Lidil

Revue de linguistique et de didactique des langues

42 | 2010

Multimodalité de la communication chez l'enfant

Transcrire et annoter la multimodalité : quand les productions des enfants sourds ré- interrogent les outils d'analyse

Agnès Millet et Isabelle Estève

Édition électronique

URL : <http://journals.openedition.org/lidil/3062>

DOI : 10.4000/lidil.3062

ISSN : 1960-6052

Éditeur

UGA Éditions/Université Grenoble Alpes

Édition imprimée

Date de publication : 30 novembre 2010

Pagination : 9-33

ISBN : 978-2-84310-184-7

ISSN : 1146-6480

Référence électronique

Agnès Millet et Isabelle Estève, « Transcrire et annoter la multimodalité : quand les productions des enfants sourds ré-interrogent les outils d'analyse », *Lidil* [En ligne], 42 | 2010, mis en ligne le 31 mai 2012, consulté le 01 mai 2019. URL : <http://journals.openedition.org/lidil/3062> ; DOI : 10.4000/lidil.3062

Transcrire et annoter la multimodalité : quand les productions des enfants sourds ré-interrogent les outils d'analyse

*Agnès Millet et Isabelle Estève**

RÉSUMÉ

La présente contribution vise à rendre compte, de façon dynamique, du cheminement qu'il nous a fallu parcourir pour parvenir à élaborer une grille d'annotations sous *Elan*[®], à même de répondre aux questions que posent les spécificités des productions langagières de jeunes enfants sourds âgés de 6 à 12 ans, lors d'une tâche narrative. Les contextes de surdité amènent les jeunes locuteurs à utiliser toutes les ressources langagières à leur disposition dans des productions qui s'inscrivent d'entrée de jeu dans la multimodalité – et la bilinguïté. Afin de rendre compte de ces pratiques spécifiques, il nous a fallu considérer l'ensemble des éléments verbaux et non verbaux, vocaux et gestuels, comme les éléments d'une production intégrée. La description de ces énoncés bimodaux nous a amenées à nous décentrer du linguistique pour prendre en compte le « vouloir-dire » de l'enfant et à reconsidérer tout à la fois la question des unités de transcription et celle de la complexité des rapports intersémiotiques.

ABSTRACT

This paper deals with the central question of transcribing deaf children's productions. We present the annotation template we have created on Elan[®], explaining in detail how and why the observation of the narrative productions of 6 to 12 year-old deaf children led to modify the previously available annotation schemes. The deaf children draw on every resource available in both modalities: voice and gesture. Thus, these productions are fundamentally multimodal and bilingual. In order to describe these specific practices, we propose to consider verbal and non-verbal, vocal and gestural materials as part and parcel of one integrated production. A linguistic-centered transcription is not efficient for describing such bimodal productions, since describing

* LIDILEM, Université de Grenoble.

bimodal utterances implies taking into account the “communicative desire” (“vouloir-dire”) of the children. For this reason, we need to reconsider both the question of the unit of transcription and the issue of the complexity of semiotic interactions in the bimodal utterances.

Nous nous proposons dans cet article de restituer, dans une perspective dynamique, la démarche que nous avons engagée afin d’élaborer une grille d’annotations sous *Elan*^{®1} qui puissent nous permettre de décrire et d’analyser les productions langagières de jeunes enfants sourds âgés de 6 à 12 ans, lors d’une tâche narrative².

Le projet « Multimodalité », soutenu par l’Agence nationale de la recherche (2005-2009) et dirigé par Jean-Marc Colletta, incluait, à des fins comparatives, un volet sur les enfants sourds. Il s’agissait de voir comment les enfants sourds, soumis à un protocole de narration identique à celui proposé à des enfants entendants³, utilisaient, au travers de la multimodalité, leurs ressources verbales et non verbales, pour restituer une histoire qu’ils venaient de visionner sous la forme d’un dessin animé de Tom et Jerry (le script du dessin animé permettant de mieux contextualiser les exemples donnés dans l’article est fourni en Annexe).

Introduction

Le volet « surdité » de la recherche s’est réalisé en fin de programme et la grille que nous souhaitons présenter et surtout commenter ici, dans les cheminements qui ont été les nôtres, s’est organisée à partir d’une grille, élaborée par l’équipe constituée, pour rendre compte de la mise en œuvre discursive de la multimodalité chez les enfants entendants (Colletta, Venouil, Kunene, Kaufman et Simon, 2009). Notre grille, en tant que telle, a déjà fait l’objet d’une publication (Millet et Estève, 2009). Cependant, il nous a paru important de compléter cette première présentation statique, par une dimension dynamique susceptible de cer-

1. *Elan*[®] est un logiciel développé par le Max Planck Institut qui permet de créer des annotations alignées complexes à partir de données video et/ou audio. Voir <<http://www.mpi.nl/tools/elan.html>>.

2. Le recueil de données a été effectué dans des Classes d’intégration scolaire (CLIS) lyonnaises et grenobloises : oraliste, « bilingue » (LSF / français écrit) et « mixte », c’est-à-dire utilisant le français et la LSF dans l’enseignement.

3. Voir dans ce numéro l’article de Djaber Fantazi : « La gestualité cohésive dans les récits d’enfants âgés de 9 à 11 ans ».

ner la façon dont l'utilisation de la multimodalité par les enfants sourds éclaire, d'un point de vue général, la mise en place et le développement des fonctions langagières, entendues dans leur acception large, c'est-à-dire concernant tant le plan linguistique que les aspects non verbaux.

Les partis pris théoriques

S'appuyant sur cette conception, notre approche s'articule en outre sur les notions de répertoire verbal et de plurilinguisme, en lien avec la sociolinguistique du contact des langues.

Une conception globale de l'activité langagière

Notre travail s'inscrit dans la perspective de McNeill (1992) qui considère que les modalités vocales et gestuelles participent toutes deux de l'expression langagière. Nous distinguons donc bien les pratiques linguistiques – expression au moyen de la langue – des pratiques langagières – construites sur le verbal et le non-verbal.

Une définition large du bilinguisme

Nous adoptons ici une définition sociolinguistique et non normative du bilinguisme et de la personne bilingue telle que, entre autres, Grosjean (1982) a pu la proposer, à savoir : une personne utilisant deux langues dans sa vie de tous les jours ; définition qui est liée à celle de MacNamara (1967) : « Le bilingue est quelqu'un qui possède une des quatre habiletés linguistiques à savoir comprendre, parler, lire et écrire, dans une langue autre que sa langue maternelle. »

Une conception plus sociale que politique du contact de langues

Comme le proposent les théories du bilinguisme développées par Ludi et Py ([1986] 2002), on considère que le lieu du contact des langues n'est pas le territoire, comme ce peut être le cas par exemple dans les théories diglossiques, mais la personne et que parce que la personne est le lieu du contact linguistique, elle développe des compétences et des activités langagières propres et originales, ce que l'on nomme « parler bilingue ».

