



**HAL**  
open science

## Transcrire et analyser les corpus d'interactions adulte-enfant

Christophe Parisse, Aliyah Morgenstern

► **To cite this version:**

Christophe Parisse, Aliyah Morgenstern. Transcrire et analyser les corpus d'interactions adulte-enfant . Edy Veneziano, Anne Salazar Orvig, Josie Bernicot. Acquisition du langage et interaction, L'Harmattan, pp.201-222, 2010. halshs-00531693

**HAL Id: halshs-00531693**

**<https://shs.hal.science/halshs-00531693>**

Submitted on 10 Jan 2012

**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.

A paraître dans « Transcrire et analyser les corpus d'enfant. In Edy Veneziano, Anne Salazar Orvig, Josie Bernicot Eds, Acquisition du langage et interaction. Paris : L'Harmattan »

## **Transcrire et analyser les corpus d'interactions adulte-enfant**

Christophe Parisse et Aliyah Morgenstern

*MoDyCo, INSERM et Université Sorbonne Nouvelle - Paris 3*

### **Introduction**

Ce chapitre a pour objectif de partager avec nos lecteurs le fruit de la réflexion que nous menons depuis des années sur le recueil, la transcription, le codage et l'analyse des productions spontanées des enfants en interaction et sur les outils informatisés qui sont utilisés dans les travaux sur l'acquisition. Nous espérons ainsi aider de jeunes collègues à entrer dans la communauté des chercheurs de terrain qui font de la linguistique de corpus et partagent des données de qualité, ce qui pourra ainsi faire avancer notre connaissance générale sur le langage de l'enfant.

Nous pensons que même si les situations expérimentales présentent un grand intérêt, le corpus spontané des productions des enfants en interaction est fondamental pour comprendre l'appropriation progressive du langage comme l'a en particulier montré Brown (1973). Ce type de corpus permet de garantir que l'on analyse les productions langagières dans des situations naturelles et de tester ainsi les limites des théories que nous utilisons et des hypothèses que nous posons. Il nous permet également de nous laisser surprendre par les productions langagières et d'observer l'imprévu, tout en n'analysant que ce qui existe réellement et non les créations du chercheur.

Par ailleurs, le travail sur les productions spontanées de l'enfant dans un contexte dialogique montre qu'il s'agit d'une situation exemplaire pour

- travailler sur un corpus ORAL avec tous les problèmes techniques et scientifiques que cela pose ;
- avoir des données vidéos et textuelles via les transcriptions et les avoir alignées ;
- comprendre la nature interprétative des codages, mais aussi du langage.

De nos jours, les chercheurs qui travaillent sur des suivis longitudinaux d'enfants en milieu naturel peuvent filmer les

A paraître dans « Transcrire et analyser les corpus d'enfant. In Edy Veneziano, Anne Salazar Orvig, Josie Bernicot Eds, Acquisition du langage et interaction. Paris : L'Harmattan »

enfants à leur domicile. Les enregistrements sont ensuite transcrits et informatisés. Mais, si l'acquisition du langage est sans doute le premier domaine dans lequel on a systématiquement utilisé des données conversationnelles spontanées, avant de pouvoir s'appuyer sur toutes ces technologies, les premiers observateurs du langage enfantin avaient adopté une autre méthode de recueil des données : la méthode du journal (Morgenstern, sous presse). Les journaux des scientifiques de la fin du dix-neuvième siècle et du début du vingtième siècle avaient l'extraordinaire avantage de permettre au chercheur, le plus souvent l'un des parents de l'enfant, de faire des descriptions quotidiennes. Certains scientifiques ont d'ailleurs tenu des journaux sur leurs propres enfants, notamment entre la fin du XIX<sup>e</sup> et le début du XX<sup>e</sup> siècle (Taine 1976, Darwin 1977). Les études de cas, détaillées, approfondies, ancrées dans la réalité quotidienne permettent de rendre compte du développement de l'enfant, de mettre au jour les liens entre le langage, la cognition, le psychisme, et le développement moteur. Depuis les années 1950, le magnétophone, puis le caméscope ont révolutionné la discipline et depuis les années 1980, s'ajoute encore la possibilité d'utiliser des logiciels pour faire des transcriptions informatisées, puis d'aligner les transcriptions avec le document source (enregistrement audio ou vidéo) et de mutualiser facilement les données par internet.

Quand on transcrit des corpus d'interactions adulte-enfant, et en particulier si l'on veut faire mettre les corpus à disposition d'autres chercheurs, il est nécessaire de prendre en compte différents domaines d'analyse et des approches variées qui concernent les productions de l'enfant, ou de l'adulte, en prenant en compte le niveau prosodique, phonologique, phonétique, morpho-syntaxique, pragmatique, discursif... mais aussi le non verbal (gestes, regards, mimiques, actions, manipulations d'objets, jeux, routines) et le contexte extra-linguistique (situations diverses). Aussi avons-nous intérêt à faire la description la plus fine possible du contexte extra-linguistique tout en gardant en parallèle l'alignement des transcriptions et des enregistrements vidéos (Morgenstern, Parisse 2007).

S'il est nécessaire de rassembler des bases de données les plus riches possibles de suivis longitudinaux d'enfants en interaction, il est également important de constituer son propre corpus afin de donner vie à sa recherche dans la relation que nous nouons avec

A paraître dans « Transcrire et analyser les corpus d'enfant. In Edy Veneziano, Anne Salazar Orvig, Josie Bernicot Eds, Acquisition du langage et interaction. Paris : L'Harmattan »

l'enfant et sa famille. Quand on pénètre dans le milieu familial, et que l'on occupe soi-même la position d'observateur, on doit chercher à s'y sentir bien, en tester les limites, être accepté et ne pas faire intrusion. Ce processus d'ajustement à la situation « de terrain » permet d'entrer concrètement dans l'activité de recherche. Le fait de donner une existence officielle à cette activité de recueil de données en faisant signer un document légal d'autorisation parentale est une manière de s'assurer une légitimité et de donner corps au « corpus » et au monde de la recherche que nous représentons auprès de la famille qui nous a acceptés. C'est ainsi que la relation que nous avons nouée avec la famille et qui s'inscrit dans le temps, s'inscrit également dans l'espace scientifique.

