

Cours 5 : Méthodologie de construction de corpus

Introduction :

Le recueil de données orales ou multimodales constitue une étape essentielle dans la recherche comme dans la pratique clinique en orthophonie. Il implique cependant une réflexion approfondie, à la fois sur le plan scientifique et clinique, mais aussi sur le plan législatif et éthique.

La déontologie du recueil :

Avant toute collecte de données orales ou multimodales, il est indispensable de respecter un certain nombre d'exigences légales et déontologiques. Deux étapes sont incontournables :

1. informer le participant du cadre dans lequel s'inscrit le recueil (objectifs scientifiques ou cliniques, modalités d'enregistrement, traitement et usage des données, anonymisation, conservation, protection, diffusion, etc.) ;
2. obtenir son consentement libre et éclairé à la lumière de ces informations (Baude et al., 2006).

L'initiation des étudiants en orthophonie à la pratique du recueil de données orales ou multimodales dès le début de leur cursus représente une occasion privilégiée de sensibilisation à ces bonnes pratiques. En effet, ils seront amenés à se confronter à ces questions à plusieurs reprises au cours de leur formation : lors d'études de cas, dans le cadre de leur mémoire de fin d'études, ou plus tard, dans leur pratique clinique quotidienne. Même si la collecte de données en orthophonie est encadrée par le secret thérapeutique, elle doit répondre à des règles précises de respect du consentement et de protection des données personnelles.

L'enregistrement est par ailleurs largement reconnu comme une activité « naturelle », « fréquente » et même « incontournable » dans la pratique orthophonique (Tran & Bécavin, 2001 ; Witko, 2010). Ces enregistrements permettent notamment de réécouter les productions pour la rédaction du compte rendu de bilan, d'analyser finement certains phénomènes langagiers, ou encore de revenir avec le patient (et éventuellement son entourage) sur des extraits précis. Ils offrent également la possibilité de conserver une trace des productions spontanées, en complément des productions élicitées.

Selon Witko (2010), ces enregistrements favorisent une observation longitudinale du langage du patient — de la forme à l'usage — ainsi qu'une

analyse de la dynamique interactionnelle, éléments essentiels pour ajuster les axes de remédiation. Ils constituent aussi un outil d'auto-évaluation pour le praticien, permettant une réflexion sur ses propres pratiques langagières (de Weck, 2003 ; Rodi, 2011 ; da Silva, 2010) et contribuent, plus largement, à la recherche scientifique sur les troubles du langage à partir de corpus authentiques.

Ces considérations légales ne sont toutefois pas sans impact sur les observations cliniques ou scientifiques. Comme le rappellent Baude et al. (2006), le simple fait d'informer le sujet sur les objectifs de l'enregistrement peut modifier la situation observée et altérer le caractère spontané ou « naturel » des données recueillies. Ce phénomène est d'autant plus marqué lorsque le patient présente une insécurité linguistique importante, souvent liée à la nature même de ses difficultés langagières (Tran & Bécavin, 2001).

Dès lors, l'établissement d'un climat de confiance entre le patient et le praticien devient une condition essentielle pour garantir la qualité du recueil. Comme le souligne Witko (2010), la régularité de la pratique de l'enregistrement facilite son acceptation, notamment lorsque celui-ci est présenté comme un outil professionnel ordinaire, voire comme un véritable participant à l'interaction (Morgenstern, 2016).

Questions de pratiques scientifiques et cliniques

Sur le plan scientifique, l'enregistrement ne constitue que la première étape technique de la constitution d'un corpus. Il s'inscrit dans un ensemble de choix méthodologiques et théoriques qui doivent être explicités en amont. Les chercheurs et praticiens doivent ainsi se poser une série de questions déterminantes :

- Pourquoi enregistrer ?
- Qui enregistrer ?
- Dans quelle situation (spontanée, dirigée, conversationnelle, en séance, etc.) ?
- Quelle durée prévoir pour l'enregistrement ?
- Faut-il privilégier un dispositif audio ou vidéo ?
- Quel matériel utiliser ?
- Le chercheur ou le clinicien doit-il être présent lors de l'enregistrement ?

