

Télé-séminaire de recherche RIICLAS
23 mars 2026

**Accélérer la transposition des grammaires scientifiques
des langues océaniques en ressources pédagogiques**

Jacques Vernaudon & Sébastien Christian

Fare Pasifika – centre de recherche sur les Sociétés et les Humanités du Pacifique

(UAR 2503)



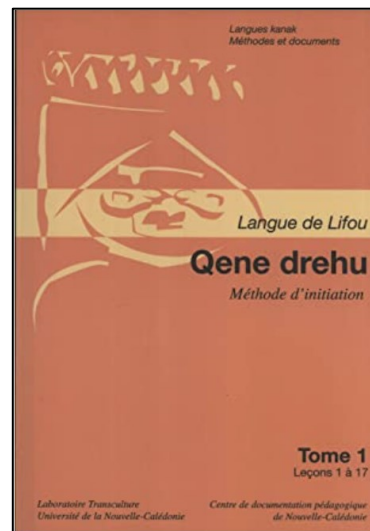
Expérience de la transposition didactique

1998



Paia, M. et Vernaudon, J. (1998). *Tahitien. Ia ora na, méthode d'initiation à la langue tahitienne*. Inalco & Bibliothèque publique d'information du Centre Georges Pompidou

2001



Lercari, C., Sam, L., Gowe, M. et Vernaudon, J. (2001). *Langue de Lifou, Qene drehu. Méthode d'initiation*. UNC & CDPNC

2013



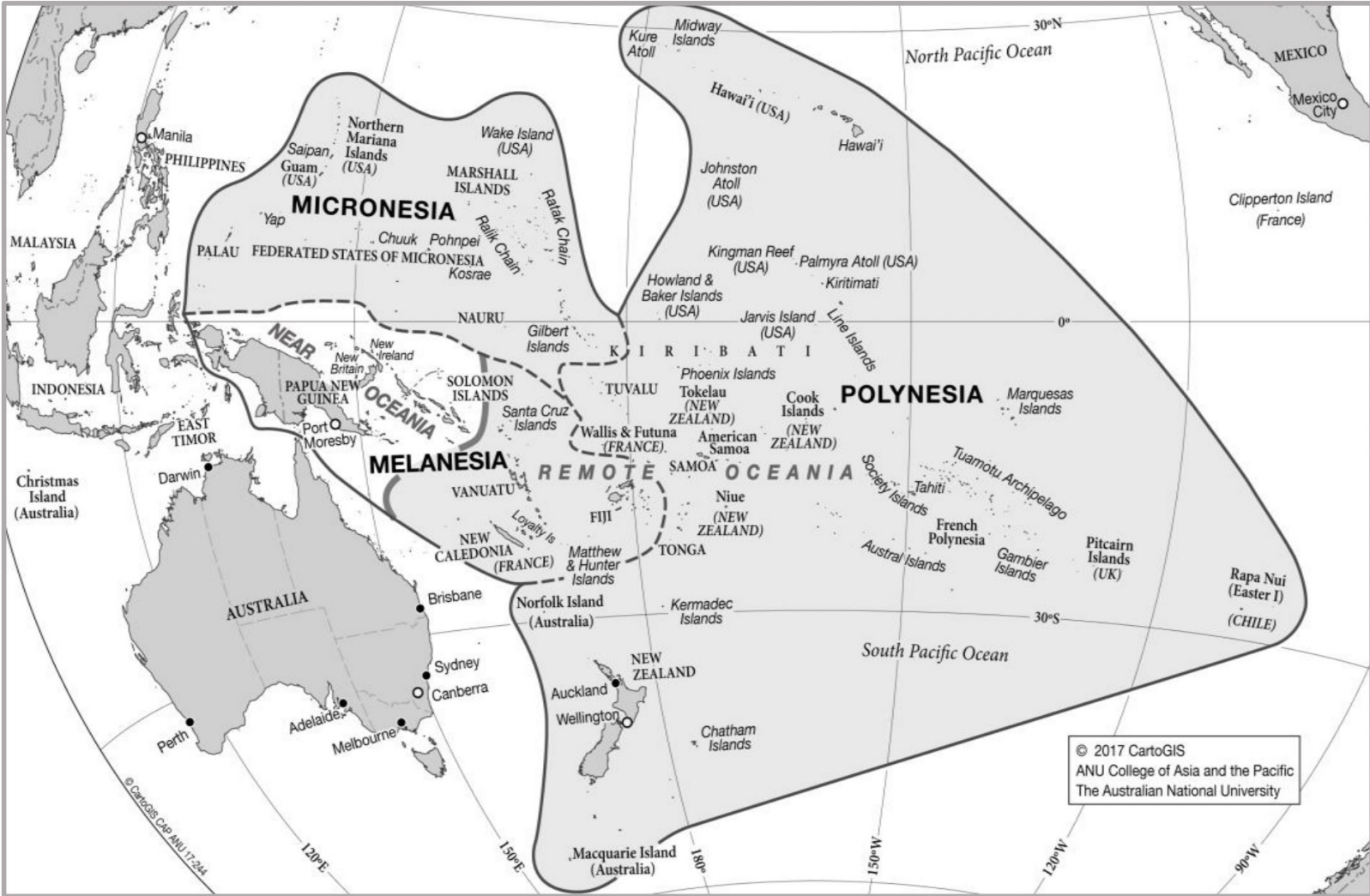
Vernaudon, J., Bearune, S., Hmae, J. et Mauberret, F. (2013). *Nengone*

