n'est sans doute pas neutre dans ce paysage : on peut faire l'hypothèse que, là où le cursus est largement francophone, le français gagne du terrain jusque dans l'expression de l'expérience vécue. C'est précisément ce qui justifie que notre échantillon couvre plusieurs filières.

2.4. Constitution du corpus : instrument, échantillon, éthique

Les données ont été recueillies par un questionnaire bilingue (français / arabe) de neuf questions, combinant trois registres d'items. Des questions fermées portent sur l'intensité du stress ressenti, les principales sources de stress, la qualité du sommeil, les stratégies adoptées et les usages attendus d'un éventuel agent conversationnel. Une question ouverte et une série de phrases à compléter sollicitent une parole spontanée sur les préoccupations actuelles et les formulations que les étudiants se répètent lorsque leur état se dégrade. Deux items, enfin, portent

sur l'indicible et sur l'attente : ce que l'étudiant aimerait pouvoir dire sans oser le faire, et la formulation qu'il souhaiterait entendre d'un agent. Un choix de conception de l'instrument mérite d'être souligné : la consigne invite explicitement à répondre dans la langue où l'on s'exprime le mieux. Le choix de langue n'est donc pas imposé, il est observé et devient ainsi une donnée, non une contrainte.

Le questionnaire, en format papier, a été administré en présentiel en avril 2026 auprès d'étudiants de l'Université Abbès Laghrour de Khenchela inscrits dans les quatre filières retenues. L'administration a été assurée par les deux étudiantes porteuses du projet, qui se sont déplacées dans les facultés concernées. La présentation orale de l'enquête a été conduite en français et en arabe, en cohérence avec le caractère bilingue de l'instrument. Elle permettait de clarifier les items lorsque cela était nécessaire. Aucun critère d'inclusion autre que l'inscription dans l'une des quatre filières et la participation volontaire n'a été appliqué.

Cet instrument se distingue des modules cliniques du produit lui-même (auto-évaluations standardisées de type PHQ-9 ou GAD-7) : sa finalité n'est pas diagnostique mais langagière. Soixante et onze réponses ont été réunies auprès d'étudiants de quatre filières, dont la répartition est donnée ci-dessous.

Filière	Nombre des étudiants
Biologie	16
Médecine	20
Sciences et technologies	20
Droit	15
Total	**71**

Tableau 1. Composition de l'échantillon (n = 71)

Il s'agit d'un échantillon exploratoire : nous n'en tirons aucune prétention épidémiologique. La diversité des filières y est traitée comme une donnée de segmentation : non comme une variable explicative des comportements, mais comme un indicateur de profils d'usages susceptibles d'orienter la conception conversationnelle du produit. La participation volontaire relève d'un échantillon de convenance. L'auto-sélection qui en résulte peut introduire un biais : les étudiants disposés à parler de leur détresse diffèrent possiblement de ceux qui s'y dérobent. Cette sélection limite la transférabilité des observations, sans toutefois invalider leur intérêt pour l'analyse des besoins. Le corpus est par ailleurs de nature déclarative et élicitée. Il documente une mise en mots sollicitée par questionnaire, non le langage spontané d'une demande d'aide en situation. Les spécifications qui en découlent valent comme hypothèses de conception, à éprouver ensuite face aux usages réels. L'administration présentielle assistée pourrait enfin avoir influencé les formulations déclarées, notamment en raison du contexte institutionnel de passation. Cette possibilité est retenue comme une limite interprétative.

La collecte a respecté un protocole éthique formalisé. Chaque participant s'est vu attribuer un code anonyme, jamais relié à des données nominatives, et a donné un consentement éclairé. Les réponses sensibles ont été stockées de manière sécurisée et chiffrée. Ce cadre n'est pas un ornement méthodologique : il conditionne la légitimité même d'une recherche qui recueille la parole de détresse. En pratique, aucune réponse n'a présenté de signaux de détresse aiguë justifiant l'activation du protocole d'alerte.

2.4. Une lecture sociolinguistique

L'analyse procède sur deux plans. Les items ouverts et les phrases à compléter forment le corpus proprement dit. Celui-ci est examiné selon six axes : langue-matrice de l'affect, hybridité et mixage, registre de consolation, registre attendu de l'agent, texture affective, indicible et confidentialité. Chacun de ces axes est conçu pour

déboucher sur une spécification de conception. La procédure complète, les conventions de transcription et la matrice de codage sont détaillées dans le protocole reproduit en annexe.

Un principe gouverne cette lecture : nous lisons le corpus en linguistes. Ce qui nous retient n'est pas ce que les réponses révéleraient de l'état psychique des étudiants, mais la manière dont la détresse y est mise en mots : quelle langue, quel registre et quelles bascules. Nous ne formulons aucune affirmation sur la prévalence de tel ou tel trouble ; notre objet est la verbalisation, non sa fréquence clinique. Compte tenu de la taille et de la nature micro-textuelle et plurilingue du corpus, l'analyse est qualitative et manuelle. Le recours à un logiciel de lexicométrie est écarté, car la faiblesse des effectifs rendrait les mesures instables et la normalisation automatique risquerait d'effacer les traits porteurs. Le critère de suffisance n'est pas la représentativité, mais la saturation. Le codage, conduit de manière itérative, a visé la stabilisation des six axes. Les dernières réponses n'introduisaient plus de catégorie nouvelle. L'analyse a été menée manuellement par une seule analyste. Faute de double codage, la fiabilité repose sur l'explicitation du protocole (annexe B) et sur une posture réflexive assumée. Analyser, depuis une position de chercheuse francophone, une parole de détresse souvent formulée en darija expose au risque de normaliser ou de surinterpréter les formes hybrides. La conservation de la graphie d'origine et le refus de la normalisation lexicométrique visent précisément à contenir ce risque.

