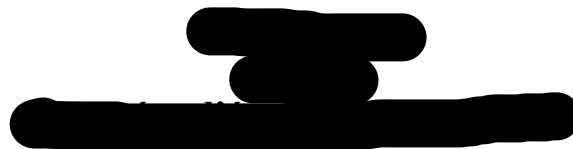


**Corpus oraux et multimodaux**

**LETEILLIER**

**Marion**

(Numéro d'étudiant – **12400026**)



**UFR LLASIC**

**Département d'Informatique intégrée en Langues, Lettres et Langage**

**Master 1 Mention Sciences du Langage Parcours Industries de la Langue**

**Enseignante responsable : Solange Rossato**

**Année universitaire  
2024 – 2025**

## Étape 0

Les enregistrements consistent en trois vidéos tournées dans une classe de grande section de maternelle. Dans l'ordre chronologique, ils durent respectivement 8 minutes et 12 secondes (104C0005), 2 minutes (136C0003) et 2 minutes (136C0004). Un groupe d'élèves – Anaïs, Precioso, Adassa, Fatma et Cyrine – interagit avec leur maîtresse autour d'une table. Ils sont invités à répondre à des questions, puis à répéter une histoire avec pour thème une moufle et des animaux. L'hypothèse de recherche est que cette activité stimule les compétences langagières des enfants et leur capacité à structurer un récit par la répétition et la reformulation des énoncés. Nous verrons des différences à la fois qualitatives et quantitatives entre les enfants. Pour analyser ces interactions, nous identifierons les phrases-clés reflétant la compréhension de l'histoire par les enfants. Les types de phrases produites seront classés en onomatopées, mots isolés, phrases simples, phrases complexes et reformulations. Nous observerons aussi les interruptions des enfants (corrections, anticipations, prises de parole, répétitions, hors contexte) ainsi que celles de la maîtresse (validation, correction, reformulation, discipline). Les résultats attendus montreraient une augmentation du nombre de phrases complexes et de reformulations chez les enfants, une baisse des mots isolés ou onomatopées. Le rôle des interventions de la maîtresse devrait être central dans l'amélioration des récits des élèves. Cette analyse permettra de mieux comprendre l'impact du suivi pédagogique sur le développement des compétences langagières des enfants.

## Étape 1

La segmentation prend en compte les groupes de souffle, délimités par des pauses plutôt longues (supérieure à 250 ms), ou bien des inspirations perceptibles. Lorsqu'une montée de l'intonation indique qu'un locuteur n'a pas terminé de parler (un pic intonatif signalant la continuité), cela est pris en compte pour éviter de segmenter prématurément. Les fins de propositions sont identifiées même si elles ne sont pas toujours marquées par des pauses ou des indices intonatifs évidents, en s'appuyant sur le contexte. La seule ponctuation utilisée est le point d'interrogation (?) pour signaler les questions. Les majuscules sont réservées aux noms propres et ne sont pas utilisées en début de phrase. Les particules et mots expressifs fréquents dans la parole spontanée (euh, ben, bah, ouais, hum) sont transcrits tels quels. Si un locuteur omet un accord grammatical ("elle s'est assis" au lieu de "elle s'est assise"), cela est transcrit tel quel. Les réductions de mots sont transcrites en ne suivant pas la prononciation observée, mais sa correction ("dessus" pour "d'ssus").

## Étape 2

Pour réaliser la transcription phonologique et l'alignement au niveau du phonème, nous avons utilisé la plateforme BAS Web Services avec l'outil "Pipeline without ASR" sur le site [clarin.phonetik.uni-muenchen.de](http://clarin.phonetik.uni-muenchen.de). Nous avons téléchargé les fichiers audios au format WAV ainsi que les fichiers texte TXT contenant les transcriptions manuelles préalablement réalisées. Le pipeline choisi était G2P-MAUS-PHO2SYL qui génère : une transcription phonémique, une segmentation en mots, et une division syllabique de la transcription phonémique. Même si ces résultats ont été générés pour démontrer la maîtrise de l'outil, les aspects de ces sorties n'ont pas été nécessaires pour nos analyses. Nous avons sélectionné le français comme langue d'analyse. Les fichiers de sortie ont été générés au format Praat (TextGrid) avec un encodage en IPA (UTF-8) afin de préserver la représentation des caractères phonétiques. En cliquant sur "Run Web Service", nous avons activé la génération automatique de la transcription phonologique et des alignements phonémiques. Nous avons téléchargé les fichiers TextGrid, qui contiennent la transcription phonologique alignée au niveau des phonèmes.

