



HAL
open science

Évaluer automatiquement la complexité syntaxique des productions orales enfantines à l'aide d'un outil en ligne

Loïc Liégeois, Christine da Silva-Genest, Christophe Benzitoun, Caroline Masson,
Christophe Parisse

► To cite this version:

Loïc Liégeois, Christine da Silva-Genest, Christophe Benzitoun, Caroline Masson, Christophe Parisse. Évaluer automatiquement la complexité syntaxique des productions orales enfantines à l'aide d'un outil en ligne. 12èmes Journées Internationales de Linguistique de Corpus (JLC), ICAR; LIDILEM; DDL; PRAXILING; CLLE, Oct 2025, Lyon (ENS LSH), France. ⟨hal-05343098⟩

HAL Id: hal-05343098

<https://hal.science/hal-05343098v1>

Submitted on 3 Nov 2025

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.



HAL Authorization

Évaluer automatiquement la complexité syntaxique des productions orales enfantines à l'aide d'un outil en ligne

Loïc Liégeois¹, Christine da Silva-Genest², Christophe Benzitoun³, Caroline Masson⁴ et Christophe Parisse⁵
¹LRL, Université Clermont Auvergne ; ²DevAH, Université de Lorraine ; ³ATILF, Université de Lorraine ; ⁴CLESTHIA, Université Sorbonne Nouvelle ; ⁵MODYCO, Université Paris Nanterre
loic.liegeois@uca.fr, christine.da-silva-genest@univ-lorraine.fr ; christophe.benzitoun@univ-lorraine.fr ;
caroline.masson@sorbonne-nouvelle.fr ; cparisse@parisnanterre.

L'évaluation du langage enfantin : état des lieux

L'évaluation de la complexité du langage enfantin est une thématique récurrente dans le domaine de l'acquisition du langage. Puisque l'une des caractéristiques principales du langage de l'enfant est qu'il se complexifie au fur et à mesure de son développement, l'évaluation de cette complexité est souvent vue comme primordiale dans le but, par exemple, de situer un enfant particulier par rapport à une cohorte. Le seul âge de l'enfant n'est pas suffisant pour déterminer son stade de développement linguistique. Ainsi, la longueur moyenne des énoncés (MLU, Brown, 1973) a très vite été utilisée comme un indice de complexité des énoncés enfantins et reste encore aujourd'hui l'étalon de mesure du développement du langage le plus utilisé. Toutefois, on sait que cette mesure, n'est plus fiable au-delà de l'âge de 48 mois (Klee & Fitzgerald, 1985 ; Klee et al., 1989 ; Blake et al., 1993). Pour autant, le langage enfantin continue de se complexifier au-delà de cet âge : la proportion d'énoncés complexes s'accroît largement (Diessel, 2004). De nombreux instruments de mesure du développement du langage ont donc été proposés comme la *clausal density* (Berman, 1996) ou le *Complex Syntax Type-Token Ratio* (Witkowska et al., 2022). Il existe également des mesures plus globales du développement morphosyntaxique telles que l'*Index of Productive Syntax* (Scarborough, 1990) ou le *Developmental Sentence Scoring* (Lee, 1974) pour l'anglais. Malgré l'utilisation de ces mesures pour évaluer le langage d'enfants plus âgés, il est difficile de s'accorder sur une mesure syntaxique pertinente et corrélée à un degré de complexité du langage.

De plus, la méthodologie de recueil des données à l'origine du calcul des indices de complexité est un facteur important à prendre en compte. En effet, elle influence directement la complexité syntaxique des énoncés observés. Par exemple, les tâches de récit et de restitution d'histoire induisent, chez des enfants de 5 à 7 ans, un langage plus complexe qu'une conversation classique recueillie en situation naturelle (Westerveld & Vidler, 2016 ; Westerveld et al., 2004).

En dépit de son intérêt, les mesures de complexité syntaxique du langage enfantin spontané sont peu utilisées et laissent la place à l'utilisation de tests de langage qui privilégient des situations de production très contraintes et dirigées. Notre objectif est double. Il s'agit tout d'abord, dans la continuité de nos précédents travaux, de présenter la mesure que nous avons développée pour évaluer la complexité des productions spontanées d'enfants âgés de plus de 4 ans. Celle-ci fera l'objet d'une validation empirique. Ensuite, notre communication s'articulera autour de la présentation de l'outil en ligne que nous développons et qui permettra le calcul automatique de la mesure de complexité à partir du dépôt d'une transcription par l'utilisateur.