Nous présenterons tout d'abord les raisons que l'on peut avoir de transcrire du corpus longitudinal avec des outils informatisés. Nous discuterons également les problèmes que posent la fiabilité des transcriptions et leur côté subjectif qu'il est impossible de nier ou d'éradiquer. Toute transcription est en effet le résultat des positions théoriques et de l'interprétation du chercheur. Nous parlerons ensuite des méthodes de transcription et des choix scientifiques qu'il est important de faire. Nous finirons ce chapitre sur la réalisation pratique de transcriptions et sur les possibilités qu'ouvre l'interopérabilité des logiciels de transcription en vue de recherches pluridisciplinaires et multimodales sur les mêmes données.

## **1. Transcrire avec des moyens modernes**

### **1.1. Pourquoi transcrire avec des moyens modernes?**

Il est de nos jours absolument nécessaire de transcrire ses enregistrements audio ou vidéo avec des outils informatisés. Ces outils peuvent sembler fastidieux au premier abord, effrayer le novice, mais ils rendent le travail du transcripateur beaucoup plus confortable et rigoureux. Ils permettent d'aller plus vite une fois qu'on les maîtrise, d'utiliser des formats de codage normalisés, de faire des analyses de fréquence et surtout d'aligner les transcriptions avec les enregistrements, ce dont on ne saurait aujourd'hui se passer. Grâce à ces outils informatisés, on se donne également les moyens d'entrer dans une communauté

A paraître dans « Transcrire et analyser les corpus d'enfant. In Edy Veneziano, Anne Salazar Orvig, Josie Bernicot Eds, Acquisition du langage et interaction. Paris : L'Harmattan »

internationale de chercheurs qui mutualisent leur données via internet, de pérenniser les données et on apporte sa pierre à la construction des connaissances sur le langage en général et le langage de l'enfant en particulier.

L'interprétation des productions enfantines est difficile et souvent subjective en raison de la distance entre les énoncés des enfants et la représentation que nous avons du langage oral, venues souvent de notre apprentissage et utilisation du langage écrit et de ses normes. Transcrire du langage enfantin est une excellente école pour comprendre la nature multi-modale de la communication langagière en situation. En effet, si la distance entre productions et transcriptions semble se réduire à mesure que l'enfant grandit, elle n'est jamais abolie et il faudra toujours se servir des indices supra-segmentaux, visuels et extra-linguistiques pour donner un sens et une forme aux données. Faut-il interpréter ? C'est une particularité forte du type de discours et d'interaction étudiés en acquisition du langage : il est fréquent que les productions langagières et/ou gestuelles ne soient pas univoques. C'est le cas par exemple des onomatopées comme « poum », « hop », etc., à propos desquelles il importe de savoir si elles ont été prononcées en tapant dans une balle ou en sautant, par exemple. On voudrait, grâce à la transcription systématique des événements et gestes accompagnant ce qui est dit, pouvoir associer la production verbale au contexte, et tenir compte de cette association dans nos analyses quantitatives et qualitatives.

### **1.2. Qu'est-ce qu'une transcription ?**

Il existe un grand nombre de formes de transcription en fonction des buts scientifiques visés, mais aussi en fonction du modèle théorique sous-jacent. Les transcriptions peuvent aussi se faire à de multiples niveaux : phonétique, phonologique, orthographe, syntaxe, sémantique, actes de langage, pragmatique, gestes, pointages, interactions avec autrui, avec la scène, etc. La situation de production de langage (interactions, récit oraux, textes écrits) joue également de manière très importante sur les formats de transcription. Enfin la nature des locuteurs (bébés, jeunes enfants, adultes) ou des conditions de recueil (laboratoire, lieu

A paraître dans « Transcrire et analyser les corpus d'enfant. In Edy Veneziano, Anne Salazar Orvig, Josie Bernicot Eds, Acquisition du langage et interaction. Paris : L'Harmattan »

ordinaire, endroit bruyant) modifie également la manière de transcrire et les codes utilisés.

La quasi-totalité des transcriptions de langage est fondée sur le modèle de la transcription écrite, c'est-à-dire de la traduction dans le système de la langue écrite des situations de parole. Ceci veut dire que les transcriptions comportent en général une ligne dite principale qui contient des éléments dont la plupart sont en format orthographique standard. Dans le cas de transcriptions de l'oral, il existe en général tout un ensemble de conventions permettant de modifier certains mots en fonction de leur prononciation effective à l'oral et d'ajouter au format écrit ordinaire un large ensemble de codes permettant de signaler des éléments non-verbaux, les pauses, les intonations, les hésitations, les reprises, les recouvrement d'énoncés entre locuteurs, etc. Enfin, il faut aussi indiquer systématiquement le nom des locuteurs. On trouve aussi dans certains systèmes des codes qui vont au-delà de la transcription et qui servent à la caractériser. Par exemple, on trouve des codes pour caractériser les écarts à la norme syntaxique, pour décrire la structure syntaxique. Il existe ainsi de nombreuses conventions parmi lesquelles on peut citer celles de CHILDES<sup>1</sup>, du GARS/DELIC<sup>2</sup>, de VALIBEL<sup>3</sup>, etc.

