

## **Offre de post-doc (durée 18 mois) : Analyse automatique de la parole enfantine**

[english version below]

**Date de début** : 1<sup>er</sup> septembre 2026

**Mission** : orientée vers la recherche

**Lieu de travail** : Laboratoire d'Informatique de Grenoble, Univ. Grenoble Alpes, France

**Mots-clefs** : reconnaissance automatique de la parole, modèles auto-supervisés, développement du langage

### **Description :**

Dans le cadre de la chaire *Socialization and Language at School: An Opportunity and a Challenge for AI in the Exploitation of Massive Data (SoLaS)* financée par MIAI Cluster (<https://miai-cluster.univ-grenoble-alpes.fr/research/chairs/socialization-and-language-at-school-an-opportunity-and-a-challenge-for-ai-in-the-exploitation-of-massive-data-1625646.kjsp>), nous proposons un contrat post-doctoral d'une durée de 18 mois s'inscrivant dans le Workpackage 2 : *Automatic transcription tools for spontaneous child speech, including tools to assess language development levels*. Les données sur lesquelles se fondent cette recherche sont issues du projet DylNet (*Language dynamics, linguistic learning, and sociability at preschool: benefits of wireless proximity sensors in collecting big data* <ANR-16-CE28-0013>; <https://dylnet.univ-grenoble-alpes.fr/>) lors duquel ≈ 30 000 heures d'enregistrements audio d'enfants de maternelle ont été collectées (Dai et al., 2022).

Dans le domaine du traitement automatique des langues, la parole enfantine reste un défi majeur. Ce projet combine deux enjeux importants : le traitement de la parole spontanée (Evain et al., 2024) et les difficultés particulières que soulève la parole enfantine (Singh et al., 2026 ; Fan et al., 2024).

L'objectif de la recherche post-doctorale sera de proposer et d'implémenter un modèle d'évaluation du niveau de développement langagier enfantin à partir de fichiers audio et de métadonnées diverses donnés en entrée.

L'équipe GETALP du LIG travaille depuis longtemps sur des modèles auto-supervisés dédiés au français : FlauBERT (Le et al., 2020) et Jargon (Segonne et al., 2024) pour l'écrit, LeBenchmark pour l'oral (Evain et al., 2021). Plus récemment, GETALP a proposé le modèle multimodal Pantagrue qui traite à la fois le langage écrit et le langage oral (Le et al., 2026). L'adaptation de ces modèles auto-supervisés aux voix d'enfants a commencé à être explorée dans le cadre de la chaire SoLaS pour des tâches de reconnaissance automatique de la parole. Dans ce cadre, le travail réalisé pendant ce post-doc s'inscrira dans les travaux de l'équipe pour réaliser une évaluation automatique du niveau de développement langagier enfantin.

Étant donné la grande variabilité dans les habiletés langagières des enfants de maternelle entre 3 et 6 ans, le modèle proposé devra en tenir compte pour positionner le niveau langagier des enfants par rapport à une cohorte du même âge. Pour cela, nous disposons d'évaluations en réception du langage (scores en lexique et en syntaxe en réception) pour tous les enfants pour lesquels nous disposons d'enregistrements audio. Les grilles de positionnement langagier de Rousset et al. (2019) associées aux données audio recueillies dans le cadre du projet PARM<sup>1</sup> (Buson et al., 2024 ; Buson et al. à paraître) seront également disponibles. Cette prédiction d'un niveau de développement langagier sera complétée par des analyses plus localisées permettant une évaluation de la production langagière plus détaillée :

- débit de parole (Colletta et al., 2008)
- longueur moyenne des énoncés (Rice et al., 2010)
- diversité lexicale (Malvern et al., 2004)
- complexité morphosyntaxique (Canut, 2014)

---

<sup>1</sup> PARM : *PARler Raconter en Maternelle*.

Ces analyses ont été réalisées automatiquement sur un ensemble de données transcrites par des linguistes et sont actuellement en cours de validation.

