HAL
open science

Mise en forme de corpus de langues tai (tai paw, tai yo, tai don) : alignement temporel entre manuscrit, enregistrement audio et annotation

Matthew Deo, Séverine Guillaume, Thi Hang Dinh, Michel Ferlus, Tri Doi Tran, Alexis Michaud

▶ **To cite this version:**

Matthew Deo, Séverine Guillaume, Thi Hang Dinh, Michel Ferlus, Tri Doi Tran, et al.. Mise en forme de corpus de langues tai (tai paw, tai yo, tai don) : alignement temporel entre manuscrit, enregistrement audio et annotation. Journée d'étude CORLI : Traitements et standardisation des corpus multimodaux et web 2.0., May 2018, Paris, France. hal-01873903

HAL Id: hal-01873903

https://hal.science/hal-01873903

Submitted on 6 Oct 2018

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.

Mise en forme de corpus de langues tai (tai paw, tai yo, tai don) : **alignement temporel entre manuscrit, enregistrement audio et annotation**

Matthew DEO, Séverine GUILLAUME, Thi Hang DINH, Michel FERLUS, Tri Doi TRAN, Alexis MICHAUD

Objectifs

- Afficher des manuscrits synchronisés avec un enregistrement audio et un appareil philologique (transcription, traduction, gloses et commentaires)
- Ouvrir les documents numérisés à de nouveaux modes d'exploration
- Ouvrir de nouvelles possibilités en traitement d'image et traitement automatique des langues

Les écritures

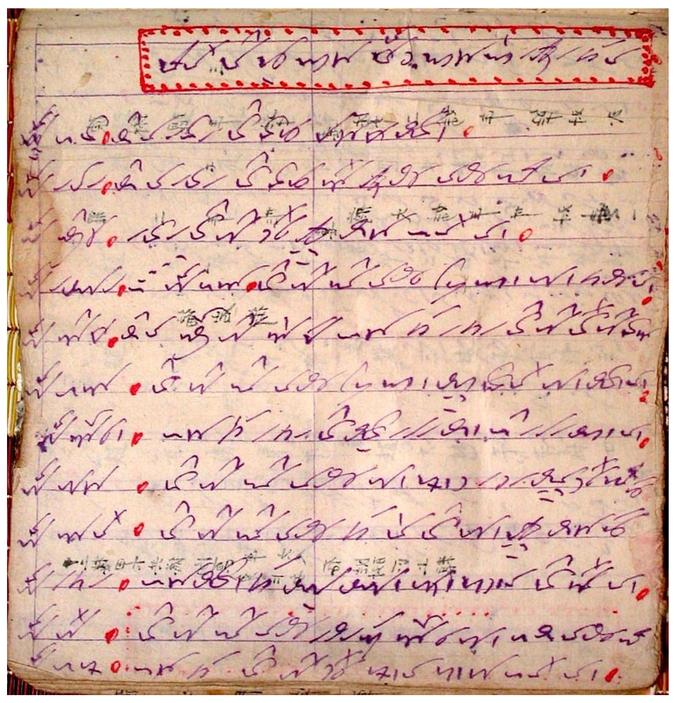
Ecritures tai (thai) : d'origine indienne ; au Vietnam : écritures conservatrices, dont le **lai pao** et l'écriture tai deng.
Transmission quasi-interrompue.

Les lectures

Lecture des manuscrits par des lettrés tai.
Enregistrements de Michel Ferlus, spécialiste des langues et écritures d'Asie du Sud-Est

Les annotations

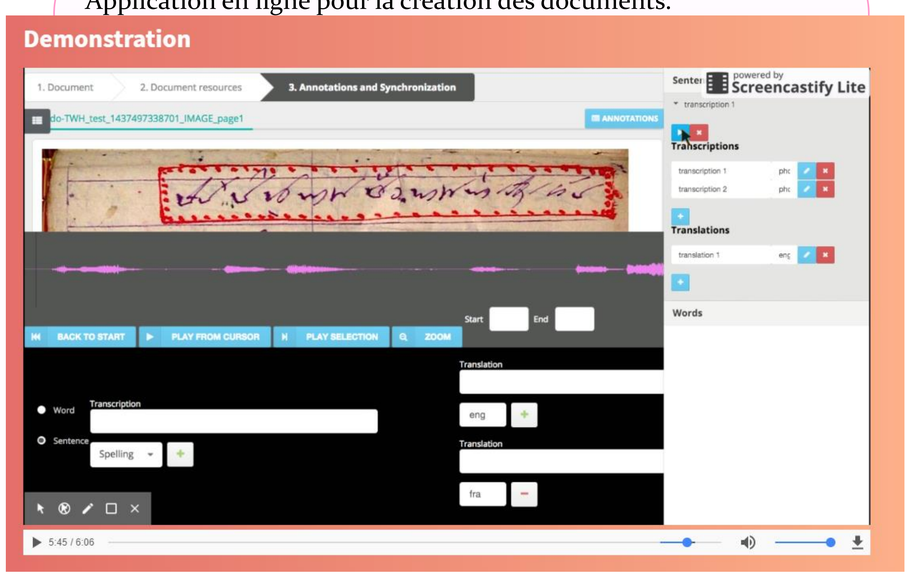
Travail de Michel Ferlus, en collaboration avec Trần Trí Dõi.
Mise en forme par Đinh Thị Hằng avec l'aide de Matthew Deo



