



HAL
open science

Contribution de l'analyse d'un corpus de comptes rendus de bilan orthophonique à la classification des termes diagnostiques des troubles du langage

Frédérique Brin-Henry

► **To cite this version:**

Frédérique Brin-Henry. Contribution de l'analyse d'un corpus de comptes rendus de bilan orthophonique à la classification des termes diagnostiques des troubles du langage. 2013. hal-01116616

HAL Id: hal-01116616

<https://hal.science/hal-01116616>

Preprint submitted on 13 Feb 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Version auteur

COLLOQUE Cultures de recherche en linguistique appliquée / Research cultures in applied linguistics,

14 au 16 novembre 2013 à Nancy.

Contribution de l'analyse d'un corpus de comptes rendus de bilan orthophonique à la classification des termes diagnostiques des troubles du langage.

Frédérique Brin-Henry¹

L'application directe de procédés d'analyse d'un corpus de textes paramédicaux à la réflexion terminologique concernant les termes diagnostiques utilisés en orthophonie est une démarche relativement nouvelle (Dubreil & Daille 2005, Plug *et al* 2009, Brin-Henry 2011 et 2014) dont l'enjeu est à la fois professionnel et linguistique.

Un enjeu professionnel :

En effet il paraît pertinent pour les professionnels de concevoir, puis d'engager une remise en cause des classifications et des représentations ontologiques² (voir par exemple les considérations sur la *dysphasie* chez Billard 2004, ou chez Gagnon et al 1997). Nous pensons que cette remise en cause est possible grâce à l'examen attentif de la langue spécialisée émanant de la pratique clinique. La place de la terminologie sur l'étape diagnostique est relativement reconnue (Jammal 1989) et ce qui nous préoccupe en premier lieu est le lien entre le terme et la réalité de la communication humaine, au travers de la question suivante: « la terminologie crée-t-elle la pathologie ? », ce qui sous-entend « en quoi la variété des termes diagnostiques contribue-t-elle à cette pathologisation d'un comportement humain (le langage) pour ce qui concerne les orthophonistes ? ». Ce questionnement émane comme très souvent d'une attitude réflexive autour de la pratique, avec un possible diagnostic pathologique de certains comportements linguistiques, qui pourrait être:

- la conséquence de l'excès de normalisation de l'analyse des réalisations linguistiques,
- malmené par les variations diachroniques et synchroniques dans l'utilisation terminologique (Canguilhem 1966, Lövestam et al, 2014).

Ce regard sur l'étape onomasiologique de la pose du diagnostic nous amène naturellement à explorer plus avant les fondements de la terminologie orthophonique et à étudier les étiquettes diagnostiques apparaissant dans des textes intégraux où cette étape est explicite (les comptes rendus de bilan orthophonique ou CRBO).

Un enjeu linguistique

¹ Orthophoniste et Docteure en Sciences du Langage, Chef de projet chargée de mission recherche au Centre Hospitalier de Bar-le-Duc (55), Membre du Laboratoire ATILF- Université de Lorraine/CNRS- UMR 7118, frederique.henry@atilf.fr

² Une ontologie est une spécification formelle d'une conceptualisation d'un domaine, partagée par un groupe de personnes, qui est établie selon un certain point de vue imposé par l'application construite (Studer et al., 1998).

En nous inspirant de l'analyse de corpus dont l'intérêt a été démontré pour des langues de spécialité (Condamines 2005 et 2007, Depierre 2006), nous avons procédé à l'annotation des CRBO et l'extraction de candidats-termes nous intéressant. La fouille et l'annotation de textes intégraux authentiques permet de nous appuyer sur des réalisations situées sur le plan temporel et géographique (identité linguistique du français de spécialité). De plus une analyse complémentaire experte par notre double compétence de chercheur et praticien permet la construction éclairée des différents classements. En effet certains tests sur les syntagmes explorés peuvent être sources d'ambiguïté : la possibilité pour l'examineur de reprendre le texte originel sous l'angle professionnel du domaine favorise alors la prise de décision.

Le corpus

Un corpus protégé a été constitué en 2007 : 436 comptes rendus authentiques de bilan orthophonique ont été regroupés, annotés et fouillés. Le comportement linguistique de 558 candidats-termes a été examiné, parallèlement à une analyse du processus diagnostique fondé sur des métadonnées. Ainsi les syntagmes extraits ont été examinés afin de déterminer, à partir de leurs propriétés, leur valeur terminologique à visée diagnostique. Des moments différents de traitement des données ont été nécessaires.