Ces décisions influencent directement la nature des données recueillies, leur interprétation et leur portée clinique ou scientifique. Former les futurs orthophonistes à cette réflexion méthodologique, dès leur entrée dans la formation, leur permet de développer une posture de praticien-chercheur, attentif à la qualité, à la validité et à l'éthique de ses observations.

C'est au moment de l'analyse des données que les étudiants peuvent véritablement mesurer les implications des choix méthodologiques qu'ils ont opérés en amont. Ils découvrent concrètement que ces choix — qu'ils concernent le type de données recueillies, les conditions d'enregistrement ou encore la méthode de transcription — influencent directement la nature des observations et des interprétations possibles. Cette prise de conscience leur permet de comprendre pourquoi ces paramètres sont habituellement définis avant la phase de recueil, en fonction des objectifs scientifiques (Sinclair, 1996, cité par Habert et al., 1997) et/ou cliniques (da Silva Genest & Masson, 2017). Elle les conduit également à saisir que les différentes étapes constitutives d'une étude de corpus sont interdépendantes et doivent être pensées comme un tout cohérent (Adolphs & Knight, 2010).

Dans le cadre de cet apprentissage, les étudiants sont invités à réaliser un enregistrement audio ou vidéo (selon l'autorisation parentale obtenue) d'un enfant de grande section de maternelle, en situation de commentaire d'un dessin qu'il vient de réaliser. Ce choix de situation présente plusieurs intérêts pédagogiques et méthodologiques. D'une part, il confronte les étudiants à la réalité du recueil de données orales avec de jeunes enfants, et à la difficulté que certains peuvent éprouver à s'exprimer spontanément face à un micro ou à une caméra (Blanche-Benveniste & Pallaud, 2001). D'autre part, il permet d'aborder, au moment de la transcription, les questions posées par le traitement de productions non standard.

Au cours de l'analyse, les étudiants sont amenés à s'interroger sur l'influence de leur présence — ainsi que de celle de l'outil d'enregistrement — sur les productions observées, les conduisant à réfléchir au fameux « paradoxe de l'observateur » (Labov, 1972). Ce questionnement constitue une étape essentielle pour les préparer à la pratique clinique, où la relation entre le praticien et le patient influence inévitablement les productions langagières.

Aucun critère strict de sélection des enfants n'est imposé, afin de laisser aux étudiants une certaine autonomie. Certains choisissent d'enregistrer des enfants qu'ils estiment « bons parleurs », d'autres, au contraire, des enfants qui semblent rencontrer des difficultés. Dans un cas comme dans l'autre, la transcription et l'analyse révèlent souvent des surprises : l'enfant perçu comme ayant un langage « peu développé » manifeste parfois des compétences fines, tandis que celui considéré comme « avancé » présente encore des formes en cours d'acquisition.

Les étudiants sont encouragés à interagir avec l'enfant durant l'enregistrement, dans une situation dyadique proche de celles qu'ils rencontreront en séance d'orthophonie. Cette modalité leur permet d'adopter une posture réflexive sur leurs propres productions langagières, qu'elles soient d'ordre structurel,

fonctionnel ou interactionnel. La situation d'échange à deux facilite également les étapes ultérieures de transcription et d'analyse, en limitant la complexité interactionnelle.

Le choix du dessin comme support d'interaction repose sur la volonté de mutualiser la collecte des données avec d'autres enseignements, notamment ceux portant sur le développement du graphisme. Ce dispositif favorise le dialogue interdisciplinaire, tout en stimulant la parole de l'enfant grâce à un support qu'il a lui-même créé. Néanmoins, cette situation présente certaines limites pour l'analyse linguistique, notamment lorsque l'enregistrement vidéo n'a pas pu être réalisé : les nombreux déictiques produits par l'enfant renvoient souvent à des éléments visuels absents de l'enregistrement audio. Ce constat conduit les étudiants à réinvestir les réflexions précédemment menées sur les avantages et les limites des différentes méthodes de recueil de données (Behrens, 2008).