3. Analyse : des besoins langagiers aux spécifications conversationnelles

L'analyse porte sur soixante et onze réponses : biologie (16), médecine (20), sciences et technologies (20), droit (15). Il s'agit de voir comment ces besoins langagiers fondent les choix de conception de l'agent : quels traits s'imposent pour l'ensemble des usagers, et lesquels doivent se régler selon la filière. Les proportions énoncées sont des tendances, non des mesures. Conformément au protocole, chaque axe conduit de l'observation langagière à une spécification de conception. Ces deux plans sont analytiquement distincts. D'une part, la description sociolinguistique documente ce que les usagers font avec la langue. D'autre part, la spécification de conception prescrit ce que le système devra faire en retour. Le passage de l'un à l'autre n'est pas automatique. Il est gouverné par le principe de la médiation à vecteur inversé. Ainsi, tout trait récurrent dans le corpus devient soit un invariant obligatoire du système, soit un paramètre configurable selon le profil de l'usager. C'est cette opération de traduction que chaque axe met en œuvre.

3.1. Un répertoire hybride et stratifié

Le questionnaire, bilingue, laissait à chacun le choix de la langue. La première chose que révèlent les réponses n'est pas la langue elle-même, mais ce qu'elle prend en charge. Dès qu'il s'agit de dire l'émotion (Q3, Q5, Q7), c'est l'arabe sous forme de darija seule ou mêlée d'arabe standard qui s'impose. Sur l'ensemble du corpus, environ les deux tiers des répondants (47/71) formulent leur vécu émotionnel en darija ou en arabe standard. Cette tendance atteint son pic dans les filières arabophones (biologie, droit), où elle est quasi exclusive, et s'atténue en médecine, où elle reste majoritaire sans plus s'imposer seule. « نخاف فرات » (j'ai peur, énormément - biologie), « الخلعة و ضيق في صدري » (la frayeur, une oppression dans la poitrine - médecine), « يضرني راسي و نولي قلقان » (ma tête me fait mal et je deviens anxieux - droit) constituent trois énoncés, issus de trois filières différentes. Ils partagent un trait commun qui dépasse le simple choix de langue. La darija n'est pas convoquée parce qu'elle est disponible puisque le français l'est tout autant pour ces répondants mais parce qu'elle est le code dans lequel l'intensité émotionnelle trouve sa formulation la plus directe. C'est précisément là que réside la distinction entre langue utilisée et langue de l'affect, à savoir que la darija est la langue de l'affect (Pavlenko, 2005 ; Dewaele, 2010).

Cette dominance n'est pourtant pas un trait stable de la population. En effet, elle varie avec la filière, et se relâche là où le cursus est le plus francophone. En médecine, environ un tiers des répondants (6-7/20) disent leur détresse directement en français, comme l'illustrent les formulations suivantes : « Je me sens beaucoup de stress au point que j'en veux faire rien que dormir », « Je veux l'immigration ou je veux

abandonner la médecine », « Un mélange de fatigue profonde et agitation intérieure ». La langue d'enseignement ne déplace donc pas seulement le mixage. Pour une partie des locuteurs, elle déplace la langue-matrice elle-même de l'émotion (Oostendorp et al., 2024). On reconnaît là un effet de la répartition diglossique (Albirini, 2011), avec le français et l'arabe standard pour la sphère savante et la darija pour la sphère intime, modulé par l'exposition curriculaire.

Le même axe gouverne une seconde variable : non plus quelle langue porte l'émotion, mais comment les langues se combinent dans l'énoncé. Lorsque le français entre, c'est rarement comme langue autonome. Il s'intègre à la phrase arabe, parfois jusqu'à la morphologie. « نـديقوتي » (je suis dégoûté, de dégoûter), « نريفيزي » (je révise), « ندير نوي بلونش » (je fais une nuit blanche) et « نديبريسي » (je déprime) illustrent cette hybridité. Le lexème est français, la conjugaison darija. Cette morphologie mixte est analytiquement décisive : le locuteur ne passe pas d'une langue à l'autre. Il opère dans un répertoire intégré unique où les ressources françaises et arabes sont activées simultanément au sein d'un même acte énonciatif. Il ne s'agit pas d'une alternance codique au sens classique. Cela montre que les frontières entre langues nommées sont moins opératoires pour l'usager que pour le système de traitement qui cherchera à l'analyser. La médecine et les sciences et technologies poussent l'hybridité jusqu'à l'anglais et la graphie latine, comme l'illustrent « Trust the process », « Stay hard », « Nrml ». De nouveau l'épaisseur du mixage croît avec la langue d'enseignement. Dans le droit, filière arabophone, elle se réduit à un résidu lexicalisé, comme « ستراس » (stress) et « ميموار » (mémoire), déjà intégrés à la darija. La médecine et les sciences ajoutent par-dessus toute une couche portée par le cursus (réviser, nuit blanche, dégoûter conjugués). En ordre de grandeur, les énoncés hybrides (darija + insertions françaises ou anglaises) représentent moins d'un tiers des réponses expressives en droit et biologie, contre plus de la moitié en médecine et sciences et technologies. La couche infra-lexicale (lexème français conjugué à la darija) concerne quasi exclusivement ces deux dernières filières (médecine + ST, soit 40/71).

Deux observations, donc, mais un seul ressort. Le répertoire mobilisé est à la fois hybride et stratifié. Il est hybride parce que le français se coule dans la matrice arabe jusqu'au niveau du *token*. Il est stratifié parce que son épaisseur, y compris dans sa fonction affective, dépend de la filière. Pour la conception, les deux niveaux se lisent ensemble. Un moteur monolingue échoue d'emblée. De même, un moteur qui se contenterait de router chaque segment vers sa langue échouerait, car la bascule est infra-lexicale et doit être traitée sur le *token* hybride lui-même. Dans les filières arabophones, l'affect se dit presque exclusivement en darija. L'agent doit donc comprendre et accueillir cette variété pour ne pas manquer le signal. Il doit également pouvoir recevoir le français-matrice là où la filière le fait apparaître. Le degré de mixage n'est dès lors pas une donnée fixe à absorber, mais un paramètre à régler selon le public visé.

3.2. Les registres de l'émotion

Savoir en quelles langues se dit la détresse ne dit pas encore comment elle s'y inscrit. Une fois le code identifié, une autre question se pose, celle de savoir par quels moyens l'émotion encode-t-elle son sens et son intensité ? Les réponses dessinent deux canaux que le lexique seul ne suffit pas à capter et qu'un système naïf tendrait précisément à effacer.