Évaluer la complexité du langage enfantin au moyen de données naturelles et spontanées

Dans le cadre d'interventions orthophoniques, les professionnelles portent un intérêt certain à l'évaluation d'un langage enfantin spontané. En effet, les situations naturelles d'interaction engendrent une meilleure implication de l'enfant et permettent à l'orthophoniste de mieux saisir ses compétences. Cela induit une intervention orthophonique plus efficace (Rodi, 2017). Comparées aux indicateurs fournis par les résultats à des tests de langage standardisés, les mesures issues de l'analyse du langage spontané offrent une forte fiabilité, une validité plus robuste et une meilleure sensibilité (Guo & Eisenberg, 2015; Pavelko & Owens, 2017).

Cependant, les orthophonistes se trouvent confrontées à plusieurs types de difficultés lorsqu'il s'agit d'évaluer le degré de complexité d'énoncés enfantins produits en situation d'interaction. Tout d'abord, il n'existe pas de mesure claire, standardisée et étalonnée de la complexité. Si les spécialistes ont une bonne intuition de ce qui est complexe chez un enfant, il n'existe pas à ce jour d'indices fiables et partagés afin de la mesurer scientifiquement. Par ailleurs, les comparaisons entre sujets au développement atypique (ou entre enfants au développement typique et atypique) s'avèrent difficiles car les mesures ne sont pas standardisées (da Silva et al., 2023). De plus, la mise en œuvre d'une évaluation fondée sur des interactions spontanées avec l'enfant s'avère particulièrement chronophage pour les orthophonistes (Klatte et al., 2022) et peu adaptée à leur pratique professionnelle.

C'est pour répondre à ces problématiques que nous avons souhaité développer, à la suite de nos précédents travaux, un outil permettant une mesure de la complexité syntaxique à partir de l'analyse automatique de corpus de productions spontanées de jeunes enfants. Cet outil se veut être facile d'utilisation, accessible en ligne, reposant sur des indices fiables et permettant une comparaison des données, tout en tenant compte des aspects éthiques et réglementaires.

Définition d'un indice de complexité

Pour définir un indice de complexité syntaxique, deux pistes ont été explorées. Une première piste « M16 (règles ad-hoc) » s'appuie sur les caractéristiques de la grammaire de l'oral (Blanche-Benveniste, 1990) et des productions enfantines orales spontanées. Nous avons créé une typologie du degré d'élaboration langagière en retenant 16 critères de complexité (présence d'un sujet nominal, d'adjectifs épithètes ou de temps verbaux spécifiques). Ensuite, cette typologie est appliquée automatiquement aux corpus de productions enfantines par des méthodes de traitement automatique du langage. Dès qu'un des 16 critères retenus apparaît dans l'énoncé d'un enfant, celui-ci est considéré comme complexe. Chaque critère apparaît donc comme un marqueur révélateur de la présence d'une caractéristique linguistique exceptionnelle, rare, ou classiquement considérée comme associée à un discours complexe par des spécialistes de l'acquisition du langage.

Une deuxième piste « Apprentissage TTK » est basée sur l'utilisation de mesures lexico-syntaxiques automatiques réalisées au moyen d'un outil développé dans le projet ANR TextToKids (Blandin et al., 2020). Ces mesures sont appliquées au corpus Colaje (Morgenstern et Parris, 2012) ainsi qu'au corpus EVALANG (da Silva-Genest, Christmann et Willaime, 2023 ; da Silva-Genest, Masson et Guigourès, 2023). Ce dernier corpus est constitué de productions spontanées d'enfants, avec et sans trouble du développement du langage (TDL), âgés de 5 à 7 ans (Liégeois et al., 2024, da Silva-Genest et al., 2023). Parmi l'ensemble des mesures, nous retenons les 20 les plus efficaces, c'est-à-dire celles qui sont corrélées avec l'âge de l'enfant. Ces 20 mesures sont ensuite utilisées pour apprécier le nombre d'indices

linguistiques inhabituellement élevés dans un énoncé (plus que le 80ème quantile). À la différence de la méthode M16, la méthode TTK n’induit donc pas de biais tirés de l’expertise des spécialistes et se fonde uniquement sur une méthode automatique d’apprentissage non-supervisé. Avec cette méthode, un énoncé est considéré comme complexe si au moins un des indices apparaît inhabituellement élevé dans l’énoncé, comparé aux données du corpus d’apprentissage.