### Compétences requises :

- Doctorat en informatique dans le domaine du Traitement Automatique des Langues
- Maîtrise des architectures état de l'art pour les modèles auto-supervisés
- Ouverture aux Sciences Humaines et Sociales
- Bonnes compétences de communication à l'oral et à l'écrit

### Environnement scientifique :

La personne recrutée sera supervisée par Solange Rossato (LIG, Univ. Grenoble Alpes), Benjamin Lecouteux (LIG, Univ. Grenoble Alpes) et Aurélie Nardy (Lidilem, Univ. Grenoble Alpes). Solange Rossato et Benjamin Lecouteux sont chercheurs dans l'équipe GETALP du LIG et travaillent notamment en reconnaissance automatique de la parole. Aurélie Nardy est chercheuse en acquisition du langage et elle a notamment supervisé le projet ANR DyLNet.

La personne recrutée sera accueillie au sein de l'équipe GETALP du LIG (<https://lig-getalp.imag.fr/>) qui possède une forte expertise et de l'expérience dans le domaine du Traitement du Langage Naturel. Cette équipe offre un environnement de travail stimulant, multinational et offre les moyens nécessaires à la réalisation de la thèse en termes d'équipements et d'échanges scientifiques. Des réunions régulières avec les trois superviseurs auront lieu tout au long du contrat post-doctoral.

### Envoi des candidatures :

Les candidatures devront contenir un CV détaillé, une copie du diplôme de doctorat, un résumé d'une page de la thèse de doctorat, une lettre de motivation et au moins deux références (personnes susceptibles d'être contactées). Une ou plusieurs lettres de recommandation peuvent être jointes au dossier.

Elles seront adressées à Solange Rossato ([solange.rossato@univ-grenoble-alpes.fr](mailto:solange.rossato@univ-grenoble-alpes.fr)), Benjamin Lecouteux ([benjamin.lecouteux@univ-grenoble-alpes.fr](mailto:benjamin.lecouteux@univ-grenoble-alpes.fr)), et Aurélie Nardy ([aurelie.nardy@univ-grenoble-alpes.fr](mailto:aurelie.nardy@univ-grenoble-alpes.fr)).

### Références

- Buson, L., Rousset, I. et Rossato, S. (2024). S'exprimer dans un langage oral syntaxiquement correct et précis. Syntaxe et variation en maternelle. *Pratiques*, 201-202. <https://doi.org/10.4000/122u3>
- Buson, L., Rousset, I. et Rossato, S. (à paraître). *Méthode Parm - PARler Raconter en Maternelle*. UGA Éditions.
- Canut, E. (2014). Acquisition des constructions syntaxiques complexes chez l'enfant français entre 2 et 6 ans. *SHS Web of Conferences*, 8, 1437-1452. <https://doi.org/10.1051/shsconf/20140801092>
- Colletta, J.-M., Pellenq, C., Rousset, I. (2008). Évolution du débit de parole chez l'enfant francophone dans des tâches narrative et conversationnelle. *27èmes Journées d'Études sur la Parole*, Association Francophone de la Communication Parlée. <https://hal.science/hal-01292877v1>
- Dai, S., Bouchet, H., Karsai, M., Chevrot, J.-P., Fleury, E., Nardy, A. (2022). Longitudinal data collection to follow social network and language development dynamics at preschool. *Scientific Data*, 9, Article 777. <https://doi.org/10.1038/s41597-022-01756-x>
- Evain, S., Nguyen, M. H., Le, H., Boito, M. Z., Mdhaïffar, S., Alisamir, S., ... , Besacier, L. (2021). Task agnostic and task specific self-supervised learning from speech with LeBenchmark. In *Thirty-fifth Conference on Neural Information Processing Systems (NeurIPS 2021)*. <https://neurips.cc/virtual/2021/29889>
- Evain, S., Rossato, S., Portet, F. (2024). Unraveling spontaneous speech dimensions for cross-corpus ASR system evaluation for French. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)* (pp. 17165-17175). <https://aclanthology.org/2024.lrec-main.1491/>
- Fan, R., Balaji Shankar, N., Alwan, A. (2024). Benchmarking Children's ASR with Supervised and Self-supervised Speech Foundation Models. In *Proc. Interspeech 2024* (pp. 5173-5177). [https://www.isca-archive.org/interspeech\\_2024/fan24b\\_interspeech.pdf](https://www.isca-archive.org/interspeech_2024/fan24b_interspeech.pdf)
- Le, P. H., Pelloin, V., Chatelain, A., Bouziane, M., Ghennai, M., Guan, Q., ... & Schwab, D. (2026). Pantagruel: Unified Self-Supervised Encoders for French Text and Speech. *arXiv preprint*. <https://arxiv.org/abs/2601.05911>