La multimodalité sourde

La multimodalité – et, partant, l'instanciation langagière – des enfants sourds présente plusieurs particularités.

Tout d'abord, nous soulignerons que, selon les sources (Gillot, 1998), 90 à 95 % des enfants sourds naissent dans des familles entendant,

qui, donc, ne connaissent à priori pas l'existence des langues gestuelles – soit, pour le contexte français, la LSF. Le choix de ces familles se porte actuellement de façon extrêmement majoritaire sur une éducation «oraliste»⁴.

Ainsi, bon nombre d'enfants sourds⁵, bien qu'ils aient indubitablement une barrière sensorielle leur empêchant un accès naturel à la langue vocale⁶, ne sont pas introduits, de par les choix éducatifs qui sont faits pour eux, dans une gestualité d'essence linguistique. Les progrès de la technique, visant à réparer l'audition *via* des appareillages de plus en plus sophistiqués, allant des prothèses auditives (ou contours d'oreilles) aux implants cochléaires, accentuent encore ce phénomène. Néanmoins, pour certains d'entre eux, cette gestualité s'investit de composants quasi-linguistiques (voire linguistiques), comme nous le verrons.

Pour d'autres enfants, moins nombreux, le choix d'une éducation incluant la LSF est fait : dès lors la gestualité est porteuse d'informations symbolisées au moyen d'un système linguistique⁷. Si ces enfants sont introduits, par le biais d'une rééducation orthophonique, dans la dimension vocale du français⁸, ils auront donc deux modalités pour exprimer des contenus véhiculés linguistiquement.

4. Le choix existe entre une éducation «oraliste» – c'est-à-dire excluant tout recours à une langue gestuelle – ou une éducation dite «bilingue» – c'est-à-dire utilisant, à des degrés très divers, une langue gestuelle.

5. Il n'existe à notre connaissance aucune statistique sur les choix parentaux. Les terrains que nous connaissons nous laissent penser que le choix oraliste est très largement majoritaire.

6. On utilise ici le terme de «vocal» en opposition à «gestuel» faisant référence à la modalité utilisée, tandis que le terme «oral» est utilisé en opposition à «écrit» pour référer à la dimension communicative, et aux différences essentielles entre le mode oral et le mode scriptural, spécialement en termes d'interactions, de négociation du sens et de gestion des inférences.

7. Toutes les recherches internationales développées depuis une cinquantaine d'années – que nous ne pouvons toutes citer ici – montrent que les langues gestuelles sont des systèmes linguistiques qui présentent des différences avec les systèmes linguistiques des langues vocales, puisqu'ils sont fondamentalement spatiaux et iconiques, mais qui présentent aussi de nécessaires ressemblances, puisqu'ils sont des systèmes linguistiques humains.

8. Face à la querelle séculaire qui oppose les tenants d'une éducation «oraliste» et d'une éducation «bilingue» de l'enfant sourd (voir note 3), il existe une autre querelle, plus récente (depuis la fin des années 1970), entre les acceptations que l'on peut avoir d'une éducation «bilingue» : certains excluent une

Ainsi, la bimodalité sourde introduit nécessairement dans la bilinguïté. Il s'agit, en outre, d'une forme de bilinguïté spécifique puisque chaque modalité est tour à tour susceptible de servir de base linguistique aux énoncés. Ceux-ci ne peuvent dès lors être envisagés comme constitués de deux modalités œuvrant en quelque sorte parallèlement, mais bien comme la mise en œuvre dynamique et conjointe des ressources langagières multimodales.

Multimodalité : productions conjointes ou productions parallèles

Transcrire les productions narratives des enfants sourds amène à interroger plus précisément le lien vocalité/gestualité, alors que, concernant les entendants, il semble qu'un postulat sous-jacent subordonne le gestuel à la parole, sans doute du fait que les modalités sont spécifiées dans chacun des deux domaines verbal d'une part (voix) et non verbal d'autre part (geste). Dans un premier temps, nous présentons les grandes lignes de la grille d'annotations, telle qu'elle a été élaborée pour la transcription des narrations des enfants entendants, pour ensuite discuter tous les aménagements qui ont dû être faits.

Une grille de départ

La première opération de codage consiste, sur une ligne *Enfant*, à «transcrire les paroles» ; cette transcription permet une notation fidèle des propos de l'enfant, avec les hésitations, les prononciations effectives, les pauses, etc. Mais la centration sur le linguistique exige une remise aux normes lexicales et syntaxiques, proches de celles de l'écrit, s'effectuant sur la ligne *proposition syntaxique* afin de procéder à l'identification des types de propositions. Cette ré-écriture⁹ a par exemple amené à négliger tout ce qui vocalement ne relève pas du linguistique – telles par exemple les onomatopées qui s'inscrivent certes, par rapport au français dans une monomodalité, mais investissent un registre langagier différent.

éducation à la dimension vocale de la langue française, tandis que d'autres l'incluent.

9. Par exemple, le segment suivant : «et après euh // l'œuf // i(l):: // l'oisillon qui est à l'intérieur d(e) l'œuf» est ramené à la proposition : «et après l'oisillon qui est à l'intérieur de l'œuf».

Parallèlement à ces annotations concernant le linguistique, sont annotés les mouvements corporels sous la ligne *Gestes*. Les deux entrées sont donc distinctes par nécessité puisqu'il s'agit de linéariser des processus globalisés dans l'expression langagière. Mais elles sont surtout indépendantes : l'identification des unités vocales et celle des unités gestuelles n'étant pas subordonnées l'une à l'autre, dans l'appréhension globale des productions de l'enfant. Nous verrons qu'un tel choix d'indépendance des lignes d'annotation n'est pas adéquat pour rendre compte de la réalité langagière des productions des enfants sourds, ce qui interroge nécessairement la pertinence de maintenir deux entrées distinctes, et de fixer le linguistique au centre du découpage des propos de l'enfant. En effet, le recours à la bimodalité dans les pratiques des enfants sourds nous a amenées très rapidement à revoir ces éléments centraux de la transcription¹⁰.

Matériaux discursifs verbaux et non verbaux : quelle(s) réalité(s) chez les enfants sourds ?

À la différence des discours des enfants entendants, dans les productions des enfants sourds, le linguistique peut être supporté par deux langues distinctes : le français ou la LSF.

