

La connotation du vocabulaire somatique : une étude de cas comparative bilingue en oncologie

Federica Vezzani

Volume 68, Number 1, April 2023

URI: <https://id.erudit.org/iderudit/1106329ar>

DOI: <https://doi.org/10.7202/1106329ar>

[See table of contents](#)

Publisher(s)

Les Presses de l'Université de Montréal

ISSN

0026-0452 (print)

1492-1421 (digital)

[Explore this journal](#)

Cite this article

Vezzani, F. (2023). La connotation du vocabulaire somatique : une étude de cas comparative bilingue en oncologie. *Meta*, 68(1), 97–118.
<https://doi.org/10.7202/1106329ar>

Article abstract

Tools from corpus linguistics and translation-oriented terminography can usefully integrate a method of analysis aimed at identifying the connotative potential of medical vocabulary. In particular, we will apply these tools to the somatic vocabulary present in specialized and popularized medical texts relating to the field of oncology. In a situation of translation from Italian into French, we will present, as a representative example, the pairs of terms *seno/mammella* and *sein/mamelle* which, beyond any obvious cross-language similarities, manifest very different connotations that the translation process must take into account.

La connotation du vocabulaire somatique : une étude de cas comparative bilingue en oncologie

FEDERICA VEZZANI

Università degli Studi di Padova, Padoue, Italie

federica.vezzani@unipd.it

RÉSUMÉ

Les outils issus de la linguistique de corpus et de la terminographie orientée traduction, peuvent utilement intégrer une méthode d'analyse visant à cerner le potentiel connotatif du vocabulaire médical. En particulier, nous les appliquerons au vocabulaire somatique présent dans les textes médicaux spécialisés ou vulgarisés relevant du domaine de l'oncologie. En situation de traduction de l'italien vers le français, nous nous servirons, à titre d'exemple représentatif, des paires de termes *seno/mammella* et *sein/mamelle*, qui, par-delà de trop évidentes ressemblances interlangues, manifestent au contraire des connotations bien différentes dont doit tenir compte le processus de traduction.

ABSTRACT

Tools from corpus linguistics and translation-oriented terminography can usefully integrate a method of analysis aimed at identifying the connotative potential of medical vocabulary. In particular, we will apply these tools to the somatic vocabulary present in specialized and popularized medical texts relating to the field of oncology. In a situation of translation from Italian into French, we will present, as a representative example, the pairs of terms *seno/mammella* and *sein/mamelle* which, beyond any obvious cross-language similarities, manifest very different connotations that the translation process must take into account.

RESUMEN

Las herramientas de la lingüística de corpus y la terminografía orientada a la traducción, pueden integrar de manera útil una metodología de análisis orientada a identificar el potencial connotativo del vocabulario médico. Dichos recursos se aplicarán, en particular, al vocabulario somático presente en textos médicos especializados y de divulgación relacionados con el campo de la oncología. Utilizaremos como ejemplo representativo de la traducción del italiano al francés los pares de términos *seno/mammella* y *sein/mamelle* que, más allá de las evidentes semejanzas entre lenguas, manifiestan al contrario connotaciones muy diferentes, algo que el proceso de traducción debe tener en cuenta.

MOTS CLÉS/KEYWORDS/PALABRAS CLAVE

vocabulaire somatique, connotation, langue médicale, terminologie, traduction
somatic vocabulary, connotation, medical language, terminology, translation
vocabulario somático, connotación, lenguaje médico, terminología, traducción

1. Introduction¹

Les études sur le vocabulaire *somatique* s'articulent autour d'un large éventail de recherches du point de vue tant lexicologique que terminologique. Dans les contextes de langue générale, le corps humain – interprété comme un « universel cognitif » (Wierzbicka 2007) – fait l'objet de nombreuses investigations visant à révéler sa

perception dans une perspective interlinguistique et interculturelle (Iordanskaja et Mel'čuk 1997; Enfield, Majid *et al.* 2006; Maalej et Yu 2011; Brenzinger et Kraska-Szlenk 2014). En particulier, les termes désignant les parties du corps que l'on retrouve dans les discours non spécialisés font l'objet de réflexions intéressantes en raison de leur potentiel phraséologique (Martin-Berthet 2007; Kim 2017; Marcol-Cacoñ 2019), *métaphorique* (Cuturi 1981; Veland 2003; Goschler 2005), *métonymique* (Hilpert 2007), *pragmatique* (Dingemanse 2009) et *polysémique* (Bertin 2018). En outre, pour ce qui est de leur connotation, certaines contributions scientifiques portent sur le risque que peut manifester le vocabulaire somatique face au tabou linguistique (Widlak 1965; Allan 2007; Chadaverian 2010; Gao 2013).

Si l'on se penche sur les études de nature terminologique, on constate plutôt un intérêt de recherche orienté vers la composante dénotative du vocabulaire somatique qui se manifeste dans les discours spécialisés. Cette tendance tient à un long processus de normalisation de la nomenclature anatomique, dont l'objectif est l'uniformisation de la langue de spécialité médicale pour favoriser la monosémie et la désambiguïsation de la signification de termes désignant des concepts potentiellement liés à la santé et au bien-être des individus (Sadegh-Zadeh 2012). Ce processus a conduit, après de nombreuses révisions, à la publication en 1998 de la nomenclature *Terminologia Anatomica* (TA)², développée par le Comité fédératif de terminologie anatomique et rassemblant les indications terminologiques concernant environ 8 000 termes et expressions en langue latine internationalement reconnues (pour un historique sur l'évolution de la nomenclature, voir entre autres: Kachlik, Baca *et al.* (2008) et Karenberg (2015). Toutefois, bien qu'elle constitue la terminologie standard de référence, la TA demande des révisions qualitatives et quantitatives constantes (Chmielewski et Strzelec 2020): à l'heure actuelle, la dernière version disponible seulement en latin et en anglais a été publiée en 2019 par le *Programme international fédératif de terminologie anatomique* (FIPAT). Par ailleurs, d'un point de vue traductif, la nomenclature internationale en latin soulève, sur le plan national, des difficultés d'adaptation en langue locale dont le traducteur doit forcément tenir compte (Rouleau 2003; Vandaele et Gingras Harvey 2013; Soubrier 2014; Renevier 2017).

L'importance de l'uniformisation dénotative de la terminologie se reflète en particulier dans les contextes où la transmission d'informations est vitale. Comme l'indiquent Galinski et Nedobity (1988), la normalisation de la terminologie n'est pas toujours nécessaire, sensée ou souhaitable dans tous les domaines des activités humaines. Toutefois, si la sécurité des êtres humains dans le cadre d'une activité spécialisée est affectée, par exemple pour les services d'urgence médicale, la normalisation est indispensable. En ce sens, l'aspect connotatif du vocabulaire somatique est censé céder la place à la rigueur dénotative de la terminologie employée dans la langue de spécialité médicale afin de véhiculer un message dépourvu de toute ambiguïté. À cet égard, de nombreuses études (voir section 2 sur l'état de la question) visent à montrer que, bien qu'elle ait été conçue dans un souci de normalisation, la terminologie médicale fait également appel au phénomène de la connotation, qui peut varier en fonction de la langue.

Dans ce contexte, notre étude vise à proposer une méthode – illustrée dans la section 3 – basée sur l'intégration d'outils de la linguistique de corpus et de la terminographie orientée traduction: ces outils permettent en effet de cerner le potentiel connotatif des termes somatiques identifiés dans les textes médicaux spécialisés et

vulgarisés relevant du domaine de l'oncologie. En particulier, dans une perspective traductive, nous verrons si et comment ce phénomène se manifeste pour le couple de langues romanes italien-français. Le cas de la figure envisagé ici implique le processus de traduction spécialisée d'une langue source L1 vers une langue cible L2, où la L1 représente la langue maternelle (dans ce cas l'italien) et la L2 la langue de travail étrangère du traducteur (dans ce cas le français). Dans ce cadre, ce que nous essayons de compenser, c'est un possible manque de sentiment linguistique en français. En ce sens, nous chercherons des différences potentielles de connotation entre les variantes terminologiques qui désignent les concepts anatomiques dans les deux langues. En adoptant la méthodologie proposée, nous présenterons, dans la section 4, l'étude de cas spécifique des termes italiens *seno/mammella* et des termes français *sein/mamelle*. Nous verrons alors combien le phénomène de connotation influe sur les correspondances traductives réellement pertinentes.

2. Connotation en langue médicale

La langue médicale est une « langue de spécialité » (Lerat 1995) utilisée dans des situations de communication (orales ou écrites) qui impliquent la transmission d'une information relevant d'un champ d'expérience particulier à un public aussi bien expert (professionnels de santé) que non expert (patients) (Cortelazzo 1990; Bianco 2002). En tant que telle, la langue de spécialité médicale devrait présenter des caractéristiques spécifiques vérifiables dans ses productions. Parmi celles-ci, Gotti (1991) considère, par exemple, les aspects de précision, de concision et de transparence de la terminologie médicale afin que les informations spécialisées soient clairement compréhensibles pour tous les destinataires concernés. En outre, Serianni (2003) identifie les traits souhaitables 1) de monoréférentialité, à savoir la caractéristique d'univocité entre le signe et le référent qui implique la non-utilisation de l'homonymie et de la synonymie; et 2) de neutralité émotionnelle, selon laquelle le ton du texte reste neutre et objectif, accordant donc plus d'importance à l'aspect dénotatif qu'à l'aspect connotatif.

