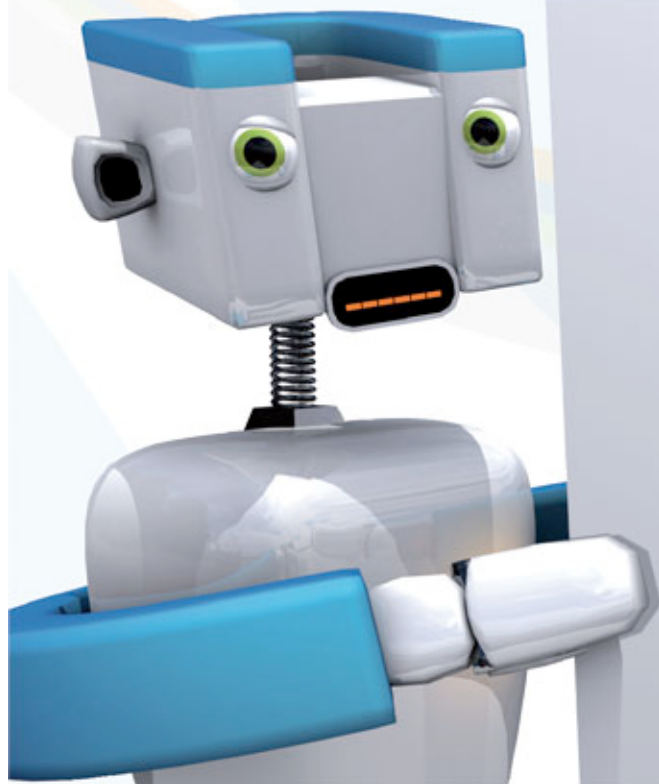


LE COLLOQUE



DU NUMÉRIQUE

État des lieux de la Traduction Automatique



Espace
Numérique
Entreprises

AR D I
AR D
NUMÉRIQUE





Hervé Blanchon & Christian Boitet

LIG – GETALP

Herve.Blanchon@imag.fr, Christian.Boitet@imag.fr





Plan

- Introduction
- Axes d'analyse des systèmes de TA
 - Architecture opérationnelle
 - Architecture linguistique
 - Architecture computationnelle
- Bilan sur les systèmes existants
- Perspectives pour les utilisateurs
- Perspectives de recherche
- Conclusion
- Références



Introduction

- La TAO (Traduction Automatisée par Ordinateur)
 - n'est pas une science
 - ni une collection de recettes
- C'est
 - une technologie scientifiquec'est-à-dire
 - un ensemble de méthodes qui progressent par
 - intégration d'idées nouvelles
 - amélioration incrémentale des savoir-faire



Introduction

Les types de traduction

- **Re-création**
 - Traduire, adapter de la poésie, de la publicité
- **Localisation**
 - Adapter un contenu à un environnement culturel donné (*site Web, logiciel*)
- **Diffusion**
 - Rendre strictement le contenu d'un document, sans ajout ni omission, même si le style sent la traduction
- **Suivi** (traduction *rapide*)
 - *Traduction dépistage* de textes écrits (*Google*)
 - Interprétation simultanée



Introduction

- Un problème **très difficile** par nature
 - **Ambiguïté** inhérente de la langue à tous niveaux
 - **Flou** : une langue naturelle ne peut pas être complètement axiomatisée (modélisée précisément)
 - toute description formelle d'une langue *complète* est sous-génératrice et sur-génératrice
 - seuls des sous-langages « petits » ou contrôlés peuvent être modélisés (axiomes + règles) avec exactitude
 - **Taille**
 - millions de concepts/termes
 - variété des phénomènes grammaticaux



Introduction

- **Traduction Humaine et Traduction Automatique** sont différentes
 - La TA n'est pas là pour remplacer les humains
 - les voitures ne sont pas des chevaux !
 - La TA n'imité pas les humains
 - les avions ne battent pas des ailes !



Introduction

- **Limite inhérente**
au problème de l'automatisation

$$\text{Couverture} \times \text{Automatisation} \times \text{Qualité} \ll 100\%$$

- l'amélioration de l'un des critères se fait forcément au détriment des deux autres
 - ↗ couverture $\Rightarrow$ ↘ automatisation et/ou qualité
 - ↗ automatisation $\Rightarrow$ ↘ couverture et/ou qualité
 - ↗ qualité $\Rightarrow$ ↘ couverture et/ou automatisation



Introduction

- La TA peut servir à
 - produire de la Traduction de Haute Qualité (diffusion)
 - depuis une langue source
 - vers une (plusieurs) langue(s) cible(s)
 - comprendre (suivi)
 - une langue étrangère inconnue ou peu connue
 - permettre une inter-communication bilingue (suivi)
- Objectifs oubliés lors de l'évaluation des systèmes de TA



Architecture Opérationnelle

- Conditions de développement et d'usage d'un système
 - Tâches et utilisateurs *(à suivre)*
 - Paire(s) de langues, volumes, genres *(à suivre)*
 - Possibilité(s) d'intervention humaine *(à suivre)*
 - Ressources disponibles
 - Données
 - gros corpus parallèles (annotés ou non) pour la TA empirique
 - Humaines
 - linguistes computationnels, lexicographes pour la TA experte



AO : Tâches et utilisateurs

- Aider des utilisateurs bilingues à produire de bonnes traductions
 - diffusion
- Aider des utilisateurs à comprendre une langue non (ou peu) connue
 - suivi (traduction dépistage, acquisition)
- Aider la communication interpersonnelle (tchat, traduction de parole)