Application des deux méthodes à un corpus de productions enfantines

Le premier de nos objectifs est de présenter les résultats obtenus suite à l’application des méthodes précédemment décrites à un corpus de productions naturelles d’enfants âgés de 5 à 7 ans. Pour le moment, les méthodes M16 et TTK ont été appliquées aux productions de 6 enfants, avec et sans TDL, recueillies dans des situations de jeu et de récit. Ces enfants sont issus du corpus EVALANG, et les données issues de leurs productions n’ont pas participé à l’apprentissage TTK. L’ensemble des données regroupent environ 1500 énoncés. L’apprentissage TTK a été réalisé avec deux jeux de données : le reste du corpus EVALANG et le corpus COLAJE. Les résultats obtenus ont été comparés avec un codage manuel réalisé par les auteurs de l’étude. En d’autres termes, nous avons comparé la classification binaire des énoncés (simple / complexe) effectuée manuellement, automatiquement via la méthode M16 et automatiquement via la méthode TTK. Les résultats sont présentés dans la table ci-dessous.

Mesure (méthode de traitement)	Rappel	Précision	F1
M16 (règles ad-hoc)	0.70	0.58	0.63
Apprentissage TTK (corpus XXX)	0.57	0.52	0.54
Apprentissage TTK (corpus COLAJE)	0.60	0.49	0.54
Combinaison de méthodes (M16 et TTK)	0.85	0.50	0.63

La mesure de Rappel indique quelle proportion d’énoncés complexes est bien repérée. La valeur 0.85 indique ainsi que 85% des énoncés complexes ont été identifiés. La mesure de Précision informe sur le nombre d’énoncés correctement identifiés comme complexes. La valeur 0.50 indique que 50% des énoncés annotés comme complexes le sont effectivement. La meilleure mesure de qualité des algorithmes, F1, tient quant à elle compte du rappel et de la précision.

Nous pouvons observer que la méthode *ad hoc* M16 est la meilleure sur notre corpus de test. Il est surtout intéressant que les deux méthodes produisent des résultats qui ne se recouvrent pas exactement. En les combinant, nous récupérerons davantage d’énoncés complexes (85%) au prix d’une méthode moins précise qui attrape plus d’énoncés non complexes. En effet, 50% des énoncés identifiés comme complexes ne sont pas considérés comme tels lors de l’annotation manuelle effectuée par des experts.

Présentation de l’outil

Notre second objectif est de présenter l’outil en ligne dcpx (Parisse, 2025) mis à disposition de la communauté (<https://dcpx.ortolang.fr/fr/>). Cet outil permet d’obtenir un indice de complexité des productions d’un enfant à partir de la transcription d’un enregistrement (10-15min). Cette transcription pourra être directement chargée par l’utilisateur ou bien générée par un outil de transcription automatique. Notre cahier des charges était le suivant : cet outil devait être simple d’utilisation, accessible en ligne, transparent au niveau des calculs effectués et respectueux des principes éthiques et législatifs. Dans cette partie de la communication, nous nous focaliserons

donc à la fois sur les problématiques techniques et éthiques que posent la mise à disposition d'un tel outil, avant d'effectuer une courte démonstration de son utilisation.

Références bibliographiques

Berman, R. A. (1996). "Form and function in developing narrative abilities: the case of 'and'," in *Social Interaction, Context, and Language: Essays in Honor of Susan Ervin-Tripp*, eds D. Slobin, J. Gerhardt, A. Kyratzis, and J. Guo (Mahwah: Lawrence Erlbaum), 343–367.

Blake, J., Quartaro, G., & Onorati, S. (1993). Evaluating quantitative measures of grammatical complexity in spontaneous speech samples. *Journal of Child Language*, 20(1), 139–152. <https://doi.org/10.1017/S0305000900009168>

Blanche-Benveniste, C. (1990). *Le français parlé*. Éditions du CNRS.

Blandin A., Lecorvé, G., Battistelli, D. et Étienne, A. (2020). Age Recommendation for Texts. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 1431–1439, Marseille, France. European Language Resources Association.