Or, d'un point de vue strictement terminologique, les traits susmentionnés sont souvent des composantes requises – en vertu d'une communication efficace – que ne confortent ou ne respectent pas nécessairement les manifestations réelles du discours médical. La terminologie médicale présente de nombreuses complexités qui se reflètent à la fois dans une perspective intralinguistique – par exemple dans la communication entre le médecin et le patient (Serianni 2005; Castro *et al.* 2007; Balliu 2010; Vecchiato et Gerolimich 2013; Magris 2016) – et dans une perspective interlinguistique lors du processus de traduction de textes spécialisés (Monin 1993; Rouleau 1995; Jammal 1999; Alarcón-Navío *et al.* 2016). Dans l'ensemble des éléments contribuant à cette complexité, le rôle joué par la connotation dans la langue médicale – et, plus généralement, dans les langues de spécialité – fait l'objet d'études à partir de la fin du xx^e siècle. Comme l'affirment Thoiron et Béjoint :

La vision selon laquelle le terme ne serait que l'étiquette d'un concept suppose que le sens du terme ne soit que dénotatif, ce qu'il est certes dans l'idéal, et peut-être dans un certain nombre de cas au moment de sa naissance, voire de certaines de ses mises en œuvre. Mais un sens connotatif peut fort bien s'attacher à un terme au fur et à mesure des usages en discours, comme à un mot. Que le terminologue choisisse de l'ignorer

lorsqu'il se penche sur une terminologie est une chose, mais cela ne signifie pas que le connotatif n'existe pas. (Thoiron et Béjoint 2010: 109)

D'un point de vue métaterminologique, les notions de dénotation et de connotation, et toutes les différentes acceptions qu'elles peuvent revêtir, ont été largement débattues en littérature: voir, par exemple, les études de Martinet (1969), Gary-Prieur (1971), Molino (1971), Kerbrat-Orecchioni (1977) et, en particulier, Mounin (1963) en ce qui concerne la perspective traductive. Dans cette étude, nous nous référons à la connotation telle que définie par Bloomfield (1933: 151-155), c'est-à-dire l'ensemble des «valeurs supplémentaires» et subjectives attribuées à un signe linguistique en raison de facteurs linguistiques, culturels et sociaux. La notion de connotation s'oppose à celle de dénotation entendue comme la définition objective d'un terme valable pour tous les locuteurs.

Déjà dans ce contexte, l'auteur affirmait que «[i]n the case of scientific terms, we manage to keep the meaning nearly free from connotative factors, though even here we may be unsuccessful [...]» (Bloomfield 1933: 152). En effet, s'éloignant de l'idée que le signifié d'un terme spécialisé doit être exempt de toute surcharge d'ordre émotionnel ou connotatif, Lefèvre souligne que «[l]a vie sémiologique des unités de la langue fait que les signifiés se chargent de sèmes issus de l'environnement social, culturel, etc. des utilisateurs de la langue» (Lefèvre 2004: 59). Dans cette lignée, l'attention de nombreux chercheurs s'est déplacée vers l'identification d'éléments connotatifs également présents dans la communication spécialisée. Parmi les diverses contributions, l'étude de Jahr (2000) présente l'analyse d'une série de textes spécialisés et postule l'existence d'un niveau émotionnel de la structure textuelle qui s'ajoute aux niveaux grammatical et pragmatique. À ce propos, Balliu affirme que:

les publications scientifiques ne sont en aucun cas le degré zéro de la connotation; la seule différence avec les œuvres littéraires réside dans le fait que la connotation y est plus pernicieuse, comme si elle désavouait le caractère objectif de l'écrit. (Balliu 1994: 24)

La connotation de la littérature médicale, que Balliu (2005b) explique par des intentions de simplification et de clarification, peut se remarquer à partir de la deuxième moitié du XIX^e siècle, qui vit le triomphe de la «médecine anatomo-clinique moderne» impliquant un paradigme différent du vivant (Donato 2004), où le patient se retrouve au centre de l'attention en faisant l'objet de nouvelles techniques de manipulation. Toujours d'après l'auteur, «[u]n texte médical ne peut faire l'économie du fait qu'en fin de compte, il concerne toujours l'homme, dans ce qu'il a de plus humain et de plus fragile, la vie» (Balliu 2005a: 75).

Pour cette raison, le texte médical implique, de manière plus ou moins latente, un potentiel connotatif remarquable. L'étude de Monin (1996) élabore, à ce propos, une typologisation intéressante des connotations des termes éponymes. Parmi ceux-ci, l'autrice considère les connotations:

- a) culturelles: le syndrome de Pheidippides;
- b) religieuses: la *pomme d'Adam*;
- c) nationales: le terme ghanéen *Kwashiorkor*;
- d) désuètes: l'expression *danse de Saint-Guy*;
- e) de notoriété scientifique: la *maladie de Parkinson*.

Enfin, une perspective diachronique de la connotation en terminologie médicale est plutôt fournie dans les contributions de Dury (2012, 2013). D'après l'auteur, la connotation péjorative attribuée à certains termes est la force motrice qui anime le « besoin néologique » des spécialistes : ceux-ci ont en effet régulièrement besoin de rebaptiser des concepts dont la dénomination s'est chargée, avec le temps, de connotation dévalorisante. L'accent y est mis sur la variation temporelle, qui éclaire le rôle de la connotation, sur l'obsolescence des termes et sur la coexistence de variantes synonymiques dans le discours spécialisé. Ainsi, le terme anglais **vegetative state**, en raison de sa connotation péjorative, s'est trouvé peu à peu évincé par les termes **unresponsive wakefulness syndrome** ou **UWS**, et **minimally conscious state**, que les spécialistes estimaient plus neutres et plus exacts du point de vue scientifique (Dury 2013 : 8).

2.1 *Connotation en contexte oncologique : objectifs, méthodes et outils d'analyse*

Le panorama décrit ci-dessus nous amène donc à constater un intérêt portant sur la place et le rôle de la connotation dans la langue médicale et dans sa terminologie. Cependant, bien qu'il soit reconnu largement en littérature qu'un terme médical puisse être connoté, à notre connaissance, le potentiel connotatif des termes désignant les parties du corps qui se manifestent dans les textes de nature médicale reste sous-estimé. En effet, comme nous l'avons mentionné dans l'introduction, la plupart des études sur le vocabulaire somatique décrivent la perception et l'interprétation de celui-ci dans la langue générale. Adoptant une perspective différente, notre article vise plutôt à identifier et à examiner la connotation du vocabulaire somatique employé tant dans les textes médicaux spécialisés (destinés aux experts du domaine) que dans les textes médicaux vulgarisés (destinés au grand public ou aux experts d'autres disciplines).

Par ailleurs, notre intérêt est orienté vers la dimension traductive de ces deux typologies textuelles : le but de cette étude est de proposer une méthodologie générale permettant de révéler le potentiel connotatif du vocabulaire somatique et, partant, de guider les choix de traduction du professionnel et/ou de l'apprenant traducteur. En particulier, la méthode proposée dans les sections suivantes vise à soutenir le processus de traduction impliquant le scénario dans lequel un texte médical est soumis à la traduction d'une langue source L1 vers une langue cible L2, où la L1 représente la langue maternelle et la L2, la langue de travail étrangère du traducteur. Le choix de ce cas de figure est dû à la pratique de plus en plus courante de la traduction vers la L2 dans le monde de la traduction spécialisée (Ferreira et Schwieter 2017 ; Mraček 2019 ; Vigier Moreno 2019) : d'où la nécessité de satisfaire les besoins formatifs et opérationnels d'un (futur) professionnel de la langue.

Or, puisque la connotation est une « valeur supplémentaire » (Bloomfield 1933) subjective attribuée à un signe linguistique, on peut supposer que pour un traducteur ce phénomène est plus perceptible et perçu dans sa propre L1 que dans la L2 de travail (voir à ce propos : Liu et Zhong 1999 ; Gablasova 2015). À cette fin, nous proposons une méthodologie générale basée sur l'intégration d'outils issus 1) de la linguistique de corpus (section 3.1 et 2) de la terminographie orientée traduction (section 3.2) afin d'identifier le potentiel connotatif dans une langue cible L2 qui peut influencer les

choix terminologiques lors du processus de traduction. L'étude de cas envisagée dans cet article prévoit la combinaison de langues romanes italien-français, où la première joue le rôle de L1 et la seconde de L2. Le domaine de travail choisi, quant à lui, est celui de l'oncologie. Les travaux en terminologie oncologique visent le plus souvent à fournir des ressources linguistiques pouvant aider les patients et leurs proches à comprendre correctement les termes spécialisés. Pour ne citer que quelques exemples d'outils de support linguistique offerts dans ce domaine, on remarque le dictionnaire *LexOnco* (Delavigne, Carretier *et al.* 2007), la ressource *OncoTerm* (Faber 2002) et le projet pilote *Cancer FrameNet* (Roberts, Si *et al.* 2018)³. Dans la lignée de ces études, nous nous intéressons en particulier aux textes médicaux relatifs au cancer du sein, en tant que pathologie oncologique ayant la plus forte incidence selon les données statistiques fournies par le Centre international de la recherche sur le cancer de l'Organisation mondiale de la santé⁴.