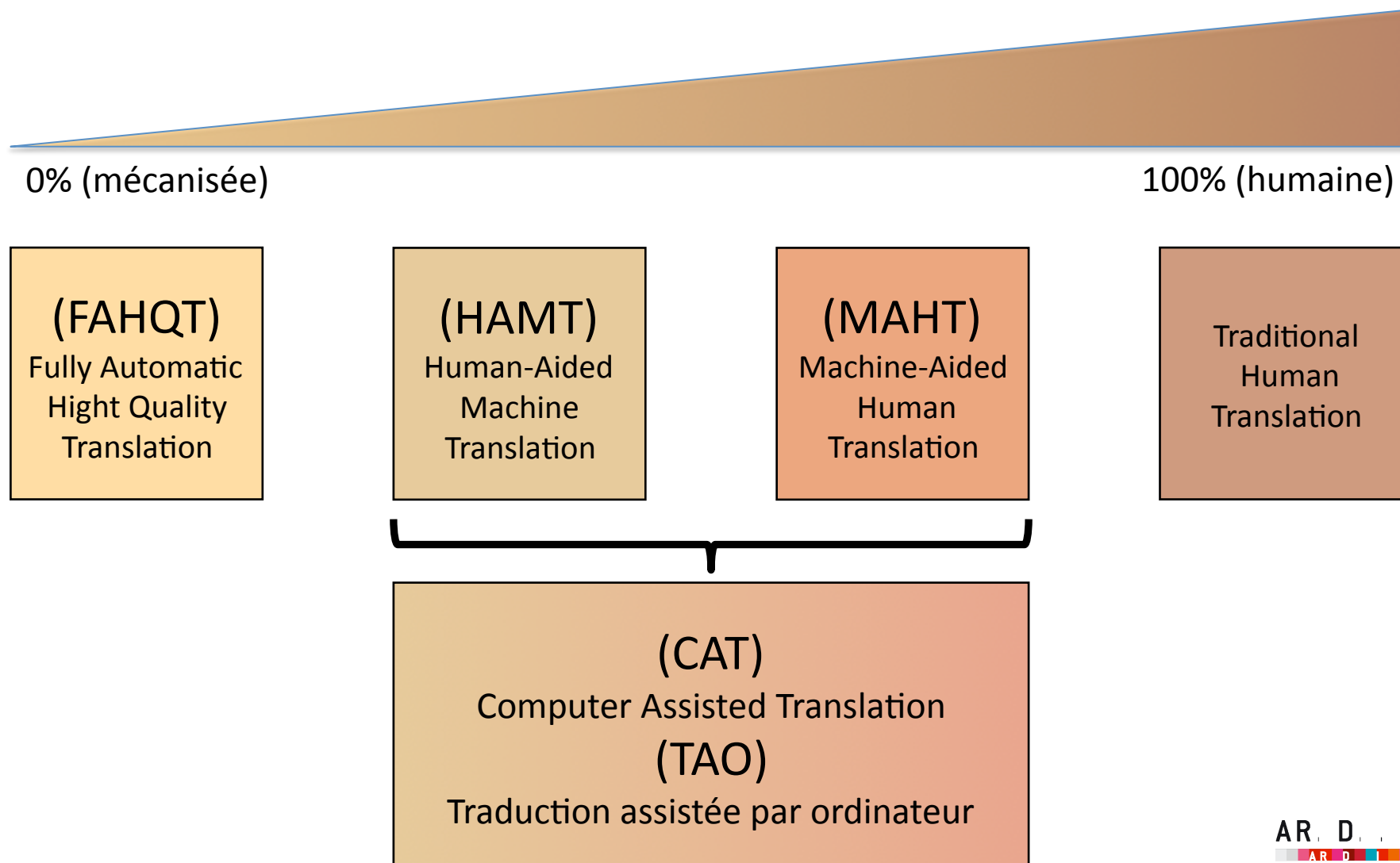


AO : Paire(s) de langues, volumes, genres

- Chemins traductionnels ou directions
 - $1 \rightarrow 1$:
 - ALT/JE (NTT)
 - $1 \leftrightarrow 1$:
 - ATLAS-II (fujitsu) : dicos 29 M mots + 5,57 M mots
 - Converser for Healthcare (Spoken Translation)
 - $1 \rightarrow N$:
 - MedSLT
 - $1 \leftrightarrow N$:
 - Phraselator (Armée US)
 - $N \leftrightarrow N$:
 - pour des débats, tchat multilingue, forces multinationales de maintien de la paix



AO : Niveau d'implication humaine





FAHQT (Fully Automatic High Quality Translation)

- Possible dans des cadres restreints
 - Il faut que le contexte soit très bien défini (petit sous-langage, petit vocabulaire)
 - Ce sont normalement des systèmes autonomes



FAHQT (Fully Automatic High Quality Translation)

- Exemples

- MÉTÉO

- Opérationnel depuis mai 1977
 - Traduction de bulletins météorologiques au Canada
 - Anglais ↔ Français ; 60000 mots/jours ; 3% de corrections (20M mots/an E → F, 10 M mots/an F → E)

- ALTFLASH

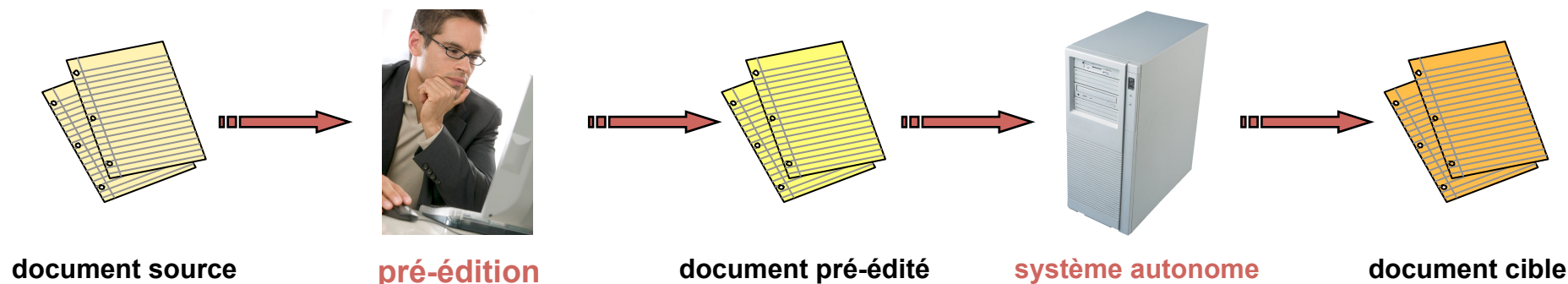
- Opérationnel depuis 2001
 - Traduction de brèves de la bourse Nikkei
 - Japonais → Anglais ; résultats bruts sans révision humaine, car qualité « pro » moins cruciale que pour la météo



HAMT (Human-Aided Machine Translation)

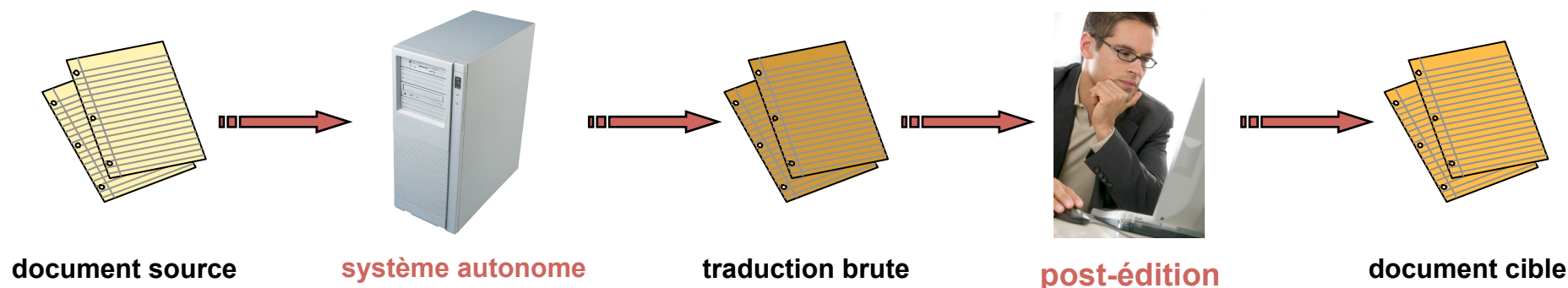
- L'utilisateur assiste le processus de traduction
 - Il faut que le contexte soit bien défini
 - Ce sont des systèmes interactifs
- Trois approches possibles
 - Pré-édition
 - Post-édition
 - Interaction

HAMT avec pré-édition



- **Pré-édition** : pour quoi faire ?
 - Rechercher les problèmes que va rencontrer le système et les supprimer
 - Identification des noms propres, des entités nommées, des hors-texte (formules...)
 - Choix des catégories pour les homographies
 - Indication des clauses enchâssées
 - Parenthésage des structures coordonnées
 - Reformuler le texte dans un langage contrôlé (extrême)