Les discours des enfants sourds laissent parfois apparaître, de manière similaire à ceux des enfants entendants, le français comme support de la production linguistique, tandis que la gestualité n'est que co-verbale ; c'est le cas, par exemple, dans l'extrait de la narration de Jalil, qui restitue la chute de l'œuf.

Exemple 1 : Jalil¹¹

Français : et après l'œuf il...il...elle bouge il tombe par terre

Gestes : MD mvt > B

10. Nos propos seront illustrés d'extraits de discours des enfants suivants : Abdel et Driss scolarisés dans une CLIS oraliste, Louise scolarisée dans une CLIS bilingue (LSF/français écrit) et Jalil, Medhi et Yanis scolarisés dans une CLIS « mixte ».

11. MD = main droite ; mvt > B = mouvement vers le bas ; stf = spécificateur de taille et de forme ; pr = proforme manuelle ; pr-corp. = proforme corporelle ; loc = locus – les termes « stf », « proforme » et « locus » seront explicités dans la section « Réaménagements » de la 3^e partie « Multimodalité et bilinguisme : économies intersémiotiques ».

Cependant, certains discours présentent l'organisation inverse, le linguistique étant alors supporté par la modalité gestuelle, tandis que la vocalité n'est que co-verbale, comme dans cet exemple produit par Louise :

Exemple 2 : Louise

Onomatopées : pup ! pup !

LSF : stf-coquille, stf-oiseau-dedans, BOUGER

LSF : stf-coquille, stf-œuf-dedans, POUSSER, A-L-INTERIEUR

Si dans ces deux types de discours, un découpage des productions de l'enfant centré sur le linguistique reste opératoire, une langue pouvant être identifiée comme support des productions de l'enfant, d'autres types de discours, beaucoup plus fréquents chez les enfants sourds, témoignent d'une organisation bimodale qui n'a pas d'équivalent chez les enfants entendants, et dont il faut bien rendre compte.

Par ailleurs, comme nous l'avons déjà souligné, le linguistique chez les enfants sourds peut être assumé par les deux modalités simultanément. Cette simultanéité interroge tout particulièrement la question du découpage de la production linguistique – et ce, de façon caractéristique dans cet extrait des productions de Mehdi, où le français et la LSF sont utilisés en continu.

Exemple 3 : Mehdi

Frs : Après ...

LSF : APRES pr.corp-maman-oiseau, SE-METTRE-DEBOUT

Frs : pour voir petit *nœuf

LSF : pr-corp-maman-oiseau, REGARDER, loc-entre-les-jambes PETIT ŒUF

Selon la langue choisie comme base pour transcrire cet extrait, le découpage amène à identifier une ou deux propositions. En effet, il semble qu'on puisse identifier une proposition infinitive en français [après pour voir petit *nœuf], tandis que deux propositions distinctes peuvent être repérées en LSF [APRES pr.corp-maman-oiseau, SE-METTRE-DEBOUT] et [pr-corp-maman-oiseau, REGARDER, loc-entre-les-jambes PETIT ŒUF].

Ainsi, les interactions observées entre vocalité et gestualité nous donnent à repenser l'appréhension des deux modalités dans les productions des enfants sourds. Dans cet extrait d'Abdel, par exemple, où le français, les onomatopées et les gestes sont utilisés, il serait bien délicat

de proposer un découpage syntaxique cohérent qui corresponde à l'organisation réelle de l'énoncé.

Exemple 4 : Abdel

Frs : et le bébé oiseau... c'est moi le bébé oiseau et puis il... il... des grosses avec

Ono. :

tchuchuchuchuchuchu chu::

Gestes :

Ainsi, confrontées aux réalités multiples que revêt la bimodalité chez les enfants sourds, dont ces quatre exemples ne donnent qu'un aperçu, nous avons dû adapter la grille afin d'être en mesure d'annoter au plus près du sens de l'énoncé les interactions diverses entre moyens linguistiques et non linguistiques. Dès lors, nous avons dû redéfinir une échelle de description qui soit à même de saisir ces dynamiques langagières dans leur ensemble.

Rendre compte d'une expression langagière intégrée

Appréhender la bimodalité dans le contexte de la surdité nécessite de repenser l'appréhension distincte des valeurs linguistique et non linguistique de l'expression du locuteur, dans la mesure où cette distinction ne s'ancre plus dans un découpage bimodal biunivoque : une modalité pour le verbal et une autre pour le non-verbal. La présence de deux langues de modalités différentes – le français et la LSF – amène la possibilité d'une expression gestuelle et vocale, chacune de ces modalités pouvant véhiculer des valeurs linguistiques ou non linguistiques¹². Cette spécificité articulatoire du bilinguisme sourd amène à nous situer dans une appréhension particulière de la bimodalité.

12. Pour ce qui concerne la gestualité, la question de la frontière entre linguistique et non linguistique est loin d'être tranchée d'une manière générale, par exemple dans les descriptions des langues gestuelles. Cette question se pose de façon encore plus cruciale pour les productions de jeunes enfants, dont la gestualité est en cours de construction linguistique. Nous avons, pour notre part, considéré comme linguistiques les productions gestuelles qui correspondent à des segments que l'on peut retrouver dans les discours d'adultes sourds, indépendamment du fait que l'enfant ait conscience (ou non) de produire un énoncé linguistique.

Impossible parallélisme des productions multimodales

Dans la grille de départ élaborée pour la description des conduites langagières des enfants entendants, la modalité vocale, et donc linguistique, était identifiée comme référence pour la transcription des propos de l'enfant. La bimodalité était donc fondamentalement appréhendée comme la production de deux productions indépendantes, ce qui s'explique aisément puisque, tant d'un point de vue articulatoire que d'un point de vue perceptif, elles peuvent être envisagées comme distinctes. Cette conception appliquée aux pratiques des locuteurs sourds reviendrait à envisager l'utilisation verbale simultanée des modalités vocale et gestuelle comme deux énoncés linguistiques distincts et indépendants relevant chacun d'une langue particulière.

La bimodalité se détaillerait alors comme une production *parallèle* de deux messages syntaxiquement et sémantiquement indépendants et conformes à chacune des langues utilisées, ce qui paraît être le cas dans l'exemple suivant, où ce qui est énoncé en français est strictement équivalent à ce qui est énoncé en LSF, les normes de chacune des deux langues étant respectées.