Le premier donne à la souffrance un sens. À la question de ce qu'ils se disent quand ça va mal (Q5f), une nette majorité des répondants, dans les quatre filières (environ les trois quarts, 53/71), mobilise un registre religieux ou providentialiste de remise à Dieu : « شدة و تفوت » (une épreuve, et ça passe), « فيها خير إن شاء الله », « ان مع العسر يسرا », « يفرج ربي », « اقرأ الادعية » (je récite des invocations). Ce que ces formules partagent dépasse la langue. Elles convoquent toutes un cadre interprétatif préexistant dans lequel l'épreuve est dotée d'un sens : elle est traversable, temporaire et inscrite dans un ordre qui la dépasse. L'étudiant ne cherche pas à supprimer la détresse. Il la recadre. Ce recadrage est une ressource de résilience déjà active, antérieure à toute intervention

externe. Ce ne sont donc pas des ornements, mais des idiomes de détresse, qui cadrent l'épreuve avant même qu'on cherche à la soulager.

Le second donne à la souffrance son intensité, et il passe par la forme plus que par le sens. L'intensité s'inscrit d'abord dans la graphie, par la répétition de lettres : « مخي شاااعل » (mon cerveau est en feeeu), « تعبت انااااا » (je suis épuisée, moiii), « نخاف بزاااف » (j'ai très peuuur). Elle passe aussi par le corps : « ضيق في التنفس » (difficulté à respirer), « وجع في الراس » (mal de tête) et « مخي نحب نطفيه » (mon cerveau, je veux l'éteindre). Elle passe parfois par le rire, qui amortit l'aveu : « تعبت يا عالم هههه » (je suis épuisée, oh les gens, hahaha). Ces formes ne sont pas des approximations graphiques. L'allongement de la chaîne de caractères constitue un système de gradation para-prosodique à l'écrit. Le locuteur module ainsi l'intensité de son affect par la longueur même du signe. De même, le rire final n'invalide pas la détresse. Il l'amortit socialement et permet l'aveu sans l'exposer. Effacer l'un ou l'autre, c'est effacer le signal lui-même.

Dans les deux cas, la même tentation guette la conception, et le même coût en découle. Un dispositif dont les scripts de soutien dériveraient d'une thérapie cognitivo-comportementale séculière entrerait en dissonance avec le registre providentialiste. Une normalisation automatique du texte ramènerait « تعبتتتت » à « تعبت », effaçant le marqueur d'intensité. Sécularisation d'un côté, assainissement de l'autre. Dans les deux cas, un cadre étranger est imposé à l'énoncé, au risque d'en manquer le signal. Le principe de conception est donc unique : préserver et résonner plutôt que recadrer. Il se décline en deux spécifications. Sur le plan du sens, les scripts de soutien doivent résonner avec le système de consolation préexistant. La spécification est ainsi culturelle autant que linguistique. Sur le plan de l'intensité, l'agent doit préserver et interpréter l'allongement graphique et l'expression somatique. Ceux-ci constituent des indices de sévérité susceptibles d'alimenter la logique d'escalade. Il doit toutefois les pondérer sans sur-réagir ni les minimiser. Faire redescendre la langue de spécialité, c'est aussi en garder la texture affective plutôt que de l'assainir.

3.3. Attentes envers l'agent

Sur la formule attendue (Q9), les répondants dessinent une demande dont le noyau est constant : ne pas être jugé et ne pas être bercé de fausses consolations. Plusieurs réponses l'illustrent : « ما يحكمش عليا » (qu'il ne me juge pas), « لا اريد ان يكذب علي و يحفزني » (qu'il ne me mente pas en me motivant) et « ما يستهزأش بمشاكلي » (qu'il ne se moque pas de mes problèmes). Le profil le plus francophone l'exprime en français : « Restez neutre, précis, no drama et surtout ne me jugez pas ».

Au-delà de ce socle partagé, environ trois répondants sur cinq (43/71) privilégient un registre franc, logique et neutre, comme l'illustre la formulation « نصائح … بدون التعاطف معي » (des conseils, sans compassion). Environ un tiers du corpus (24/71) exprime au contraire une attente de soutien et de réconfort : « يديرلي نحب يواسيني », « كوراج » (qu'il me donne du courage). Ce que les répondants partagent, ce n'est donc pas le refus de l'empathie, mais le refus du jugement et de la consolation factice. Ainsi, la franchise apparaît comme une préférence majoritaire, non comme une exigence unanime.

À cela s'ajoutent deux faits. La confidentialité émerge comme une exigence explicite chez environ un répondant sur quatre (18/71), surtout en médecine et en sciences et technologies. Elle apparaît notamment dans les réponses suivantes : « المحافظة على الخصوصيات », « عدم التدخل في الخصوصيات ». Par ailleurs, l'indicible affleure dans plusieurs réponses : « لا استطيع وصف حالتي النفسية » (je ne peux pas décrire mon état), « ما خفي في قلبي اعظم » (ce qui est caché dans mon cœur est plus grand), tandis que d'autres affirment au contraire tout pouvoir dire.

Les spécifications qui en découlent sont triples. Il convient d'abord de privilégier un registre franc et non jugeant plutôt que compatissant. Il faut ensuite prévoir une garantie de confidentialité explicite et visible. Enfin, pour l'indicible, l'agent doit proposer des amorces et une aide à la mise en mots, comme le faisait déjà le questionnaire avec ses phrases à compléter.

Le tableau suivant synthétise, filière par filière, les profils langagiers dégagés aux sections 3.1 à 3.3. Les qualificatifs retenus expriment des tendances observées dans le corpus. Ils ne constituent pas des mesures. Ce récapitulatif vise à rendre visible d'un coup d'œil ce qui varie selon la filière et ce qui demeure transversal. La section 3.4 en tirera les conséquences pour la conception.