HAMT avec post-édition



- **Post-édition** : pour quoi faire ?
 - Corriger la traduction produite par la machine en fonction de la qualité à obtenir
 - Très peu s'il s'agit de recherche d'informations
 - Beaucoup et par un spécialiste familier du domaine s'il s'agit de textes qui doivent être publiés et diffusés
- Presque toutes les solutions opérationnelles

HAMT avec interaction



- **Interaction** pour quoi faire ?
 - Aider le système au cours du processus de traduction
- **Interaction** pour qui ?
 - **Traducteur** : interaction pour spécialiste
 - **Auteur monolingue** : interaction pour non-spécialiste
 - Interaction durant la rédaction
 - Interaction durant la traduction



MAHT (Machine Aided Human Translation)

- L'utilisateur est assisté dans le processus de traduction par des outils informatiques
 - Correcteurs
 - Orthographiques, grammaticaux, stylistiques
 - Dictionnaires
 - Monolingues, bilingues, & encyclopédiques
 - Concordanciers
- Idée : Intégrer les services dans un
Poste de Travail du Traducteur



Architectures linguistiques

- Trois types d'architectures
 - Traduction directe
 - Traduction par analyse, transfert, génération
 - Traduction par pivot interlingue (assez abstrait)



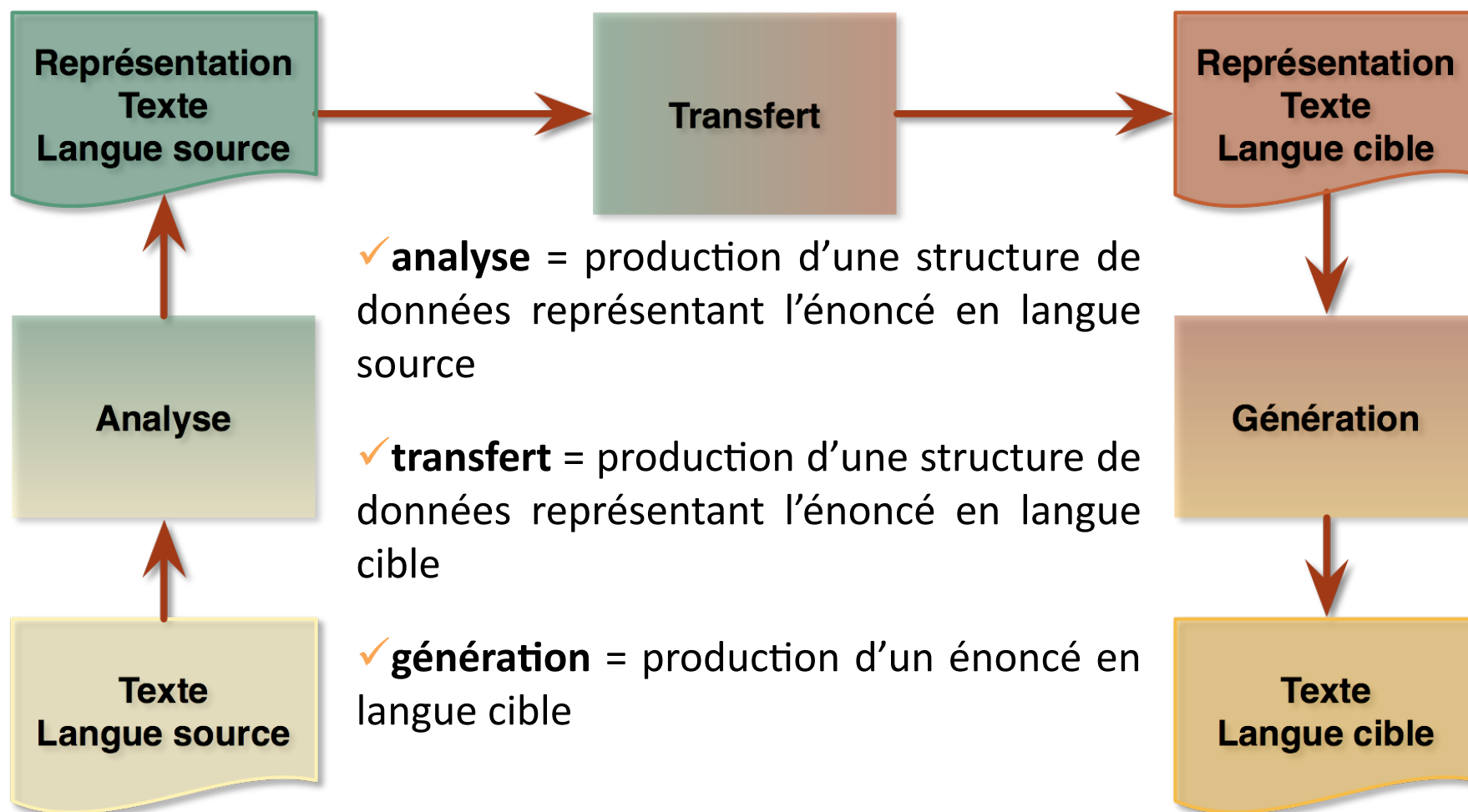
Traduction directe

- 1 système = 1 boîte noire (1 étape)
 - succession de traitements de la langue source vers la langue cible sans production de structure intermédiaire