Malgré ces contraintes, cette expérience constitue une introduction précieuse à la collecte et au traitement de productions spontanées, que Rondal (2003) appelle de ses voeux dans le cadre de l'évaluation orthophonique. Lors de l'analyse, les formateurs insistent sur le caractère partiel des données recueillies : il ne s'agit que d'un fragment des compétences langagières de l'enfant. Les étudiants sont ainsi sensibilisés à la prudence interprétative et à l'importance d'éviter toute généralisation abusive dans la rédaction de leurs analyses.

Aucune consigne verbale précise n'est fournie pour orienter l'enfant, ce qui permet aux étudiants de s'interroger sur l'influence de la consigne donnée sur le type de discours produit (Blanche-Benveniste & Pallaud, 2001). La durée minimale de l'enregistrement est fixée à trois minutes. Bien que le temps ne constitue pas une mesure particulièrement pertinente en linguistique de corpus (Baude et al., 2006 : 28), cette contrainte temporelle offre un cadre rassurant aux étudiants. Elle peut toutefois devenir source de difficulté lorsque le débit de l'enfant est lent, ce qui fournit une nouvelle occasion pédagogique d'aborder la question de la taille et de la représentativité du corpus, en fonction des phénomènes linguistiques étudiés.

Ainsi, cette activité permet non seulement d'introduire les étudiants aux principes fondamentaux du recueil et de l'analyse de corpus, mais aussi de les initier à une démarche réflexive sur leurs pratiques professionnelles futures.

L'analyse des corpus recueillis offre également l'occasion de rappeler aux étudiants qu'un corpus n'a pas besoin d'être volumineux pour être pertinent. Comme le rappellent Tomasello et Stahl (2004), la densité d'un corpus ne se mesure pas uniquement à la quantité de paroles produites, mais à la richesse des informations qu'il contient. Même si l'enfant s'exprime peu, il est possible d'en

tirer de nombreuses observations sur la structure des énoncés, sur les stratégies communicatives employées, ou encore sur les interactions entre langage verbal et non verbal. Les travaux de Debras (2018) et de Canut et Vertalier (2008) montrent d'ailleurs que des analyses qualitatives approfondies peuvent être menées sur des productions individuelles limitées, pour peu que l'observation soit fine et contextualisée.

Cette approche rejoint les réalités auxquelles les étudiants seront confrontés dans leur future pratique clinique. En orthophonie, il est fréquent de rencontrer des patients – enfants ou adultes – dont les productions verbales sont restreintes, fragmentaires ou absentes. Ces situations ne constituent pas un obstacle à l'évaluation linguistique, mais invitent au contraire à développer une attention particulière aux manifestations langagières non verbales : gestes, regards, mimiques, attitudes corporelles, ainsi qu'aux indices interactionnels qui témoignent de la compréhension ou de l'intention communicative du patient. En travaillant sur des corpus de faible densité, les étudiants apprennent ainsi à repérer la valeur informative de chaque fragment produit et à interpréter les comportements communicationnels dans leur globalité.

Pour les besoins de cette activité, il est proposé aux étudiants de réaliser eux-mêmes des enregistrements des productions d'un enfant à l'aide de leur téléphone portable. Ce choix n'est pas anodin : il vise à placer les étudiants dans des conditions de recueil proches de celles qu'ils rencontreront sur le terrain, tout en réduisant la charge technique. Le téléphone est un outil qu'ils utilisent quotidiennement, et il leur permet de se concentrer sur la qualité de l'interaction plutôt que sur la manipulation d'un dispositif d'enregistrement complexe. De plus, les enfants sont aujourd'hui largement familiarisés avec la présence de cet appareil dans leur environnement, ce qui contribue à rendre la situation plus naturelle. Comme le note Habert (2004), l'usage d'un dispositif simple et familier atténue la dimension anxiogène parfois associée à la technique et favorise une interaction plus authentique. L'objectif n'est pas seulement de simplifier la tâche, mais aussi de montrer aux étudiants qu'un outil accessible peut être utilisé à bon escient dans leur pratique future, à condition de respecter les principes éthiques et méthodologiques liés à la collecte de données.