Cependant, dans le cas de données orales ou signées, il est très difficile de résumer une transcription à une simple ligne orthographique, même si celle-ci est largement enrichie de notes diverses. En effet, il se pose à un certain moment des problèmes de lisibilité, de suivi des modifications et de contrôle des formats. C'est pourquoi on trouve souvent des « lignes » d'annotations supplémentaires. Par exemple, pour les jeunes enfants on trouvera souvent une ligne phonétique et/ou phonologique, une ligne situation, éventuellement une ligne geste, pointage, regard, intonation, etc. Ces lignes, largement utilisées dans le système CHILDES et reprises dans presque tous les autres systèmes, sont appelées en général « tire », de l'anglais tier, signifiant niveau.

---

<sup>1</sup> <http://childes.psy.cmu.edu/>

<sup>2</sup> <http://sites.univ-provence.fr/delic/corpus/conventions.html>

<sup>3</sup> [http://www.uclouvain.be/cps/ucl/doc/valibel/documents/conventions\\_valibel\\_2004.PDF](http://www.uclouvain.be/cps/ucl/doc/valibel/documents/conventions_valibel_2004.PDF)

A paraître dans « Transcrire et analyser les corpus d'enfant. In Edy Veneziano, Anne Salazar Orvig, Josie Bernicot Eds, Acquisition du langage et interaction. Paris : L'Harmattan »

Elles sont systématiquement apparues avec la généralisation des logiciels de transcription.

### **1.3. Les limites et les dangers des transcriptions**

Il nous est donc imposé d'opérer des choix théoriques importants. Ces choix sont une grande responsabilité car ils auront ensuite un effet considérable sur les analyses scientifiques menées. Tout corpus est une construction, au sens où il est toujours déjà le produit des analyses du chercheur (Ochs, 1979). Ainsi les procédures de transcription puis de codage des données méritent une attention toute particulière, si l'on veut éviter de dénaturer les productions.

Les méthodes de compilation, de transcription et de présentation des données ont été très discutées depuis la fin des années soixante dix, en raison des nombreuses variables intervenant à chaque étape du processus. On peut se demander si la transcription doit être de nature phonétique (pour plus d'authenticité), phonologique, lexicale, orthographique (pour plus de lisibilité et pour permettre les analyses automatiques comme sur la morphosyntaxe en utilisant des outils « standards » basés sur les conventions orthographiques) et jusqu'où aller dans les détails étant donné que l'on peut toujours recourir, grâce à l'alignement, aux données sources. Par ailleurs, faut-il opérer une sélection, et si oui selon quels critères ? À quelles conditions les informations ajoutées, et notamment celles qui ne concernent pas les productions verbales stricto sensu, sont-elles un enrichissement plutôt qu'une gêne ? Les choix techniques sont des mises en applications des choix scientifiques et c'est le va et vient entre technique et recherche scientifique qui doit être constamment en cours.

On peut remettre en question la notion de ligne principale et son contenu qui conditionne les analyses et celles des chercheurs qui utilisent des données déjà recueillies et transcrites. Dans quelle mesure l'interprétation des productions est-elle fiable ? Faut-il faire appel à un accord inter-juge ? Est-il possible d'avoir un accord suffisamment « bon » sur des données d'enfants très jeunes dont les productions sont phonologiquement très instables et ne peuvent se comprendre qu'en utilisant à la fois les informations situationnelles, contextuelles et nos connaissances sur l'enfant ? Il

A paraître dans « Transcrire et analyser les corpus d'enfant. In Edy Veneziano, Anne Salazar Orvig, Josie Bernicot Eds, Acquisition du langage et interaction. Paris : L'Harmattan »

ne semble possible de construire une compréhension suffisamment fine des productions de l'enfant qu'à partir d'observations fréquentes qui nous permettent de nous acclimater à la musique propre à SA langue, aux situations particulières dans lesquelles il se trouve et à ses capacités cognitives, motrices, pratiques, intellectuelles... qui conditionnent sa pratique langagière. Est-ce à dire qu'il est impossible pour un chercheur qui ne connaît pas l'enfant et n'a que la vidéo de transcrire ? Certainement pas, mais il nous faut être humble face à la multitude d'informations nécessaires pour s'assurer de la fiabilité d'une transcription et être bien conscient que nous ne pourrions jamais totalement reconstruire ce que l'enfant aura voulu dire... d'autant plus qu'il lui arrive de produire une musique, des formes sans autre intention que le fait d'interagir avec autrui ou tout simplement de se faire plaisir. Même dans le monde idéal du psycholinguiste, l'activité de transcription ne pourra pas être une science exacte, et peut-être est-ce ce qui fait la magie du langage : les signaux acoustiques et visuels émis par les locuteurs ne sont pas représentables sans un minimum d'interprétation qui correspond à des décisions auxquelles il est important de réfléchir. Chaque décision a un effet sur les analyses manuelles ou automatiques qui suivront. Il est donc important d'une part de bien avoir en tête son objet d'étude pour faire les bons choix de transcriptions et d'autre part, si l'on n'a pas transcrit soi-même les données, de bien connaître les transcriptions sur lesquelles on veut travailler avant de décider de les utiliser pour un projet de recherche bien particulier.

Il reste néanmoins que, quelle que soit la technique utilisée, il n'y a pas une seule transcription possible mais une multitude, toutes « justes » d'un certain point de vue. Pour s'affranchir de cette limitation intrinsèque, la seule méthode efficace est la description de la manière dont s'est déroulée la transcription et du choix des conventions et des techniques utilisées. L'adoption d'un standard normalisé ou connu, d'un outil informatique partageable et l'utilisation conjointe à la transcription du son ou de la vidéo, aide à garantir cette transparence, et est aussi un garant de qualité du travail scientifique réalisé avec les corpus.