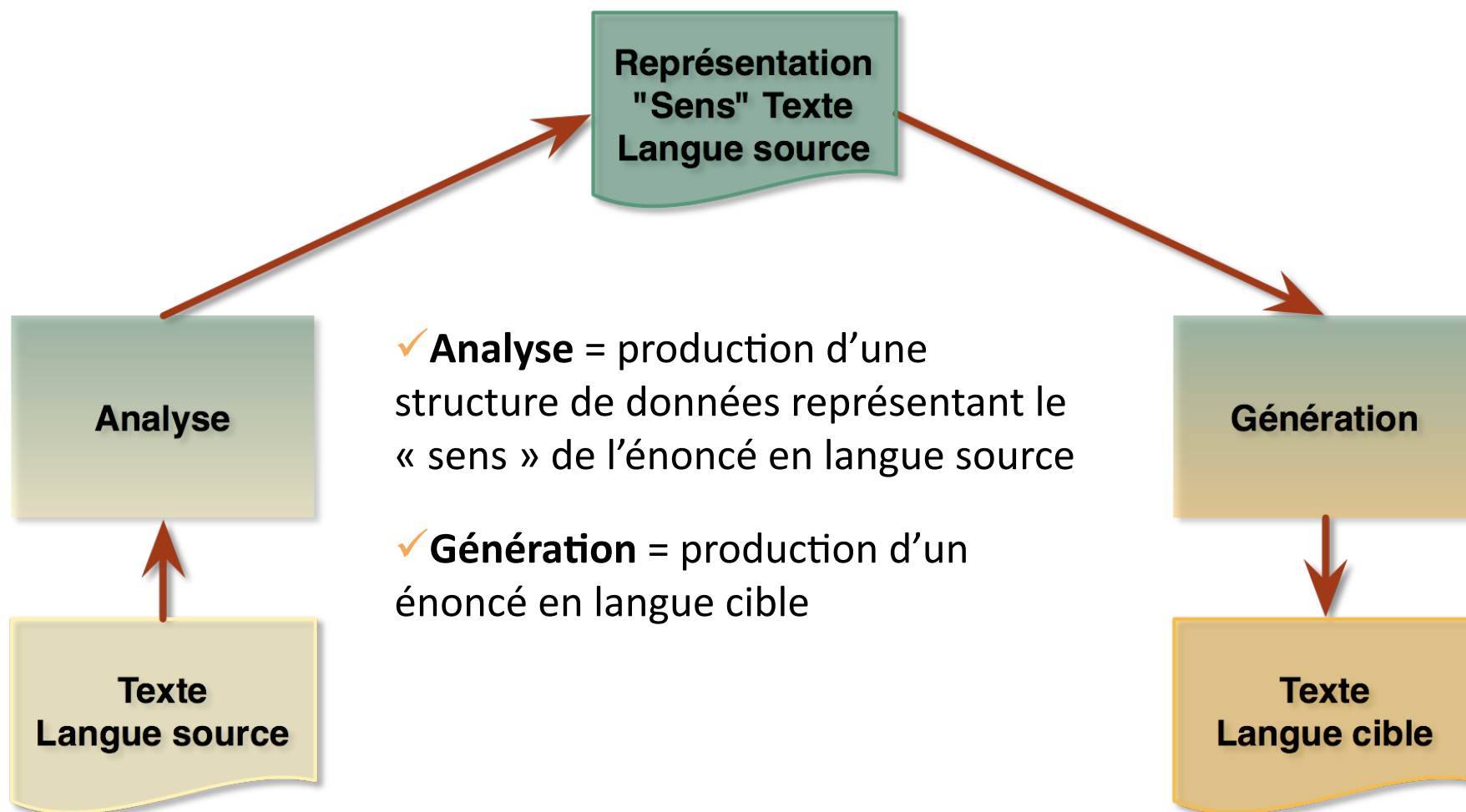
Analyse, transfert & génération

- 1 système = 3 boîtes noires (3 étapes)

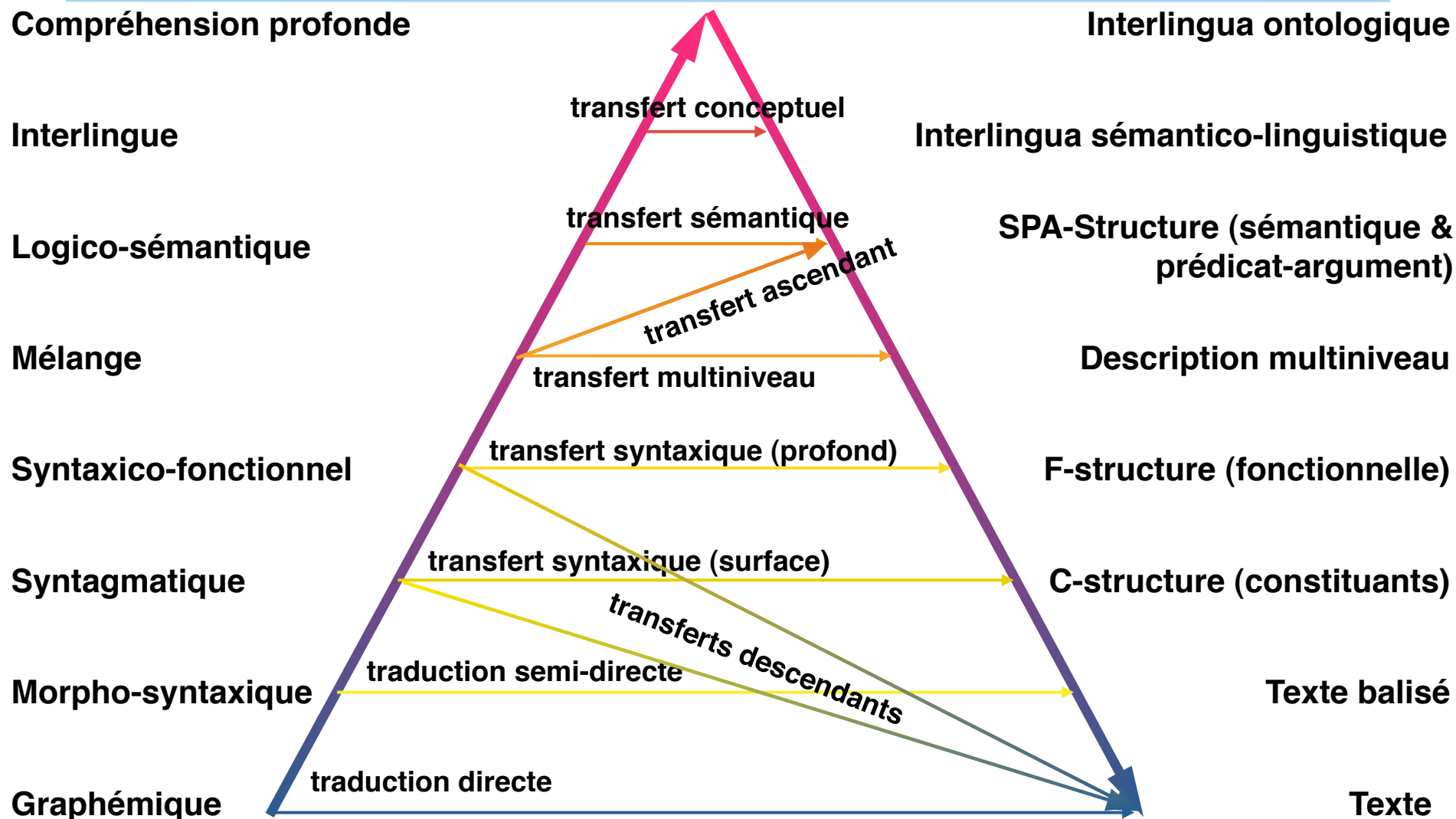


Traduction par pivot

- 1 système = 2 boîtes noires (2 étapes)