Exemple 5 : Adulte

Français : oui oui j'étais là

LSF : OUI PTE-1¹³ LA

En réalité, cette mise en œuvre de la bimodalité n'est que rarement rencontrée dans les pratiques des locuteurs sourds bilingues, et ce, même dans des situations où l'utilisation des deux langues est pourtant contrainte par la situation de communication – en présence d'un locuteur sourd et entendant par exemple (Estève, 2006, 2007). Ceci s'explique vraisemblablement par le fait que, dans une perspective expressive, il est peu réaliste de penser que l'expression d'un locuteur bilingue mettrait en œuvre les deux façons de dire qu'il possède en se bornant à véhiculer une information identique dans les deux langues, mettant ainsi en parallèle deux messages identiques. Une telle conception d'une mise en œuvre *parallèle* de la bimodalité néglige tout à la fois les potentialités langagières combinatoires et l'utilisation relative des modalités.

13. PTE-1 = pointé de première personne du singulier.

La multimodalité au service d'un énoncé unique

Ainsi, comme nous en avons convaincu le corpus recueilli auprès d'enfants sourds, appréhender l'expression du locuteur sourd à l'échelle de sa production globale comme un tout composé d'éléments vocaux et gestuels nous paraît donc plus à même de cerner son essence. On parlera alors, par opposition à l'idée de parallélisme que l'on vient d'évoquer, de productions *conjointes*, dans lesquelles l'expression ne peut être saisie qu'à l'échelle de la bimodalité puisque productions vocales et gestuelles co-construisent les propos du locuteur, comme c'est le cas dans l'exemple suivant.

Exemple 6 : Adulte

Frs : Mais euh... ma famille un petit peu fait un petit peu c'est mieux

LSF : FAMILLE UN-PEU SIGNER

Dans cet exemple, emprunté à un locuteur adulte, en effet, le sens de l'énoncé ne se construit que dans la perception et la compréhension des deux modalités – utilisées ici dans leur dimension linguistique. L'élément sémantiquement central [SIGNER] n'est produit qu'en LSF, tandis que « c'est mieux » n'est exprimé qu'en français. Ceci nous amène à reconnaître que l'énoncé est composé de deux messages : un message vocal et un message gestuel, qui sont interreliés au service de l'expression du sujet parlant. Au-delà de cet exemple, dans les faits, chacune des modalités peut être utilisée par le locuteur sourd dans ses dimensions verbales et non verbales : français et onomatopées pour la modalité vocale, LSF et gestes pour la modalité gestuelle. La structure théorique de l'énoncé peut ainsi être représentée comme suit :

Figure 1. – Structure de l'énoncé.

Partant de cette conception discursive, il nous a fallu proposer une catégorisation des pratiques qui rende compte de toutes les conjugaisons possibles des langues et des modalités au sein d'un énoncé.

Organisation des productions à l'échelle de la bimodalité

Au sein de l'énoncé, les possibilités de combinaisons sont diverses : elles peuvent être bimodales ou monomodales d'une part, et verbales, non verbales ou mixtes d'autre part. Notre première catégorisation visait donc à distinguer les structures des énoncés sur la base de deux critères : bimodalité et bilingualité (bimodal bilingue, bimodal monolingue, monomodal monolingue, etc.) [Millet, 2007 ; Millet et Estève, 2008]. Il nous est vite apparu que les interactions observées entre ces différents niveaux outrepassaient, par leur diversité, cette première catégorisation qui, si elle permet d'étiqueter la nature des combinaisons, n'autorise pas une description suffisamment fine de la manière dont intervient chaque élément dans la structure des énoncés. En effet, la structure de l'énoncé peut être déterminée par une langue, les deux langues ou des éléments non verbaux. Nous proposons donc de rendre compte de la structure de l'énoncé autour de la notion de « base » – l'entrée par la langue n'étant pas toujours effective – en proposant une catégorisation des types de pratiques à la croisée de ces trois critères :

- le choix de la base : la langue ou la modalité organisatrice du discours ;
- le choix du mode : monomodal vs bimodal ;
- la nature de la combinaison : monolingue, bilingue, non verbal.

À partir de l'observation de pratiques langagières d'adultes sourds (Millet et Estève, 2008), nous avons pu identifier trois types de bases permettant la catégorisation des pratiques suivantes.

- *Les pratiques à base français* regroupent les énoncés structurés par la langue française dans sa réalisation sonore. On trouve dans ces pratiques une utilisation variable de la modalité gestuelle : gestes, insertions ponctuelles de LSF, insertion continue de LSF.
- *Les pratiques à base LSF* regroupent les énoncés dont la structure relève de la LSF. On trouve dans ces pratiques une utilisation variable de la modalité vocale : onomatopées, insertions ponctuelles de français, insertion continue de français, la modalité gestuelle pouvant être investie également par des labialisations¹⁴.

14. Les labialisations consistent à articuler des mots du français sans émission sonore.

- *Les pratiques à base bilingue* correspondent aux pratiques dont la structure n’est pas attribuable à une langue en particulier, mais à l’interaction des deux langues.

L’observation des pratiques enfantines nous a amenées à enrichir cette première catégorisation (Millet et Estève, 2009). En effet, confrontées à des productions de locuteurs chez qui les systèmes linguistiques sont en cours de construction, il nous a fallu reconnaître un quatrième type de pratiques liées à une nouvelle base.

- *Les pratiques à base non verbale* correspondent aux énoncés dont la structure est déterminée par des éléments non linguistiques. On notera que l’on inclut dans ce type de pratiques des énoncés non verbaux, bien qu’ils soient introduits par des connecteurs du type « après » « ensuite » ou des syntagmes introducteurs du type « comme ça », énoncés en français et/ou en LSF. En effet, à notre sens, ces éléments « linguistiques » jouent un rôle d’impulseur énonciatif, qui atténue leur portée linguistique.

Le schéma suivant récapitule l’ensemble des types de pratiques.

		Pratiques à base français	Pratiques à base bilingue	Pratiques à base LSF	Pratiques à base non verbale
Monolingue	monomodal	français seul		LSF seule	
	bimodal	français + gestes		LSF+onomatopées	
Bilingue	monomodal			LSF+labialisations	
	bimodal	français + insertions <small>punctuelles LSF continues LSF</small>	français + LSF superposées	LSF + insertions <small>punctuelles frs labialisations /vocalisations</small>	
Non-verbal	monomodal				Onomatopées Gestes
	bimodal				Onomatopées+gestes

Figure 2. – Outils de catégorisation des types de pratiques (Estève, 2009).

De même qu’il nous a fallu sortir d’une approche par trop «linguis-tico-centrée » pour rendre compte des pratiques, de même la question de l’unité de découpage s’est posée, un peu dans les mêmes termes, puisque l’unité déterminant le découpage dans la grille prévue pour les entendants était la proposition.