A paraître dans « Transcrire et analyser les corpus d'enfant. In Edy Veneziano, Anne Salazar Orvig, Josie Bernicot Eds, Acquisition du langage et interaction. Paris : L'Harmattan »

## 2. Comment transcrire ?

### 2.1. Formats et présentation de données.

En dehors de la difficulté inhérente au processus de transcription, il existe plusieurs manières de présenter les données ainsi transcrites qui présentent toutes des avantages et des inconvénients. Il n'existe pas de présentation et de format « parfait », mais toute une série de modèles qui sont plus ou moins adaptés aux différents usages que l'on peut en faire. Dans la mesure où nous recommandons d'utiliser des logiciels de transcription car ils sont les seuls à offrir des garanties sur le respect d'un format et la possibilité de partage des données, nous nous contenterons ici de présenter les deux principaux formats disponibles dans les logiciels gérant des corpus de langage oral et qui sont utilisés dans les travaux sur le langage de l'enfant.

#### Format texte

Le format texte est une présentation verticale des transcriptions comme un texte classique. L'axe vertical, de haut en bas, représente l'axe du temps ou la succession temporelle. L'espace d'écriture de chaque ligne est parfois utilisé comme moyen complémentaire de codage ou d'alignement temporel. Chaque ligne principale est en général suivie d'un ou plusieurs tires qui contiennent des informations dépendantes de la ligne principale. Un exemple classique de ce type de format est celui du logiciel CLAN (CHILDES), mais c'est aussi le cas de Transcriber ou de PHON.

```
*CHI: je [ʃ] j' appuie sur vert ? •
%mor: pro:subj|je v|appuyer prep:art|sur adj|vert ?
%pho: zə apɥi sɥv vɛʁ
%add: à FAT .
%act: CHI appuie sur un bouton de la machine à café .
*FAT: oui vas+y . •
%mor: adv:yn|oui v:mdllex|vas+y .
```

**Exemple de partition issu du programme CLAN (les « • » représentent des indicateurs d'alignement temporel texte/vidéo).**

A paraître dans « Transcrire et analyser les corpus d'enfant. In Edy Veneziano, Anne Salazar Orvig, Josie Bernicot Eds, Acquisition du langage et interaction. Paris : L'Harmattan »

Certains logiciels comme PHON permettent de lier directement les annotations (par exemple la transcription phonétique) à des parties de la ligne principale (par exemple les mots puis les phonèmes). D'autres plus nombreux comme CLAN ne font pas de lien explicite, le seul lien existant étant celui de la représentation choisie par l'utilisateur (dans l'exemple ci-dessus l'utilisation d'espaces).

Le format texte est très bien adapté à l'analyse textuelle, en particulier les récits, mais aussi les successions d'interactions lorsqu'une analyse s'intéresse à des suites d'échanges qui sont plus facilement visualisés en format texte qu'en format partition (voir ci-dessous). Le format texte permet également d'afficher un grand nombre de lignes dépendantes en même temps. Le format texte présente par contre des limitations importantes dans la représentation des dialogues et en particulier des chevauchements entre locuteurs. Il est également moins bien adapté à une situation d'interaction impliquant plus de deux ou trois personnes.

#### **Format partition**

Le format partition est une présentation horizontale des transcriptions qui ressemble à une partition de musique. Le déroulement du temps est représenté par l'horizontale de gauche à droite. Selon les logiciels, cet axe horizontal est proportionnel au déroulement temporel ou à la taille du texte transcrit. Dans ce format, à chaque interlocuteur correspond une ligne de la partition. Des lignes supplémentaires permettent de coder les lignes dépendantes de chaque interlocuteur. Ce format est largement utilisé dans les logiciels servant à coder les interactions de langage (par exemple ELAN ou EXMARaLDA), mais c'est aussi la présentation utilisée dans un logiciel comme PRAAT (bien que spécialisé dans le traitement du son, PRAAT permet la transcription et l'alignement temporel).

A paraître dans « Transcrire et analyser les corpus d'enfant. In Edy Veneziano, Anne Salazar Orvig, Josie Bernicot Eds, Acquisition du langage et interaction. Paris : L'Harmattan »



Exemple de partition issue du programme ELAN

Comme le format texte, ce type de format présente des avantages et des inconvénients. Les avantages sont une précision extrême dans les chevauchements d'énoncés et dans les alignements temporels en général, la possibilité de visualiser facilement les interactions de nombreux interlocuteurs et la liberté de maniement des annotations dépendantes par rapport aux annotations principales. Les longs énoncés ou les tours de paroles et les annotations secondaires, si elles sont en grand nombre, peuvent cependant être assez peu lisibles.

## 2.2. Choix d'un modèle de transcription

De manière idéale, le choix d'un système de transcription ne devrait être fait que sur des critères d'efficacité ou de besoin de la recherche. On ne saurait qu'encourager les usagers à connaître et apprendre de nombreux modes de transcription et présentation de corpus afin de faire le meilleur choix. Néanmoins, l'utilisation d'un type de système ou d'un autre relève souvent de l'habitude ou de l'appartenance à une communauté (par exemple : acquisition du langage oral, interactions ou langues signées). Les utilisateurs peuvent aussi être amenés à choisir un format parce qu'il n'existe qu'un seul logiciel qui remplisse exactement les besoins des recherches qu'ils mènent. Dans les deux cas, ceci conduit parfois à compenser d'une manière ou d'une autre les faiblesses d'un système imposé et pousse parfois les utilisateurs à utiliser ou continuer d'utiliser un système qui n'est pas toujours optimal pour leurs besoins. L'idéal en la matière serait une meilleure compatibilité entre logiciels qui permette de passer plus facilement de l'un à l'autre, et de même utiliser deux logiciels différents pour la même recherche lorsque le besoin existe. A cet égard de nombreux progrès ont été réalisés pour atteindre une plus grande

A paraître dans « Transcrire et analyser les corpus d'enfant. In Edy Veneziano, Anne Salazar Orvig, Josie Bernicot Eds, Acquisition du langage et interaction. Paris : L'Harmattan »

interopérabilité des logiciels et il est souvent possible de convertir effectivement des transcriptions. Cet exercice peut, hélas, parfois imposer une limite à la nature des éléments convertis et il est important de les connaître et d'être prudent avant d'entreprendre une conversion.