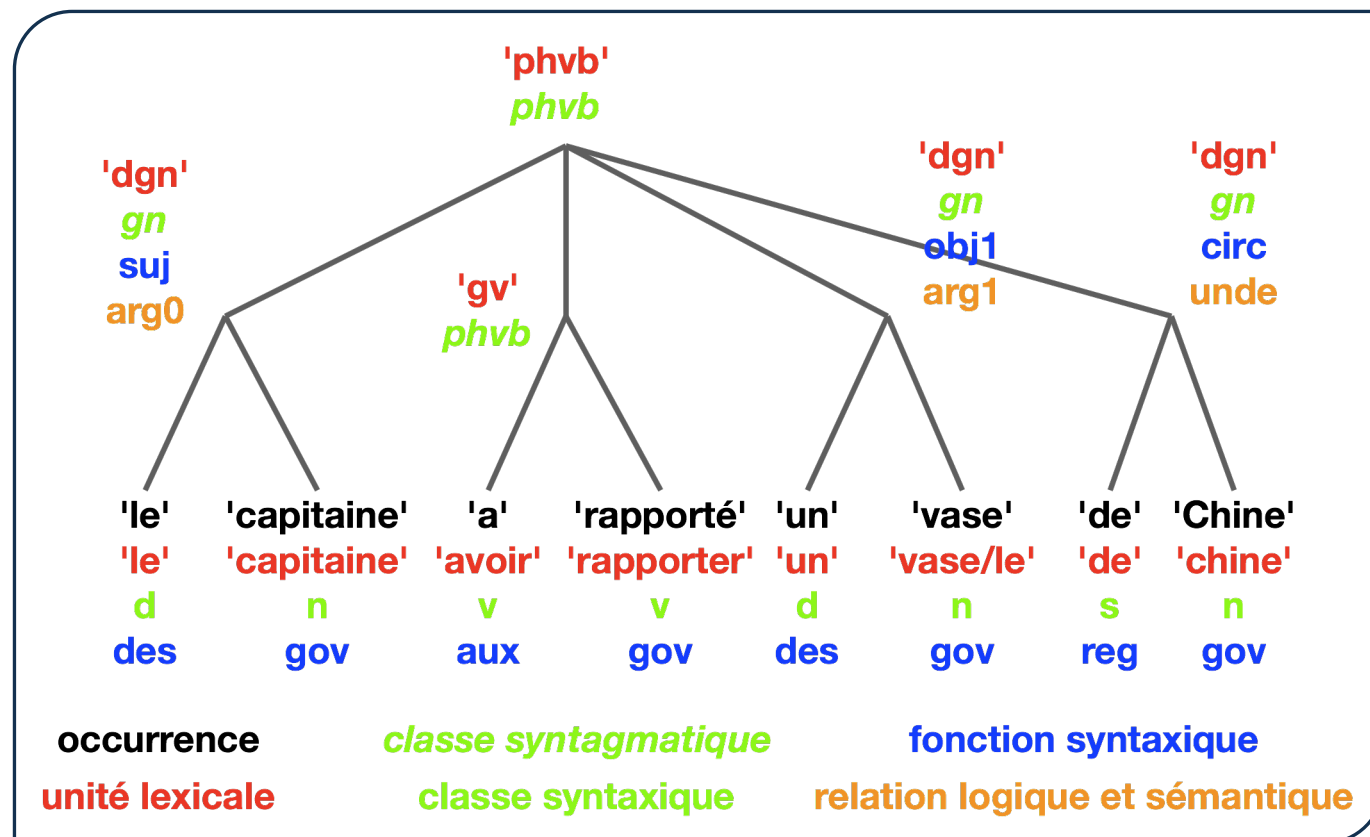


Architectures linguistiques (triangle de Vauquois)



Exemples de structure (1/3)

- Structure arborescente multiniveau
 - Le capitaine a rapporté un vase de Chine

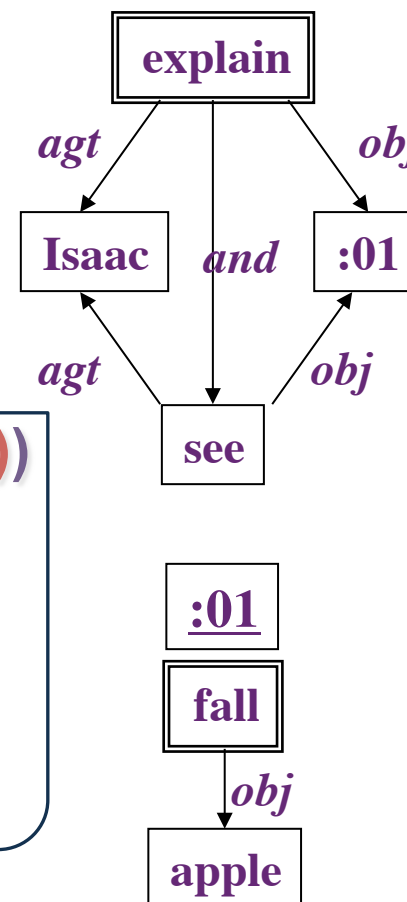


Exemples de structure (2/3)

- Interlingua sémantico-linguistique
 - UNL

Isaac sees that an apple falls and he explains it.

```
agt(explain(icl>do).@entry, Isaac(icl>proper noun))
obj(explain(icl>do).@entry, :01)
obj:01(fall(icl>occur).@entry, apple.@undef)
and(explain(icl>do).@entry, see(icl>do))
agt(see(icl>do), Isaac(icl>proper noun))
obj(see(icl>do), :01)
```



UNL (hyper) graph



Exemples de structure (3/3)

- Interlingua sémantico-pragmatique (préontologique)

<client> et je voudrais une chambre du dix au quinze septembre à trento

```
{c:give-information+disposition+room (  
  conjunction=discourse,  
  disposition=(desire, who=i),  
  room-spec=(identifiability=no, quantity=1, bedroom, location=name-trento),  
  time=(start-time=(md=10), end-time=(md=15, month=9)))}
```

- IF de Nespole!
- Acte de parole : give-information
- Concepts : disposition, room
- Acte de dialogue : give-information+disposition+room
- Arguments :
 - disposition=(desire, who=i)
 - room-spec=(identifiability=no, quantity=1, bedroom, location=name-trento)
 - time=(start-time=(md=10), end-time=(md=15, month=9))



Architecture computationnelle

- Il s'agit ici de la méthode de calcul de chacune des phases constituant les étapes ...
 - analyse: graphémique, morphologique, syntaxique
- ... cette dimension est dans les faits orthogonale à l'architecture linguistique
 - on trouve toutes les combinaisons [Boitet TALN08]
- Trois types d'approches
 - Approches expertes
 - Approches empiriques
 - Approches hybrides



Approches expertes

- Approche directe, procédurale
 - « Artisanat linguiciel »
 - Grammaires, automates et dictionnaires « codés » directement dans un langage de programmation
- Approche à base de règles
 - Inspirée par la linguistique formelle et la compilation
 - Grammaires formelles, automates et dictionnaires
 - +
 - Moteur