Dimension	Biologie (n=16)	Médecine (n=20)	ST (n=20)	Droit (n=15)
Langue-matrice de l'affect (Q3/Q5/Q7)	Darija dominante	Darija + français coprésents	Darija dominante	Darija exclusive
Hybridité et mixage	Résiduelle (couche lexicalisée)	Dense (couche curriculaire + infralexicale)	Dense (couche curriculaire + anglais)	Minimale (emprunts figés)
Registre de consolation (Q5f)	Religieux / providentialiste	Religieux / providentialiste	Religieux / providentialiste	Religieux / providentialiste
Attente envers l'agent (Q9)	Franchise, non-jugement	Neutralité, précision	Franchise, non-jugement	Franchise, non-jugement
Confidentialité (Q7)	Présente	Explicite et récurrente	Explicite et récurrente	Présente

Tableau 2. Profils langagiers par filière — synthèse qualitative (n = 71)[1]

Ce tableau fait apparaître deux lignes de partage nettes. La ligne horizontale qui sépare le registre de consolation des autres dimensions montre que le providentialisme est le seul trait strictement transversal, identique dans les quatre filières. La ligne verticale qui sépare la médecine des filières arabophones montre que c'est là que se concentrent les variations les plus significatives : langue-matrice, densité de l'hybridité et attente envers l'agent. C'est précisément cette double articulation, entre invariant transversal et variable curriculaire, que la section suivante traduit en contraintes hiérarchisées de conception.

3.4. Invariant et variable de conception

La comparaison des filières permet de répondre à la question posée, à condition de calibrer le statut de chaque trait. Un trait domine sans être strictement invariant : la darija comme langue de l'affect. Elle est quasi exclusive dans les filières arabophones et s'atténue en médecine, où plusieurs étudiants disent leur détresse en français. Ces traits fondent des contraintes de conception communes à toutes les filières.

En revanche, ce qui suit l'exposition curriculaire varie. Il s'agit notamment du degré de mixage français/anglais et de la langue-matrice de l'émotion elle-même, qui se déplace vers le français en médecine. Ces éléments constituent autant de paramètres réglables par segment. L'analyse des besoins identifie ainsi ce que l'agent doit faire pour tous et ce qu'il doit pouvoir moduler selon le public.

Observation langagière	Statut	Spécification de conception
Détresse dite en darija — dominante, quasi exclusive en filières arabophones ; français-matrice en médecine	Tendance forte modulée par la filière	Comprendre et produire la darija (et le français en milieu francophone) ; condition de perception du signal
Formulations religieuses ou providentialistes de réassurance	Convergence transversale forte	Bibliothèque de scripts compatible avec ce registre, non avec une TCC séculière

[1] Les qualificatifs expriment des tendances observées dans le corpus ; ils ne constituent pas des mesures.

Refus du jugement et de la fausse consolation ; franchise majoritaire, soutien pour une partie	Socle partagé + hétérogénéité	Registre non jugeant et authentique ; adaptativité franchise / soutien
Mixage français/anglais croissant avec la langue d'enseignement	Variable	Traiter le *token* hybride ; régler le mixage par segment
Allongements et idiomes somatiques	Transversal	Préserver comme indices de sévérité ; ne pas normaliser ; escalade
Exigence de confidentialité ; aveux d'indicible	Transversal	Garanties de confidentialité explicites ; amorces de verbalisation

Tableau 3. Des observations langagières aux spécifications de conception (n = 71 ; biologie 16, médecine 20, ST 20, droit 15)

4. Discussion

Au terme de cette analyse, une idée se dégage : la conception d'un agent conversationnel plurilingue relève moins du traitement automatique des langues que d'un travail de médiation linguistique, fondé sur l'analyse des besoins des usagers. Cette idée dépasse largement le cas de NAFSIATI. Dans cette configuration, le traitement automatique apparaît comme l'instrument de mise en œuvre, tandis que la médiation constitue le principe structurant de la conception. Cette inversion de perspective peut être discutée selon quatre registres : méthodologique, théorique, sociolinguistique et entrepreneurial.

Sur le plan méthodologique, l'étude prolonge tout en déplaçant la tradition du FOS. Là où l'analyse des besoins conduit habituellement à des dispositifs pédagogiques (Mangiante et Parpette, 2004), elle débouche ici sur des spécifications de système : architecture interactionnelle, choix linguistiques, paramétrage des échanges. Le geste analytique quitte ainsi le champ curriculaire pour entrer dans celui de la conception de produits numériques. Plus encore, l'opération ne se limite pas à une description des usages : elle institue une hiérarchisation des contraintes, distinguant invariants obligatoires et variables configurables. L'analyse cesse alors d'être un préalable descriptif pour devenir un opérateur de décision.

La distinction entre invariant et variable ne produit son plein effet de conception que lorsqu'elle est confrontée à un arbitrage. Par exemple, le registre franc et non jugeant (invariant transversal) et le registre religieux de consolation (convergence forte) peuvent coexister dans une bibliothèque de scripts, mais ils entrent en tension sur la posture interactionnelle de l'agent : doit-il adopter une neutralité séculière ou un ethos culturellement situé ? Le cadre conçu ici ne fournit pas directement la réponse. Il fournit une règle de priorité selon laquelle un invariant l'emporte sur une convergence, qui l'emporte elle-même sur une variable.

Ce déplacement engage directement la dimension entrepreneuriale du dispositif. Dans cette perspective, les choix langagiers apparaissent comme des choix de conception stratégiques. Ils ne déterminent pas seulement la qualité de l'interaction, mais également les conditions d'appropriation du dispositif par ses usagers. Pour les produits conversationnels fondés sur l'échange verbal, la langue constitue ainsi une ressource entrepreneuriale, inscrite dans le produit lui-même. Sa capacité à reconnaître les usages effectifs des utilisateurs peut représenter un facteur de différenciation, favoriser son adoption et contribuer, à terme, à sa viabilité. L'analyse des besoins ne relève donc plus d'un raffinement méthodologique mais d'un levier de conception stratégique.

Sur le plan théorique, la notion de médiation linguistique se trouve également déplacée. Le modèle classique repose sur une logique descendante : un savoir expert est reformulé en direction d'un public non spécialiste. Le corpus analysé suggère au contraire une inversion du point d'origine. Ce ne sont plus les usages qui s'ajustent au savoir, mais le dispositif expert qui se reconfigure à partir des pratiques ordinaires.

L'adaptation ne vise plus l'usager en tant qu'apprenant, mais le système en tant qu'entité ajustable aux pratiques langagières effectives.

Ce renversement donne son contenu empirique à la notion de médiation linguistique à vecteur inversé proposée en amont : l'orientation de l'adaptation y est redirigée vers le dispositif lui-même. L'analyse des besoins y occupe une fonction structurante, dans la mesure où elle rend visibles les régularités interactionnelles, les registres mobilisés et les formes hybrides auxquelles le système doit se conformer.