### **2.3. Planifier et organiser des transcriptions**

Le temps consacré à réaliser des transcriptions peut être très long. Il varie largement selon l'entraînement des transcribers, le type de format exigé, le niveau de qualité visé, la situation de recueil de l'enregistrement sonore ou vidéo, le type de production langagière des locuteurs ainsi que leur débit. Le temps nécessaire pour une transcription s'évalue en mesurant la proportion du temps mis à transcrire par rapport à la durée de l'enregistrement.

Pour une simple transcription de ligne principale orthographique, on pourra avoir dans les meilleurs cas des rapports de cinq à six pour des transcribers rapides, un langage parfaitement audible et bien enregistré, une diction claire. Ceci veut dire qu'il faudra cinq à six heures pour transcrire une heure d'enregistrement. Ces performances sont rarement atteintes et nécessitent une grande pratique et une utilisation optimale des matériels ou logiciels à sa disposition, comme par exemple le « SoundWalker » de CLAN qui reproduit l'utilisation d'une pédale de contrôle du défilement du son. Le rapport effectif est plus souvent de dix à quinze pour un travail standard. Dans le cas où le milieu sonore est mauvais ou la langue produite difficile à percevoir (par exemple dans le cas de jeunes enfants, de troubles de langage, ou d'accents spécifiques), les temps peuvent s'allonger, les rapports monter à vingt fois le temps de l'enregistrement.

Lorsque la transcription nécessite des lignes secondaires en plus de la ligne principale, le temps de transcription peut aisément doubler. Ainsi l'ajout d'une ligne phonétique peut prendre beaucoup de temps, surtout si elle doit être réalisée avec précision sur tous les locuteurs. Il faut aussi faire attention dans ce cas de figure à utiliser les logiciels les plus adaptés pour l'exercice. Pour les études de dialogue enfant-adulte, ou d'interactions, on est

A paraître dans « Transcrire et analyser les corpus d'enfant. In Edy Veneziano, Anne Salazar Orvig, Josie Bernicot Eds, Acquisition du langage et interaction. Paris : L'Harmattan »

souvent amené à ajouter beaucoup de codes supplémentaires (non-verbal, situations, gestes, etc.).

Enfin le temps global peut aussi se modifier en fonction du temps consacré au contrôle des transcriptions. Dans le cas de double transcription complètement réalisée en aveugle, il faut compter deux fois le temps total plus le temps nécessaire pour obtenir un accord entre transcrip-teurs. Ce travail est trop lourd pour être réalisé pour toute transcription et il est en général réservé à quelques échantillons permettant de mesurer la qualité de transcription. La pratique la plus courante consiste à faire vérifier les transcriptions, soit par l'investigateur de l'étude, soit par un spécialiste. Une telle vérification prend un temps important, variable selon la quantité de corrections à apporter à la première transcription. Un temps de deux à quatre fois le temps original n'est pas exceptionnel, surtout dans le cas de transcriptions délicates à réaliser.

Lorsque l'on tient compte de l'ensemble des éléments nécessaires pour réaliser du corpus de langage oral (en incluant aussi le temps de tournage et de digitalisation), on se rend compte du temps que nécessite une étude complète. Sur la base d'un temps (tout compris) allant de quinze à quatre-vingt dix heures par heure de corpus selon les conditions du corpus, on comprend qu'un corpus qui représente un suivi mensuel de deux ans coûte entre trois cent et mille huit cent heures de travail, c'est-à-dire environ deux à douze mois de travail à temps plein.

Ces chiffres expliquent combien il est important de mutualiser les données de corpus, à la fois pour des raisons économiques mais aussi tout simplement pour permettre de réaliser certains travaux dans des temps raisonnables. La mutualisation ne doit malgré tout pas limiter la réalisation de nouveaux recueils de corpus qui sont importants pour obtenir de nouvelles données, couvrir des champs complémentaires ou vérifier des résultats précédents, mais aussi pour apprendre réellement ce qu'est la production de langage, que ce soit celle d'enfants, d'adultes ou de personnes souffrant de troubles du langage. En ce sens, le recueil et la transcription de corpus reste la meilleure école qui soit, mais il convient, en regard du temps nécessaire pour obtenir un corpus exploitable, de correctement planifier le recueil et la transcription, faute de quoi la

A paraître dans « Transcrire et analyser les corpus d'enfant. In Edy Veneziano, Anne Salazar Orvig, Josie Bernicot Eds, Acquisition du langage et interaction. Paris : L'Harmattan »

transcription peut occuper tout le temps de travail consacré à un mémoire ou même à une thèse.

### **3. Réaliser des transcriptions**

Le travail sur un corpus de langage oral passe par quatre phases : (a) l'enregistrement (vidéo ou son), (b) la digitalisation (transformation du film ou du son en un format exploitable par un logiciel de transcription), (c) la transcription (pour laquelle il existe de nombreux outils), (d) l'exploitation du corpus (fonctions de recherche, statistique et visualisation des outils de corpus).