2025



Vernaudon, J., Paia, M. (2025). *'Ia Reo!*. UPF

L'Océanie

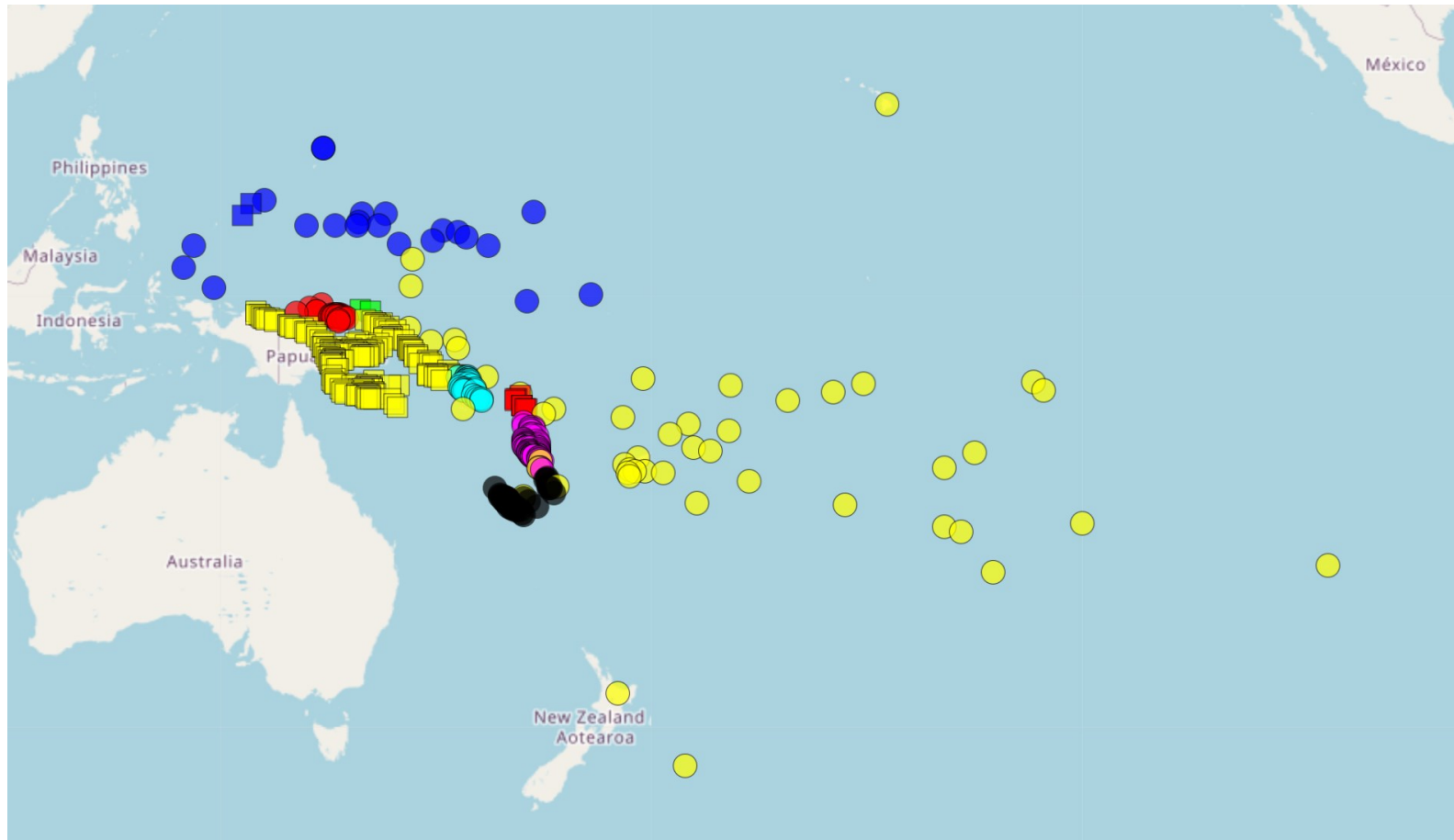


La diversité des langues d'Océanie

Régions		Familles linguistiques	# langues
Mélanésie Nouvelle-Guinée, Salomon	Océanie proche	"Papoue"	862
Salomon du sud-est, Vanuatu, Nouvelle-Calédonie, Fidji		Austronésienne	345
Polynésie	Océanie lointaine		212
Micronésie	38		
		21	

1478

Les langues océaniques (~ 600)



<https://glottolog.org/resource/languoid/id/ocea1241>

Politiques linguistiques et éducatives en Océanie



- L'école, produit de la **colonisation**