Sur le plan sociolinguistique, les résultats obtenus indiquent l'existence d'une stratification interne du français dans le répertoire des répondants. Ce français se distribue entre une couche lexicalisée, largement partagée, et une couche curriculaire, dépendante des trajectoires de formation et absente dans certains espaces institutionnels.

Plus profondément, les matériaux interrogent la pertinence même des frontières linguistiques. Les formes hybrides observées dans le corpus, telles que « نديبريسي », « نريفيزي » ou « نـدير نوي بلونش », combinent, au sein d'une même unité, des ressources lexicales françaises et des schémas morphologiques issus de la darija. De telles productions donnent à penser que les répondants ne mobilisent pas leurs langues comme des codes séparés, mais comme un répertoire intégré. Dans celui-ci différentes ressources peuvent être activées simultanément selon les besoins de l'interaction. Cette observation rejoint les approches du translanguaging (García & Wei, 2014 ; Otheguy, García, & Reid, 2015), qui invitent à considérer les pratiques plurilingues comme des usages dynamiques d'un répertoire unique plutôt que comme une juxtaposition de langues étanches. Dans cette perspective, les frontières linguistiques apparaissent moins comme une réalité directement observable dans les usages que comme une nécessité de modélisation propre aux systèmes de traitement. Pour la conception d'un agent conversationnel, l'enjeu n'est donc pas seulement de reconnaître plusieurs langues, mais de traiter des productions qui circulent continuellement entre elles.

Cette convergence avec le translanguaging soulève toutefois une tension que la conception ne peut éluder. Dans sa formulation forte, le translanguaging ne se borne pas à décrire des pratiques hybrides : il conteste la pertinence même des « langues nommées », tenues pour des constructions sociopolitiques plutôt que pour des réalités psycholinguistiques (Otheguy, García, & Reid, 2015). Or, l'architecture d'un agent conversationnel repose précisément sur ces objets nommés. Elle implique l'identification de la langue, des modèles et segmenteurs propres à chaque langue, ainsi que des données d'entraînement étiquetées par langue. Le cadre descriptif dissout ainsi ce dont le dispositif technique a besoin pour fonctionner. Cette contradiction n'est pas un obstacle, mais un révélateur. Le translanguaging caractérise le répertoire de l'usager, unitaire et intégré, tandis que la chaîne de traitement relève des contraintes de l'outil. L'enjeu de conception est alors de subordonner la machinerie des langues nommées à un usager qui ne les sépare pas. Il ne s'agit pas de filtrer l'entrée par une étape d'identification qui rabattrait l'énoncé sur une seule langue. Il s'agit plutôt de tolérer le mixage intra-énoncé et infralexical et d'entraîner le système sur des données réellement plurilingues et arabizi (Amazouz et al., 2017 ; Abainia, 2020). Ce déplacement constitue une nouvelle instance de la médiation à vecteur inversé : c'est le système qui plie vers la pratique hybride, et non l'inverse. À l'échelle de l'implémentation, les ressources étiquetées par langue demeurent incontournables. Le translanguaging y opère donc moins comme une architecture que comme une orientation régulatrice de la conception. Il impose au produit de s'adapter au mixage des usagers sans les contraindre à trancher entre leurs langues.

Ces résultats doivent enfin être rapportés aux limites de l'étude, qu'il convient de rassembler. L'échantillon, restreint à 71 répondants et recruté par convenance, ne permet aucune généralisation statistique. Le corpus, déclaratif et élicité, ne documente pas le langage spontané d'une demande d'aide réelle. Le codage, conduit par une seule analyste, n'a pas fait l'objet d'une validation intercodeur. L'administration présentielle assistée constitue une limite interprétative supplémentaire. Enfin, les spécifications produites n'ont pas encore reçu de validation expérimentale. Celle-ci relève de la quatrième phase du cycle de conception (section 2.2).

Conclusion

La question initiale portait sur la manière dont une analyse des besoins langagiers pouvait fonder les choix de conception d'un agent conversationnel plurilingue dans un contexte entrepreneurial. Cette analyse a été menée auprès des usagers visés. L'étude montre que cette opération repose sur un geste central : la transformation des pratiques langagières en un système de contraintes hiérarchisées, distinguant invariants structurants et variables paramétrables.

Ce geste, désigné comme médiation linguistique à vecteur inversé, consiste à déplacer l'axe d'adaptation : non plus faire converger l'usager vers une norme linguistique, mais reconfigurer le dispositif à partir des pratiques plurilingues ordinaires.

Trois lignes de force en découlent. Sur le plan méthodologique, l'analyse des besoins se détache de son ancrage pédagogique pour devenir un instrument de conception. Sur le plan théorique, la médiation cesse d'être descendante et redistribue l'autorité du savoir vers l'usage. Enfin, les résultats suggèrent que le français se distribue selon plusieurs couches d'usage. Les frontières linguistiques semblent davantage opératoires pour les systèmes de traitement que pour les pratiques observées chez les répondants.

Cette recherche ne présente pas l'agent implémenté, mais le cadre de spécification linguistique situé en amont de son développement. C'est dans cet espace que se déterminent les conditions d'adoption d'un dispositif numérique. Elle relève en cela d'une ingénierie linguistique exploratoire : les spécifications produites ont le statut d'hypothèses de conception fondées sur l'analyse des besoins, non de prescriptions validées par l'usage. La validation interviendra lors de la quatrième phase du cycle décrit en 2.2, consacrée aux tests utilisateurs après l'implémentation. Cette confrontation permettra d'éprouver les spécifications proposées dans la présente étude.

Des travaux ultérieurs pourront confronter ces spécifications à un agent effectivement implémenté, évaluer leur acceptabilité en interaction réelle et examiner leur transférabilité à d'autres contextes sociolinguistiques et à d'autres dispositifs numériques plurilingues.