#### **3.1. L'enregistrement**

L'enregistrement de corpus est une étape fondamentale car elle est délicate et ne peut être recommencée comme une prise de vue de cinéma. Plusieurs éléments fondamentaux doivent être contrôlés avec soin.

La situation doit être préparée au mieux. De nombreux éléments rentrent en compte, notamment les besoins scientifiques (situation, milieu, nombre de locuteurs, etc.). S'il est clair que la réalisation doit se plier aux aspects scientifiques, il faut chercher à se mettre dans les meilleures conditions de réalisation : lumière correcte, absence de bruits annexes, espace pour déplacer aisément la caméra. Les éléments techniques peuvent dans certains cas compenser certaines limitations de la situation, comme l'utilisation d'un grand angle dans des espaces réduits ou celle d'un micro directionnel externe à la caméra dans le cas de bruits environnants. On peut aussi multiplier les caméras ou les micros lorsque le contrôle de la situation est fondamental. Dans une certaine mesure, la qualité des matériels vidéo ou audio récents (surtout lorsque le matériel acheté est de bonne facture, c'est-à-dire un ordre de prix allant de 500 à 800€<sup>4</sup>) facilite le recueil de nouveaux enregistrements en assurant une bonne qualité vidéo ou sonore

---

<sup>4</sup> Nous avons tenu à donner des prix pour fixer les ordres de grandeur du coût du matériel à utiliser pour une étude de corpus, hors coût de l'ordinateur lui-même. Ces prix correspondent à l'année 2009 et sont susceptibles d'évoluer, le plus souvent à la baisse, avec l'amélioration des techniques et la diffusion de ce type de matériel.

A paraître dans « Transcrire et analyser les corpus d'enfant. In Edy Veneziano, Anne Salazar Orvig, Josie Bernicot Eds, Acquisition du langage et interaction. Paris : L'Harmattan »

même sans l'utilisation de matériel spécialisé comme des micros externes. Ceux-ci, s'ils sont de bonne qualité, sont toutefois indispensables dans tous les travaux exigeants du point de vue acoustique.

La plupart des caméras actuelles peuvent convenir, mais il vaut mieux éviter les formats de type DVD qui sont le plus souvent difficiles à convertir. Les formats tout numérique sont une bonne option, ainsi que les formats Mini-DV. La qualité haute-définition est un plus, mais n'est pas nécessaire.

### **3.2. La digitalisation**

L'ère du tout numérique a grandement facilité cette part du travail. Pour des recueils anciens, dits analogiques, on est obligé de passer par du matériel spécialisé (cartes d'acquisition numérique pour ordinateur, ou caméra numérique externe intermédiaire). Les cartes (internes ou externes) pour ordinateur ne sont plus très coûteuses (moins de 150€ en général) mais tous les matériels n'offrent pas une très bonne qualité de digitalisation, ou ne permettent pas tout type de transformation.

Les caméras numériques de type Mini-DV ou de type support mémoire (format MPEG-4, AVCHD) permettent une efficacité maximale car le processus est une simple conversion et non une digitalisation. Ce procédé s'opère sans perte de qualité. La multiplication des systèmes et des modèles, alliée à la grande mobilité de ce secteur économique, nous interdit de proposer des solutions clés en mains. Il faut simplement savoir que d'une façon générale il faut aller de la caméra à l'ordinateur (par la prise Firewire dans le cas de Mini-DV, par transfert de données pour les mémoires). Il faut ensuite réaliser une conversion de données, du format origine de la caméra vers le format utilisé par le logiciel de transcription (son, vidéo Quicktime, vidéo Windows Media Vidéo, selon le cas). Ces deux étapes nécessitent souvent l'achat d'outils peu coûteux comme un logiciel de montage vidéo (I-Movie est gratuit sur les Mac d'Apple, de nombreux logiciels payant sont disponibles sur les PCs pour des prix allant de 50 à 100€) et comme un programme de conversion vers Quicktime. La capacité de conversion vers Windows Media Vidéo est en général

A paraître dans « Transcrire et analyser les corpus d'enfant. In Edy Veneziano, Anne Salazar Orvig, Josie Bernicot Eds, Acquisition du langage et interaction. Paris : L'Harmattan »

gratuitement assurée par les logiciels de montage vidéo. Le format DVD, peu recommandé, nécessite des programmes spécifiques de conversion (de type « DVD Ripper ») disponibles sur Internet à des coûts de l'ordre de 30€ à 50€.

### 3.3. La transcription

Il existe une dizaine de logiciels permettant la transcription de dialogues oraux, avec alignement à la vidéo ou au son. Chaque logiciel, au-delà de son type de présentation (voir ci-dessus présentation en mode texte ou partition), vise un usage spécifique et présente des avantages et des inconvénients. Nous ne décrivons pas l'ensemble de ces logiciels ici (voir <http://icar.univ-lyon2.fr/projets/corinte/confection/alignement.htm> pour une liste très complète) mais listerons simplement les six logiciels les plus utilisés pour travailler sur le langage de l'enfant (voir aussi Rohlfing & al., 2006, pour une description des logiciels au format partition). Tous ces logiciels sont gratuits.