Une fois les enregistrements effectués, une étape essentielle du travail consiste à amener les étudiants à réfléchir de manière critique sur leurs propres choix méthodologiques. Ce moment de recul leur permet de prendre conscience du rôle actif de l'observateur dans la construction du corpus. Ils peuvent ainsi s'interroger sur plusieurs aspects : le moment choisi pour déclencher l'enregistrement (ont-ils enregistré la consigne adressée à l'enfant ou seulement la réponse ?), les éléments intégrés dans le champ en cas d'enregistrement vidéo (ont-ils filmé uniquement l'enfant, ou aussi le dessin qu'il réalise ? ont-ils choisi

une prise de vue incluant l'adulte ?), ou encore la durée et la continuité du recueil (ont-ils interrompu ou poursuivi l'enregistrement selon la dynamique de l'échange ?). Ces choix, parfois effectués sans réflexion consciente, influencent pourtant la nature et l'interprétation des données.

L'analyse collective de ces pratiques en cours permet alors de souligner l'importance du contexte dans l'étude du langage. Les productions langagières ne peuvent être comprises indépendamment du cadre interactionnel et des comportements non verbaux qui les accompagnent. Il est souvent nécessaire de disposer d'indices visuels ou prosodiques pour interpréter correctement certaines formes ou certains silences. Cette prise de conscience est fondamentale, car elle conduit les étudiants à concevoir l'acte de langage comme un phénomène multimodal et situé, et non comme une simple suite de mots.

Enfin, dans la perspective de la rédaction de leurs analyses, les étudiants sont invités à examiner les forces et les limites de leur démarche, et à proposer une méthodologie alternative. Ils apprennent ainsi à identifier les biais possibles, à justifier leurs choix et à formuler des pistes d'amélioration. Cette démarche critique constitue un apprentissage essentiel de la posture de chercheur-praticien : elle les prépare à concevoir, évaluer et ajuster leurs méthodes dans les futures situations d'évaluation ou de recherche. En s'exerçant à observer, analyser et remettre en question leurs propres pratiques, les étudiants développent une compétence réflexive qui leur sera indispensable tout au long de leur carrière d'orthophoniste.

La transcription des productions orales constitue une étape décisive dans la formation à l'analyse linguistique et clinique. L'un des premiers constats effectués avec les étudiants concerne la tendance à attribuer à l'enfant certaines caractéristiques du français oral ordinaire : élisions, dislocations, omission du « ne » de négation, hésitations, reprises, reformulations ou encore pauses remplies. Ces phénomènes, pourtant typiques du français parlé, sont parfois interprétés à tort comme des marqueurs d'immaturité linguistique ou de difficulté langagière. Cet écart d'interprétation révèle un enjeu fondamental : le choix du point de référence adopté par l'étudiant lorsqu'il évalue la parole d'autrui. S'appuie-t-il sur un modèle de langue écrite, normée et stabilisée, ou sur une conception du langage oral, spontané et contextualisé ? Cette question, centrale dans l'enseignement de la linguistique appliquée à l'orthophonie, engage une réflexion épistémologique sur la nature même du langage et sur la diversité de ses formes.

Pour les aider à dépasser ces représentations normatives, il est proposé aux étudiants de transcrire leurs propres productions, issues d'interactions réelles ou simulées. Cet exercice, souvent déstabilisant, leur permet de faire l'expérience directe de la complexité de la langue orale dite « standard ». Ils découvrent ainsi

que les disfluences, les ellipses ou les reprises ne sont pas des anomalies, mais des manifestations naturelles d'un langage en action, soumis aux contraintes de l'interaction et de la cognition. En prenant conscience de ces phénomènes dans leur propre discours, ils développent une écoute plus fine et moins jugeante des productions d'autrui, qu'il s'agisse d'enfants, d'adultes ou de patients présentant des troubles du langage.