Définir une unité de base dans le cadre de la multimodalité sourde

En effet, compte tenu de notre perspective théorique, il nous a fallu définir une unité de découpage qui ne restreigne pas au syntaxique l'appréhension des productions langagières des enfants. Il nous fallait donc proposer un découpage en unités discursives qui prenne en compte dans leur relation de dépendance les éléments vocaux et gestuels qui participent à l'élaboration de l'énoncé – que ces éléments soient verbaux ou non verbaux. Cependant, repérer des unités discursives cohérentes, de fait moins immédiatement objectivables qu'une proposition, exige de définir des critères explicites permettant une délimitation au plus près des productions réelles.

Si nous prenons un extrait des productions d'Abdel où celui-ci restitue le déplacement de l'œuf jusqu'à la maison de la souris – restitution qui se construit autour d'éléments linguistiques en français et d'éléments non verbaux gestuels – nous pouvons observer que la centration sur le linguistique (fig. 3) produit un découpage artificiel qui ne dégage sur la ligne « Enfant » qu'une seule unité autour du verbe exprimé en français dans une structure : thématisation – indépendante :

Figure 3. – Découpage « linguistico-centré ».

En revanche, la prise en compte de l'ensemble des éléments langagiers – ici, spécialement, le non-verbal gestuel – conduit à identifier sur la ligne « Enfant » trois unités de sens décrivant distinctement : l'une, le trajet de l'oiseau dans sa coquille ; l'autre, re-pointant ce trajet et la dernière achevant la description du trajet vers son but : la maison de Jerry, comme le montre la transcription réalisée dans la figure 4.

Figure 4. – Découpage selon le *vouloir-dire* de l'enfant.

Ainsi, la prise en compte des pratiques effectives des enfants sourds et le constat de l'importance de l'expression non linguistique, particulièrement gestuelle, dans leurs expressions, nous ont amenées à proposer de centrer la délimitation des unités autour du *vouloir-dire* de l'enfant sourd, qu'il s'ancre ou non dans une structure linguistique. Nous proposons donc d'adopter conjointement une perspective linguistique et sémiotique pour délimiter les unités au plus près de l'intention communicative du locuteur envisagé dans son expression multimodale.

C'est donc sur la ligne *Enfant* que s'effectue un premier découpage temporel des productions langagières qui n'a d'autres fonctions que la démarcation de chaque unité. À partir de ce premier repérage, on renseigne les quatre lignes majeures : *Français*, *Onomatopées*, *LSF*, *Gestes*, dépendantes de ce premier découpage sémantico-syntaxique. La structure de l'énoncé est ainsi repérée dans la dépendance des productions verbales et non verbales, vocales et gestuelles.

Dans cette perspective de transcription, l'annotation des dynamiques sémiotiques participant à l'élaboration de l'énoncé a dû être également ré-envisagée afin de mettre en relief les spécificités des rapports intersémiotiques dans ces discours multimodaux et bilingues.

Multimodalité et bilinguisme : économies intersémiotiques

Multimodalité : complexité des rapports sémiotiques

Pour l'étude des discours narratifs des enfants entendants, l'annotation sémantique du lien vocalité/gestualité était rattachée à la ligne *Gestes* et appréciée par rapport au contenu véhiculé par le linguistique. Si, dans un premier temps, à des fins de comparaison, nous avons conservé l'annotation proposée pour les entendants, la complexité des interactions sémiotiques attestées dans les énoncés des enfants sourds – qui s'expriment bien au-delà d'une relation duelle, bimodale, biu-

nivoque – nous a amenées à situer l’annotation au niveau de l’unité sémantico-syntaxique, et donc à revoir la catégorisation des relations intersémiotiques.

Multimodalité et vouloir-dire

Dans les productions des enfants sourds, chaque ressource joue un rôle tout à la fois spécifique et complémentaire dans l’organisation de l’énoncé. Par exemple, dans l’extrait suivant, la mobilisation de la vocalité vient étayer l’élaboration linguistique réalisée en LSF, où français et onomatopées participent à l’étayage du schéma spatial construit en LSF pour représenter l’épisode où l’œuf enfonce la porte de la maison de Jerry. Le mot « porte » intervient en complémentarité de la description du référent réalisée en LSF « stf-surface-plate », les onomatopées spécifient par ailleurs l’aspect brusque de l’action.

Exemple 7 : Yanis

Français : après porte

Onomatopées : chpouf !

LSF : après stf-surface-plate, pr-œuf, RENTRER-DANS

Chaque ressource adopte un rôle spécifique dans la construction de cette symbolisation supportée par la LSF, qui ne serait complète sans la prise en compte de l’information véhiculée par le français, puisque la référence du spécificateur de taille et de forme (STF) n’est pas explicite en LSF.

De la même manière, toutes les ressources disponibles peuvent être mobilisées pour construire une symbolisation complexe, comme dans l’énoncé suivant, produit par Driss, qui restitue ici l’épisode où l’oiseau fait un trou dans le mur. Dans l’unité 1, le français, les gestes ainsi que quelques signes de la LSF sont mobilisés ensemble pour inscrire la référence du mur dans l’espace, et pouvoir, dans l’unité suivante, décrire l’action dans une symbolisation non verbale.

Exemple 8 : Driss

Unité 1 :

Frs : Maintenant sur.. tu te rappelles le carré comme ça marron

LSF : SAVOIR MARRON

Gestes :

Unité 2 :	
Frs :	<i>après</i>
Ono. :	tl llllllllllt
Gestes :

Ces deux exemples montrent bien la nécessité de placer l'annotation des relations intersémiotiques à l'échelle de la production globale, émanation du *vouloir-dire* de l'enfant sourd.

Appréhender la complexité des liens entre vocal et gestuel

Puisque, comme nous l'avons vu, les ressources vocales et gestuelles fondent une seule unité sémantique structurée par deux architectures sémiotiques distinctes¹⁵, il a été nécessaire d'annoter – à l'échelle de l'unité – la nature des relations qu'entretiennent toutes les ressources mobilisées pour l'élaboration de l'énoncé dans une construction du sens globalisée/intégrée. Par ailleurs, repérer les relations intersémiotiques à l'échelle de l'unité exige de revoir la catégorisation proposée pour les entendants.

L'ensemble des catégories a été refondu : certains éléments – « redondant », « complémentaire » et « contradictoire » – ont donc été redéfinis ; d'autres – « substitution » et « élaboration » – abandonnés ; en nous inspirant des propositions faites par Volterra, Caselli, Capirci et Pirchio (2005), nous avons proposé de caractériser les rapports intersémiotiques de cinq façons différentes¹⁶.

- *Redondant* : la redondance s'applique aux combinaisons où les deux modalités véhiculent une information identique. Pour les combinaisons français/LSF, on notera que la redondance s'accompagne d'une structure unique conforme à l'une des deux langues (par exemple dans le cas du français signé).