3. Mise en forme du corpus comparable bilingue

Dans le scénario impliquant la traduction d'un texte médical, la première étape de la méthodologie proposée consiste en la compilation d'un corpus comparable bilingue qui collige les textes spécialisés et vulgarisés relevant du domaine à l'étude. Reprenant les notions de décodage et de transcodage du texte médical à des fins traductives, telles que définies par Jammal (1999), la constitution du corpus répond aux besoins suivants :

1. la recherche documentaire méthodique pour la compréhension correcte des informations spécialisées à traduire ;
2. l'analyse terminologique systématique pour la familiarisation et la maîtrise de la terminologie (source et cible) utilisée dans ces types de textes.

Par ailleurs, pour l'étude de cas envisagée ici, nous adoptons une approche basée sur les corpus (Tognini-Bonelli 2001) et observons l'usage de la terminologie médicale – en particulier du vocabulaire somatique – dans son environnement naturel afin de vérifier son potentiel connotatif tant dans la langue source que dans la langue cible.

La constitution d'un corpus relevant du domaine de l'oncologie a été menée au moyen du logiciel d'analyse de textes *Sketch Engine*⁵ (version 2.36.5). Le système accessible en ligne permet l'étude de corpus de textes dans une perspective multilingue et est doté d'un nombre considérable de fonctionnalités (par exemple *Word sketch*, *Thesaurus*, *Keyword and term extraction*, etc.) qui le distinguent des autres outils de gestion de corpus disponibles sur le marché (Kilgarriff, Baisa *et al.* 2014).

Pour la mise en forme d'un corpus, le logiciel offre deux possibilités : 1) l'importation de textes propres à l'utilisateur, ou 2) la collecte automatisée de textes disponibles sur le Web. En supposant que cette méthodologie s'adresse également à ceux qui n'ont pas de connaissances préalables du domaine de travail, nous avons adopté la deuxième option offerte pour accéder à un nombre suffisamment vaste de textes, consultables en ligne, à la fois spécialisés et vulgarisés sur le sujet. La collecte de documents sur le Web peut s'effectuer de trois manières :

1. recherche par mots clés pertinents au domaine d'étude ;
2. insertion d'une liste d'URL renvoyant à des pages spécifiques préalablement sélectionnées ;
3. insertion du lien renvoyant vers un site web spécifique dans le cas de recherches ciblées.

Parmi ces trois possibilités, nous avons retenu la recherche par mots clés car, en vue d'une recherche exhaustive, cette solution offre l'avantage de collecter tous les documents du Web contenant toutes les combinaisons possibles de mots clés saisis et caractérisant la pathologie à l'étude. À cet égard, il est évident qu'un inconvénient probable découlant de l'adoption de ce type d'approche se traduit par l'inclusion de documents qui ne sont pas parfaitement pertinents pour le domaine d'analyse. Cependant, d'un point de vue statistique et sur de grandes quantités de données, nous estimons que ces documents auront un impact minime sur l'analyse quantitative. Un autre inconvénient peut se manifester : c'est l'inclusion de documents provenant de sources potentiellement non fiables. À ce stade, le travail du traducteur consistera précisément à vérifier quel poids attribuer aux données obtenues à partir de ces sources selon le sujet à l'étude.

Une fois définie la méthodologie, nous avons compilé un corpus en italien, langue source et langue maternelle du traducteur pour le cas de figure envisagé dans cette étude. Le choix des mots clés italiens utilisés pour la recherche documentaire suit le critère de variation dénomminative (Freixa 2006) des deux termes italiens désignant, d'une part, la pathologie et, d'autre part, la région anatomique concernée. À partir de la consultation croisée des ressources terminologiques en oncologie citées plus haut, nous avons donc recueilli une liste de termes de type nominal appartenant au même champ lexical – et donc pas forcément synonymes entre eux – pour les deux notions susmentionnées. En particulier, pour la désignation de la pathologie, nous avons sélectionné les termes italiens cancro, tumore, carcinoma et neoplasia⁶ ; pour la région du corps considérée, les désignations retenues ont été seno et mammella.

Une fois les variantes terminologiques identifiées, nous avons saisi dans le logiciel Sketch Engine toutes les combinaisons possibles obtenues à partir de l'association des mêmes variantes (voir tableau 1). Les prépositions en italien qui relient les termes désignant, d'une part, la pathologie et, d'autre part, la région anatomique (c'est-à-dire al/alla et del/della) ont été choisies en cherchant à rendre compte des tournures possibles.

TABLEAU 1

Variantes terminologiques en italien de la pathologie et de la région anatomique et leurs combinaisons employées comme mots clés.

Pathologie	Région anatomique	Mots clés Sketch Engine
cancro	seno	cancro al seno, cancro del seno
cancro	mammella	cancro alla mammella, cancro della mammella
tumore	seno	tumore al seno, tumore del seno
tumore	mammella	tumore alla mammella, tumore della mammella
carcinoma	seno	carcinoma al seno, carcinoma del seno
carcinoma	mammella	carcinoma alla mammella, carcinoma della mammella
neoplasia	seno	neoplasia al seno, neoplasia del seno
neoplasia	mammella	neoplasia alla mammella, neoplasia della mammella

À ce stade, sur la base des entrées fournies, le système récupère automatiquement les documents disponibles sur le Web contenant toutes les combinaisons possibles des termes complexes saisis comme mots clés. Les critères sélectionnés sur le logiciel pour la recherche de documents sont essentiellement deux et concernent 1) le nombre

de mots clés combinés par recherche, fixé à deux, et 2) le nombre maximum d'URL rapatriées par le moteur de recherche, dans notre cas fixé à six, pour chaque paire de mots clés⁷.

En ce qui concerne les sources de repérage des documents, nous avons volontairement étendu la recherche à l'ensemble du Web – sans restreindre le champ d'investigation à des sites spécifiquement sélectionnés – car cela nous permet 1) de collecter tant des textes spécialisés que des textes vulgarisés, 2) d'identifier, dans ces textes, la présence éventuelle de connotations attribuées au vocabulaire somatique, et 3) d'examiner comment ce phénomène se répartit sur un corpus hétérogène.

Avec ces critères de sélection, le système a récupéré un total de 138 documents (environ 1 400 000 tokens) en italien; ceux-ci ont ensuite été annotés automatiquement d'un point de vue morphosyntaxique. En ce qui concerne les sources récupérées automatiquement, le système a extrait des documents à la fois de pages web spécialisées, comme la revue italienne *Attualità in Senologia*⁸, et de pages conçues pour la consultation des patients, comme le site de la *Fondazione AIRC per la Ricerca sul Cancro*.

Une fois le corpus italien compilé, nous avons utilisé les mêmes fonctionnalités de *Sketch Engine* pour la mise en forme du corpus comparable en français (L2 du traducteur). Or, adoptant une approche volontairement naïve pour le cas de figure envisagé ici, la sélection des mots clés pour cette langue a été menée par l'insertion des candidats traduisants français des mots clés identifiés pour l'italien. L'approche naïve consiste précisément à se mettre à la place d'une personne utilisant un dictionnaire proposant ces équivalences. La source utilisée pour la sélection des candidats traduisants est le *Dizionario Medico* (Manuila, Manuila *et al.* 1997)⁹. Cette ressource a été spécifiquement sélectionnée non seulement parce qu'elle est spécialisée dans le domaine médical, mais aussi parce qu'elle envisage les deux langues de travail de cette étude: la version consultée est en fait la première édition italienne traduite de la septième édition française de 1996.

Sur la base des traduisants attestés dans ce dictionnaire spécialisé, les variantes terminologiques françaises utilisées pour se référer à la pathologie sont: *cancer*, *tumeur*, *carcinome* et *néoplasie*. En ce qui concerne la région anatomique, les traduisants français proposés sont *sein* et *mamelle*. Il faut préciser que, à ce stade de notre méthode, nous nous appuyons uniquement sur les données fournies par cette ressource terminologique spécialisée. L'analyse des équivalences traductives et du comportement des termes dans leur environnement naturel sera menée avec l'intégration des outils proposés dans les sections suivantes.

Comme pour l'italien, le tableau 2 présente les variantes terminologiques utilisées pour la génération des mots clés en français et, même dans ce cas, les prépositions en français (*du/de la* et *au/à la*) ont été choisies en cherchant à rendre compte des tournures possibles.

TABLEAU 2

Variantes en français de la pathologie et de la région anatomique et leurs combinaisons employées comme mots clés.

Pathologie	Région anatomique	Mots clés Sketch Engine
cancer	sein	cancer du sein, cancer au sein
cancer	mamelle	cancer de la mamelle, cancer à la mamelle
tumeur	sein	tumeur du sein, tumeur au sein
tumeur	mamelle	tumeur de la mamelle, tumeur à la mamelle
carcinome	sein	carcinome du sein, carcinome au sein
carcinome	mamelle	carcinome de la mamelle, carcinome à la mamelle
néoplasie	sein	néoplasie du sein, néoplasie au sein
néoplasie	mamelle	néoplasie de la mamelle, néoplasie à la mamelle

Le corpus français obtenu se compose finalement de 175 documents (environ 1 700 000 tokens) et se configure, en termes quantitatifs, comme raisonnablement comparable au corpus établi précédemment pour l'italien. Pour ce qui est des sources de repérage des documents y contenus, on remarque, même dans ce cas, tant des pages web spécialisées, comme le *Collège National des Gynécologues et Obstétriciens Français* (CNGOF), que des sites vulgarisant des informations spécialisées, comme *l'Institut national du cancer* (INCa).