Le, H., Vial, L., Frej, J., Segonne, V., Coavoux, M., Lecouteux, B., ..., Schwab, D. (2020). FlauBERT: unsupervised language model pre-training for French. In *Proceedings of the twelfth language resources and evaluation conference* (pp. 2479-2490). <https://aclanthology.org/2020.lrec-1.302/>

Malvern, D., Richards, B., Chipere, N., Durán, P. (2004). *Lexical diversity and language development: quantification and assessment*. Palgrave Macmillan.

Rice, M.L., Smolik, F., Perpich, D., Thompson, T., Rytting, N., Blossom, M. (2010). Mean length of utterance levels in 6-month intervals for children 3 to 9 years with and without language impairments. *Journal of Speech, Language, and Hearing Research*, 53(2), 333-349. [https://doi.org/10.1044/1092-4388\(2009/08-0183\)](https://doi.org/10.1044/1092-4388(2009/08-0183))

Rousset, I., Rossato, S., Lequette, C., Latapie, E. (2019). Exploration des compétences langagières des enfants d'écoles maternelles en zone d'éducation prioritaire. *10èmes Journées Internationale de la Linguistique de Corpus*, Nov 2019, Grenoble, France. <https://hal.science/hal-02457055v1>

Segonne, V., Mannion, A., Canul, L. C. A., Audibert, A. D., Liu, X., Macaire, C., ..., Schwab, D. (2024). Jargon: a suite of language models and evaluation tasks for French specialized domains. In *Proceedings of the 2024 joint international conference on computational linguistics, language resources and evaluation (lrec-coling 2024)* (pp. 9463-9476). <https://aclanthology.org/2024.lrec-main.827/>

Singh, V. P., Sahidullah, M., Kinnunen, T. H. (2026). Causal analysis of ASR errors for children: quantifying the impact of physiological, cognitive, and extrinsic factors. *Computer Speech & Language*, 95, 101859. <https://doi.org/10.1016/j.csl.2025.101859>

---

### **Postdoctoral position (18 months): Automatic Analysis of Child Speech**

**Start date:** September 1, 2026

**Type of position:** Research-oriented

**Location:** Laboratoire d'Informatique de Grenoble (LIG), Université Grenoble Alpes, France

**Keywords:** Automatic Speech Recognition (ASR), self-supervised models, language development

#### **Description:**

This 18-month postdoctoral position is offered within the framework of the SoLaS Chair (*Socialization and Language at School: An Opportunity and a Challenge for AI in the Exploitation of Massive Data*), funded by MIAI Cluster (<https://miai-cluster.univ-grenoble-alpes.fr/research/chairs/socialization-and-language-at-school-an-opportunity-and-a-challenge-for-ai-in-the-exploitation-of-massive-data-1625646.kjsp>).

The position is part of Work Package 2: *Automatic transcription tools for spontaneous child speech, including tools to assess language development levels*.

The research relies on data from the DyLNet project (*Language dynamics, linguistic learning, and sociability at preschool: benefits of wireless proximity sensors in collecting big data* <ANR-16-CE28-0013>; <https://dylnet.univ-grenoble-alpes.fr/>), which collected approximately 30,000 hours of preschool children's audio recordings (Dai et al, 2022).

Child speech remains a major challenge in Natural Language Processing due both to spontaneous speech phenomena (Evain et al., 2024) and to the specific acoustic and linguistic variability of young speakers (Singh et al., 2026; Fan et al., 2024).