15. « [...] when gesture and speech combine, they bring into one meaning system two distinct semiotic architectures. » (McNeill et Duncan, 2000, p. 144.)

16. Nous reprenons ici la présentation que nous en avons faite dans le *Journal of Language Contact* (Millet et Estève, 2009).

- *Équivalent* : cette relation d'équivalence ne s'applique qu'aux combinaisons français/LSF, dans lesquelles les deux messages sont sémantiquement identiques, mais restent syntaxiquement conformes respectivement à chacune des deux langues. Une relation qui, si elle est attestée chez les adultes, ne l'est que rarement chez les enfants observés jusqu'ici (voir exemple 5).
- *Complémentaire* : la complémentarité s'applique aux combinaisons où les deux modalités sont porteuses d'une information différente. Chaque modalité apporte une information spécifique qui n'est pas présente dans l'autre modalité. Le sens n'est alors accessible qu'au niveau de l'unité sémantico-syntaxique prise dans sa globalité (voir exemples 2, 4, 6, 7, 8).
- *Renforçant* : ce concept s'applique aux combinaisons où un seul des deux messages est porteur de l'ensemble de l'information. L'autre message n'étant porteur que d'une information partielle au regard de l'information transmise à l'échelle de l'unité (voir exemples 1, 3).
- *Contradictoire* : la contradiction s'applique aux combinaisons où ponctuellement une unité entre en contradiction avec l'information véhiculée par la modalité organisatrice. Par exemple, lorsque le français organise l'énoncé et que les signes de la LSF produits ponctuellement ne sont pas conformes au contenu sémantique de ce qui est prononcé. Il s'agit donc le plus souvent d'erreurs ou de lapsus, liés à la gestion simultanée de deux systèmes linguistiques.

Au demeurant, quel que soit le rapport sémiotique entre le gestuel et le vocal, il convient également d'envisager la dynamique intra-modalité et tout particulièrement, au sein de la modalité gestuelle, de pouvoir envisager, au sein du développement du bilinguisme sourd, les tensions entre les gestes et les signes, ce par quoi nous terminerons cette présentation.

Du geste au signe : catégorisation et pertinence langagière

Dans la grille d'annotation prévue pour les narrations des enfants entendants, chaque geste a été affecté d'une valeur. L'entrée de cette ligne d'annotation était une entrée fonctionnelle qui distinguait sept valeurs pour les gestes, que ces gestes soient manuels, céphaliques ou plus globalement corporels. L'organisation de la catégorisation des gestes en sept classes représente une réduction et une réorganisation des niveaux de hiérarchie par rapport à d'autres grilles d'interprétation des gestes

– par exemple celle proposée par Colletta (2004, p. 159-177) – ; mais représente une augmentation des catégories par rapport à la classification établie, par exemple par J. Cosnier (1993, p. 111). Étant hautement interprétatives, ces catégorisations fonctionnelles sont sujettes à de nombreuses fluctuations, ainsi qu'à quelques difficultés d'application/interprétation, même si, comme le montre Colletta (2004, p. 178-186), certaines catégories sont bien reconnues¹⁷. Quoi qu'il en soit de ces difficultés, nous exposerons brièvement les sept catégories proposées au départ pour l'analyse afin d'expliquer ensuite les modifications que nous avons proposées pour rapprocher gestualité verbale et non verbale.

Les sept catégories de départ

Les sept valeurs qui avaient été affectées aux gestes dans la grille élaborée par Colletta *et al.* (2009) sont les suivantes :

- *valeur représentationnelle* : cette valeur est liée au processus narratif et à l'investissement corporel du locuteur dans la narration – représentation des objets de la narration, mimiques et mimes adoptant le point de vue du personnage ;
- *valeur performative* : les gestes à valeur performative peuvent être autonomes ou co-verbaux et sont liés à l'expression d'un acte de langage ou au renforcement ou la modification de la valeur illocutoire de l'énoncé ;
- *valeur de cadrage* : il s'agit des gestes « qui expriment un état émotionnel ou un état mental du locuteur » (Colletta *et al.*, 2008, p. 23) ;
- *valeur discursive* : seront affectés de cette valeur tous les gestes qui participent à la structuration de la parole – segmentation, cohésion ;
- *valeur déictique* : cette valeur concerne exclusivement les pointages exophoriques ;
- *valeur interactive* : cette valeur est liée à l'interlocution et concerne tous les gestes relevant de la fonction phatique, décrite par Malinowski, et redéfinie par Jakobson (1963)¹⁸ ;

17. Les tests menés par Colletta, et analysés du point de vue statistique, lui permettent de conclure globalement à la consistance et à la fiabilité de sa classification.

18. « Il y a des messages qui visent essentiellement à établir, prolonger ou interrompre la communication, à vérifier si le circuit fonctionne [...] à attirer

- *valeur énonciative* : cette valeur s'exprime dans tout « geste manuel ou expression faciale qui manifeste que le locuteur cherche un mot ou une expression » (Colletta *et al.*, 2008, p. 23).

Si nous avons gardé les trois dernières catégories qui ne posaient pas de problèmes par rapport à notre objectif, il nous a fallu, compte tenu du fait qu'ils étaient sur-représentés massivement, affiner la catégorie des gestes représentationnels.

Réaménagements

La gestualité des jeunes enfants sourds, même lorsqu'ils ne sont pas exposés formellement à la LSF et reçoivent une éducation oraliste, présente des proximités remarquables avec la LSF, que l'on ne retrouve chez les enfants entendants ni dans les mêmes degrés d'utilisation ni dans la même qualité des gestes (Millet et Batista, 2009), même si l'on sait que la LSF utilise les mêmes procédés de base que la gestualité entendant en les systématisant (Colletta et Millet, 1998). En outre, la visée développementale de la recherche nécessitait de pouvoir repérer très précisément l'évolution de la gestualité chez ces enfants sourds, que cette gestualité soit linguistique ou non. Il fallait donc que l'on puisse repérer les procédés non verbaux se rapprochant des procédés verbaux, c'est pourquoi nous avons subdivisé la catégorie des représentationnels en distinguant tout d'abord les trois catégories suivantes :

- *les pointages endophoriques* : ils comprennent les gestes manuels ou céphaliques de pointage ayant une valeur de référence discursive anaphorique ou cataphorique ; ces pointages, qui sont des procédés linguistiques assurant la référence et la cohésion discursive en LSF, sont nécessairement liés à la création d'un locus – défini comme portion d'espace sémantisée de manière à pouvoir servir la référence de l'un ou l'autre des actants du procès mis en discours ;
- *les illustratifs* : il s'agit des gestes qui représentent les caractéristiques du référent (taille, forme, etc.) soit par configuration manuelle, soit par le dessin des caractéristiques du référent dans l'espace. Là encore il s'agit d'un procédé très productif en LSF, qui utilise la gestualité pour spécifier la taille et la forme des

l'attention de l'interlocuteur ou à s'assurer qu'elle ne se relâche pas [...]» (Jakobson, 1963, p. 217.)