3.1. Extraction et analyse terminologique bilingue

Une fois notre corpus comparable construit, la méthodologie proposée prévoit l'extraction et, ensuite, l'analyse de la terminologie source et cible qui y est contenue à l'aide des outils terminographiques orientés traduction.

L'extraction de la terminologie relevant du sous-domaine d'investigation a été menée grâce à la fonctionnalité de *Sketch Engine* appelée *Keywords and terms extraction*. D'un point de vue méthodologique, cette fonction repose sur une analyse contrastive de deux corpus distincts :

1. le corpus de travail, c'est-à-dire le corpus spécialisé dont les termes sont extraits (dans notre cas, d'abord le sous-corpus de textes en italien, puis, le sous-corpus en français) ;
2. le corpus de référence, utilisé comme sous-corpus de comparaison rassemblant des textes de langue générale, habituellement d'une plus grande taille que le focus corpus et fourni par défaut par *Sketch Engine* (dans notre cas *l'Italian Web 2020* (*itTenTen20*) et le *French Web 2017* (*frTenTen17*)).

L'opération est basée sur la formule mathématique simple, utilisée pour calculer le degré de « termhood » (Kageura et Umino 1996) d'une séquence de caractères extraite. La formule identifie donc les séquences qui apparaissent plus fréquemment dans le corpus de travail que dans le corpus de référence. Les fréquences relatives servent à comparer des corpus de tailles inégales. Le résultat de l'extraction automatique fournit une liste de candidats termes, organisés en deux pages d'affichage portant respectivement sur les candidats termes simples (constitués d'un seul mot-forme) et sur les candidats termes complexes (constitués de plusieurs mots-formes). À ce point de la méthode, le traducteur-terminologue est censé dépouiller ponctuellement la liste obtenue afin d'identifier la terminologie source et cible pertinente à son domaine

d'investigation. Il en mènera ensuite une analyse systématique afin de maîtriser la langue de spécialité (voir les objectifs définis dans Jammal (1999) mentionné au début de la section 3). Dans le cadre de cette étude, nous nous limitons à la narration des analyses menées sur les termes somatiques extraits afin d'identifier leur potentiel connotatif dans les deux langues de travail. En premier lieu, *seno/mammella* et *sein/mamelle* qui ont servi de mots clés pour la constitution du corpus.

L'analyse terminologique se fondera sur le modèle de fiche terminologique bilingue de la ressource *TriMED* (Vezzani et Di Nunzio 2020a; 2020b). À ce propos, il nous semble important de préciser que l'étude comparative de la connotation dans le vocabulaire somatique du domaine de l'oncologie s'inscrit dans un projet de recherche plus vaste qui a débuté en 2018 et qui vise spécifiquement l'implémentation d'une ressource terminologique multilingue et polyvalente pour le domaine médical (Vezzani, Di Nunzio *et al.* 2018). Le projet conduisant à la mise en œuvre de cet outil découle de la nécessité de mettre à la disposition de différents utilisateurs – patients, traducteurs et médecins – un outil qui favorise la diffusion d'informations médicales selon différents registres de langue (spécialisé, semi-spécialisé et non spécialisé) (Vezzani, Di Nunzio *et al.* 2018: 4368-4369). L'une des nouveautés de la ressource *TriMED* réside dans le modèle de fiche terminologique fourni qui vise à répondre aux besoins d'information des diverses catégories d'utilisateurs. En particulier, la fiche de *TriMED* comprend jusqu'à 42 catégories de données afin d'illustrer un cadre global sur le comportement morphosyntaxique, sémantique et phraséologique des termes analysés. Le tableau 3 résume la liste des catégories de données prises en compte dans ce modèle de fiche et qui ont servi à l'analyse de la terminologie somatique.

TABLEAU 3

Catégories de données de TriMED

Axe d'analyse	Catégorie de données
Morphologie	partie du discours, genre grammatical, nombre grammatical, formes dérivées
Phonétique	transcription API
Étymologie	dérivation, composition
Variation	variante orthographique, abréviation, expansion, acronyme
Sémantique	définition, analyse sémique, synonyme, hyponyme, hyperonyme
Phraséologie	unité phraséologique, collocation
Pragmatique	contexte d'utilisation
Registre	nom populaire, nom scientifique, codes de classification internationaux
Domaine	domaine, sous-domaine

Comme la fiche terminologique a été également conçue pour encadrer les aspects variationnistes et de registre de la langue de spécialité médicale, les sources consultées pour la compilation des champs sont à la fois des ressources lexicographiques et terminographiques. Le but, en ce sens, est de faire dialoguer les dictionnaires de langue spécialisée et générale afin d'obtenir le maximum de bénéfice en termes de complémentarité des informations.

4. Enquête contrastive sur la connotation : le cas des variantes terminologiques seno/mammella et sein/mamelle

L'exploitation du corpus comparable supportée par la fiche terminologique *TriMED* permet d'observer des informations pertinentes du point de vue quantitatif et qualitatif du vocabulaire somatique employé en contexte oncologique. Dans une optique complémentaire, la combinaison des outils proposés avec cette méthode, conçue afin de soutenir le processus de traduction médicale d'une L1 vers une L2, nous a permis d'isoler un premier cas de connotation différente, pour la paire de langues considérée, attribuée aux termes désignant la région anatomique concernée.

L'examen du corpus comparable a permis de constater, en termes quantitatifs, une fréquence d'apparition différente des deux paires de termes lorsqu'ils sont associés à toutes les variantes dénominatives de la pathologie – précédemment illustrées dans les tableaux 1 et 2 – tant dans le corpus italien que dans le corpus français. Les tableaux 4 et 5 présentent les résultats de l'analyse menée grâce à la fonctionnalité offerte par *Sketch Engine* pour l'identification des n-grammes, c'est-à-dire des séquences de tokens. Le critère de recherche utilisé concerne la longueur des séquences, fixée dans notre cas à un minimum de 2 et un maximum de 3 tokens par n-gramme. De plus, grâce aux fonctionnalités avancées, nous avons défini les critères concernant le terme initial – c'est-à-dire les variantes terminologiques désignant la pathologie dans les deux langues – et le terme final (d'abord seno/sein, puis mammella/mamelle) combinés dans la séquence. Cette méthode nous a permis d'inclure toutes les combinaisons de séquences contenant les prépositions (*du/de la* et *au/à la* pour le français, et *del/della* et *al/alla* pour l'italien) généralement utilisées – avec des fréquences très variées – parmi les deux éléments du syntagme dans les deux langues¹⁰.

Les tableaux ci-dessous illustrent les différentes combinaisons de variantes terminologiques et le nombre de leurs occurrences dans les deux sous-corpus de textes en italien et en français.

TABLEAU 4

Occurrences des variantes dans le sous-corpus de textes en italien

Variants dénominatives	Sous-corpus en italien
Cancro + seno	508 occurrences
Cancro + mammella	438 occurrences
Tumore + seno	639 occurrences
Tumore + mammella	409 occurrences
Carcinoma + seno	24 occurrences
Carcinoma + mammella	148 occurrences
Neoplasia + seno	22 occurrences
Neoplasia + mammella	32 occurrences

TABLEAU 5

Occurrences des variantes dans le sous-corpus de textes en français

Variantes dénominatives	Sous-corpus en français
Cancer + sein	1 886 occurrences
Cancer + mamelle	16 occurrences
Tumeur + sein	105 occurrences
Tumeur + mamelle	15 occurrences
Carcinome + sein	34 occurrences
Carcinome + mamelle	0 occurrence
Néoplasie + sein	0 occurrence
Néoplasie + mamelle	0 occurrence

D'un point de vue quantitatif, nous pouvons constater que les n-grammes comprenant les termes *seno/sein* sont généralement plus fréquents que les séquences comprenant *mammella/mamelle* dans les deux sous-corpus¹¹. En particulier, cette différence est plus marquée pour le sous-corpus en français, dans lequel le terme *mamelle* a une fréquence beaucoup plus faible. La fonction *concordance* de *Sketch Engine* a permis de récupérer les contextes des n-grammes en français contenant les termes *cancer + mamelle* (16 occurrences) et *tumeur + mamelle* (15 occurrences). Une analyse manuelle ciblée des 31 occurrences a montré que, dans 11 cas (1/3 des occurrences), le terme *mamelle* est utilisé dans des contextes oncologiques faisant référence aux animaux. En effet, le corpus contient des documents récupérés par exemple sur le site *AniCura*, réseau européen de cliniques vétérinaires spécialisées. En revanche, interrogé sur les n-grammes contenant *cancro/tumore/carcinoma/neoplasia + mammella* (au total 1027 occurrences), le corpus italien ne livre aucune référence au monde animal. Nous rappelons à ce moment que, sur la base des données attestées dans le dictionnaire médical spécialisé précédemment cité dans la section 3, notre point de départ reposait sur les équivalences traductives entre le terme italien *seno* et le terme français *sein*, et entre le terme italien *mammella* et le terme français *mamelle*, respectivement, proposées par ce dictionnaire. Or, le contraste entre les deux sous-corpus, pour ce qui est de ces deux derniers termes, laisse supposer pour les natifs non francophones une dénotation à extension limitée – absence du trait/humain/ (= /- humain/), et présence du trait /non-humain/ (= /+ non-humain/) – pour le terme *mamelle* qui pourrait influencer son équivalence traductive avec le terme italien *mammella*.