Architecture logicielle :

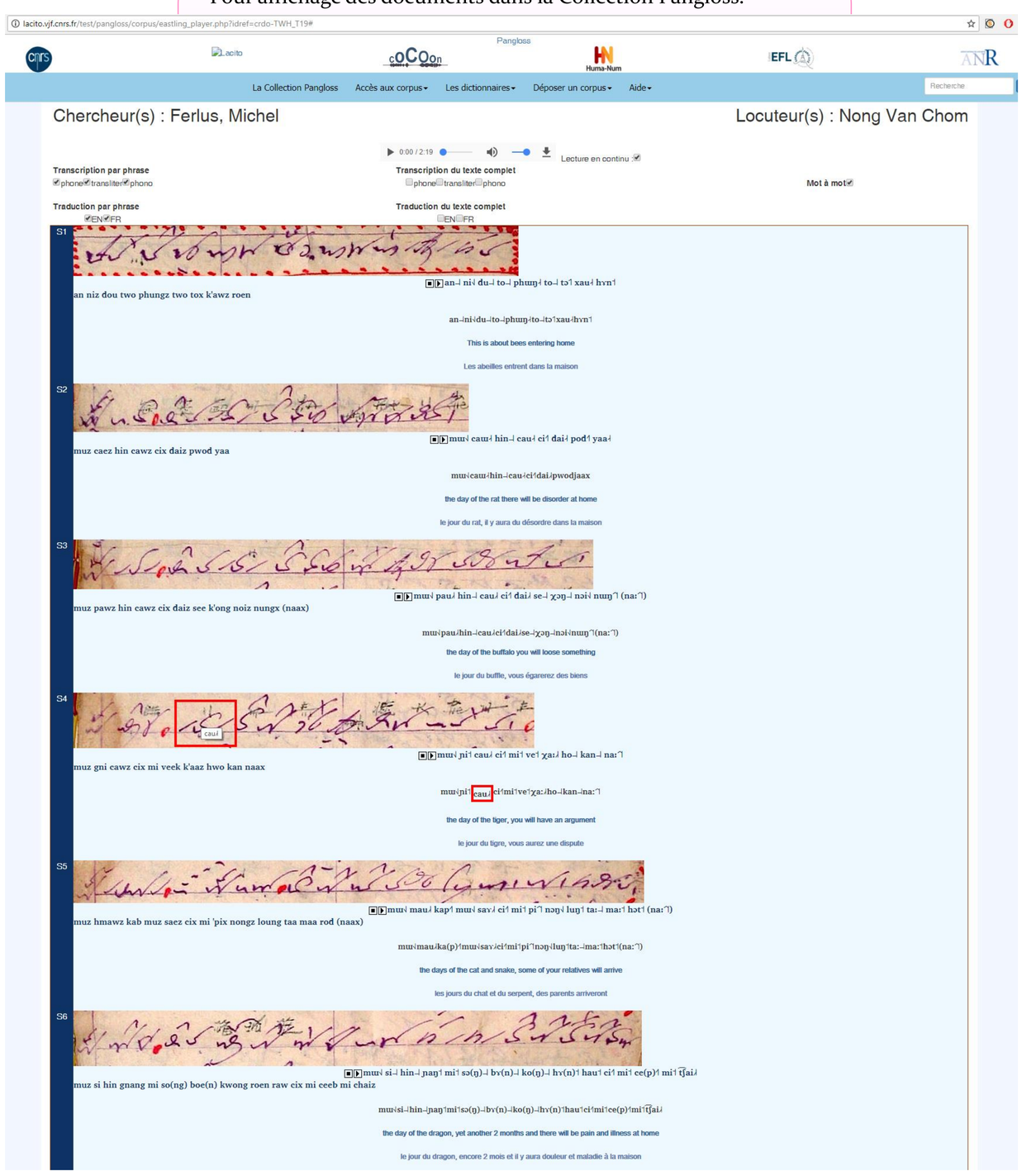
Eastling - éditeur

Application en ligne pour la création des documents.



Eastling - lecteur

Pour affichage des documents dans la Collection Pangloss.



Documents en libre accès, en ligne

Hébergement dans la Collection Pangloss
http://lacito.vjf.cnrs.fr/pangloss/

Collection Pangloss

*La Collection Pangloss fait partie de l'Archive ouverte **Cocoon**, Collection de Corpus Oraux Numériques*
Contact: pangloss.cnrs@gmail.com

Code source disponible

https://github.com/CNRS/Pangloss/tree/master/eastling
Ouverture du repositoire au public prévue à l'automne 2018
D'ici là : adresser vos demandes à matthewdeo@gmail.com

Démo en ligne : http://lacito.vjf.cnrs.fr/test/pangloss/tools/eastling/demo/
Site de Matthew Deo : http://moon-light.fr/manuscrit/les-abeilles-entrent-dans-la-maison/

Pour en savoir plus

Ferlus, Michel. 1999. Sur l'ancienneté des écritures thai d'origine indo-khmère. Bangkok. https://halshs.archives-ouvertes.fr/halshs-00922719/
Ferlus, Michel; Trần, Trí Dõi. Présentation de l'écriture Lai Pao des Thai de Tương Dương (Nghệ An). https://halshs.archives-ouvertes.fr/halshs-00923025/