 - Les "langues indigènes" étaient **bannies** des classes
 - > 1980 : L'école s'ouvre aux langues locales

- enjeu **pédagogique** : Favoriser le développement personnel et la réussite scolaire
- enjeu **patrimoniale** : Valoriser et transmettre un patrimoine linguistique et culturel
- enjeu **socio-politique** : Favoriser la compréhension entre les groupes, en répondant à une demande de réparation des préjudices de la période coloniale

Grammaire explicite et apprentissage des langues

Grammaire implicite
et inductive

vs

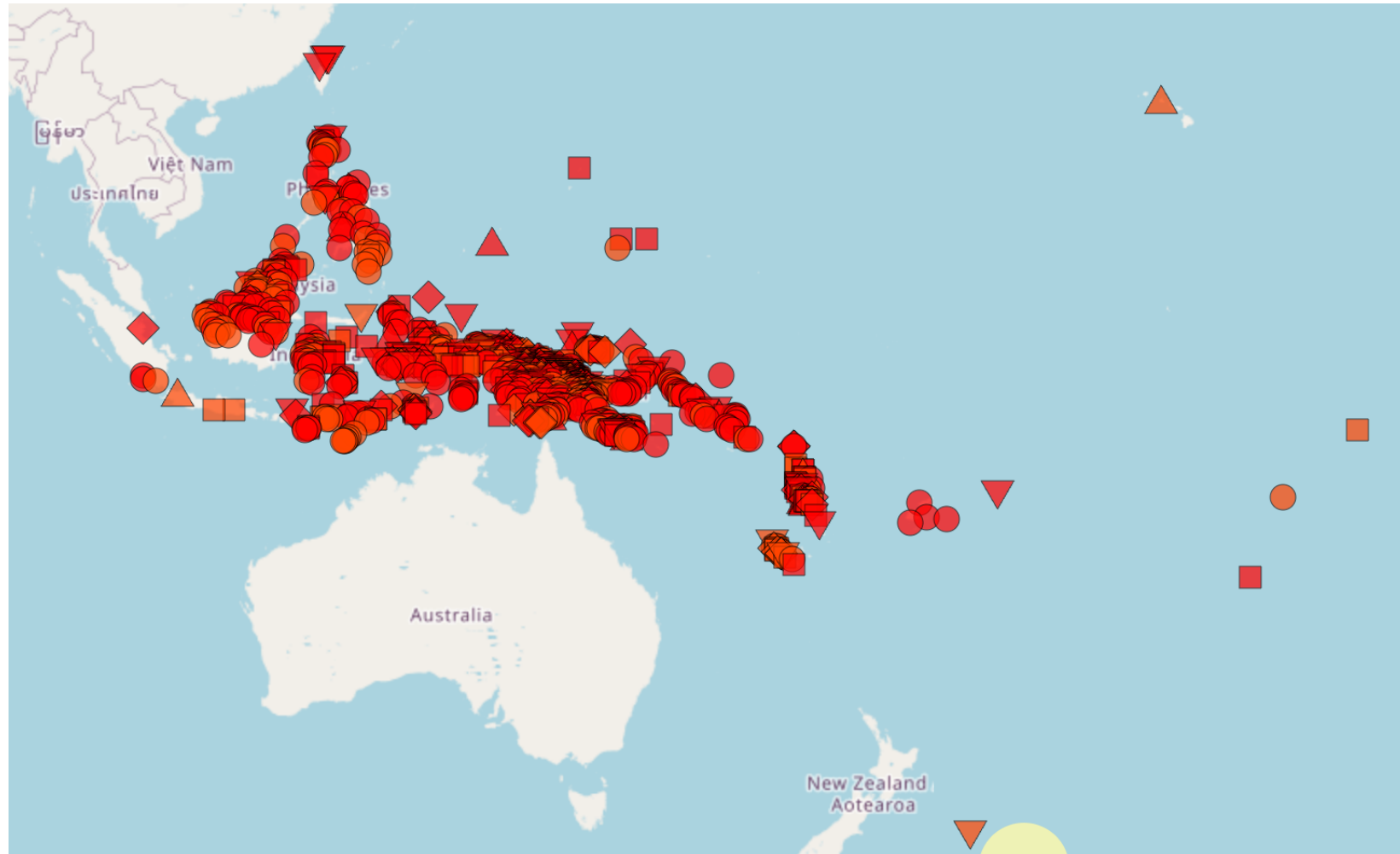
Grammaire explicite
(déductive ou inductive)