La méthodologie prévoit, à ce stade, l'étude qualitative sur les propriétés linguistiques des deux paires de variantes terminologiques de la région anatomique à l'aide de la fiche *TriMED*. Les champs des fiches terminologiques de *TriMED* nous permettent de réclamer des informations pertinentes sur la signification, l'usage et l'équivalence traductive de ces termes tant en langue de spécialité qu'en langue générale. Pour leur compilation, nous nous sommes appuyée sur les données issues de l'analyse croisée de trois types de ressources (précisées et énumérées dans les tableaux suivants):

1. de type médical spécialisé et donc orienté vers les professionnels de la santé;
2. de type médical vulgarisé s'adressant également au grand public;
3. de langue générale.

En se concentrant d'abord sur l'italien, les tableaux 6 et 7 illustrent les définitions fournies respectivement pour les termes seno et mammella pour chaque type de ressource.

TABLEAU 6

Définitions du terme seno

Typologie	Ressource	Définition
1	<i>L'enciclopedia della medicina</i> (Drago et Boroli 1996: 704)	Formazione anatomica o parte di organo caratterizzati da cavità, infossamenti, dilatazioni, perlopiù contenente liquido o altri tessuti.
2	<i>Dizionario di medicina</i> (Aite 1966: 189)	Nome generico dato a diverse cavità del corpo.
3	<i>Treccani</i> ¹²	[...] Spazio o solco mediano, verticalmente disposto e corrispondente in anatomia alla regione sternale, che nel petto della donna in età fertile è compreso tra le mammelle [...]. <u>Nel linguaggio com., le mammelle stesse.</u> ¹³

TABLEAU 7

Définitions du terme mammella

Typologie	Ressource	Définition
1	<i>L'enciclopedia della medicina</i> (Drago et Boroli 1996: 533)	Ognuna delle due ghiandole caratteristiche dei mammiferi, ipotrofiche nel maschio e particolarmente sviluppate nelle femmine, destinate alla secrezione del latte [...]; <u>ognuna è separata da quella corrispondente da un ampio solco, detto seno.</u>
2	<i>Dizionario di medicina</i> (Aite 1966: 148)	Organo femminile pari, situato nella regione anteriore del torace, formato dalle ghiandole mammarie [...].
3	<i>Treccani</i>	In anatomia, organo ghiandolare, proprio dei mammiferi euteri e marsupiali, che secerne il latte [...].

L'analyse croisée des ressources en italien permet d'affirmer que :

1. Le terme seno est générique et désigne différentes cavités corporelles. En l'occurrence, il désigne la cavité corporelle située entre les mamelles. Enfin, dans la langue courante, le terme est attesté comme synonyme de mammella.
2. Le terme mammella désigne les deux glandes mammaires séparées par un espace concave appelé seno.

En italien, bien que mammella et seno désignent deux parties anatomiques distinctes (les deux glandes, d'une part, et la cavité qui les divise, d'autre part), les deux termes peuvent être utilisés comme synonymes dans la langue courante lorsqu'on se réfère aux femmes.

Si l'on passe à la langue française, les tableaux 8 et 9 illustrent les définitions fournies respectivement pour les termes *sein* et *mamelle* dans les trois typologies de ressources.

TABLEAU 8

Définitions du terme *sein*

Typologie	Ressource	Définition
1	<i>Dictionnaire médical de l'Académie nationale de médecine</i>	Cf: Mamelle
2	<i>Dictionnaire médical</i> (Quevauvilliers 2009: 854)	Chacune des mamelles de la femme. Les seins sont des organes glandulaires responsables de la sécrétion lactée, situés dans la région supérieure, antérieure et latérale du thorax.
3	<i>Dico en ligne Le Robert</i>	Chacune des mamelles de la femme.

TABLEAU 9

Définitions du terme *mamelle*

Typologie	Ressource	Définition
1	<i>Dictionnaire médical de l'Académie nationale de médecine</i>	Ensemble constitué chez les mammifères par la glande mammaire, la peau qui la recouvre et la couche cellulo-adipeuse intermédiaire. <u>Synonymes: sein (dans l'espèce humaine)</u>
2	<i>Dictionnaire médical</i> (Quevauvilliers 2009: 570)	Cf: Sein
3	<i>Dico en ligne Le Robert</i>	1) Organe des femelles des mammifères, sécrétant le lait. 2) [vieux] Sein de femme

L'analyse croisée des ressources en français permet d'observer que :

1. Le terme *sein* désigne chacune des deux mamelles de la femme et, dans le dictionnaire médical spécialisé, renvoie à la définition donnée pour le terme *mamelle*.
2. Le terme *mamelle* désigne l'organe qui sécrète le lait. En ce qui concerne sa variation, on remarque que, dans les dictionnaires médicaux spécialisés et vulgarisés, le synonyme attesté pour l'espèce humaine est *sein* alors que, pour le dictionnaire de langue générale, la synonymie avec ce terme est attestée comme obsolète lorsqu'on se réfère aux femmes.

En français, les deux termes *sein* et *mamelle* peuvent être employés dans la langue de spécialité comme des synonymes, alors que la langue générale atteste désormais comme désuet l'emploi de *mamelle* au lieu de *sein*.

Compte tenu de cette indication diachronique, nous avons également consulté le *Dictionnaire historique de la langue française* (Rey 2011) : *mamelle*,

employé dès les premiers textes en parlant du sein de la femme, a vieilli en ce sens sauf dans des expressions comme enfant à la mamelle (1675), et la locution disparue dès la mamelle (XVII^e s.), littérairement « dès l'enfance ». C'est en revanche le terme normal en anatomie. Tout autre emploi commun a une connotation plaisante ou péjorative. (Rey 2011)

À ce point, en combinant les diverses informations issues de l'observation du corpus et de l'analyse terminologique, on peut affirmer que l'usage commun du terme *mamelle* a évolué au fil du temps, acquérant aujourd'hui une valeur supplémentaire connotative – potentiellement renvoyant à l'espèce animale selon les occurrences du corpus – qui en conditionne son usage lorsqu'on se réfère à l'espèce humaine.

Laissant de côté l'aspect diachronique (au moins aux fins de cette première étude), nous vérifions la référence synchronique à la sphère animale à partir de l'analyse 1) des articles disponibles sur *Wikipédia*, 2) de la prédiction automatique des mots clés sur les moteurs de recherche, et 3) des données fournies dans la version française de la ressource *Medical Subject Headings (MeSH terms)*. D'un point de vue strictement terminologique et interlinguistique, il est intéressant de noter une utilisation différente des termes désignant la partie anatomique que l'on retrouve dans les pages de l'encyclopédie libre *Wikipédia*¹⁴. Du côté italien, l'article consacré à la description de la pathologie, mis à disposition sur la ressource, porte le nom *Tumore alla mammella*. Accédant à la contrepartie française – en sélectionnant la langue directement à partir de cette page –, on peut constater que l'hyperlien renvoie à un article qui porte, en revanche, le titre de *Cancer du sein*. Si nous limitons l'enquête à la seule partie anatomique, la page *Wikipédia* italienne *Mammella* renvoie à l'article français *Sein*. Or, en recherchant le terme *Mamelle* dans la version française de la ressource, on peut, sans aucun doute, noter que les informations contenues dans l'article se réfèrent exclusivement à la sphère animale. De plus, il est intéressant de noter qu'à partir de cette page, il n'y a pas de lien intertextuel avec la contrepartie italienne¹⁵. Le fait que le terme *mamelle* est plus fréquemment associé à la partie anatomique des animaux se constate également grâce à la prédiction automatique de mots clés sur les moteurs de recherche (Chen et Lin 2010). Par exemple, en saisissant la séquence des mots clés *cancer* et *mamelle* dans la boîte de *Google*, le moteur de recherche prédit, sur une base statistique, la séquence de termes potentiellement d'intérêt pour l'utilisateur qui, dans ce cas, sont « chat » et « chien ». Enfin, pour la compilation du champ de la fiche de *TriMED* relatif aux codes de classification internationaux, nous avons également consulté le vocabulaire *MeSH terms* généralement utilisé pour l'indexation d'articles (bio)médicaux contenus dans des bases de données telles que *PubMed* (Coletti et Bleich 2001). L'indexation consiste dans l'identification des sujets d'une publication à l'aide des mots clés qui sont définis comme « descripteurs » (Sadegh-Zadeh 2012). La version française de la ressource a été mise à disposition par l'Institut national de la santé et de la recherche médicale (INSERM) : le descripteur français « Tumeurs du sein » présente, comme variantes dénominatives, les termes complexes « Cancer du sein » et « Tumeurs mammaires humaines ». En ce sens, la spécification de l'adjectif *humain* nous amène à confirmer que le terme *mamelle* – ou dans ce cas spécifique l'adjectif qui en dérive – demande à être désambiguïsé quant à son référent.

Les données issues de cette analyse supplémentaire permettent donc de confirmer qu'à ce jour, le terme français *mamelle* – utilisé à l'origine pour les mammifères en général – perd le trait /humain/ (=/-humain/) et conserve le trait /+non-humain/, rétrécissant son champ sémantique d'application aux animaux. La raison de cette élimination peut être liée à la perception d'une aura péjorative du terme et donc l'apparition d'une prosodie sémantique péjorative dans son application médicale aux humains. Autrement dit, *sein* remplace *mamelle* par euphémisme, au prix d'une métonymie par contiguïté originelle. En revanche, ce phénomène d'élimination de trait dénotatif et d'ajout de connotation ne se produit pas pour le terme italien *mammella* qui est normalement employé tant en référence aux femmes qu'aux animaux (/+ humain//+ non-humain/).