Nom	Praat	Phon	Clan
Origine	www.praat.org	phon.ling.mun.ca/phontrac/wiki/	childes.psy.cmu.edu
Spécificité	Phonétique, signal sonore	Phonologie	Récit, Interactions
Système	PC/Mac/Linux	PC/Mac	PC/Mac
Présentation	Partition	Texte	Texte
Vidéo/Son	Son	Vidéo	Vidéo ou Son
Format	Wav, Aif, MP3, Au	Quicktime	Quicktime
Type de fichier	TextGrid	Phon et TalkbankXml	Chat - TalkbankXml
Importation		Clan, Praat (prévision)	Elan, Exmaralda, Praat
Exportation		Clan, Praat (prévision)	Elan, Exmaralda, Praat, Texte



A paraître dans « Transcrire et analyser les corpus d'enfant. In Edy Veneziano, Anne Salazar Orvig, Josie Bernicot Eds, Acquisition du langage et interaction. Paris : L'Harmattan »

			(avec commande COMBO)
Commandes annexes	Langage de script très puissant	Recherche	Large bibliothèque de fonctions

**Tableau 1a : Description synthétique des logiciels de transcription**

Nom	Elan	Exmaralda	Transcriber
Origine	www.lat-mpi.eu/tools/elan	exmaralda.org	trans.sourceforge.net
Spécificité	Interaction multi-locuteurs	Interactions, textes	Récit, Interactions
Système	PC/Mac/Linux	PC/Mac/Linux	PC/Mac/Linux
Présentation	Partition	Partition	Texte
Vidéo/Son	Vidéo et Son	Vidéo ou Son	Son
Format	Quicktime, Windows Media, Java Media	Java Media Framework	Tous formats courants
Type de fichier	EAF (Xml)	XML	Trs XML
Importation	Clan, Transcriber, Shoebox	TASX, Praat	Clan, autres
Exportation	Shoebox, Clan, Texte (plusieurs types), autres	TASX, Praat, formats Html	Texte
Commandes annexes	Recherche	Logiciels complémentaires puissants	

**Tableau 1b : Description synthétique des logiciels de transcription**

Les six logiciels présentés dans les tableaux ci-dessus permettent tous de présenter des transcriptions en relation avec du

A paraître dans « Transcrire et analyser les corpus d'enfant. In Edy Veneziano, Anne Salazar Orvig, Josie Bernicot Eds, Acquisition du langage et interaction. Paris : L'Harmattan »

son ou de la vidéo, ce qui les rend adaptés à l'analyse du langage oral et des interactions. Ils ne sont pas concurrents dans la mesure où ils remplissent tous des fonctions différentes. Ils ont également tous une grande panoplie de fonctions qui leur offrent une certaine adaptabilité et n'oblige pas à systématiquement changer de logiciel lorsque l'on change de mode de travail. Cette souplesse ne devrait toutefois pas empêcher l'utilisateur averti de changer de logiciel lorsque cela peut lui apporter des solutions techniques efficaces. On peut profiter des capacités de conversion entre logiciels pour passer d'un logiciel à l'autre et utiliser intelligemment le logiciel le plus approprié en fonction de ses besoins du moment. L'utilisateur non-spécialiste peut choisir son logiciel en fonction des caractéristiques décrites dans la troisième ligne des tableaux 1a et 1b ci-dessus. Certains logiciels très spécialisés comme PRAAT et PHON ne sont pas concurrents des autres logiciels car l'utilisateur qui désire faire de la phonétique ou de la phonologie les utilisera naturellement. Les autres logiciels sont parfois choisis selon des préférences de laboratoire ou des habitudes. Ce n'est pas toujours une bonne solution car certains logiciels ont des fonctions spécifiques qui sont utiles ou parfois même vitales pour certaines études. Ainsi, ELAN est incomparable lorsqu'il s'agit d'aligner plusieurs lignes de transcription (plusieurs interlocuteurs ou/et plusieurs événements) de manière temporelle précise. CLAN est particulièrement adapté au traitement de récit (comme Transcriber) mais aussi à la description des interruptions, reprises, erreurs de performance ou variantes de prononciation grâce à son riche système de codage. Enfin, EXAMARaLDA présente une configuration intermédiaire entre ces deux extrêmes.

### **3.4. L'exploitation des corpus**

Les logiciels informatisés présentent l'avantage de permettre l'exploitation automatisée des données qu'ils peuvent créer. Les possibilités offertes par les différents logiciels sont très variables et fonction de l'objectif scientifique pour lequel ils ont été conçus.

Par exemple CLAN, conçu pour traiter les corpus servant à étudier l'acquisition du langage permet de calculer la LME (Longueur Moyenne d'Enoncé), mesure proposée par Brown (1973) et systématiquement utilisée dans la littérature sur

A paraître dans « Transcrire et analyser les corpus d'enfant. In Edy Veneziano, Anne Salazar Orvig, Josie Bernicot Eds, Acquisition du langage et interaction. Paris : L'Harmattan »

l'acquisition comme permettant d'évaluer le niveau approximatif de développement de langage d'un enfant. PRAAT, d'abord conçu pour traiter et afficher du signal de parole permet d'automatiser et modifier les traitements sur ce signal. EXMARaLDA, conçu pour traiter des données textuelles de volume important, dispose de logiciels annexes pour traiter et éditer des corpus entiers.

Pour le traitement du langage de l'enfant, le logiciel aujourd'hui le plus adapté est CLAN. Non seulement il permet comme tous les autres logiciels de réaliser des recherches dans les corpus, mais il permet aussi de calculer des fréquences lexicales, des fréquences phonologiques, les différences entre formes produites et attendues, des rapports type-token, la LME, de faire de l'analyse morphologique et syntaxique, et aussi de calculer des indices de développement syntaxique comme le DSS (Developmental Sentence Score - Lee, 1974). Il dispose également de toute une série de commandes annexes permettant de vérifier et éditer les corpus. Enfin, le résultat de toute commande peut être enchaîné avec une autre commande et également stocké dans un fichier de format texte. Cette dernière possibilité permet d'utiliser des outils comme un traitement de texte ou un tableur pour éditer ces résultats ou les utiliser pour faire des calculs supplémentaires ou des statistiques. Le défaut de CLAN est que ces commandes annexes passent par une interface utilisateur en mode commande qui, bien que puissante, reste difficile à maîtriser pour beaucoup d'utilisateurs. Par exemple, la commande de recherche de chaîne de caractères présente d'immenses possibilités mais au prix d'un usage complexe. On ne peut que recommander à l'utilisateur de procéder par essais et erreurs en allant de commandes simples à des commandes complexes de manière progressive (cette recommandation qui s'applique à tous les programmes informatiques est particulièrement importante ici).