- Les apprenants qui suivent un enseignement formel atteignent généralement des niveaux de maîtrise de la L2 plus élevés que ceux qui suivent un enseignement non explicite, et progressent plus rapidement, à condition que la grammaire soit présentée dans un ordre pertinent (*inter alia* Purpura 2004 ; Ellis, 2006).
- Un enseignement qui articule la grammaire de la L2 à celle de la L1 est préférable (*inter alia* McManus & Marsden 2019)

La grammatisation

"Processus qui conduit à **décrire** et à **outiller** une langue sur la base des deux technologies qui sont encore aujourd'hui les piliers de notre savoir métalinguistique : la **grammaire** et le **dictionnaire**" (Auroux 1994: 109).

Langues de "Papunesia" sans description grammaticale



<https://glottolog.org/langdoc/status/browser?macroarea=Papunesia&focus=sdt#3/-8.23/188.70>

Situation descriptive des langues de Polynésie française

<i>Langue</i>	<i>Grammaire approfondie (> 300 pages)</i>	<i>Esquisse grammaticale</i>	<i>Dictionnaire</i>	<i>Liste de mots</i>
tahitien	oui	oui	oui	oui
marquisien	non	oui	oui	oui
pa'umotu	non	non	oui	oui
mangarévien	non	oui	oui	oui
rapa	non	oui	non	oui
rurutu	non	non	non	oui
ra'ivavae	non	non	non	oui