En guise de conclusion, la méthodologie adoptée permet donc d'avancer les considérations suivantes :

1. Le terme italien *mammella* (et son synonyme *seno* entendu dans cette acception spécifique)¹⁶ devrait être traduit par le terme français *sein*, car le choix traductif de *mamelle* serait plutôt perçu comme connotativement péjoratif et renvoyant à l'espèce animale.
2. Cette attribution connotative du terme français *mamelle* semble n'être attestée que pour la langue générale, alors que le terme semble être normal dans la langue de spécialité (voir le dictionnaire historique de la langue française).
3. Cependant, cette diversité d'usage ne se confirme pas dans le corpus français, où même pour les textes médicaux spécialisés on observe la prédominance du terme *sein*.
4. En ce sens, on peut donc constater que la connotation qui se manifeste dans la langue générale migre aussi dans la langue de spécialité, influençant l'emploi de la terminologie, bien que, en vertu des traits souhaitables mentionnés dans la section 2 (notamment la neutralité émotionnelle), cela devrait accorder plus d'importance à la composante dénotative.

5. Conclusion et perspectives

Dans cette étude, nous avons proposé une méthode d'analyse générale qui permet, d'une part, d'observer les termes dans leur environnement naturel et, d'autre part, d'analyser et de documenter leurs propriétés linguistiques de manière systématique et suffisamment exhaustive, combinant les informations obtenues à partir de ressources spécialisées et de langue générale. Intéressée par la perspective traductionnelle, nous avons testé cette méthodologie pour le cas de figure impliquant le processus de traduction de textes médicaux spécialisés et vulgarisés d'une langue source L1 vers une langue cible L2, où la L1 représente la langue maternelle (dans ce cas l'italien) et la L2 la langue de travail étrangère du traducteur (dans ce cas le français).

La complémentarité des deux outils utilisés s'avère valable pour soutenir des réflexions contrastives bilingues sur la connotation attribuée aux termes somatiques employés en contexte oncologique. L'étude de cas des termes *seno/mammella* et *sein/mamelle* montre à quel point la connotation joue un rôle fondamental dans l'équivalence traductive et influence les choix terminologiques des professionnels et/ou apprenants traducteurs – en particulier les natifs non francophones – afin d'éviter une traduction perçue, aujourd'hui, comme péjorative et/ou décontextualisée par rapport à la sphère de l'être humain.

Les données obtenues permettent d'établir un point de départ pour l'étude diachronique du vocabulaire somatique. À l'heure actuelle, l'une des limites de la présente étude est de n'avoir fait qu'effleurer la perspective temporelle sur l'évolution de la dénotation et de la connotation du terme *mamelle*. Ce que nous entendons donc faire comme travaux futurs, c'est approfondir cet aspect pour éclairer les raisons qui ont conduit à la variation de l'usage de ce terme. En outre, sur la base des données obtenues à ce jour, nous envisageons de poursuivre l'analyse afin de traiter en profondeur d'autres cas particuliers pour ces deux langues romanes.

Pour conclure, nous entendons également mener une enquête terminologique bilingue à grande échelle en nous adressant à la fois aux spécialistes du domaine médical et au public non expert. À ce propos, nous voudrions chercher à connaître 1) les raisons qui incitent les médecins à utiliser préférentiellement un terme plutôt qu'un autre, et 2) la perception, dans le grand public, de ces variantes dénominatives. En ce sens, l'enquête a le but de vérifier si, en termes de variation diastratique, le spé-

cialiste est conditionné dans le choix de la terminologie lorsqu'il s'adresse au patient. Enfin, le retour sur la perception du grand public nous permettrait d'obtenir des données pertinentes pouvant ultérieurement confirmer ou infirmer les connotations identifiées.

REMERCIEMENTS

L'autrice tient à remercier la professeure Geneviève Marie Henrot et le professeur Giorgio Maria Di Nunzio pour leurs remarques et suggestions concernant la version préliminaire de la présente étude.

NOTES

1. Cet article est une extension de la communication suivante : VEZZANI, Federica (2021) : La connotation du vocabulaire somatique : une étude de cas comparative bilingue en oncologie. Colloque international *Lexique et corps humain*, Centre de recherche Europes-Eurasie (CREE) de l'INaLCO – axe « Dictionnaires, corpus, réseaux. Changements linguistiques » et ATILF CNRS – Université de Lorraine, événement en ligne, 11-12 juin 2021.
2. Toutes les références aux « Ressources terminologiques » mentionnées dans cet article se trouvent dans l'annexe 1. Cette catégorie hétérogène comprend des ressources de type : bases de données terminologiques, thésaurus, nomenclatures et glossaires.
3. En étendant le champ d'investigation au domaine médical au sens large, le panorama est peuplé d'un nombre considérable de ressources multilingues qui, en termes de variation dénominative diastratique (Freixa 2006), fournissent le terme technique médical et son équivalent populaire. Pour une liste de ces ressources, voir Vezzani et Di Nunzio (2020a : 272-273).
4. *International Agency for Research on Cancer* (2022) : Cancer Fact Sheets. World Health Organization. Consulté le 21 juillet 2022, <<https://gco.iarc.fr/today/fact-sheets-cancers>>.
5. Toutes les références aux fonctionnalités de *Sketch Engine* utilisées pour cette étude se trouvent dans l'annexe 2, consacrée aux « Systèmes de gestion de corpus ».
6. Bien que l'objet de cette étude se concentre spécifiquement sur l'analyse du vocabulaire somatique, il est néanmoins nécessaire de préciser que les variantes terminologiques sélectionnées pour la pathologie (tant en italien qu'en français) se réfèrent toutes à des concepts médicaux différents mais liés. Pour plus d'informations, nous invitons le lecteur à consulter, entre autres, les deux pages d'informations suivantes : SCC (2022) : Types de tumeurs. *Société canadienne du cancer*. Consulté le 21 juillet 2022, <<https://cancer.ca/fr/cancer-information/what-is-cancer/types-of-tumours>>; *Fondation contre le Cancer* (2022) : Les types de cancers. *Fondation contre le Cancer*. Consulté le 21 juillet 2022, <www.cancer.be/le-cancer/jeunes-et-cancer/les-cancers/les-types-de-cancers#:~:text=les%20carcinomes%20sont%20issus%20de,os%2C%20du%20cartilage%2C%20etc>.
7. Ces critères ont été définis pour des raisons d'espace disponible sur le compte *Sketch Engine*.
8. Toutes les références aux pages web (tant spécialisées que de vulgarisation) mentionnées dans cet article se trouvent dans l'annexe 3.
9. Les références de tous les dictionnaires papier utilisés dans le cadre de cette étude sont disponibles dans l'annexe 4.
10. Aux fins de cette enquête visant uniquement les aspects sémantiques de la terminologie médicale, nous n'aborderons pas les questions morphosyntaxiques relatives aux variations des prépositions qui feront l'objet d'études ultérieures.
11. La seule exception se trouve dans le corpus italien où le terme *mammella* est plus fréquemment utilisé en co-occurrence avec les termes spécialisés *carcinoma* et *neoplasia*.
12. Toutes les références des entrées consultées dans des dictionnaires en ligne se trouvent dans l'annexe 5.
13. Tous les soulignements sont les nôtres.
14. Toutes les références des entrées de *Wikipédia* commentées dans le texte se trouvent dans l'annexe 6, consacrée aux « Entrées d'encyclopédie électronique ».
15. Pour aller plus loin, nous avons vérifié si ce phénomène se retrouve également pour la langue anglaise. En effet, la page *Tumore alla mammella* renvoie à l'anglais **Breast cancer**, l'article

Mammella renvoie à la page **Breast** et, à partir de la page française *Mamelle*, nous pouvons accéder, dans ce cas, à la contrepartie anglaise qui porte le nom **Udder** afférent, de nouveau, à la sphère animale: *Wikipedia* (2022): Udder. *Wikipedia*, the free encyclopedia. Consulté le 21 juillet 2022, <<https://en.wikipedia.org/wiki/Udder>>.

16. Pour des raisons d'espace et de pertinence par rapport à l'objet d'étude, nous n'approfondissons pas dans cet article les questions traductives relatives au couple de traduisants *seno/sinus* au sens de cavité corporelle.