Cette démarche introspective les amène également à comprendre que les « accidents de la parole » (Tran & Bécavin, 2001) sont constitutifs d'un oral sain, et qu'il importe de distinguer ce qui relève d'une variation typique de la modalité orale de ce qui pourrait signaler un trouble. Comme le souligne Loufrani (2000), l'orthophoniste doit être capable d'interpréter les productions d'un sujet non pas en fonction d'un idéal linguistique abstrait, mais en fonction de la situation de communication et des conditions de la tâche proposée. Ainsi, un même énoncé pourra être jugé « atypique » dans un contexte scolaire, mais tout à fait attendu dans un échange spontané ou émotionnel.

L'analyse de leurs propres transcriptions conduit les étudiants à s'interroger sur la fréquence et la régularité de phénomènes qu'ils percevaient jusque-là comme marginaux. Le caractère systématique de la transcription – qui impose de noter chaque mot, chaque pause, chaque reprise – met en évidence la récurrence de structures telles que les dislocations ou les amorces interrompues. Cette prise de conscience joue un rôle formatif essentiel : elle permet de déconstruire l'idée d'une langue orale « défectueuse » par rapport à l'écrit et d'adopter une vision plus réaliste de la compétence langagière. Or, cette vision est déterminante pour la pratique clinique future : accompagner un patient, c'est l'aider à développer un usage de la langue fonctionnel et socialement adapté, non à imiter une norme artificielle. Le rôle de l'orthophoniste devient alors celui d'un médiateur entre des usages linguistiques situés et les exigences communicationnelles de l'environnement social du patient (Tran & Bécavin, 2001).

Par ailleurs, la transcription ouvre un espace de réflexion sur les représentations que les étudiants se font de la syntaxe et de son acquisition. Selon les conventions retenues, la transcription peut inclure des annotations spécifiques pour signaler des constructions figées, des formules routinières ou des expressions multi-mots produites par l'enfant. Ce travail d'annotation rend visible la dimension graduelle et dynamique de la syntaxe, telle qu'elle est envisagée par les modèles constructivistes. Behrens (2008) montre par exemple que la mise en évidence de ces segments figés éclaire la manière dont l'enfant construit progressivement des schémas syntaxiques à partir d'unités préformées. Cette perspective peut être mise en lien avec les travaux de Goldberg (1995) sur la grammaire de constructions, selon laquelle la compétence linguistique repose sur l'accumulation et la généralisation d'expressions ancrées dans l'usage.

Ces approches théoriques trouvent un écho direct dans la pratique orthophonique. Comprendre que les productions de l'enfant relèvent souvent de combinaisons partielles ou routinisées, plutôt que de règles abstraites mal appliquées, permet d'adopter une posture d'évaluation et d'intervention plus nuancée. Dans le domaine de la pathologie du langage, cette vision est défendue par Leroy (2013), qui propose d'interpréter certaines erreurs syntaxiques non comme des manques, mais comme des indices du fonctionnement constructif du langage en développement ou en réorganisation.

Enfin, cette expérience de transcription réflexive peut être prolongée par une discussion sur la place des métadonnées dans la constitution du corpus. En cherchant à comprendre pourquoi tel phénomène apparaît ou se répète, les étudiants sont amenés à formuler des hypothèses explicatives qui nécessitent des informations contextuelles : âge du locuteur, situation d'énonciation, type d'interlocuteur, durée de l'échange, thème abordé, degré de familiarité, etc. Cette réflexion sur les variables externes ouvre la voie à une véritable démarche scientifique, dans laquelle la transcription n'est plus une simple étape technique, mais un outil d'observation et d'analyse ancré dans un cadre méthodologique rigoureux.