Les autres logiciels ont en général des interfaces utilisateurs plus modernes (qui ne riment pas toujours avec simplicité d'utilisation). Si aucun logiciel n'a la même variété de commandes que CLAN, les fonctions de recherche peuvent être plus puissantes. C'est le cas d'ELAN, mais encore plus de PHON qui utilise un véritable langage de programmation pour faire ses recherches.

A paraître dans « Transcrire et analyser les corpus d'enfant. In Edy Veneziano, Anne Salazar Orvig, Josie Bernicot Eds, Acquisition du langage et interaction. Paris : L'Harmattan »

EXMARaLDA présente l'avantage d'offrir des programmes annexes (Corpus Manager et EXAKT) qui permettent de manier des bases de données, d'éditer des métadonnées, de réaliser des recherches sur les annotations et de calculer des concordances.

Apprendre à connaître et à manier les commandes d'exploitation des corpus est certainement coûteux en termes d'efforts et de temps. Ce temps est toutefois largement compensé par les résultats que cela permet d'obtenir. Pour ceux qui veulent aller plus loin, il est possible d'utiliser les corpus, éventuellement après transformation, avec des outils spécifiques plus puissants destinés au traitement automatique du langage et des textes.

### **Conclusion**

Il n'est plus possible aujourd'hui d'envisager des travaux sur l'acquisition du langage qui n'aient pas un lien avec des corpus de productions attestées ou des situations expérimentales. Ceci est aujourd'hui beaucoup plus facile qu'autrefois grâce au nombre important de logiciels et d'outils disponibles. Pour ceux qui ne peuvent constituer leur propre corpus, des bases de données comme CHILDES (USA) ou le CRDO (France) permettent une première approche des corpus. On peut espérer que la généralisation de normes communes permettant d'améliorer l'interopérabilité des systèmes apportera encore plus de données et de résultats scientifiques dans un futur proche.

Malgré ses énormes avantages, il faut toutefois rester conscient des limites du travail sur corpus. Si les hypothèses scientifiques sont mal posées, si le corpus est mal conçu ou utilisé à mauvais escient, les résultats obtenus seront sans valeur. Travailler sur plus de 10 000 mots n'est pas forcément synonyme de qualité ou d'intérêt scientifique.

C'est pourquoi les auteurs de ce chapitre, en accord avec la plupart des chercheurs et spécialistes du domaine, préconisent, pour l'étudiant comme pour le chercheur confirmé, d'utiliser à la fois les deux stratégies.

-1) Recueillir du corpus quand on commence ses recherches et continuer à en recueillir tout au long de son travail scientifique. Seul le travail de recueil et surtout de transcription

A paraître dans « Transcrire et analyser les corpus d'enfant. In Edy Veneziano, Anne Salazar Orvig, Josie Bernicot Eds, Acquisition du langage et interaction. Paris : L'Harmattan »

permet de bien appréhender les phénomènes réels, de mesurer leur dimension effective, comme un moment unique et mystérieux du développement d'un être humain mais aussi comme un moment où des milliers d'informations se combinent.

-2) Donner ses propres corpus à la communauté, en respectant les règles d'anonymat et de paternité des travaux, et regarder les corpus des autres pour les comparer aux siens. Ceci présente un double intérêt : augmenter le matériel disponible pour arriver peut-être un jour à disposer de quantité d'information comparable à celle que l'enfant reçoit dans les premières années de sa vie ; inciter à la comparaison entre études et à la multiplication des points de vue sur de mêmes enfants, afin de faire avancer la compréhension de ce domaine de recherche pour lequel beaucoup de travail reste encore à accomplir.

### Références

- Brown, R., 1973, *A first language : The early stages*. Cambridge, Mass., Harvard University Press.
- Darwin, C., 1877, A Biographical Sketch of an Infant, *Mind*, vol. 2 : 285-294.
- Lee, L. (1974). *Developmental Sentence Analysis*. Evanston, IL: Northwestern University Press.
- Morgenstern, A. en collaboration avec Sandra Benazzo, Marie Leroy, Emmanuelle Mathiot, Christophe Parisse, Anne Salazar Orvig, Martine Sekali (sous presse). *L'enfant dans la langue : de l'observation du naturaliste à l'analyse du linguiste*. Paris : Presses de la Sorbonne Nouvelle.
- Morgenstern, A., Parisse, C. (2007). Codage et interprétation du langage spontané d'enfants de 1 à 3 ans. Corpus n°6 "Interprétation, contextes, codage", pp. 55-78.
- Projet Corinte: ICAR (CNRS, Lyon 2, ENS-LSH)  
<http://icar.univ-lyon2.fr/projets/corinte/confection/alignement.htm>
- Rohlfing & al. (2006). Comparison of multimodal annotation tools – workshop report. *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* (ISSN 1617-1837) Ausgabe 7 (2006), 99-123.
- Taine, H., 1876, L'acquisition du langage chez les enfants et dans l'espèce humaine. *Revue Philosophique de la France et de l'Etranger*, 1, 357-395.