- a) Étapes manuelles de traitement du corpus:
 - Anonymisation
 - Prise d'indices et de filtres
 - Sélection et constitution des métadonnées
- b) Etapes semi-automatiques
 - Etablissement de listes de fréquence par concordancier: AntConc, Lexico3
 - Codage XML grâce au projet CLARIN-DATIST permettant une interface informatique (<http://www.clarin.eu/external/>) sur serveur protégé
 - Analyses quantitatives et qualitatives (AFC) avec R sur 1 puis 2 variables
 - Extraction et analyse formelle et classement des termes diagnostiques et de leurs propriétés (caractères)

Résultats :

La terminologie orthophonique et la langue des orthophonistes sont constituées de mots de la langue française courante, organisés selon les modalités linguistiques, phraséologiques et discursives du français, soumis aux contraintes intra et extralinguistiques habituelles. Il faudra également y ajouter des termes de spécialité (unités simples et complexes), et une organisation collocationnelle³ spécifique au domaine pour ce qui concerne le « jargon technique ». Les acceptions et les usages des niveaux de langue différents permettent de constituer une langue de spécialité, langue orthophonique, comprenant à la fois des termes du français courant mais également les syntagmes d'une langue scientifique (Baylon et Mignot, 2005), comme par exemple des termes affixés (dysphonie), des composés (anosognosie) ou des collocations (retard de parole). De façon assez inattendue, nous avons observé une fréquence remarquable de mots appartenant au français courant, comme « difficultés » « et troubles ».

Difficulté(s)	- est utilisé en référence à des fonctions organiques (alimentation, respiration...), fonctions cognitives et langagières (mémoire, langage,
---------------	--

³ de façon générale, une collocation représente l'apparition consécutive et fréquente de deux mots (Halliday, 1962, Dubois *et al*, 2007, Charaudeau et Mainguenu, 2002). Polguère (2003) en donne une définition plus spécifique en évoquant une combinaison syntagmatique (expression semi-idiomatique), modifiable par choix sémantique, et que cette combinaison modifie le sens originel de la base de la collocation

	temps...) - est souvent suivi d'une préposition (en, de, d') et un nom ou d'un verbe
Trouble(s)	- Est utilisé pour mentionner une atteinte de l'intégrité du fonctionnement linguistique et communicationnel - Souvent suivi d'une préposition (collocation) et d'un nom - souvent qualifié (adjectifs)

Tableau 1: présentation des usages de mots très fréquents

On pourra alors isoler au cœur de cette langue orthophonique les termes diagnostiques, qui sont des unités terminologiques utilisées pour désigner les termes simples (*aphasie*) ou les collocations (*trouble d'articulation*) ayant servi à l'orthophoniste pour étiqueter les troubles ou difficultés mis en évidence au cours du bilan orthophonique. Ces termes peuvent être regroupés selon leur morphologie et également selon leur usage en lien avec une pathologie spécifique. Ils sont parfois apparentés à des termes non spécifiques (*retard, difficultés*) mais sont traités au même titre que ceux ayant une connotation plus directe à la pathologie (*dysphonie, bégaiement*).

Les résultats les plus marquants de cette étude concernent les métadonnées obtenues, les analyses concernant les termes fréquents et les termes diagnostiques (candidats-termes), et enfin les propriétés des termes.

Les résultats quantitatifs et qualitatifs obtenus à partir des métadonnées permettent une connaissance du corpus (profils des orthophonistes, des patients, répartition des pathologies). Ainsi les analyses factorielles de correspondance (AFC) montrent que les circonstances les plus favorables à la pose d'un diagnostic orthophonique à l'issue du bilan sont :

- Pour un patient âgé de 3 à 7 ans
- Ou pour des troubles du langage écrit et troubles acquis des fonctions supérieures
- Lorsque le diagnostic médical n'est pas identifié (étiologie plus large non précisée)
- Lorsque l'orthophoniste a précisé avoir utilisé une évaluation standardisée

Les résultats qualitatifs de l'étude obtenus grâce à l'examen terminologique (et dans une perspective d'analyse de discours) concernent des indices sémanticosyntaxiques relevés concernant les termes diagnostiques :

- la typologie des termes diagnostiques : ils se répartissent en deux catégories exclusives d'unités terminologiques, la première regroupant des termes décrivant une atteinte du fonctionnement (processus cognitif ou linguistique), et l'autre une atteinte de la structure (système cognitif ou linguistique).
- Les syntagmes peuvent être décrits comme des collocations (ex : *troubles de l'acquisition du langage écrit*), des locutions (ex : *trouble d'articulation*), des dérivés et affixés (ex : *bégaiement, dyslexie*) ou des composés (ex : *anosognosie*)...
- Ils possèdent une configuration sémanticosyntaxique riche (Knittel 2005, 2009) grâce à l'association avec des adjectifs ou des syntagmes prépositionnels.

Conclusion :

L'intérêt de la démarche nous semble évident pour la meilleure compréhension du processus d'étiquetage des pathologies du langage. Nous pensons également qu'elle est utile en ce que cette approche d'analyse de corpus diffère de l'analyse de la production pathologique d'un discours oral

ou écrit dans la connaissance de la pathologie du langage et de la communication, ou bien encore de l'étude de la construction et de l'évolution de la terminologie (Oliveira 2009). Nous nous intéressons également aux implications à déduire pour la description du processus onomasiologique spécifique aux orthophonistes français dans l'étiquetage des troubles du langage (Roche 2005). Ainsi nous obtenons des informations sur les représentations des pathologies chez le praticien et les conséquences possibles sur la frontière entre normal et pathologique, et sur les raisons pour lesquelles les variantes employées montrent une interprétation différente de la nature du problème présenté par le patient.