La transcription

Une fois l'enregistrement réalisé, la transcription constitue une étape centrale dans la constitution d'un corpus et dans la formation des étudiants en orthophonie.

Parce que le travail de transcription peut être long, minutieux et source de questionnements fructueux, nous proposons aux étudiants de travailler en binôme. Après avoir écouté ou visionné chacun leurs enregistrements, ils sélectionnent celui qui leur semble le plus pertinent, puis le transcrivent séparément. Lors de la mise en commun, ils découvrent souvent qu'ils n'ont pas perçu les mêmes éléments (Cappeau, 2008). Cette divergence d'écoute les conduit à prêter une attention renouvelée aux données, à s'interroger sur la subjectivité de la perception auditive, sur leurs propres représentations du langage et de la langue, ainsi que sur leurs attentes vis-à-vis du sujet enregistré. Ils constatent également qu'ils n'ont pas toujours représenté de la même manière les mêmes observables, ce qui les amène à réfléchir aux choix implicites qui interviennent dans tout acte de transcription (Ochs, 1979).

La transcription, en nécessitant des écoutes ou visionnages multiples, permet souvent de révéler des détails auxquels les étudiants n'avaient pas prêté attention lors du premier visionnage : un geste accompagnant un énoncé, une mimique, une réaction non verbale de l'interlocuteur, etc. Ces découvertes les amènent à

réfléchir à leurs représentations du langage et à la place qu'ils accordent à la dimension non verbale dans l'interaction.

En parallèle, les étudiants prennent conscience de leurs attentes initiales vis-à-vis du sujet enregistré. Ainsi, certains, persuadés d'avoir choisi un enfant « qui ne parle pas très bien », réalisent à la transcription que, malgré certaines particularités phonologiques pouvant altérer l'intelligibilité, l'enfant produit des énoncés syntaxiquement complexes. Cette prise de conscience les invite à réfléchir à la nature de leur écoute, souvent orientée vers la détection d'indices pathologiques (cf. Tran & Bécavin 2001 : 31). Grâce à la transcription et à l'analyse de corpus, cette écoute « orientée » peut être confrontée à un ensemble d'observables mesurables et contextualisés dans le système linguistique du sujet cible, ce qui renforce la crédibilité et la rigueur de l'analyse qualitative (Witko 2010 : §18).

Dans ce cadre, il est fréquent que les étudiants attribuent au langage de l'enfant des caractéristiques relevant en réalité du français oral ordinaire : élisions, dislocations, omission du « ne » de négation, hésitations, reprises ou recherches lexicales. Ce constat constitue un point de départ privilégié pour les amener à s'interroger sur le point de référence qu'ils adoptent lorsqu'ils étudient des productions d'enfants : se fondent-ils sur le langage adressé à l'enfant ou sur un modèle écrit, plus normatif ? Transcrire leurs propres productions leur permet de prendre conscience, « de l'intérieur », des spécificités du français oral standard, vivant et naturel. Ils découvrent également que les « accidents de la parole » (Tran & Bécavin 2001) font partie intégrante d'un oral non pathologique, et qu'il importe, dans l'analyse des productions de patients — enfants comme adultes — de distinguer ce qui relève du trouble de ce qui relève simplement de la modalité orale d'une tâche (cf. Loufrani 2000).

L'analyse qui suit la transcription amène les étudiants à identifier les facteurs explicatifs possibles des phénomènes observés. En réfléchissant aux paramètres externes, ils comprennent l'importance de collecter des métadonnées au moment du recueil — telles que la situation de communication, le contexte interactionnel ou la familiarité entre interlocuteurs — afin de formuler des hypothèses solides.

Le caractère systématique de la transcription permet également de mesurer la fréquence réelle de certains phénomènes, comme les dislocations, souvent sous-estimés. Questionner leurs représentations du français parlé devient alors une étape essentielle de leur formation. Dans leur future pratique, les orthophonistes accompagneront des patients vers un usage de la langue « réaliste », adapté à leur environnement social et communicationnel, plutôt que vers un modèle strictement normatif (Tran & Bécavin 2001).