The postdoctoral researcher will design and implement a model for assessing children's language development levels from audio recordings and associated metadata. This work builds upon the long-standing expertise of the GETALP team (LIG) in self-supervised models for French, including FlauBERT (Le et al., 2020) and Jargon (Segonne et al., 2024) for written language, LeBenchmark for speech (Evain et al., 2021), and more recently Pantagruel, a multimodal model processing both text and speech (Le et al., 2026). The adaptation of these self-supervised models to children's speech has already begun within the framework of the SoLaS Chair for automatic speech recognition tasks. In this context, the work carried out during the postdoctoral project will build on these ongoing efforts, with the objective of developing an automatic assessment framework for children's language development levels.

Given the substantial variability in language abilities among preschool children aged 3 to 6, the proposed model will need to account for age-related differences in order to position each child's language level relative to an age-matched cohort. To this end, we have access to receptive language assessments (lexical and

syntactic comprehension scores) for all children for whom audio recordings are available. Language positioning grids (Rousset et al., 2019) associated with the audio data collected within the PARM<sup>2</sup> project (Buson et al., 2024; Buson et al., forthcoming) will also be available. This global prediction of language development level will be complemented by more fine-grained analyses enabling a detailed assessment of language production, including:

- Speech rate (Colletta et al., 2008)
- Mean Length of Utterance (Rice et al., 2010)
- Lexical diversity (Malvern et al., 2004)
- Morphosyntactic complexity (Canut, 2014)

These analyses were automatically computed on a dataset transcribed by linguists and are currently undergoing validation.

#### **Required qualifications:**

- PhD in Computer Science with a specialization in Natural Language Processing or Speech Processing
- Strong expertise in state-of-the-art self-supervised architectures
- Openness to interdisciplinary collaboration with the Humanities and Social Sciences
- Excellent written and oral communication skills

#### **Scientific Environment:**

The recruited researcher will be supervised by Solange Rossato and Benjamin Lecouteux (researchers at the GETALP team, Laboratoire d'Informatique de Grenoble, Université Grenoble Alpes), whose work focuses in particular on automatic speech recognition, and by Aurélie Nardy (Lidilem, Université Grenoble Alpes), a researcher in language acquisition who notably led the ANR DyLNet project.

The postdoctoral researcher will join the GETALP team at LIG, which has strong expertise and extensive experience in Natural Language Processing and speech technologies. The team provides a stimulating, international research environment and offers the necessary computational resources and scientific interactions to ensure the successful completion of the project. Regular meetings with the three supervisors will be held throughout the duration of the postdoctoral contract.

#### **Application procedure:**

Applications must include: a detailed CV, a copy of the PhD diploma (or certificate of completion), a one-page summary of the doctoral thesis, a cover letter and at least two references (referees who may be contacted). One or more recommendation letters may also be included in the application file. Applications should be sent to Solange Rossato ([solange.rossato@univ-grenoble-alpes.fr](mailto:solange.rossato@univ-grenoble-alpes.fr)), Benjamin Lecouteux ([benjamin.lecouteux@univ-grenoble-alpes.fr](mailto:benjamin.lecouteux@univ-grenoble-alpes.fr)), et Aurélie Nardy ([aurelie.nardy@univ-grenoble-alpes.fr](mailto:aurelie.nardy@univ-grenoble-alpes.fr)).

#### **References**

- Buson, L., Rousset, I. et Rossato, S. (2024). *S'exprimer dans un langage oral syntaxiquement correct et précis. Syntaxe et variation en maternelle. Pratiques*, 201-202. <https://doi.org/10.4000/122u3>
- Buson, L., Rousset, I. et Rossato, S. (forthcoming). *Méthode Parm - PARler Raconter en Maternelle*. UGA Éditions.
- Canut, E. (2014). Acquisition des constructions syntaxiques complexes chez l'enfant français entre 2 et 6 ans. *SHS Web of Conferences*, 8, 1437-1452. <https://doi.org/10.1051/shsconf/20140801092>
- Colletta, J.-M., Pellenq, C., Rousset, I. (2008). Évolution du débit de parole chez l'enfant francophone dans des tâches narrative et conversationnelle. *27èmes Journées d'Études sur la Parole*, Association Francophone de la Communication Parlée. <https://hal.science/hal-01292877v1>
- Dai, S., Bouchet, H., Karsai, M., Chevrot, J.-P., Fleury, E., Nardy, A. (2022). Longitudinal data collection to follow social network and language development dynamics at preschool. *Scientific Data*, 9, Article 777. <https://doi.org/10.1038/s41597-022-01756-x>