RÉFÉRENCES

- ALARCON-NAVÍO, Esperanza, LÓPEZ-RODRÍGUEZ, Clara et TERCEDOR-SÁNCHEZ, Maribel (2016): Variation dénomminative et familiarité en tant que source d'incertitude en traduction médicale. *Meta: Journal des traducteurs/Meta: Translators' Journal*. 61(1):117-144.
- ALLAN, Keith (2007): The pragmatics of connotation. *Journal of Pragmatics*. 39(6):1047-1057.
- BALLIU, Christian (1994): L'enseignement de la traduction médicale: pour une nouvelle pragmatique. *Meta: Journal des traducteurs/Meta: Translators' Journal*. 39(1):15-25.
- BALLIU, Christian (2005a): La didactique de la traduction médicale, deux ou trois choses que je sais d'elle. *Meta: Journal des traducteurs/Meta: Translators' Journal*. 50(1):67-77.
- BALLIU, Christian (2005b): Le nouveau langage de la médecine: une affaire de socioterminologie. *Meta: Journal des traducteurs/Meta: Translators' Journal*. 50(4).
- BALLIU, Christian (2010): Le traducteur, le médecin et le patient. *Meta: Journal des traducteurs/Meta: Translators' Journal*. 55(1):15-22.
- BERTIN, Thomas (2018): La polysémie des noms de parties du corps humain en français: analyse sémantique d'artère, bouche, cœur, épaule et pied. Thèse de doctorat. Normandie: Normandie Université.
- BIANCO, Anna (2002): *Lingue speciali e traduzione specializzata*. Lecce: Adriatica salentina.
- BLOOMFIELD, Leonard (1933): *Language*. Londres: George Allen & Unwin Ltd.
- BREZINGER, Matthias et KRASKA-SZLENK, Iwona (2014): *The body in Language: Comparative Studies of Linguistic Embodiment*. Leyde: Brill.
- CASTRO, Cesar, WILSON, Clifford, WANG, Frances *et al.* (2007): Babel babble: physicians' use of unclarified medical jargon with patients. *American Journal of Health Behavior*. 31(1):67-77.
- CHADAVERIAN, Carine (2010): La médiation en milieu médical: la gestion des tabous linguistiques et culturels dans le monde arabe. In: Mathieu GUIDÈRE, dir. *Traduction et médiation humanitaire*. Paris: Le Manuscrit, 153-172.
- CHEN, Ping-I. et LIN, Shi-Jen (2010): Automatic keyword prediction using Google similarity distance. *Expert Systems with Applications*. 37(3):1928-1938.
- CHMIELEWSKI, Piotr Paweł et STRZELEC, Bartłomiej (2020): Should Terminologia Anatomica be revised and extended? A critical literature review. *Folia Morphologica*. 79(1):1-14.
- COLETTI, Margaret H. et BLEICH, Howard L. (2001): Medical subject headings used to search the biomedical literature. *Journal of the American Medical Informatics Association*. 8(4):317-323.
- CORTELAZZO, Michele (1990): *Lingue speciali: la dimensione verticale*. Padoue: Unipress.
- CUTURI, Flavia (1981): Metafore, proiezioni e rideterminazione nella terminologia anatomica. *La Ricerca Folklorica*, 25-32.
- DELAVIGNE, Valérie, CARRETIER, Julien, BRUSCO, Sylvie *et al.* (2007): Comprendre les mots du cancer: LexOnco, dictionnaire d'oncologie SOR SAVOIR PATIENT pour les personnes malades et leurs proches. Aspects méthodologiques. *Bulletin du Cancer*. 94(4):405-410.
- DINGEMANSE, Mark (2009): The Selective Advantage of Body-Part Terms. *Journal of Pragmatics*. 41(10):2130-2136.
- DONATO, Maria Pia (2004): La medicina clinica moderna in Europa e l'eredità di Foucault. *Storica*. 29:45-60.
- DURY, Pascaline (2012): Le sentiment d'un besoin néologique chez l'expert pour remplacer un terme à connotation péjorative- Quelques exemples tirés du domaine médical. *Neologica*. 6:81-93.

- DURY, Pascaline (2013): Que montre l'étude de la variation d'une terminologie dans le temps? Quelques pistes de réflexion appliquées au domaine médical. *Debate Terminológico*. 9:2-10.
- ENFIELD, Nick J., MAJID, Asifa et VAN STADEN, Miriam (2006): Cross-linguistic categorisation of the body: Introduction. *Language Sciences*. 28(2-3):137-147.
- FABER, Pamela (2002): Oncoterm: sistema bilingüe de información y recursos oncológicos. In: Amparo ALCINA, dir. *La traducción científico-técnica y la terminología en la Sociedad de la información*. Castelló de la Plana: Publicacions de la Universitat Jaume, 177-188.
- FERREIRA, Aline et SCHWIETER, John (2017): Directionality in translation. In: Aline FERREIRA et John SCHWIETER, dir. *The handbook of Translation and Cognition*. New York: John Wiley & Sons, 90-105.
- FREIXA, Judit (2006): Causes of denominative variation in terminology: A typology proposal. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*. 12(1):51-77.
- GABLASOVA, Dana (2015): Learning technical words through L1 and L2: Completeness and accuracy of word meanings. *English for Specific Purposes*. 39:62-74.
- GALINSKI, Christian et NEDOBITY, Wolfgang (1988): Special languages, terminology planning and standardization. In: R. A. STREHLOW, dir. *Standardization of Technical Terminology: Principles and Practices (Second Volume)*. Philadelphie: ASTM, 4-13.
- GAO, Chunming (2013): A sociolinguistic study of English taboo language. *Theory and Practice in Language Studies*. 3(12):2310-2314.
- GARY-PRIEUR, Marie-Noëlle (1971): La notion de connotation(s). *Littérature*. 4:96-107.
- GOSCHLER, Juliana (2005): Embodiment and body metaphors. *Metaphorik*. 9:33-52.
- GOTTI, Maurizio (1991): I linguaggi specialistici: caratteristiche linguistiche e criteri pragmatici. Scandicci: La Nuova Italia.
- HILPERT, Martin (2007): Chained metonymies in lexicon and grammar: a cross-linguistic perspective on body part terms. In: Günter RADDEN, Klaus-Michael KÖPCKE, Thomas BERG et Peter SIEMUND, dir. *Aspects of Meaning Construction*, Amsterdam/Philadelphia: John Benjamins Publishing Company, 77-98.
- IORDANSKAJA, Lidia et MEL'ČUK, Igor (1997): Le corps humain en russe et en français - Vers un Dictionnaire explicatif et combinatoire bilingue. *Cahiers de lexicologie*. 70:105-137.
- JAHR, Silke (2000): *Emotionen und Emotionsstrukturen in Sachtexten*. Berlin/New York: De Gruyter.
- JAMMAL, Amal (1999): Une méthodologie de la traduction médicale. *Meta: Journal des traducteurs/Meta: Translators' Journal*. 44(2):217-237.
- KACHLIK, David, BACA, Vaclav, BOZDECHOVA, Ivana et al. (2008): Anatomical terminology and nomenclature: past, present and highlights. *Surgical and Radiologic Anatomy*. 30(6):459-466.
- KAGEURA, Kyo et UMINO, Bin (1996): Methods of automatic term recognition: A review. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*. 3(2):259-289.
- KARENBERG, Axel (2015): *Fachsprache Medizin im Schnellkurs: Für Studium und Berufspraxis*. Stuttgart: Thieme.
- KERBRAT-ORECCHIONI, Catherine (1977): *La connotation*. Lyon: Presses universitaires de Lyon.
- KILGARRIFF, Adam, BAISA, Vít, BUŠTA, Jan et al. (2014): The Sketch Engine: ten years on. *Lexicography*. 1:7-36.
- KIM, Mi Hyun (2017): Étude contrastive de la phraséologie des noms d'éléments du corps en coréen et en français. Thèse de doctorat. Lorraine: Université de Lorraine.
- LEFÈVRE, Michel (2004): Terminologie et discours «empratique». *Cahiers du CIEL: Des fondements théoriques de la terminologie*. 7:53-70.
- LERAT, Pierre (1995): *Les langues spécialisées*. Paris: Presses universitaires de France.
- LIU, Dilin et ZHONG, Shouman (1999): Acquisition of culturally loaded words in EFL. *Foreign Language Annals*. 32(2):177-187.

- MAALEJ, Zouheir A. et YU, Ning (2011): *Embodiment via body parts: Studies from various languages and cultures*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- MAGRIS, Marella (2016): Verso un uso consapevole della terminologia nell'informazione medica rivolta al paziente. In: Luciana T. SOLIMAN, dir. *Lingue speciali: dinamiche di ricerca*, Padova: CLEUP.
- MARCOL-CACOŃ, Lucyna (2019): Le parti del corpo nelle espressioni fraseologiche e il loro legame con la terminologia medica: l'italiano e il polacco a confronto. *Linguistica Silesiana*. 273-282.
- MARTIN-BERTHET, Françoise (2007): Verbes de coups et parties du corps. *Verbum*. 1(2):67-80.
- MARTINET, André (1969): *La linguistique. Guide alphabétique*. Paris: Denoël.
- MOLINO, Jean (1971): La connotation. *La linguistique*. 7(1):5-30.
- MONIN, Sylvie (1993): La siglaison en langue médicale et problèmes de traduction. *ASp. la revue du GERAS*. 2:29-53.
- MONIN, Sylvie (1996): Termes éponymes en médecine et application pédagogique. *ASp. la revue du GERAS*. 11(14):217-237.
- MOUNIN, Georges (1963): *Les problèmes théoriques de la traduction*. Paris: Gallimard.
- MRAČEK, David (2019): The Role(s) of Native Speakers in L2 Translation. *Studie z aplikované lingvistiky*. 10(1):7-25.
- RENEVIER, Nathalie (2017): La traduction médicale: des nomenclatures à l'orthographe, petit florilège des pièges et difficultés. *Traduire. Revue française de la traduction*. 236:5-12.
- ROBERTS, Kirk, Si, Yuqi, GANDHI, Anshul et al. (2018): A FrameNet for cancer information in clinical narratives: Schema and annotation. In: Nicoletta CALZOLARI et al., dir. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. European Language Resources Association (ELRA), 272-279.
- ROULEAU, Maurice (1995): La langue médicale: une langue de spécialité à emprunter le temps d'une traduction. *TTR: traduction, terminologie, rédaction*. 8(2):29-49.
- ROULEAU, Maurice (2003): La terminologie médicale et ses problèmes. *Panacea*. 4(12):143-152.
- SADEGH-ZADEH, Kazem (2012): *Handbook of Analytic Philosophy of Medicine*. Dordrecht: Springer.
- SERIANNI, Luca (2003): *Italiani scritti*. Bologne: Il mulino.
- SERIANNI, Luca (2005): Un treno di sintomi. I medici e le parole: percorsi linguistici nel passato e nel presente. Milano: Garzanti Libri.
- SOUBRIER, Jean (2014): Traduction et langues de spécialité: aspects de la traduction médicale. *Équivalences*. 41(1):119-153.
- THOIRON, Philippe et BÉJOINT, Henri (2010): La terminologie, une question de termes?. *Meta: Journal des traducteurs/Meta: Translators' Journal*. 55(1):105-118.
- TOGNINI-BONELLI, Elena (2001): Corpus linguistics at work. *Corpus Linguistics at Work*, 1-236.
- VANDAELE, Sylvie et GINGRAS HARVEY, Mariane (2013): Les nomenclatures anatomiques: histoire et traduction. *Panacea*. 14(38):222-234.
- VECCHIATO, Sara et GEROLIMICH, Sonia (2013): Le langage médical est-il « trop complexe »? *Nouvelles perspectives en sciences sociales*. 9(1):81-122.
- VELAND, Reidar (2003): Structure lessico-semantiche dei nomi di parti del corpo. *Studia Neophilologica*. 75(2):126-132.
- VEZZANI, Federica, DI NUNZIO, Giorgio Maria et HENROT, Geneviève (2018): TriMED: A Multilingual Terminological Database. In: Nicoletta CALZOLARI, Khalid CHOUKRI, Christopher CIERI et al., dir. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. European Language Resources Association (ELRA), 4367-4371.
- VEZZANI, Federica et DI NUNZIO, Giorgio Maria (2020a): Methodology for the standardization of terminological resources: design of TriMED database to support multi-register medical communication. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*. 26(2):266-298.