Le projet *ANAREO*



Délégation à la Recherche
de Polynésie française



Tearere Teikitohe, **Fatu Hiva**, 2024



Albert Hugues, **Rikitea**, 2024



Herenui Vanaa, **Reao**, 2024



Takurua Parent, **Rapa**, 2024

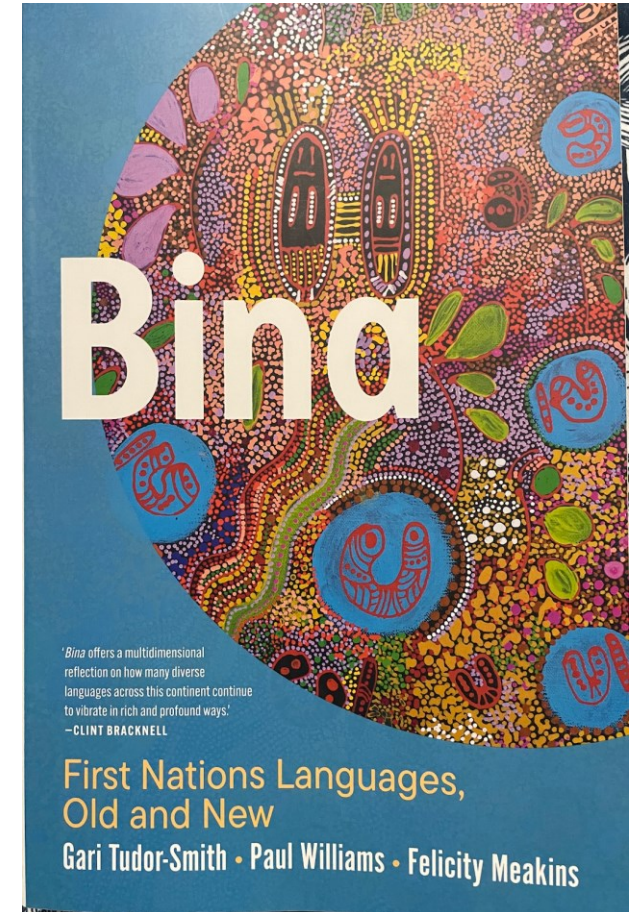
La grammatisation

- **QUI ?** Missionnaires, autodidactes passionnés, linguistes...
- **POURQUOI ?** Pour traduire la bible, pour enseigner une L2, pour enseigner une L1, pour comprendre le langage humain...
- **QUELLE FORME ?**
Livres, articles, mémoires de master/doctorat (publiés ou non)...
- **COMBIEN DE TEMPS ?**
3 à 8 ans pour une esquisse grammaticale, plusieurs dizaines d'années pour une grammaire exhaustive

Grammaires de linguistes et communautés

« Many Indigenous people wanting to learn (and learn about) their languages find grammars hard going. **Linguistic terminology is a major barrier.** [...] These technical descriptions of languages can frustrate Indigenous people who are already facing many barriers to accessing their languages. »

(Tudo-Smith, G., Williams, P., Meakins, F., 2024, *Bina, First Nations Languages, Old and New*, p. 57)



La transposition didactique

Yves Chevallard (1985)

Savoirs savants



Savoirs à enseigner



Savoirs enseignés

Grammaire de Jaussen (1861)

ARTICLE 4. — **Déclinaison.**

Noms communs.

SINGULIER	PLURIEL
NOM. <i>te ro</i> , la fourmi.	<i>te mau ro</i> , les fourmis.
GÉN. <i>no te ro</i> , de la fourmi.	<i>no te mau ro</i> , des fourmis.
DAT. <i>i te ro</i> , à la fourmi.	<i>i te mau ro</i> , aux fourmis.
ACC. <i>i te</i> (ou <i>te</i>) <i>ro</i> , la fourmi.	<i>i te</i> (ou <i>te</i>) <i>mau ro</i> , les fourmis.
VOC. <i>e te ro é</i> , ô fourmi!	<i>e te mau ro e</i> , ô fourmis!
ABL. <i>na te ro</i> , par la fourmi.	<i>na te mau ro</i> , par les fourmis.

Manuel scolaire de Vernier (1882)

LEÇON I

L'ARTICLE (1)

TE PARAU FAATAA (1)