---

<sup>2</sup> PARM: *PARler Raconter en Maternelle* (Speaking and Storytelling in Preschool).

- Evain, S., Nguyen, M. H., Le, H., Boito, M. Z., Mdhaffar, S., Alisamir, S., ... , Besacier, L. (2021). Task agnostic and task specific self-supervised learning from speech with LeBenchmark. In *Thirty-fifth Conference on Neural Information Processing Systems (NeurIPS 2021)*. <https://neurips.cc/virtual/2021/29889>
- Evain, S., Rossato, S., Portet, F. (2024). Unraveling spontaneous speech dimensions for cross-corpus ASR system evaluation for French. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)* (pp. 17165-17175). <https://aclanthology.org/2024.lrec-main.1491/>
- Fan, R., Balaji Shankar, N., Alwan, A. (2024). Benchmarking Children's ASR with Supervised and Self-supervised Speech Foundation Models. In *Proc. Interspeech 2024* (pp. 5173-5177). [https://www.isca-archive.org/interspeech\\_2024/fan24b\\_interspeech.pdf](https://www.isca-archive.org/interspeech_2024/fan24b_interspeech.pdf)
- Le, P. H., Pelloin, V., Chatelain, A., Bouziane, M., Ghennai, M., Guan, Q., ... & Schwab, D. (2026). Pantagruel: Unified Self-Supervised Encoders for French Text and Speech. *arXiv preprint*. <https://arxiv.org/abs/2601.05911>
- Le, H., Vial, L., Frej, J., Segonne, V., Coavoux, M., Lecouteux, B., ..., Schwab, D. (2020). FlauBERT: unsupervised language model pre-training for French. In *Proceedings of the twelfth language resources and evaluation conference* (pp. 2479-2490). <https://aclanthology.org/2020.lrec-1.302/>
- Malvern, D., Richards, B., Chipere, N., Durán, P. (2004). *Lexical diversity and language development: quantification and assessment*. Palgrave Macmillan.
- Rice, M.L., Smolik, F., Perpich, D., Thompson, T., Rytting, N., Blossom, M. (2010). Mean length of utterance levels in 6-month intervals for children 3 to 9 years with and without language impairments. *Journal of Speech, Language, and Hearing Research*, 53(2), 333-349. [https://doi.org/10.1044/1092-4388\(2009/08-0183\)](https://doi.org/10.1044/1092-4388(2009/08-0183))
- Rousset, I., Rossato, S., Lequette, C., Latapie, E. (2019). Exploration des compétences langagières des enfants d'écoles maternelles en zone d'éducation prioritaire. *10èmes Journées Internationale de la Linguistique de Corpus*, Nov 2019, Grenoble, France. <https://hal.science/hal-02457055v1>
- Segonne, V., Mannion, A., Canul, L. C. A., Audibert, A. D., Liu, X., Macaire, C., ..., Schwab, D. (2024). Jargon: a suite of language models and evaluation tasks for French specialized domains. In *Proceedings of the 2024 joint international conference on computational linguistics, language resources and evaluation (lrec-coling 2024)* (pp. 9463-9476). <https://aclanthology.org/2024.lrec-main.827/>
- Singh, V. P., Sahidullah, M., Kinnunen, T. H. (2026). Causal analysis of ASR errors for children: quantifying the impact of physiological, cognitive, and extrinsic factors. *Computer Speech & Language*, 95, 101859. <https://doi.org/10.1016/j.csl.2025.101859>