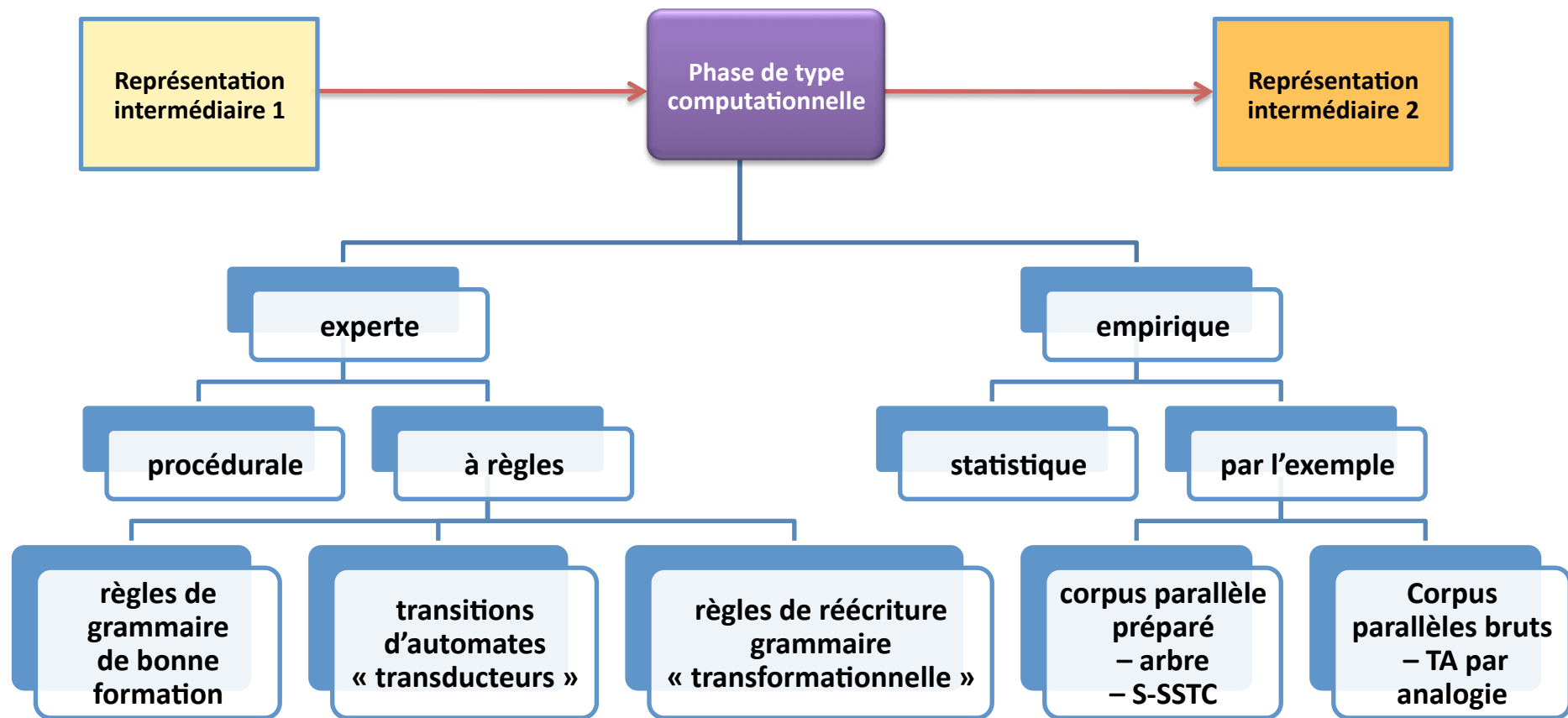


Approches empiriques

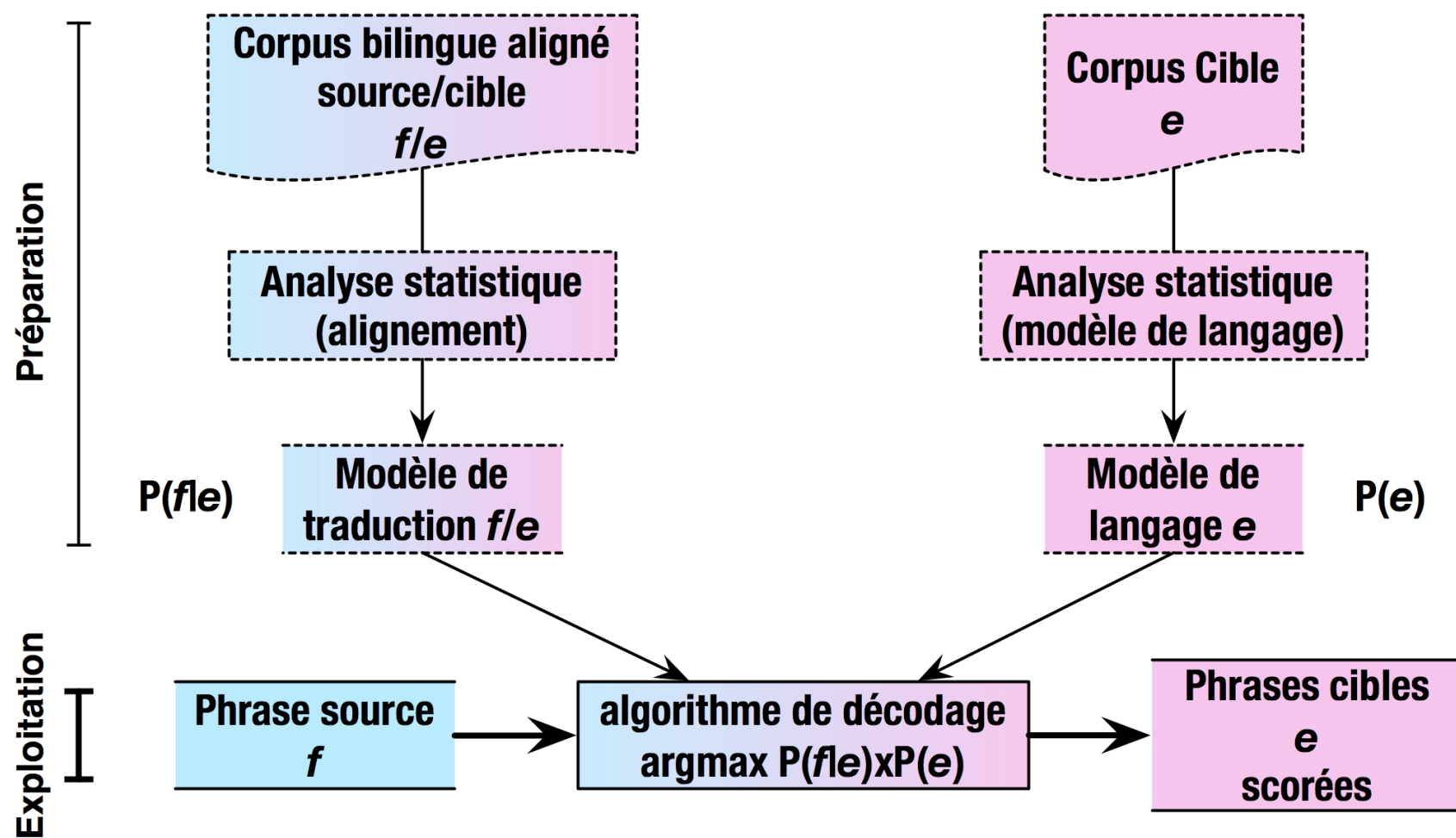
- **Apprentissage**
 - supervisé ou non
 - de *correspondances* à partir de données plus ou moins annotées
 - texte → texte
 - texte → structure
 - structure → structure
 - structure → texte

Architecture computationnelle

- Approches computationnelles possibles



Exemple : TA empirique directe





Bilan sur les systèmes existants

- **Nombreux systèmes disponibles**
 - sur le Web
 - intégrables aux éditeurs de textes les plus courants
- **Retour d'expériences industrielles assez difficiles à trouver**
 - cf. MT Summit (*user track*)
 - cf. LISA (*Localization Industry Standards Association*)
- **TA utilisée pour la diffusion avec des systèmes personnalisés + post-édition**
 - Expériences chez Symantec, Adobe, Microsoft, ...
 - Gains de productivité observés très divers: 10 à 40 %