Brown, R. W. (1973). *A first language: The early stages*. Harvard University Press.

Diessel, H. (2004). *The Acquisition of Complex Sentences*. Cambridge University Press.

Guo, L.Y. & Eisenberg, S. (2015). Sample Length Affects the Reliability of Language Sample Measures in 3-Year-Olds: Evidence From Parent-Elicited Conversational Samples. *Language, Speech, and Hearing Services in Schools* 46: 141-153.

Klatte I., van Heugten V., Zwitserlood R., et Gerrits E., 2022, « Language sample analysis in clinical practice: speech-language pathologists' barriers, facilitators, and needs », *Language, Speech, and Hearing Services in Schools*, 53(1), p.1-16.

Klee, T., Schaffer, M., May, S., Membrino, I., & Mougey, K. (1989). A comparison of the age-MLU relation in normal and specifically language-impaired preschool children. *Journal of Speech and Hearing Disorders*, 54, 226–233.

Klee, T., & Fitzgerald, M. D. (1985). The relation between grammatical development and mean length of utterance in morphemes. *Journal of Child Language*, 12, 251–269.

Lee, L. L. (1974). *Developmental sentence analysis: A grammatical assessment procedure for speech and language clinicians*. Northwestern University Press.

Liégeois, L., da Silva, C., Benzitoun, C., Masson, C. et Parisse, C. (2024). Méthodologie(s) de segmentation des transcriptions de l'oral en vue de la création d'une mesure d'évaluation des productions verbales enfantines. *Colloque International La linguistique de l'oral spontané à travers les langues : création, annotation et analyse de corpus, segmentation du discours*. INALCO, 23-24 mai 2024, Paris.

Morgenstern, A., & Parisse, C. (2012). The Paris Corpus. *Journal of French Language Studies*, 22(Special Issue 01), 7–12. <https://doi.org/10.1017/S095926951100055X>

Parisse, C. (2025). dcpX. [Logiciel]. <https://dcpX.ortolang.fr/fr/>

Pavelko S. L. et Owens J. R. E., 2017, « sampling utterances and grammatical analysis revised (SUGAR): New normative values for language sample analysis measures », *Language, Speech, and Hearing Services in Schools*, 48, p. 197-215.

Rodi M., 2017, « Les interactions du quotidien comme ressources pour l'évaluation des capacités langagières », *GLOSSA*, 120, p. 75-98.

Scarborough, H. S. (1990). Index of productive syntax. *Applied Psycholinguistics*, 11, 1-22.

da Silva-Genest, C., Christmann, A.-L. et Willaime, P., 2023. Ressources en acquisition et pathologie de l'acquisition du langage. 11ème Journée de Linguistique de Corpus, Juillet 2023, Grenoble, France.

da Silva, C., Liégeois, L., Masson, C., Benzitoun, C. et Le Mené, M., 2023. Création d'un référentiel lexical à partir des productions verbales d'enfants à développement typique et atypique en situation de jeu. 11e Journée de Linguistique de Corpus, Université de Grenoble, 3-7 Juillet 2023, Grenoble. [hal-04153646]

da Silva-Genest, C., Masson, C. et Le Mené Guigourès, M., 2023. Évaluer les compétences morphosyntaxiques et discursives d'enfants d'âge scolaire : analyse d'une situation de jeu libre. *Études de linguistique appliquée : revue de didactologie des langues-cultures et de lexiculurologie*, 2023, 210 (2), pp.179-195. [10.3917/ela.210.0053].

Westerveld, M. F., Gillon, G. T., & Miller, J. F. (2004). Spoken language samples of New Zealand children in conversation and narration. *Advances in Speech Language Pathology*, 6(4), 195–208. <https://doi.org/10.1080/14417040400010140>

Westerveld, M. F., & Vidler, K. (2016). Spoken language samples of Australian children in conversation, narration and exposition. *International journal of speech-language pathology*, 18(3), 288–298. <https://doi.org/10.3109/17549507.2016.1159332>

Witkowska, D., Lucas, L., Jelen, M. B., Kin, H., & Norbury, C. (2022). Development of complex syntax in the narratives of children with English as an Additional Language and their monolingual peers. *PsyArXiv*. <https://doi.org/10.31234/osf.io/96dx5>