référents par le biais de ce qu'on a justement coutume d'appeler «spécificateur de taille et de forme» (STF) – en anglais Size and Shape Specifier (SAAS) [Schembri, 2003]. Ces STF peuvent être intégrés en LSF dans des structures verbales, changeant alors de statut puisqu'ils deviennent des éléments référant à des nominaux précités, que l'on a nommés dans nos descriptions de la LSF «proformes», dans la mesure où la valeur n'est plus descriptive de type adjectival, mais pronominal (Bras, Millet et Risler, 2004 ; Millet, 2006). Cette reprise, de type pronominal, se retrouve dans la gestualité non linguistique et nous l'avons donc mise en relief en créant la catégorie spécifique suivante ;

- *les spatiographiques* renvoyant aux gestes qui ont pour fonction de donner une représentation topologique de l'univers référentiel.

Par ailleurs, étant donné l'importance de la spatialité et du corps dans l'expression gestuelle, il nous a paru important de créer deux catégories de gestes référentiels supplémentaires :

- *les mimétiques trajectoires* qui désignent donc uniquement les gestes manuels caractérisant le trajet du référent ;
- *les mimétiques actions*, qui comprennent les gestes mimant l'action ou le comportement d'un référent par une mise en jeu globale du corps, un procédé là encore très productif en LSF – et quasiment obligatoire dans les narrations – que l'on a longtemps nommé « prise de rôle » (Moody, 1983) et que nous nommons « proforme corporelle ».

Pour le reste des valeurs des gestes, il nous a semblé que, face aux interactifs et aux énonciatifs que nous avons conservés, deux catégories suffisaient, nous avons donc regroupé les valeurs «performative» et «de cadrage» sous la catégorie des gestes «expressifs», quant aux gestes à valeur discursive, il nous est apparu plus clair de les nommer gestes «syntaxiques».

Ces catégorisations nous ont permis de faire un lien entre les procédés non linguistiques et les procédés linguistiques de la gestualité et d'observer finement, spécialement dans les cas où nous avons pu mener une étude longitudinale chez certains sujets, les évolutions des répertoires langagiers, en lien avec le développement des procédés verbaux et non verbaux dans les deux modalités.

Éléments de conclusion

Il apparaît, en fin de parcours, que les spécificités de la bimodalité sourde, dans la diversité des formes qu'elle peut prendre, permettent de ré-interroger les fonctions de la bimodalité entendante, et spécialement lorsque l'on analyse les productions des enfants sourds parce qu'ils sont des êtres en devenir bilingue et que la modalité gestuelle leur est le plus naturellement accessible. Ainsi, dans ce cadre, on a vu qu'une conception d'une bimodalité où chacune des modalités serait envisagée en parallèle avec des fonctions bien réparties autour de chacune des deux modalités, n'était que peu pertinente, car, in fine, beaucoup trop centrée sur le linguistique, qui n'est pas nécessairement, et loin s'en faut parfois, la base organisatrice des discours narratifs recueillis auprès des enfants sourds. Le fait d'envisager la modalité gestuelle comme possiblement organisatrice de l'énoncé amène alors à revenir sur la notion d'unité propositionnelle comme unité de base de la description. La notion d'unité syntaxico-sémantique, prenant en compte toutes les réalités des pratiques langagières paraît beaucoup plus adéquate, car plus proche de l'analyse du *vouloir-dire* de l'enfant que de son dire linguistique effectif. Elle permet de rendre compte d'une unité langagière intégrée.

Ces remises en question proviennent bien sûr du fait que, dans le cas de la bimodalité sourde, la modalité gestuelle peut s'investir d'une valeur linguistique. Elles proviennent aussi du fait que, l'accès aux réalités de la modalité vocale étant complexe et souvent différé, les jeunes enfants sourds investissent très largement la modalité gestuelle dans ses dimensions non linguistiques également. De ce fait, il faut reconnaître la non-subordination du non-verbal au verbal, considérer l'ensemble des ressources mises en œuvre dans un énoncé, admettre la possibilité que le non-verbal puisse structurer le discours et donc redéfinir les unités de transcription. Tous ces éléments amènent alors à réenvisager la valeur des rapports entre modalité vocale et modalité gestuelle et à redéfinir plus précisément les valeurs des gestes au sein de la modalité gestuelle en les rapprochant des valeurs et des procédés linguistiques qu'ils peuvent acquérir dans les langues gestuelles.

L'application de cette nouvelle grille à des productions de jeunes enfants entendants amène alors à réenvisager les découpages et les interprétations des productions langagières en les décentrant de l'unité propositionnelle et, partant, des éléments linguistiques eux-mêmes. Ceci permet, en dernière analyse nous semble-t-il, d'envisager les productions comme de réelles productions langagières intégrées – et non

comme l'utilisation parallèle de modalités –, ce qui répond bien à la conception large du langage, telle que proposée par McNeill.

RÉFÉRENCES BIBLIOGRAPHIQUES

- BRAS G., MILLET A. et RISLER A. (2004) : « Anaphores et deixis en LSF – Tentative d'inventaire des procédés », *Silexicales*, n° 4, Université de Lille, p. 57-64.
- COLLETTA J.-M. (2004) : *Le développement de la parole chez l'enfant âgé de 6 à 11 ans*, Hayen, Mardaga.
- COLLETTA J.-M. (dir.) (2008) : *Transcription et annotation des données multimodales sous ELAN – Manuel de codage*, exemplaire photocopié, laboratoire LIDILEM.
- COLLETTA J.-M. et MILLET A. (1998) : « Gestualité référentielle des entendants et LSF : des matrices communes ? », *Congrès international des linguistes* (20-25 juillet 1997), CD-ROM, Pergamon/Elsevier éd.
- COLLETTA J.-M., KUNENE R. N., VENOUIL A., KAUFMANN V. et SIMON J.-P. (2009) : « Multitrack annotation of child language and gestures », dans M. Kipp, J.-C. Martin, P. Paggio et D. Heylen (éds), *Multimodal Corpora, From Models of Natural Interactions to Systems and Applications*, LNAI 5509, Berlin, Springer, p. 54-72.
- COSNIER J. (1993) : « Étude de la mimogestualité », dans R. Pléty, *Éthologie des communications humaines. Aide-mémoire méthodologique*, Lyon, Presses universitaires de Lyon, p. 103-115.
- ESTÈVE I. (2006) : *Les pratiques communicatives des jeunes adultes sourds – Introduction à l'analyse*, mémoire de Master 1 (A. Millet dir.), Grenoble, Université Stendhal.
- ESTÈVE I. (2007) : *Analyse des conduites langagières de jeunes adultes sourds*, mémoire de Master 2 (A. Millet dir.), Grenoble, Université Stendhal.
- ESTÈVE I. (2009) : « Diversité langagière d'une classe d'enfants sourds » [en ligne], dans *Actes du Colloque Enfance et plurilinguisme* (26-27 juin 2008, Montpellier). Disponible en ligne sur <http://www.msh-m.fr/article.php3?id_article=709>.
- GILLOT D. (1998) : *Le droit des sourds : 115 propositions*, Rapport parlementaire.
- GROSJEAN F. (1982) : *Life with two Languages: an Introduction to Bilingualism*, Harvard University Press.
- JAKOBSON R. (1963) : *Essais de linguistique générale*, Paris, Éditions de Minuit.