Perspectives pour les utilisateurs

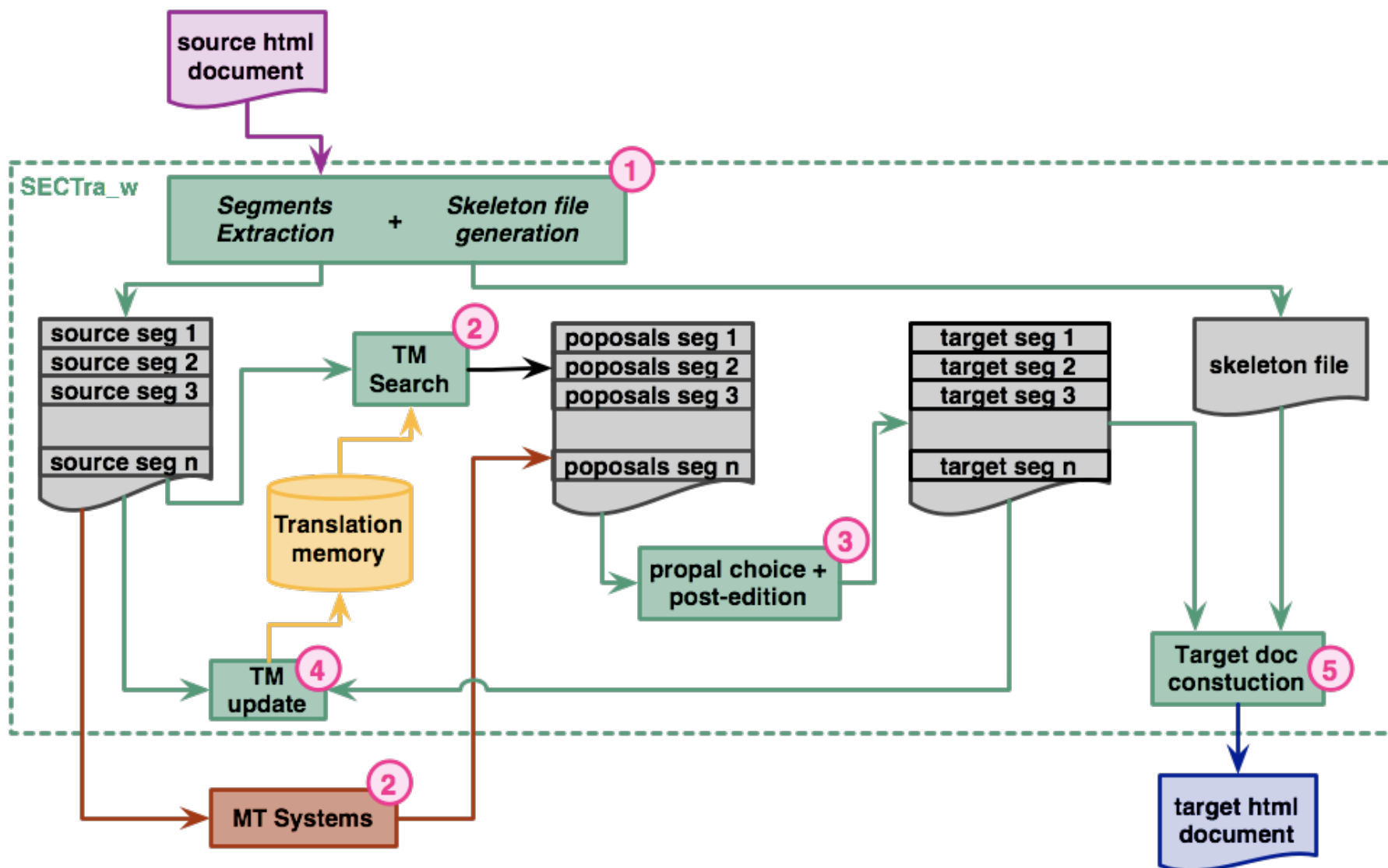
- Construction coopérative de ressources lexicales et terminologiques
 - « Jeux sérieux » : jeudemots (<http://www.lirmm.fr/jeuxdemots/>)
 - communauté d'intérêts : grandes causes (PaxHumana), centres d'intérêts (Wikipedia), logiciel libre (W3C, Mozilla)
- Construction de systèmes pour paires de langues peu dotées
 - expérience en vietnamien-français au GETALP
- Environnements coopératifs de traduction, post-édition & évaluation
 - accessible via le web
 - permettant une évaluation fondée sur la tâche
- Passerelles contributives d'accès multilingue
 - avec mémoire de traductions, dictionnaires et éventuellement systèmes de TA dédiés



Environnements coopératifs de traduction, post-édition, évaluation

- SECTra_w
 - « Système d'Exploitation Contributive de Corpus de Traductions sur le Web »
- Objectifs
 - Évaluation de systèmes de TA
 - Production de traductions de qualité via post-édition (collaborative)

SECTra_w (architecture)





SECTra_w (TA et mémoire de traductions)

- Système(s) de TA
 - utiliser n'importe quel(s) système(s) de TA disponible(s)
- Mémoire de traductions
 - Appariement parfait seulement, car
 - la TA est plus utile que l'appariement flou
 - l'appariement flou ne permet pas de gagner plus de 50% du temps par rapport à une traduction à partir de zéro
 - la post-édition de sorties Systran permet de gagner au moins 50% du temps d'édition (55% en général, 65% pour l'un des post-éditeurs de l'expérience EOLSS/UNL)
 - cf. aussi le site de Jeffrey Allen
 - <http://www.geocities.com/mtpostediting/>



SECTra_w (EOLSS/UNL)

- Online Encyclopaedia of Life Support Systems
 - <http://www.eolss.net/>
- Quelques chiffres
 - 6 600 articles en anglais
 - 20 domaines
 - 1 article
 - $\approx$ 30 pages standard (250 mots/page)
 - $\approx$ 300-400 phrases $\approx$ 7500 mots
 - total
 - 250 000 pages
 - 62,5 M mots



SECTra_w (expérience EOLSS/UNL)

- Traduire en français ...
 - 25 articles $\approx$ 880 pages $\approx$ 220 K mots
- Exemples de titres d'articles
 - Ethics and science
 - Ground and soil water characteristics
 - Human interaction with land and water: a hydrologist's conception
 - Tsunamis: causes, consequences, prediction and response
 - Composition and structure of the atmosphere
 - Water as a factor in socio-economic development: future trends



SECTra_w (expérience EOLSS/UNL)

- Traduire en français ...
- ... avec un mélange de ...
 - ... techniques automatiques ...
 - TA
 - Recherche dans des mémoires de traductions
 - ... travail contributif
 - post-édition effectuée en ligne

SECTra_w (démonstration EOLSS/UNL)

- Segments

Source (english)	Postedit (french) (515/515=100.0 %)	Suggestions	Page 3
The wave propagates from the source with a velocity of long gravity water waves in accordance with the equation	L'onde se propage de la source avec une vitesse de longues vagues de gravité selon l'équation	<< Reverso The vague propage de la source avec une vélocité de longues vagues de l'eau de la gravité conformément à l'équation << Systran La vague propage de la source avec une vitesse de longues vagues d'eau de pesanteur selon l'équation	
where g is the acceleration due to gravity, and H is the depth of the basin.	★★★★☆ Done by Margot Bergerand in 106 s. 10 Dans laquelle g est l'accélération due à la gravité, et H est la profondeur du bassin.	<< Reverso where g est l'accélération dû à gravité et l'H est la profondeur de la cuvette. << Systran là où g est l'accélération due à la pesanteur, et H est la profondeur du bassin.	
Because the average depth of the world ocean is 4 km, the typical velocity of tsunami in the ocean is 200 m s ⁻¹ or 720 km h ⁻¹ .	★★★★☆ Done by Margot Bergerand in 27 s. 10 Puisque la profondeur moyenne de ces océans du monde est de 4 kilomètres, la vitesse typique d'un tsunami dans l'océan est de 200 m s⁻¹ ou de 720 kilomètres h⁻¹.	<< Reverso Because la profondeur moyenne de l'océan mondial est 4 km, la vélocité typique de tsunami dans l'océan est s de 200 m < soupez > -1 < / soupez > ou h de 720 kms < soupez > -1 < / soupez >. << Systran Puisque la profondeur moyenne de l'océan du monde est de 4 kilomètres, la vitesse typique du tsunami dans l'océan est de 200 m s⁻¹ ou de 720 kilomètres h⁻¹.	
	★★★★☆ Done by Margot Bergerand in 25 s. 10		

SECTra_w (démonstration EOLSS/UNL)

- Segments

The screenshot displays the SECTra_w interface with three main panels: Source (english), Postedit (french), and Suggestions. The Source panel contains the text: "The wave propagates from the source with a velocity of long gravity water waves in accordance with the equation". The Postedit panel shows the French translation: "L'onde se propage de la source avec une vitesse de longues vagues de gravité selon l'équation". The Suggestions panel offers two options: "<< Reverso" and "<< Systran".