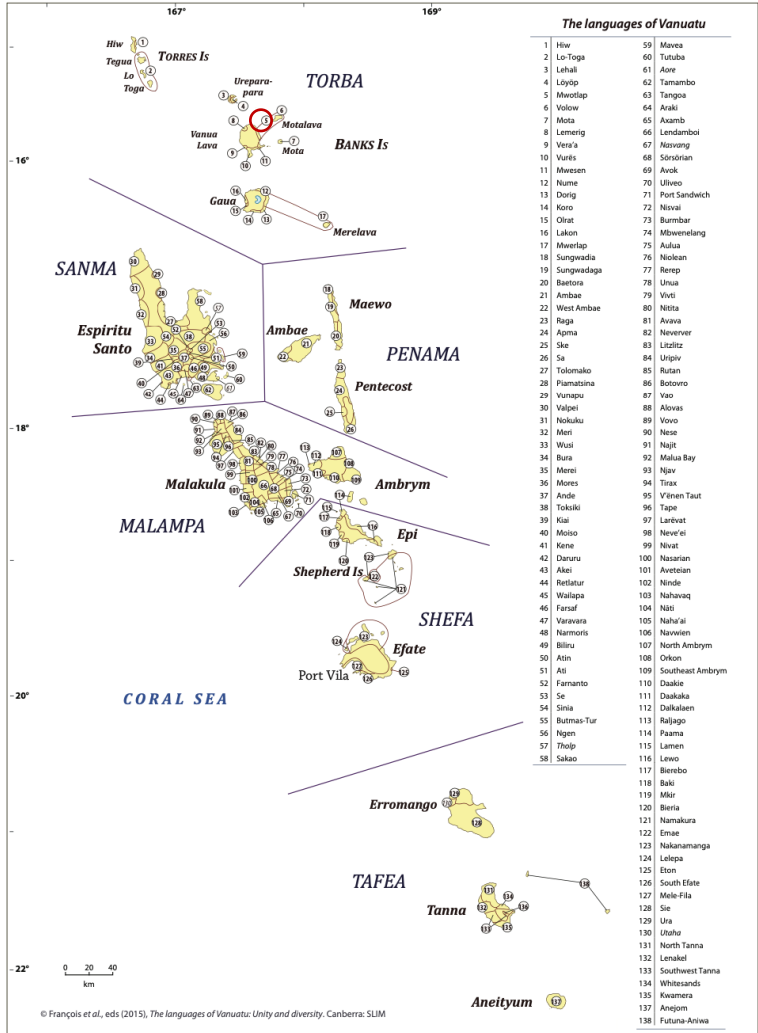
	SINGULIER
Nom. l'homme	<i>te taata</i>
Gén. de la femme	<i>no te vahine</i>
Dat. à l'enfant	<i>i te tama, tamarii</i>
Acc. le garçon	<i>i te tameroa, tamaiti</i>
Voc. ô fils	<i>e te tamaiti e</i>
Abl. par la fille	<i>na te tamahine</i>

Étapes de la transposition didactique

- Sélectionner les connaissances grammaticales à enseigner, les simplifier, les exemplifier
- Concevoir une progression pédagogique
- Traduire les explications dans la langue principale de l'école... et dans la langue océanienne
- Établir des ponts comparatifs entre la langue océanienne et la langue principale d'enseignement
- Concevoir des exercices d'application

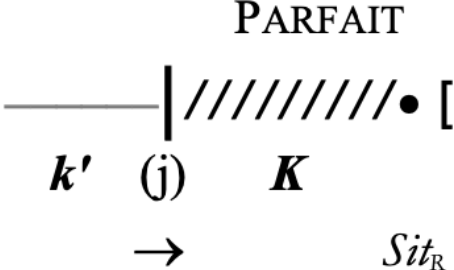
Simplifier, traduire

François, A. (2001). *Contraintes de structures et liberté dans l'organisation du discours. Une description du mwotlap, langue océanienne du Vanuatu*
Thèse de linguistique, Paris, Sorbonne, **1075 pages.**



PARFAIT – En admettant comme point de départ (temporel ou modal) une propriété k' , j'affirme que le basculement vers la propriété opposée k a effectivement eu lieu dans la réalité, de telle façon que cette propriété k caractérise désormais la situation de référence Sit_R .

