



HAL
open science

Formalisation de la temporalité en vietnamien pour la traduction et veille stratégique

Nicolas Boffo

► **To cite this version:**

Nicolas Boffo. Formalisation de la temporalité en vietnamien pour la traduction et veille stratégique. Journées 2013 des doctorants du laboratoire PRAXILING, Dec 2013, Montpellier, France. 2013. hal-04374839

HAL Id: hal-04374839

<https://hal.science/hal-04374839v1>

Submitted on 5 Jan 2024

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.

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/375030522>

BOFFO N., 2009, Formalisation de la temporalité en vietnamien pour la traduction automatique, MICA/PRAIXILING

Poster · January 2009

CITATIONS

0

1 author:



Boffo Nicolas

Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier (LIRMM)

12 PUBLICATIONS 3 CITATIONS

SEE PROFILE



Formalisation de la temporalité en vietnamien pour la traduction automatique



PRAXILING

Directeurs de thèse : Jacques BRES (PRAXILING) - Eric CASTELLI (MICA)

MICA

Sujet de thèse

Notre projet de recherche concerne la formalisation du temps et de l'aspect en vietnamien avec une mise en place d'étapes (lexico-syntaxique, sémantique, formalisation du temps et de l'aspect : étape pragmatique/praxématique) dans le traitement automatique de la temporalité en vue de la traduction automatique.

Le traitement automatique de la temporalité en vietnamien en est à ses débuts. De ce fait, notre tâche s'avère complexe puisque les outils de TAL sont rares et peu performants, et les ressources lexicales et sémantiques manquent de précision. Par contre nous avons une bibliographie linguistique assez complète qui traite de la temporalité en vietnamien.

Pour surmonter les difficultés du traitement temporel et avancer dans notre tâche nous avons été amenés à développer certaines étapes du traitement simultanément.

Enfin nous envisageons de créer un outil de calcul et de traduction automatique des relations temporelles qui aura une importance décisive pour le traitement automatique de la langue vietnamienne (TA, TAO, résumé automatique, compréhension automatique, recherche d'entités nommées,...).

Mots clés:

Formalisation, temporalité, temps et aspect, annotation temporelle, Traitement Automatique des Langues, traduction automatique, langue vietnamienne.

Travaux de la première année

Création d'un dictionnaire vietnamien/français des marqueurs temporels (MT)

1. Recherche des marqueurs

- Recherche sur l'ensemble de la bibliographie linguistique vietnamienne et française.
- Recherche avec le système de traduction automatique (utilisation de « phrase-table » et des scores de probabilités obtenus).
- Recherche d'équivalences à partir des marqueurs temporels français.
- Recherche à partir des dictionnaires vietnamien/français, français/vietnamien, vietnamien/vietnamien.

• Traduction des marqueurs temporels vietnamiens

- Pour 1 entrée nous proposons différentes traductions en français de ces marqueurs.

Ex: {bao_giò} <quand / à quel moment> <depuis longtemps / depuis on ne sait quand > <jamais> <

Dictionnaire disponible à la rubrique « documents » sur : <http://recherche.univ-montp3.fr/praxiling/spip.php?article218>

• Classification par catégories lexicales

- Première classification effectuée avec l'aide de la linguiste

Classification disponible à la rubrique « documents » sur : <http://recherche.univ-montp3.fr/praxiling/spip.php?article218>

Etiquetage des marqueurs temporels

1. Etiquetage en valeur de Temps et d'Aspect

Exemple:

Hôm qua khi tôi **đang xem** phim thì **điện thoại reo**.

Etiquetage avec le vnTagger de Lê Hồng Phương :

```

<w pos="N">Hôm qua</w> <w pos="N">khi</w>
<w pos="P">tôi</w> <w pos="R">đang</w>
<w pos="V">xem</w> <w pos="N">phim</w>
<w pos="C">thì</w> <w pos="V">điện</w>
<w pos="N">thoại</w> <w pos="V">reo</w>
<w pos=".">.</w>

```

Introduction d'étiquettes temporelles:

```

<w pos="N"> <T="P"> <A=""> Hôm qua</w>...
<w pos="R"> <T="N"> <A="INC", "C"> đang</w>...

```

Etiquettes:

<T="P"> temps Passé

<A="INC", "C"> aspect Inaccompli et Continue

Les procès **xem** et **reo** sont localisés dans le passé par le *circostant temporel* **Hôm qua** à valeur temporelle « passé ».

Le marqueur d'aspect **đang** exprime que que **xem** est en train de se dérouler lorsque le procès **reo** survient.

Règle:

Si MT<T="P"> + MT<A="INC", "C"> Alors le verbe qui suit sera à l'imparfait

Travaux en cours

- Mise à jour du dictionnaire des marqueurs temporels vietnamien/français.
- Poursuite de l'étiquetage Aspectuo-temporel.
- Traitement de la temporalité en vietnamien: SdT de Gosselin et/ou S-Langages de Schwer? (avec utilisation du VnTagger et Vntokenizer de Lê Hồng Phương) + création de règles de temporalité.
- Développement d'un module « traitement de temporalité » (entités nommées temporelles) pour NOOJ en vietnamien.