### Étape 3

Pour faire les annotations et mesures, nous avons défini des critères pour garantir une reproductibilité et permettre à d'autres annotateurs de produire des résultats similaires. Les annotations ont été réalisées à l'aide du logiciel Praat. Tout d'abord, les phrases-clés en tant que narration de l'histoire sont annotées en identifiant les phrases essentielles à la compréhension de l'histoire. Ensuite, les types de phrases sont catégorisés en plusieurs classes : onomatopées, mots isolés, phrases simples (sujet + verbe conjugué + complément), phrases complexes (plusieurs propositions), reformulations (énoncés repris ou modifiés), questions fermées (nécessitant une réponse oui ou non) et questions ouvertes (nécessitant une réponse développée). Les interactions des enfants sont aussi annotées en fonction de leur comportement. Cela inclut les corrections, les anticipations (réponses avant que l'interlocuteur n'ait terminé), les prises de parole, la participation spontanée, les comportements distrayants, ainsi que les réponses selon leur contexte (réponses à une instruction, à une demande ou à une question). Les interventions de la maîtresse sont faites en fonction de leur rôle : validation (approbation des réponses des enfants), corrections (modifications des énoncés), incitations ou demandes de participation (encouragement à la prise de parole).

### Étape 4

Dans l'audio 104C0005, l'activité commence par la présentation de l'histoire. La maîtresse pose majoritairement des questions ouvertes pour aiguiller les élèves (35) « *comment il s'appelle ?* » et 20 questions fermées « *Est-ce qu'il peut rentrer le lièvre ?* ». Precioso montre dès le départ une bonne compréhension de l'histoire et une structuration de ses énoncés avec des phrases simples, en répondant correctement aux questions (70.36s d'interaction). Il anticipe même une réponse en interrompant avec « *viens te réchauffer* ». Cependant, il se montre dissipé, avec des onomatopées récurrentes « *oh, wow, eh, wah* » ou des interventions non pertinentes « *moi je sais* », pouvant déranger les autres enfants. Lorsqu'il parle spontanément, la syntaxe peut être maladroite « *je peux pour refaire une* ». Il est souvent sollicité par la maîtresse pour aider ses camarades. Adassa, quant à elle, ne participe pratiquement pas, avec seulement une réponse « *oui* » et une onomatopée (0.8s). Fatma produit des phrases simples avec une bonne compréhension de l'histoire et une structuration de ses énoncés avec des phrases simples, en répondant correctement aux questions (29.42s). Cyrine reste peu impliquée, avec quelques mots isolés comme « *souris, renard* » ou en répondant avec affirmation « *oui oui* » ou sans former de phrases « *les pieds* » pour un total de 18.40s secondes. La maîtresse joue un rôle clé avec 189.52s de temps de parole, avec des incitations à répondre « *allez vas-y* » « *tu fais l'ours ?* ».

Dans l'audio 136C0003, les récits deviennent plus détaillés. Les questions de la maîtresse restent majoritairement ouvertes (8) pour valider la compréhension avec reformulation pour aider les élèves « *avec la moufle qu'est-ce qui se passe ?* » « *qu'est-ce qu'on nous dit de la moufle ?* » avec un total de 47.32s. Elle répète mot à mot l'histoire pour que les enfants retiennent bien. Souvent, elle doit intervenir avec des « *chut* » pour recentrer l'attention des élèves. Anaïs fait des phrases complexes et simples mais qui n'ont pas forcément le bon sens « *il n'y a personne qui répète* » au lieu de « *qui réponds* » ou hors contexte « *la souris j'aime bien* » avec une participation de 7.79 s. Precioso est de plus en plus dissipé (présence de beaucoup d'interjections, d'onomatopées) même s'il participe activement à la narration de l'histoire, en anticipant les réponses « *que c'est sa moufle* ». Son temps de parole s'élève à 11.44 secondes. Adassa s'investit plus dans l'histoire en répétant correctement, en s'affirmant avec une phrase complexe « *non c'est moi qui commence* ». Mais elle est aussi distraite, avec des interventions hors contexte « *moi j'ai trois barrettes* », en répétant plusieurs fois « *c'est encore moi* ». Avec une prise de parole spontanée « *parce que j'avais mal au ventre quand j'étais pas là* », nous voyons que la maîtrise de la syntaxe est assez avancée. Fatma avec 6.2s répète les énoncés de la maîtresse et Cyrine, participe peu (1.72s) en répétant aussi les phrases de l'histoire.

Dans l'audio 136C0004, la fin de l'activité se concentre toujours sur la répétition complète de l'histoire. La maîtresse (47.32s) a posé 1 question fermée « *Le lièvre se faufile tu m'as dit ?* » et des réponses ouvertes « *Le lièvre, il arrive et il fait quoi ?* ». La reformulation « *Donc le lièvre arrive, et ensuite ?* » permet de guider les jeunes élèves malgré des interventions pour recentrer l'attention « *T'as pas besoin de voir ce que j'écris, Adassa.* ». Adassa progresse également, avec des interventions pertinentes « *t'as oublié quelque chose* » en répondant par un élément de l'histoire « *viens te réchauffer* ». Elle est aussi intriguée par la caméra avec « *elle voit ce que je fais* » montrant toujours une bonne maîtrise de la syntaxe. Fatma (18.5s) a une participation active, entre narration et répétitions, en posant une question fermée « *t'as écrit lièvre ?* » et des réponses affirmatives « *oui, oui* ». Anaïs fait seulement une intervention hors contexte avec « *est-ce que la caméra elle nous voit ?* » mais grammaticalement correcte. Precioso (9.74s) fait une réponse complète avec « *il y a la souris et le lièvre* » mais est aussi perturbé par la caméra « *Elle voit tout le monde ?* ». Cyrine (2.87s) est intervenue une seule fois avec une phrase simple « *elle est toute seule* ».