Figure 7.3 – *Le Parfait*

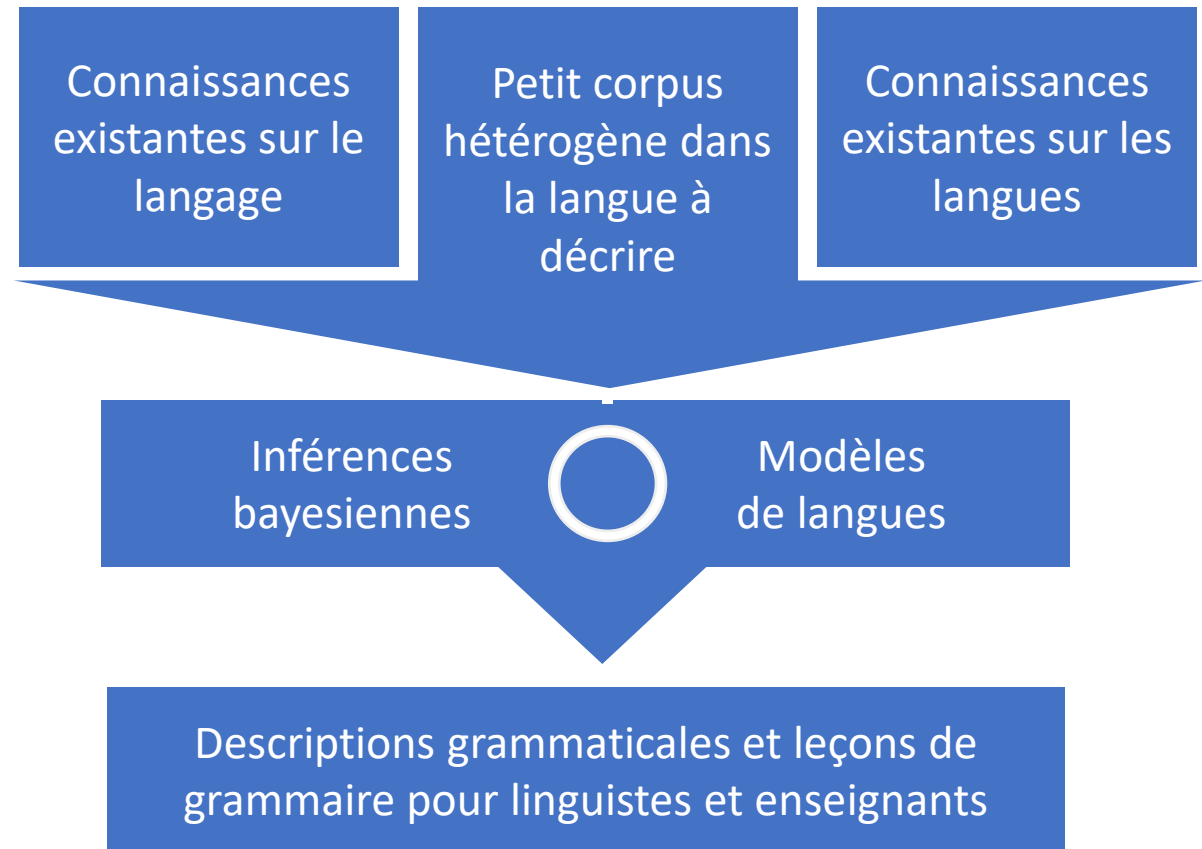


Accélérer le processus de transposition didactique



DIG4EL: Digital Inferential Grammars for Endangered Languages

- **Objectifs:** Accélérer la description de langues en danger et permettre la génération de matériel d'enseignement de leur grammaire.
- **Moyens:** Logiciel pour linguistes et grand public associant Mathématiques, Natural Language Processing (NLP) et Large Language Models (LLM).
- **Progrès:** Prototype fonctionnel en accès libre.



Principes généraux

- **DIG4EL** [“Digital Inferential Grammar for Endangered Languages”], permet d’agréger des informations éparses et hétérogènes sur une langue pour générer automatiquement des documents accélérant la description grammaticale.
- L’outil consolide différentes sources documentaires : paires de phrases, connaissances typologiques générales sur les langues du monde, et publications scientifiques disponibles.
- Il s’appuie ensuite sur une méthode hybride associant modélisation mathématique, inférence bayésienne et raisonnement par grands modèles de langage (*Large Language Model*, LLM). Le système génère des descriptions grammaticales éditables à destination de linguistes ou d’enseignants pour revue et utilisation, accélérant leur travail de documentation ou de création de leçons de grammaire.
- Sur certains sujets fondamentaux (ex. constituants de la phrase canonique, structure d’actance, négation), des trames grammaticales guident le logiciel dans la production des explications.
- L’outil a été prétesté sur 6 langues typologiquement différentes, dont 5 d’Océanie, et pour lesquelles des experts ont pu évaluer la pertinence des explications générées : iaai (Nouvelle-Calédonie), nafsán et mwotlap (Vanuatu), ngen (Côte d’Ivoire), uruanguirin (Papouasie-Nouvelle-Guinée) et tahitien (Polynésie française).

DiG4EL: Digital Inferential Grammars for Endangered Languages

- Descriptions et leçons dans la langue choisie par l'utilisateur.
- Adaptation au niveau d'expertise de l'utilisateur et des destinataires du document.
- Aucune intrusion ou pollution par des exemples ou connaissances externes.
- Documents à disposition d'experts de la langue, locuteurs, enseignants ou linguistes.