- LÜDI G. et PY B. ([1986] 2002) : *Être bilingue*, Berne, Peter Lang.
- MACNAMARA J. (1967) : « The bilingual's linguistic performance: a psychological overview », *Journal of Social Issues*, vol. XXIII, n° 2, p. 67-71.
- MCNEILL D. (1992) : *Hand and Mind, what gestures reveal about thought*, University of Chicago Press.
- MCNEILL D. et DUNCAN S. (2000) : « Growth points in thinking-for-speaking », dans D. McNeill (éd.), *Language and Gesture*, Cambridge, Cambridge University Press, p. 141-161.
- MILLET A. (2006) : « Le jeu syntaxique des proformes et des espaces dans la cohésion narrative en LSF » [en ligne], *Glottopol*, n° 7, revue en ligne sur <http://www.univ-rouen.fr/dyalang/glottopol/numero_7.html#Sommaire>, article disponible sur <http://www.univ-rouen.fr/dyalang/glottopol/telecharger/numero_7/gpl7_05millet.pdf> [consulté le 4 octobre 2010].
- MILLET A. (2007) : « Bilingual cross-modal communicative practices of young deaf adults », Paper presented at the *6th International Symposium on Bilingualism* (Hambourg, 30 mai-2 juin 2007).
- MILLET A. et BATISTA A. (2009) : « Étude qualitative de la multimodalité chez des enfants sourds et des enfants entendants de 6 ans », *Colloque Multimod2009* (9-10-11 juillet 2009, Toulouse).
- MILLET A. et ESTÈVE I. (2008) : « Pratiques langagières bilingues et multimodales de jeunes adultes sourds » [en ligne], *Actes du Colloque « Journées d'étude de la Parole »*. Disponible sur <http://www.afcp-parole.org/doc/Archives_JEP/2008_XVII_JEP_Avignon/PDF/avignon2008_pdf/JEP/029_jep_1564.pdf> [consulté le 4 octobre 2010].
- MILLET A. et ESTÈVE I. (2009) : « Contacts de langues et multimodalité chez des locuteurs sourds : concepts et outils méthodologiques pour l'analyse » [en ligne], *Journal of Language Contact* (revue disponible en ligne sur <<http://www.jlcjournal.org/>>).
- MOODY B. (1983) : *Introduction à l'histoire et à la grammaire de la langue des signes*, Vincennes, IVT, diffusion Ellipses.
- SCHEMBRI A. (2003) : « Rethinking "classifiers" in signed language », dans K. Emmorey (éd.), *Perspectives on classifier constructions in sign Languages*, Lawrence Erlbaum Associates Publishers.
- VOLTERRA V., CASELLI M. C., CAPIRCI O. et PIRCHIO S. (2005) : « Le rôle des gestes dans l'acquisition du langage chez les enfants entendants, les enfants non entendants et les enfants au développement atypique », dans C. Transler, J. Leybaert et J. E. Gombert (éds), *L'acquisition du langage par l'enfant sourd*, Marseille, Solal, p. 89-120.

ANNEXE

Script du dessin animé

Code	Description du micro épisode
A1	La maman tricote
A2	La maman regarde l'œuf
A3	La maman tricote
A4	La maman regarde l'heure
A5	La maman pose son tricot
A6	La maman borde l'œuf
A7	La maman contemple l'œuf
A8	La maman s'en va
B1	L'œuf saute
B2	L'œuf tombe dans la toile d'araignée
B3	La toile se rompt
B4	L'œuf tombe dans la fleur
B5	La fleur dépose l'œuf sur la feuille
B6	L'œuf roule sur la feuille jusqu'à la maison
B7	L'œuf pousse la porte
B8	L'œuf roule jusqu'au lit
C1	La souris se retourne sur l'œuf
C2	L'œuf réveille la souris
C3	La souris découvre l'œuf
C4	L'œuf fait tomber la souris du lit
D1	L'œuf éclot
D2	Le bébé oiseau court avec sa coquille sur la tête
D3	La souris enlève la coquille
D4	Le bébé oiseau court en cercle
D5	Le bébé oiseau prend la souris pour sa mère
D6	Le bébé oiseau fait un câlin à la souris
D7	La souris caresse la tête du bébé oiseau
E1	Le bébé oiseau a vu quelque chose
E2	Le bébé oiseau court et grimpe sur la commode
E3	Le bébé oiseau attaque la commode avec son bec
E4	Le bébé oiseau détruit le pied du lampadaire
E5	La souris veut attraper le bébé et reçoit l'abat-jour sur la tête
E6	Le bébé oiseau fait un trou dans le mur avec son bec
E7	La souris bloque l'oiseau par le bec et vibre

E8	La souris pose l'oiseau
F1	La souris a une idée
F2	La souris va chercher quelque chose à manger
F3	La souris tend un morceau de nourriture
F4	Le bébé oiseau mange
F5	La souris tend un autre morceau de nourriture
F6	Le bébé oiseau mange le morceau de nourriture et le bras de la souris
F7	La souris se libère
F8	L'oiseau se retrouve planté
F9	La souris dé plante l'oiseau
F10	La souris s'éponge le front et veut s'asseoir
F11	Le bébé oiseau détruit le tabouret et la souris tombe
G1	La souris regarde l'oiseau avec colère
G2	La souris prend le bébé oiseau dans ses bras
G3	La souris emporte l'oiseau dehors
G4	La souris regarde en l'air et cherche en tournant la tête à droite puis à gauche
G5	La souris a vu quelque chose et sourit
G6	La souris monte sur l'arbre jusqu'au nid
G7	La souris met le bébé oiseau au lit
G8	La souris fait au revoir
G9	La souris s'en va