La transcription conduit les étudiants à repenser leurs représentations de la syntaxe et de son acquisition. Selon les conventions adoptées, elle peut inclure l'annotation de constructions potentiellement figées produites par l'enfant (Behrens 2008). Cette approche permet d'introduire les hypothèses constructivistes appliquées à la syntaxe adulte (Goldberg 1995) et à la pathologie du langage (Leroy 2013), en montrant que le développement linguistique repose sur des combinaisons progressives d'unités préformées, issues de l'usage et de l'expérience.

De même que de nombreux choix sont posés en amont de l'enregistrement, la transcription suppose elle aussi un ensemble de décisions préalables — concernant les outils, les conventions ou encore le niveau de détail retenu —, étroitement orientées par l'approche théorique adoptée et par la nature des analyses à venir (qu'elles soient qualitatives, quantitatives ou mixtes). Ces choix techniques ne sont jamais neutres : ils traduisent une conception particulière du langage et influencent directement les observables qui pourront être dégagés du corpus.

La transcription confronte également les étudiants à la question délicate de l'interprétation des productions du sujet cible, notamment lorsque le transcriveur et le locuteur ne partagent pas la même variété de langue, ou lorsque le sujet est un enfant en cours d'acquisition de sa langue maternelle (Morgenstern & Parisse 2007). Cette situation est d'autant plus fréquente dans le cas de pathologies du langage, où la production peut s'éloigner considérablement des normes attendues (Tran & Bécavin 2001).

Lorsqu'ils transcrivent à l'aide de l'orthographe standard, les étudiants se trouvent souvent amenés à interpréter les productions du sujet : décider, par exemple, si une suite sonore correspond à un mot mal articulé, à une forme enfantine, ou à un segment inintelligible. Ils peuvent aussi choisir, au contraire, de suspendre leur interprétation lorsqu'un doute subsiste. Ces choix, conscients ou non, ont un impact direct sur les analyses linguistiques et développementales qui seront ensuite menées.

Afin d'accompagner cette dimension interprétative du travail, il est essentiel de créer des espaces de réflexion et d'échange — en cours, en atelier ou sur des forums en ligne — où les étudiants peuvent expliciter les problèmes rencontrés lors de la transcription. Ces moments collectifs favorisent la mobilisation des concepts et de la terminologie linguistique acquis dans les unités d'enseignement antérieures, tout en développant une posture analytique.

La mise en mots des difficultés rencontrées les conduit à produire des descriptions précises et non prescriptives, témoignant d'une réelle prise en compte des spécificités de l'oral et du langage de l'enfant. Ce travail réflexif

participe à la construction d'un regard critique et distancié sur les différents choix opérés, constituant ainsi une étape formatrice majeure dans l'apprentissage de la démarche scientifique en orthophonie.

L'utilisation du corpus dans la pratique orthophonique :

L'orthophonie connaît actuellement en Algérie un essor particulièrement dynamique. Bien que cette discipline soit relativement récente et encore peu connue du grand public, elle se structure progressivement, portée par la création de nouvelles filières universitaires, la diversification des modules d'enseignement et la mise en place de stages cliniques mieux encadrés. Ces avancées témoignent d'une volonté institutionnelle et scientifique de répondre à un besoin réel : celui de la prise en charge spécialisée des troubles du langage et de la communication dans un pays aux réalités linguistiques et culturelles complexes.

En effet, l'Algérie se caractérise par un paysage linguistique pluriel où coexistent plusieurs langues : l'arabe dialectal (darija), utilisé dans la vie quotidienne ; les langues berbères, dont le kabyle, fortement implanté dans certaines régions ; et le français. Cette situation de plurilinguisme et de diglossie **diglossie** influence directement le développement langagier des enfants, la dynamique de communication familiale et, par conséquent, la pratique orthophonique.