Enfin, nous pensons qu'il existe une place de ce type d'étude dans le travail des classifications (Demazeux 2013), et la mesure du degré de spécialisation de la langue orthophonique. Ainsi nous travaillons actuellement à des perspectives d'annotation des textes par couches successives et la création d'un guide d'annotation (Kister & Jacquey 2012) visant à alimenter un projet futur de plate-forme collaborative multilingue européenne permettant d'étendre la méthodologie à un corpus plus important. Notre recherche termino-ontologique (Mondary et al 2008, Muresan 2013) se poursuit actuellement en ce sens.

- BAYLON C., MIGNOT X. (2005). *La Communication*. Paris : Armand Colin.
- BILLARD C. (2004). Définition des dysphasies de développement: de quoi s'agit-il? *A.N.A.E.*, 76-77 : 23-25.
- BRIN-HENRY F. (2014). Using Corpus-Based Analyses in Specialised Paramedical French. *Revue Française de Linguistique Appliquée* XIX(1) : 103-115
- BRIN-HENRY, F. (2011). *La terminologie crée-t-elle la pathologie ? le cas de la pratique clinique de la pose du diagnostic orthophonique*. Thèse pour l'obtention du Doctorat en Sciences du Langage, Université de Lorraine. Disponible en ligne sur <http://tel.archives-ouvertes.fr/tel-00655952/>
- CANGUILHEM, G., (1966) (11ème édition quadrigée 2009). *Le normal et le pathologique*. Paris: Presses Universitaires de France.
- CONDAMINES, A. (2005). Linguistique de corpus et terminologie. *Langages* 57, 36-47.
- CONDAMINES A. (2007). L'interprétation en sémantique de corpus : le cas de la construction de terminologies. *Revue Française de Linguistique Appliquée*, XII(1) : 39-52.
- DEMAZEUX S. (2013). *Qu'est-ce que le DSM?* Paris, Ithaque
- DEPIERRE, A. (2006). De l'utilisation de corpus de textes spécialisés pour l'enseignement de la terminologie dans le domaine de l'anglais médical. In LEGLISE et al, 2006. *Applications et implications en Sciences du Langage*. Paris : L'Harmattan, 253-264.
- DUBREIL, E. & DAILLE, B. (2005). Analyse sémantico-discursive des collocations lexicales en corpus spécialisé : la base CONNAISSANCE-S. [en ligne] *Mots, Termes et Contextes*, Actes des 7e journées scientifiques LTT, 6-10 septembre 2005, Bruxelles. <http://perso.univ-lyon2.fr/~thoiron/JS%20LTT%202005/pdf/Dubreil.pdf> , consulté le 31 mai 2014.
- GAGNON L., MOTTRON L., JOANETTE Y. (1997). Questioning the validity of the semantic-pragmatic syndrome diagnosis. *Autism: international Journal of Research and Practice*. Vol 1, n°1, p. 37-55.
- JAMMAL A. (1989). Etiologie de certains brouillages terminologiques de la langue médicale. *Meta* 34 (4), p. 764-69
- KISTER L. & JACQUEY E. (2012). Relations syntaxiques entre lexiques terminologique et transdisciplinaire : analyse en texte intégral, *CMLF*.
- KNITTEL, M.L. (2009). Le statut des Compléments du Nom en [de NP], *Revue Canadienne de Linguistique* 54(2), 255-290.
- KNITTEL, M. L. (2005). Some remarks on adjective placement in the French NP. *Probus*, 17(2), 185-226.
- LÖVESTAM E., FJELLSTRÖM C., KOOCHEK A., & ANDERSSON A. (2014). The power of language on patient-centredness: linguistic devices in the dietetic notes of patient records. *International Journal of Applied Linguistics* (manuscript auteur) doi: 10.1111/ijal.12064.

MONDARY T., DESPRES S., NAZARENKO A., SZULMAN S. (2008) Construction d'ontologies à partir de textes: la phase de conceptualisation. Manuscrit auteur, publié dans "19èmes Journées Francophones d'Ingénierie des Connaissances (IC 2008), Nancy.

MURESAN S. & KLAVANS J.L. (2013). Inducing Terminologies From Text: A Case Study for the Consumer Health Domain, *Journal of The American Society for Information Science and Technology*, 64(4):727–744.

OLIVEIRA, I. (2009). *Nature et fonctions de la métaphore en science ; l'exemple de la cardiologie*. Paris : L'Harmattan.

PLUG, L., SHARRACK, B. & REUBER M. (2009). Seizure, Fit or Attack? The use of diagnostic labels by patients with Epileptic or Non-Epileptic Seizures. *Applied Linguistics* 31(1), 94–114.

ROCHE C., 2005. Terminologie & Ontologie. *Langages* 157 :48-62.