Au-delà du cas étudié, cette recherche invite à reconsidérer la place des sciences du langage dans les écosystèmes contemporains d'innovation. Lorsque les produits numériques reposent sur l'interaction verbale, la langue ne constitue plus seulement un moyen de communication ou un objet d'apprentissage ; elle devient une dimension de conception à part entière. Dans cette perspective, l'analyse des besoins apparaît non plus comme l'étape initiale d'un parcours de formation, mais comme le point de départ d'une ingénierie linguistique orientée vers l'usage. L'enjeu n'est plus seulement d'enseigner une langue de spécialité, mais de concevoir des dispositifs capables d'entrer dans les pratiques langagières réelles de leurs utilisateurs. C'est peut-être là que se situe aujourd'hui l'un des prolongements les plus féconds de l'héritage du Français sur Objectif Spécifique.

Références

Abainia, K. (2020). DZDC12: A new multipurpose parallel Algerian Arabizi–French code-switched corpus. *Language Resources and Evaluation, 54*(2), 419–455.https://doi.org/10.1007/s10579-019-09454-8

Abdaoui, A., Berrimi, M., Oussalah, M., & Moussaoui, A. (2021). DziriBERT: A pre-trained language model for the Algerian dialect (arXiv:2109.12346). arXiv. https://doi.org/10.48550/arXiv.2109.12346

Albirini, A. (2011). The sociolinguistic functions of codeswitching between Standard Arabic and Dialectal Arabic. *Language in Society, 40*(5), 537–562.

Algumaei, A., Yaacob, N. M., Doheir, M., Al-Andoli, M. N., & Algumaie, M. (2025). Symmetric therapeutic frameworks and ethical dimensions in AI-based mental health chatbots (2020–2025): A systematic review of design patterns, cultural balance, and structural symmetry. *Symmetry, 17*(7), 1082. https://doi.org/10.3390/sym17071082

Amazouz, D., Adda-Decker, M., & Lamel, L. (2017). Addressing code-switching in French/Algerian Arabic speech. In *Proceedings of Interspeech 2017* (pp. 62–66). ISCA. https://doi.org/10.21437/Interspeech.2017-1373

Brown, J. D. (2016). *Introducing needs analysis and English for specific purposes*. Routledge.

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). Language models are few-shot learners. Advances in Neural Information Processing Systems, 33, 1877–1901.

Conseil de l'Europe. (2001). *Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer*. Conseil de l'Europe.

Conseil de l'Europe. (2018). *Cadre européen commun de référence pour les langues : volume complémentaire avec de nouveaux descripteurs*. Conseil de l'Europe.

De Costa, P. I., Park, J. S.-Y., & Wee, L. (2016). Language learning as linguistic entrepreneurship. *The Asia-Pacific Education Researcher, 25*(5–6), 695–702.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers) (pp. 4171–4186). Association for Computational Linguistics. https://doi.org/10.18653/v1/N19-1423

Dewaele, J.-M. (2010). *Emotions in multiple languages*. Palgrave Macmillan.

Elwahsh, S., Stern, N., Singh, A., & Ayobi, A. (2025). Linguistic diversity and mental well-being: Co-designing custom AI chatbots with multilingual mothers. In *Proceedings of the 7th ACM Conference on Conversational User Interfaces (CUI '25)*. Association for Computing Machinery. https://doi.org/10.1145/3719160.3736615

Følstad, A., & Brandtzæg, P. B. (2017). Chatbots and the new world of HCI. Interactions, 24(4), 38–42. https://doi.org/10.1145/3085558

García, O., & Wei, L. (2014). *Translanguaging: Language, bilingualism and education*. Palgrave Macmillan.

Heller, M. (2010). The commodification of language. *Annual Review of Anthropology, 39*, 101–114. https://doi.org/10.1146/annurev.anthro.012809.104951

Long, M. H. (2005). *Second language needs analysis*. Cambridge University Press.

Mangiante, J.-M., & Parpette, C. (2004). *Le Français sur Objectif Spécifique : de l'analyse des besoins à l'élaboration d'un cours*. Hachette.

Norman, D. A., & Draper, S. W. (Éds.). (1986). User centered system design: New perspectives on human-computer interaction. Lawrence Erlbaum Associates.

Oostendorp, M., Little, T., & Berghoff, R. (2024). First language primacy in multilingualism and emotion research: A view from Africa. *International Journal of Bilingual Education and Bilingualism, 27*(10), 1360–1374. https://doi.org/10.1080/13670050.2024.2360113

Otheguy, R., García, O., & Reid, W. (2015). Clarifying translanguaging and deconstructing named languages: A perspective from linguistics. *Applied Linguistics Review, 6*(3), 281–307. https://doi.org/10.1515/applirev-2015-0014

Pavlenko, A. (2005). *Emotions and multilingualism*. Cambridge University Press.

Sabour, S., Zhang, W., Xiao, X., Zhang, Y., Zheng, Y., Wen, J., Zhao, J., & Huang, M. (2023). A chatbot for mental health support: Exploring the impact of Emohaa on reducing mental distress in China. *Frontiers in Digital Health, 5*, 1133987. https://doi.org/10.3389/fdgth.2023.1133987

Shneiderman, B. (2022). Human-centered AI. Oxford University Press.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30, 5998–6008.

Annexe A

NAFSIATI — Questionnaire de collecte de données

استبيان جمع البيانات

Bien-être psychologique des étudiants

الرفاهية النفسية للطلاب

Anonyme — Confidentiel

Informations importantes avant de commencer

معلومات مهمة قبل البدء

• Ce questionnaire est strictement anonyme. Ne pas écrire votre nom.

• هذا الاستبيان مجهول الهوية تمامًا. لا تكتب اسمك.

• Vos réponses serviront uniquement à construire un outil d'aide psychologique pour les étudiants algériens.

• ستُستخدم إجاباتك فقط لبناء أداة دعم نفسي للطلاب الجزائريين.

• Vous pouvez vous arrêter à tout moment et ne pas répondre à certaines questions.

• يمكنك التوقف في أي وقت وعدم الإجابة على بعض الأسئلة.

• Il n'y a pas de bonne ou de mauvaise réponse. Exprimez-vous librement, dans la langue de votre choix.

• لا توجد إجابة صحيحة أو خاطئة. عبّر بحرية، بأي لغة تريدها.