| Enfants  | Phrases simples | Mots isolés | Onomatopées | Interruptions | Phrases Complexes | Reformulations | Réponses correctes (questions fermées) | Réponses correctes (questions ouvertes) | Phrases incomplètes | Questions Ouvertes | Questions Fermées | Répétitions |
|----------|-----------------|-------------|-------------|---------------|-------------------|----------------|--|---|---------------------|--------------------|-------------------|-------------|
| Anaïs    | 2               | 0           | 0           | 0             | 1                 | 0              | 0                                      | 0                                       | 0                   | 0                  | 0                 | 0           |
| Precioso | 5               | 21          | 23          | 3             | 1                 | 0              | 2                                      | 0                                       | 15                  | 0                  | 0                 | 2           |
| Adassa   | 9               | 4           | 2           | 2             | 5                 | 0              | 0                                      | 0                                       | 1                   | 0                  | 0                 | 1           |
| Fatma    | 1               | 5           | 0           | 3             | 0                 | 1              | 0                                      | 4                                       | 6                   | 2                  | 5                 | 0           |
| Cyrine   | 2               | 14          | 5           | 1             | 0                 | 1              | 1                                      | 12                                      | 2                   | 0                  | 1                 | 0           |

### Interventions de la maîtresse

|                           |    |
|---------------------------|----|
| Nombre total de questions | 45 |
| Questions ouvertes        | 43 |
| Questions fermées         | 25 |
| Validations               | 26 |
| Corrections               | 8  |
| Reformulations            | 14 |

L'analyse des vidéos donne des réponses plutôt nuancées aux hypothèses. L'activité autour de l'histoire de la moufle semble contribuer au développement des compétences langagières. Adassa et Fatma affichent des progrès mais d'autres élèves comme Precioso, Cyrine et Anaïs montrent une progression plus limitée. Precioso, malgré une bonne compréhension de l'histoire, est freiné par un comportement souvent dissipé avec des onomatopées et des interventions hors sujet. Cyrine et Anaïs participent peu et se limitent à des phrases simples ou à des mots isolés sans évolution. Les différences entre les enfants avec le niveau de leur participation et la qualité de leurs interventions sont identifiées. Les temps de parole varient : Precioso et Fatma contribuent davantage tandis que Cyrine et Anaïs restent en retrait. Adassa et Fatma montrent des progrès qualitatifs avec des énoncés mieux structurés alors que Cyrine et Anaïs ont du mal à dépasser leurs réponses pré-construites. En ce qui concerne l'utilisation de phrases complexes et de reformulations, une progression est observée chez certains enfants comme Adassa et Fatma mais elle reste absente chez Cyrine et Anaïs. Pour Precioso, ses énoncés restent simples malgré une participation active. Pour les onomatopées et les mots isolés, Precioso continue d'en utiliser fréquemment et Cyrine et Anaïs s'appuient principalement sur des mots isolés. Le rôle de la maîtresse est central dans le déroulement de l'activité. Elle joue un rôle clé dans la structuration des échanges, en

validant les interventions des enfants, en reformulant les énoncés pour guider le récit et en maintenant la discipline. Toutefois, l'approche basée sur la répétition et les questions fermées, limite la spontanéité des enfants pour voir l'évolution de leur syntaxe. Un accompagnement personnalisé serait bénéfique pour les enfants moins impliqués, comme Anaïs et Cyrine, afin de renforcer leur participation et leurs progrès.

### **Étape 5**

L'analyse des enregistrements a permis de mettre en lumière l'impact de l'activité narrative sur les compétences langagières des enfants tout en montrant des progrès variés selon les participants. L'activité, centrée sur la répétition et la reformulation, a montré son efficacité pour améliorer les capacités de certains enfants. Mais l'environnement dans lequel les enregistrements ont été réalisés a posé des contraintes techniques. Le bruit des autres élèves de la classe a affecté la qualité sonore. Cela a influé sur la précision des transcriptions et annotations. De plus, la présence de la caméra a également perturbé le déroulement de l'activité. Certains enfants ont été distraits en posant des questions hors contexte telles que « *Est-ce que la caméra elle nous voit ?* ». De manière générale, les étapes de 0 à 4 ont permis d'aborder chaque aspect de l'analyse de manière structurée. Cela a facilité l'interprétation des données.