- VEZZANI, Federica et DI NUNZIO, Giorgio Maria (2020b): On the Formal Standardization of Terminology Resources: The Case Study of TriMED. In: Nicoletta CALZOLARI, Frédéric BÉCHET, Philippe BLACHE *et al.*, dir. *Proceedings of the twelfth International Conference on Language Resources and Evaluation (LREC 2020)*. European Language Resources Association (ELRA), 4903-4910.
- VIGIER, MORENO, Francisco J. (2019): Corpus-assisted translation of specialised texts into the L2. From the classroom to professional practice. *Trans-kom*. 12(1):90-106.
- WIDLAK, Stanislas (1965): L'interdiction linguistique en français d'aujourd'hui. *Revue belge de philologie et d'histoire*. 43(3):932-945.
- WIERZBICKA, Anna (2007): Bodies and their parts: An NSM approach to semantic typology. *Language Sciences*. 29(1):14-65.

ANNEXES

Annexe 1 : Ressources terminologiques

- FIPAT (2022): Home page. Federative International Programme for Anatomical Terminology. Consulté le 21 juillet 2022, <<https://fipat.library.dal.ca/>>
- Institut national du cancer (2022): Dictionnaire. Institut national du cancer. Consulté le 21 juillet 2022, <www.e-cancer.fr/Dictionnaire/>
- Le MeSH bilingue (2022): Medical Subject Headings – version française de l'Institut national de la santé et de la recherche médicale (Inserm). Consulté le 21 juillet 2022, <<http://mesh.inserm.fr/FrenchMesh/view/loadSheet.jsp?sheetId=D001943>>
- MeSH (2022): Search. Medical Subject Headings – National Library of Medicine. Consulté le 21 juillet 2022, <<https://meshb-prev.nlm.nih.gov/search>>
- OncoTerm (2022): Home page. OncoTerm: Sistema Bilingüe de Información y Recursos Oncológicos. Grupo de Investigación OncoTerm. Consulté le 21 juillet 2022, <www.ugr.es/~oncoterm/>
- Terminologia Anatomica 1998 on-line version (2022): Home page. Federative International Programme for Anatomical Terminology. Consulté le 21 juillet 2022, <<https://ifaa.unifr.ch/Public/EntryPage/ViewTA98New.html>>
- TriMED (2022): Home page. TriMED - Medical Termbase. The multilingual medical terminological database for everyone. Consulté le 21 juillet 2022, <<http://purl.org/TriMED>>

Annexe 2 : Systèmes de gestion de corpus

- Sketch Engine (2022): Concordance – lesson. Sketch Engine. Consulté le 21 juillet 2022, <www.sketchengine.eu/quick-start-guide/concordance-lesson/>
- Sketch Engine (2022): Corpus annotation and structures. Sketch Engine. Consulté le 21 juillet 2022, <www.sketchengine.eu/blog/corpus-annotation-and-structures/>
- Sketch Engine (2022): Home page. Sketch Engine. Consulté le 21 juillet 2022, <www.sketchengine.eu/>
- Sketch Engine (2022): Keywords and term extraction. Sketch Engine. Consulté le 21 juillet 2022, <www.sketchengine.eu/guide/keywords-and-term-extraction/>
- Sketch Engine (2022): N-grams – multiword expressions, lexical bundles. Sketch Engine. Consulté le 21 juillet 2022, <www.sketchengine.eu/guide/n-grams-multiword-expressions/>

Annexe 3 : Sites web médicaux

- AIRC (2022): Accueil. Fondazione AIRC per la Ricerca sul Cancro. Consulté le 21 juillet 2022, <www.airc.it>
- AIS (2022): Accueil. Attualità in Senologia – Scuola Italiana di Senologia ONLUS. Consulté le 21 juillet 2022, <www.senologia.it/ppp/>
- AniCura (2022): Accueil. AniCura. Consulté le 21 juillet 2022, <www.anicura.fr>
- CNGOF (2022): Accueil. Collège National des Gynécologues et Obstétriciens Français. Consulté le 21 juillet 2022, <www.cngof.fr>

Institut national du cancer (2022): Accueil. Institut national du cancer. Consulté le 21 juillet 2022, <www.e-cancer.fr>

PubMed (2022): Home page. National Library of Medicine. Consulté le 21 juillet 2022, <<https://pubmed.ncbi.nlm.nih.gov>>

Annexe 4: Dictionnaires papier

AITE, Maria Cristina (dir.) (1966): *Dizionario di medicina*. Bologna: Zanichelli.

DRAGO, Marco et BOROLI, Andrea, dir. (1996): *L'enciclopedia della medicina*. Novara: Istituto geografico De Agostini.

MANUILA, Ludmila, MANUILA, Alexandre et NICOLIN, Monique, dir. (1996): *Dictionnaire médical* (7^e édition). Paris: Masson.

MANUILA, Ludmila, MANUILA Alexandre, SANTINI, Gino *et al.*, dir. (1997): *Dizionario Medico* (1^a edizione italiana dalla 7^a edizione francese 1996- Masson-Parigi i. Roma: Editore Marpese.

QUEVAUVILLIERS, Jacques, dir. (2009): *Dictionnaire médical. Avec atlas anatomique*. 6^e édition. Paris: Masson.

REY, Alain, dir. (2011): *Dictionnaire historique de la langue française: contenant les mots français en usage et quelques autres délaissés, avec leur origine proche et lointaine*. Paris: Le Robert.

Annexe 5: Entrées de dictionnaires en ligne

DICO EN LIGNE LE ROBERT (2022): Mamelle. *Dictionnaire en ligne Le Robert*. <<https://dictionnaire.lerobert.com/definition/mamelle>>

DICO EN LIGNE LE ROBERT (2022): Sein. *Dictionnaire en ligne Le Robert*. <<https://dictionnaire.lerobert.com/definition/sein>>

DICIONNAIRE DE L'ACADÉMIE NATIONALE DE MÉDECINE (2022): Mamelle. *Dictionnaire de l'Académie nationale de médecine*. <<http://dictionnaire.academie-medicine.fr/search/results?titre=mamelle>>

DICIONNAIRE DE L'ACADÉMIE NATIONALE DE MÉDECINE (2022): Sein. *Dictionnaire de l'Académie nationale de médecine*. <<http://dictionnaire.academie-medicine.fr/search/results?titre=sein>>

TRECCANI (2022): Mammella. *Enciclopedia Italiana di Scienze, Lettere e Arti*. <www.treccani.it/vocabolario/mammella/>

TRECCANI (2022): Seno. *Enciclopedia Italiana di Scienze, Lettere e Arti*. <www.treccani.it/vocabolario/seno1/>

Annexe 6: Entrées d'encyclopédies électroniques

Wikipédia (2022): Cancer du sein. Wikipédia, L'encyclopédie libre. Consulté le 21 juillet 2022, <https://fr.wikipedia.org/wiki/Cancer_du_sein>

Wikipédia (2022): Mamelle. Wikipédia, L'encyclopédie libre. Consulté le 21 juillet 2022, <<https://fr.wikipedia.org/wiki/Mamelle>>

Wikipedia (2022): Mammella. Wikipedia, L'enciclopedia libera. Consulté le 21 juillet 2022, <<https://it.wikipedia.org/wiki/Mammella>>

Wikipédia (2022): Sein. Wikipédia, L'encyclopédie libre. Consulté le 21 juillet 2022, <<https://fr.wikipedia.org/wiki/Sein>>

Wikipedia (2022): Tumore alla mammella. Wikipedia, L'enciclopedia libera. Consulté le 21 juillet 2022, <https://it.wikipedia.org/wiki/Tumore_alla_mammella>