Dire “non” en tahitien: 'aita, e'ita, e'ere, 'eiaha et 'ore

En grammaire, dire “non” peut toucher des choses différentes: dire qu’une action n’a pas eu lieu, qu’une chose n’existe pas, que quelqu’un n’est pas quelque chose, ou encore interdire. Dans les traductions françaises du corpus, on voit des formes comme “ne ... pas”, “il n’y a pas”, “ne ... plus”, “ne ... jamais”. En tahitien, le mot négatif choisi dépend du type de phrase: actions, existence/possession, identité/qualité, interdiction, etc. Très souvent, le mot négatif vient au début et porte l’idée principale de la phrase. Les marqueurs essentiels qui reviennent dans vos documents sont: 'aita (négation factuelle/existentielle, et réponse “non”), e'ita (négation du futur, de l’irréel, de la capacité), e'ere (négation d’identité/attribut), 'eiaha (interdiction), et 'ore (surtout en relatives et dans certains souhaits). Des petits mots comme a'e nei (“jamais jusque maintenant”) et fa'ahou (“plus, plus jamais”) modifient le sens du temps et de la fréquence sous la négation.

'aita: dire qu’un fait n’a pas eu lieu; dire “il n’y a pas”; possession et lieu

'aita sert à nier des faits (souvent passés/réels), l’existence, la possession, ou le lieu, et peut être employé seul pour répondre “non”. Schémas utiles observés: 'aita + SUJET + i + VERBE; 'aita e + NOM; 'aita tā/tō + POSSESSIF + NOM; 'aita i/tō + LIEU. Avec a'e nei: “jamais (jusqu’à maintenant)”; avec fa'ahou: “plus / plus jamais”.

'aita vau i fa'aro'o a'e

Je ne savais pas.

Génération d’une leçon sur la négation en tahitien, 19 novembre 2025.

Risques et stratégies d'atténuation

1. Le système peut générer des affirmations non fondées (hallucinations) et des omissions. Si ces problèmes ne sont pas détectés lors de la révision humaine, ils risquent de se retrouver dans les supports pédagogiques.
 - > La principale mesure d'atténuation consiste à informer les utilisateurs, tout au long de leur utilisation, que les résultats peuvent contenir des erreurs.
2. Les résultats peuvent sous-représenter la variation linguistique naturelle et ne pas tenir compte de la diversité dialectale.
 - > La principale mesure d'atténuation consiste à informer et sensibiliser les utilisateurs à propos de ce risque.

Une visite de la plateforme

dig4el.org



Workshop DIG4EL, Port-Vila, Vanuatu, 28 octobre 2025

Indications bibliographiques

- Auroux, S. (1994). *La révolution de la grammatisation*. Bruxelles : Mardaga.
- Besse, H., Porquier, R. (1991). *Grammaires et didactique des langues*. Paris : Hatier/Didier.
- Chevallard, Y. (1985). *La Transposition didactique: du savoir savant au savoir enseigné*. La Pensée Sauvage.
- Christian, S., (2025). Enhancing grammatical documentation for endangered languages with graph-based meaning representation and Loopy Belief Propagation. *Natural Language Processing Journal*, Vol. 12, <https://doi.org/10.1016/j.nlp.2025.100164>
- François, A. (2001). *Contraintes de structures et liberté dans l'organisation du discours. Une description du mwotlap, langue océanienne du Vanuatu*, Thèse de linguistique, Paris, Sorbonne, 1075 p.
- François, A., et al., eds (2015). *The languages of Vanuatu: Unity and diversity*. Canberra: SLIM
- Monneret, P., Poli, F., 2021, *La grammaire du français. Terminologie grammaticale*, Direction générale de l'enseignement scolaire
- Nocus, I., Vernaudon, J., Paia, M. (dir.) (2014) *L'école plurilingue en outre-mer : Apprendre plusieurs langues, plusieurs langues pour apprendre*, Rennes, Presses universitaires de Rennes.
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 Instruction: A Research Synthesis and Quantitative Meta-analysis. *Language Learning*, 50(3), 417–528.
- Vargas, C. (2009). Peut-on inventer une grammaire pour la réussite scolaire ? *Repères*, 39, 17-39.
- Vernaudon, J. (2024), La métalangue scolaire à l'épreuve des faits de langues en contexte bilingue, *Contextes et didactiques*, 24. <https://doi.org/10.4000/1396p>
- Vernaudon, J., François, A. (à paraître). Two centuries of describing Tahitian: From Latin-based to emic categories. Dans R. Rodríguez, O. Zwartjes (dir.), *The Routledge Handbook of the History of Descriptive Linguistics*.