I. PROFIL DE L'ÉTUDIANT(E) / الملف الشخصي للطالب/الطالبة

Filière / التخصص : Biologie ☐ Médecine ☐ Sciences et technologies ☐ Droit ☐

Niveau / المستوى : L1 ☐ L2 ☐ L3 ☐ M1 ☐ M2 ☐ Doctorat ☐

Genre / الجنس : Féminin ☐ Masculin ☐ Non précisé ☐

Résidence / السكن : Cité univ. ☐ Domicile familial ☐ Logement privé ☐

Langue(s) parlée(s) / اللغات المستعملة : Darija ☐ Français ☐ Tamazight ☐ Les trois ☐ Autre : _____

II. VIE UNIVERSITAIRE ET BIEN-ÊTRE GÉNÉRAL / الحياة الجامعية والرفاهية العامة

1. En ce moment, à quel point vous sentez-vous stressé(e) dans votre vie universitaire ?

١. في الوقت الحالي، إلى أي مدى تشعر بالضغط في حياتك الجامعية؟

1 Pas du tout ☐ 2 Un peu ☐ 3 Modérément ☐ 4 Beaucoup ☐ 5 Extrêmement ☐

2. Quelles sont vos principales sources de stress à l'université ? (Plusieurs réponses possibles)

٢. ما هي المصادر الرئيسية للضغط لديك في الجامعة؟ (يمكن اختيار أكثر من إجابة)

☐ Les examens et les partiels — الامتحانات والاختبارات
☐ La charge de travail (cours, devoirs) — عبء العمل (دروس، واجبات)
☐ Les relations avec les enseignants — العلاقة مع الأساتذة
☐ Les relations avec les camarades — العلاقة مع الزملاء

☐ Les difficultés financières — الصعوبات المادية
☐ Le logement / transport — السكن /التنقل
☐ L'avenir professionnel / orientation — المستقبل المهني /التوجيه
☐ Autre : _____ — أخرى : _____

3. Décrivez avec vos propres mots ce que vous ressentez le plus souvent pendant les périodes de stress. (Utilisez la langue dans laquelle vous vous exprimez le mieux : français, darija, arabe…)

٣. صِف بكلماتك الخاصة ما تشعر به في أغلب الأحيان خلال فترات الضغط. (استخدم اللغة التي تعبّر بها أكثر)

4. Ces dernières semaines, comment évaluez-vous la qualité de votre sommeil ?

٤. في الأسابيع الأخيرة، كيف تقيّم جودة نومك؟

1 Très mauvais ☐ 2 Mauvais ☐ 3 Moyen ☐ 4 Bon ☐ 5 Très bon ☐

III. ÉMOTIONS ET VÉCU PERSONNEL / المشاعر والتجربة الشخصية

Cette partie est la plus importante pour notre projet. Répondez le plus spontanément possible.

هذا الجزء هو الأهم لمشروعنا. أجب بأكبر قدر ممكن من التلقائية.

5. Complétez librement ces phrases / أكمل هذه الجمل بحرية

Répondez dans la langue dans laquelle vous pensez naturellement — français, darija, arabe classique ou un mélange.

5a. En ce moment, ce qui me pèse le plus c'est…

٥أ. في الوقت الحالي، ما يثقل كاهلي أكثر هو…

5b. Quand j'ai un examen difficile, je ressens…

٥ب. عندما يكون لديّ امتحان صعب، أشعر بـ…

5c. La nuit avant un partiel, je…

٥ج. في الليلة التي تسبق الامتحان، أنا…

5d. Ce qui me donne envie d'abandonner mes études, c'est…

٥د. ما يجعلني أفكر في التخلي عن دراستي هو…

5e. Quand je me sens seul(e) à l'université, je…

٥هـ. حين أشعر بالوحدة في الجامعة، أنا…

5f. La phrase que je me dis souvent quand ça va mal : …

٥و. الجملة التي أقولها لنفسي كثيرًا حين تسوء الأمور : …

IV. STRATÉGIES D'ADAPTATION / استراتيجيات التكيّف

6. Quand vous vous sentez dépassé(e), que faites-vous ? (Plusieurs réponses possibles)

٦. حين تشعر بالإرهاق، ماذا تفعل؟ (يمكن اختيار أكثر من إجابة)

☐ J'en parle à un(e) ami(e) ou à la famille — أتحدث مع صديق/ة أو أحد أفراد العائلة
☐ Je me renferme sur moi-même — أنسحب إلى نفسي

☐ Je me plonge dans le travail — أنغمس في العمل والدراسة
☐ Je dors ou je me repose — أنام أو أستريح
☐ Je fais du sport / une activité physique — أمارس الرياضة أو نشاطًا بدنيًا
☐ J'utilise les réseaux sociaux — أستخدم وسائل التواصل الاجتماعي
☐ Je cherche des informations en ligne — أبحث عن معلومات على الإنترنت
☐ Je n'en parle à personne — لا أتحدث مع أحد
☐ Autre : _____ — أخرى : _____

7. Y a-t-il quelque chose que vous aimeriez pouvoir dire à quelqu'un en cas de détresse, mais que vous n'osez pas exprimer ? Si oui, écrivez-le ici librement.

٧. هل هناك شيء تودّ قوله لشخص ما في لحظات الضيق، لكنك لا تجرؤ على التعبير عنه؟ إن كان الأمر كذلك، اكتبه هنا بحرية.

V. ATTENTES VIS-À-VIS D'UN OUTIL DE SOUTIEN / التوقعات من أداة الدعم

8. Si un chatbot confidentiel était disponible pour vous aider en cas de mal-être, vous l'utiliseriez pour :

٨. إذا توفّر روبوت محادثة سري لمساعدتك في أوقات الضيق، فستستخدمه من أجل :

☐ Parler de mon stress sans être jugé(e) — التحدث عن ضغطي دون أن أُحكَم عليّ
☐ Obtenir des conseils pratiques — الحصول على نصائح عملية
☐ Être orienté(e) vers un professionnel — التوجيه نحو متخصص
☐ Simplement avoir quelqu'un à qui écrire — مجرد وجود شخص أكتب له
☐ Je ne l'utiliserais pas — لن أستخدمه
☐ Autre : _____ — أخرى : _____

9. En quelques mots, quel message ou quelle phrase aimeriez-vous qu'un chatbot vous dise quand vous ne vous sentez pas bien ?

٩. في كلمات قليلة، ما هي الرسالة أو الجملة التي تودّ أن يقولها لك روبوت المحادثة حين لا تشعر بخير؟

Merci pour votre participation / شكرًا على مشاركتك

Vos réponses contribuent directement à la construction d'un outil conçu pour vous et pour vos pairs.