Pour l'orthophoniste, cette diversité linguistique constitue à la fois une richesse et un défi. Elle suppose une adaptation constante à la langue du patient et à celle de son environnement, afin de garantir une évaluation pertinente et une prise en charge efficace. Un trouble du langage ne peut être compris qu'en tenant compte de la ou des langues dans lesquelles il s'exprime. Par conséquent, une approche unilingue ou calquée sur des modèles importés risque de produire des évaluations biaisées.

Dans ce contexte, le recours au corpus linguistique apparaît comme un outil essentiel. Le corpus — entendu comme un ensemble structuré de productions langagières recueillies dans des situations naturelles — permet d'observer et d'analyser le langage tel qu'il se manifeste réellement chez le patient, plutôt que de se limiter à des tests normatifs. Il offre une vision plus fine et contextualisée des capacités langagières, en tenant compte des interactions, des codes sociaux et des variations linguistiques propres à chaque locuteur.

Les avantages du corpus dans la pratique orthophonique sont nombreux :

1)- Une observation authentique du langage : le corpus permet de saisir les productions spontanées, les hésitations, les reformulations et les gestes

accompagnateurs du discours, donnant ainsi accès à la dynamique réelle de la communication (cf. Habert, 2004).

2)- Une meilleure différenciation entre trouble et variation linguistique : dans un contexte multilingue, il aide à distinguer ce qui relève du fonctionnement normal d'une langue ou d'un contact de langues, et ce qui correspond à un véritable trouble du langage (cf. Tran & Bécavin, 2001).

3)- Un outil de diagnostic et de suivi évolutif : en comparant les productions au fil du temps, le corpus permet d'évaluer les progrès réalisés par le patient, d'ajuster les objectifs thérapeutiques et de documenter les effets des interventions.

4)- Un support de réflexion pour les étudiants et les praticiens : la constitution et l'analyse de corpus favorisent le développement d'une posture scientifique, critique et réflexive. Les orthophonistes apprennent à observer, transcrire, coder et interpréter les données langagières selon des méthodologies rigoureuses, tout en prenant conscience de leurs propres représentations du langage (cf. Ochs, 1979).

5)-Un levier pour la recherche locale : les corpus oraux constitués dans le contexte algérien peuvent alimenter des bases de données nationales, utiles à la fois pour la recherche linguistique et pour l'élaboration d'outils d'évaluation adaptés aux langues locales (darija, kabyle, chaoui, etc.).

Malgré ces avantages, l'usage du corpus reste encore faiblement intégré à la pratique orthophonique en Algérie. Sur le terrain, on observe que peu d'orthophonistes l'utilisent systématiquement, souvent faute de formation spécifique, de ressources adaptées ou de temps dédié à cette démarche. La plupart des évaluations reposent encore sur des tests standardisés conçus pour d'autres contextes linguistiques, ce qui peut entraîner des erreurs d'interprétation et limiter la pertinence des diagnostics.

Pour remédier à cela, il serait souhaitable d'accorder plus d'importance au corpus dans la formation universitaire des futurs orthophonistes. Cela impliquerait non seulement un enseignement théorique sur la constitution, la transcription et l'analyse de corpus, mais aussi des ateliers pratiques où les étudiants apprendraient à recueillir et analyser des données langagières issues de patients ou d'enfants en situation naturelle. Un tel apprentissage favoriserait le développement d'une compétence d'observation fine, indispensable pour distinguer les marqueurs linguistiques du trouble des simples variations propres à la langue du patient.

Ainsi, promouvoir une culture du corpus dans la formation et la pratique orthophonique en Algérie, c'est ouvrir la voie à une orthophonie plus contextualisée, plus rigoureuse scientifiquement et plus respectueuse de la diversité linguistique du pays. Cela contribuerait à renforcer la qualité du diagnostic et à optimiser la prise en charge des patients, tout en participant à la construction d'un savoir scientifique local sur les particularités du langage en contexte multilingue.