إجاباتك تُسهم مباشرةً في بناء أداة مصمّمة من أجلك ومن أجل زملائك.

Projet NAFSIATI — Université Abbès Laghrour de Khenchela, 2024–2025

Annexe B

Protocole d'analyse du corpus

1. Objet et finalité

Ce protocole fixe la méthode d'analyse du corpus de réponses étudiantes. Il part de faits de langue observés et en déduit des spécifications que le produit devra satisfaire. Il prolonge l'analyse des besoins du FOS en en réorientant la finalité : non plus l'élaboration d'un cours, mais la conception d'un produit linguistique numérique.

2. Nature du corpus et posture

Réponses ouvertes à un questionnaire bilingue, réparties sur quatre filières ; énoncés courts, plurilingues (darija en graphie arabe, français, arabe standard, anglais, Arabizi) et souvent hybrides. Posture de linguiste : on analyse la mise en mots de la détresse, non l'état psychique. Analyse qualitative et manuelle ; logiciel de lexicométrie écarté ; suffisance par saturation thématique.

3. Matériau et unité d'analyse

Unité : la réponse à un item ouvert (Q3 ; Q5a–f ; Q7 ; Q9). Items fermés en contexte. Variable de segmentation : la filière.

4. Grille d'analyse

Axe	Indices à relever	Question d'analyse	Spécification visée
1. Langue-matrice de l'affect	Langue dominante des réponses expressives (Q3, Q5, Q7) ; graphie ; part français / darija / arabe standard / anglais	Dans quelle langue l'émotion se dit-elle quand le choix est libre ?	Langue(s) que l'agent doit comprendre et produire
2. Hybridité et niveau de mixage	Emprunts intégrés à la morphologie darija, alternance infralexicale, Arabizi, mèmes anglais ; couche lexicalisée vs curriculaire	Quel degré et type de mixage, et varie-t-il selon la filière ?	Traitement du *token* hybride ; paramétrage par segment
3. Registre de consolation	Formules d'auto-réconfort et idiomes de détresse, notamment religieux ou providentialistes (surtout Q5f)	Quel cadre de sens l'usager mobilise-t-il déjà pour endurer ?	Tonalité et contenu de la bibliothèque de scripts : résonner, ne pas contredire
4. Registre attendu de l'agent	Attentes en Q9 : franchise, logique, neutralité, non-jugement / soutien, consolation, motivation	Quel registre est réclamé, et est-il homogène ?	Persona et registre de l'agent ; caractère adaptatif
5. Texture affective	Allongements graphiques, idiomes somatiques, rire-exutoire	Comment l'intensité et le corps s'inscrivent-ils dans l'écrit ?	Indices de sévérité ; non-normalisation ; logique d'escalade
6. Indicible et confidentialité	Aveux d'incapacité à dire, refus, domaines tus ; demandes de vie privée	Qu'est-ce qui résiste à la mise en mots, et quelles garanties ?	Amorces de verbalisation ; garanties de confidentialité explicites

Deux couches de français à distinguer (axe 2) : une couche lexicalisée, présente partout (par ex. « stress », « mémoire »), et une couche curriculaire, portée par l'enseignement en français (par ex. « réviser », « nuit blanche » conjugués), qui n'apparaît que dans les filières francophones.

5. Procédure de codage

- Anonymiser et coder chaque fichier (BIO-, MED-, ST-, DROIT- + numéro).
- Transcrire fidèlement les réponses ouvertes, graphie d'origine conservée.
- Étiqueter la langue-matrice et les insertions.
- Coder selon les six axes.
- Relever les verbatims probants, avec glose française.
- Compter par filière sur l'ensemble du corpus, sans inférence statistique.
- Trier, axe par axe, ce qui est invariant et ce qui varie entre filières.
- Traduire chaque régularité en spécification de conception.
- Relire chaque spécification comme une instance de la médiation à vecteur inversé.

6. Conventions de transcription et de glose

- Graphie d'origine conservée ; allongements et formes hybrides non normalisés.
- Glose française entre parenthèses ; marquage : [DAR], [AR], [FR], [EN], [ARABIZI].
- Verbatims cités sans correction orthographique.
- Validation des gloses dialectales par un locuteur natif.

7. Opération centrale : invariant et variable

La comparaison entre filières distingue : un trait dominant (la darija comme langue de l'affect — quasi exclusive en filières arabophones, atténuée en médecine), des convergences transversales fortes (consolation religieuse ou providentialiste ; refus du jugement et de la fausse consolation, avec une préférence majoritaire pour la franchise), et des traits variables avec la langue d'enseignement (le mixage français/anglais, et la langue-matrice elle-même, qui se déplace vers le français en médecine). Les traits dominants et les convergences fondent des contraintes de conception ; les traits variables, des paramètres réglables par segment.

8. Fiabilité et critères de suffisance

- Saturation thématique comme critère de suffisance.
- Double codage non réalisé dans la présente étude ; inscrit comme perspective pour la validation ultérieure.
- Transparence : chaque énoncé appuyé par son verbatim.
- Segments inégaux assumés ; filières à faible effectif traitées comme illustratives.

9. Sorties attendues

- Matrice de codage réponse × axe × filière (ci-dessous).
- Tableau de synthèse besoin langagier → spécification de conception.
- Résultat invariant / variable, cœur des résultats.

Matrice de codage (gabarit)

Colonne	Contenu
Code répondant	BIO-, MED-, ST-, DROIT- + numéro
Filière	Biologie / Médecine / Sciences et technologies / Droit
Item	Q3, Q5a–f, Q7 ou Q9
Verbatim	Texte d'origine, graphie conservée, sans correction
Langue-matrice	DAR / AR / FR / EN / mixte
Type de mixage	aucun / lexicalisé / curriculaire / infralexical / Arabizi / anglais
Registre de consolation	religieux / providentialiste / laïque / absent
Registre attendu	franchise-logique / soutien / mixte / non renseigné
Texture affective	allongement / somatique / rire / aucune
Indicible · confidentialité	aveu / refus / domaine tu / demande de confidentialité
Spécification rattachée	Texte libre : contrainte de conception déduite