Below the main panels, a segment summary is shown with the following details:

- Source segment**: where g is the acceleration due to gravity, and H is the depth of the basin.
- MT pretranslations**: Because the average depth of the ocean is 4 km, the typical velocity of tsunami in the ocean is 200 m s⁻¹ or 720 km h⁻¹.
- Post-editing field**: Dans laquelle g est l'accélération due à la gravité, et H est la profondeur du bassin. Puisque la profondeur moyenne de l'océan mondial est 4 km, la vitesse typique d'un tsunami dans l'océan est de 200 m s⁻¹ ou de 720 kilomètres h⁻¹.
- Quality level of the translation (stars)**: 4 stars.
- Name of the post-editor + time spent**: Done by Margot Bergerand in 106 s.
- Score of the translation (number)**: 10.

The interface also includes a "Suggestions" panel with buttons for "<< Reverso" and "<< Systran", and a "Page" indicator showing "Page 3".

SECTra_w (démonstration EOLSS/UNL)

- Trace de la post-édition
 - texte inséré
 - ~~texte effacé~~

Postedit (french) (515/515=100.0 %)	Suggestions	Page 3
L'onde se La vague propage de la source avec une vitesse de longues vagues d'eau de pesanteur gravité selon l'équation	<< Reverso The vague propage de la source avec une vélocité de longues vagues de l'eau de la gravité conformément à l'équation << Systran La vague propage de la source avec une vitesse de longues vagues d'eau de pesanteur selon l'équation	
★★★★☆ Done by Margot Bergerand in 106 s. 10		
Dans laquelle là où g est l'accélération due à la pesanteur, gravité, et H est la profondeur du bassin.	<< Reverso where g est l'accélération dû à gravité et l'H est la profondeur de la cuvette. << Systran là où g est l'accélération due à la pesanteur, et H est la profondeur du bassin.	

SECTra_w- System of corpus management and exploitation - ViewDocument

http://eolss.imag.fr/xwiki/bin/view/Corpus/ViewDocument

es plus visités ▾ IGC CNRS SECTra_w GETALP CiteULike Intranet du LIG Liste IUT

SECTra_w- System of corpus m... SECTra_w- System of corpus m... SECTra_w- System of corpus m... SECTra_w- System of corpus m...

Show Print Watch Documentation

hchpaph Log-out

Home Import Evaluation Post-edition Multimodal TM Admin Translation Contact us

Corpus name: EOLSS Document name: D5_E4_06_01_06_TXT Language pair: english_french (515/515 = 100.0 %) Export

☒ english ☒ french

english document	french document
Tsunami (translated from Japanese for "a big wave in a harbor") is a sequence of long gravity water waves, propagating with high velocity from a tsunami source in the ocean (a large-scale zone of short-lived disturbances on the water surface) to the coastline. Local rising or sinking of the water-level in the ocean may be rather small, about 10 meters. At the same time, the large horizontal size of the geophysical disturbance, which is in the order of 100 km above the bottom earthquake, may give rise to a gravity wave about 100 km in length. The wave propagates from the source with a velocity of long gravity water waves in accordance with the equation $C_G = (g H)^{1/2}, \quad (1)$ where g is the acceleration due to gravity, and H is the depth of the basin. Because the average depth of the world ocean is 4 km, the typical velocity of tsunami in the ocean is 200 m s^{-1} or 720 km h^{-1}. Such a wave, propagating with the velocity of an airplane, may traverse the Pacific ocean in 10-12 hours and bring down a wall of water 10 m high with a velocity of more than 70 km h^{-1} upon a calm ocean beach. The wave velocity is decreased near the coastline because of shallower water and the slowing of the wave by the roughness of the bottom. The steepness of the wave is increasing because the crest of the wave moves faster than the trough of the wave, whose motion is delayed by the bottom impact.	Un tsunami (traduit du Japonais et signifiant "grande vague dans un port") est un série de longues vagues de gravité, se propageant à grande vitesse à partir de sa source dans l'océan (une grande zone de perturbations de courte durée sur la surface de l'eau) jusqu'au littoral. La montée ou la descente locale du niveau d'eau dans l'océan peut être plutôt faible, environ 10 mètres. Cependant, la largeur horizontale de la perturbation géophysique, qui peut être dans les 100 kilomètres au-dessus du tremblement de terre sous-marin, peut provoquer une onde de gravité d'environ 100 kilomètres de longueur. L'onde se propage de la source avec une vitesse de longues vagues de gravité selon l'équation $C_G = (g H)^{1/2}, \quad (1)$ Dans laquelle g est l'accélération due à la gravité, et H est la profondeur du bassin. Puisque la profondeur moyenne de ces océans du monde est de 4 kilomètres, la vitesse typique d'un tsunami dans l'océan est de 200 m s^{-1} ou de $720 \text{ kilomètres h}^{-1}$. Une telle vague, se propageant à la vitesse d'un avion, peut traverser l'océan pacifique en 10-12 heures et abaisser un mur d'eau de 10m de haut avec une vitesse de plus de $70 \text{ kilomètres h}^{-1}$ sur une plage calme. La vitesse de la vague est diminuée près du littoral car l'eau devient moins profonde et la vague est ralentie par la rugosité du fond. L'escarpement de la vague augmente parce que la crête de la vague

en attente de eolss.imag.fr...



SECTra_w (résultat EOLSS/UNL)

- Performance des traducteurs professionnels
 - Pour une page standard $\approx$ 250 mots
 - 1 h pour le premier jet
 - + 20 mn de révision pour obtenir la Haute Qualité
 - (source Sonovision, <http://www.sonovision-itep.fr/>)
 - (Adobe donne à peu près les mêmes chiffres)
- Performance des post-éditeurs amateurs
 - 15-25mn de post-édition
 - résultat comparable au premier jet d'un traducteur professionnel junior
 - + 20 mn de révision par un pro si on veut la HQ



SECTra_w (résultat EOLSS/UNL)

SECTra_w mesure le temps de modification des "cellules de post-édition", pas celui de l'accès à des ressources lexicales ni du passage d'un segment à l'autre, et est donc optimiste. On a cherché à estimer le temps complet de PE.

- Performances avec SECTra_w
 - Parmi les 3 contributeurs principaux ...
 - l'un d'entre eux a mesuré un temps de **15-20 mn/page** (gain de 65%) pour le temps TOTAL de post-édition
 - Extrapolation des données des autres post-éditeurs
 - postédition: **20-30 mn/page**
 - familiarisation avec l'environnement : **2-3 pages (1h)**
 - + recherche de terminologie : **5 à 30 mn/page**
- nécessité d'une étude plus poussée*
(plus on en fait, plus c'est rapide)



Passerelles coopératives d'accès multilingue

- **Principe**
 - Permettre un accès multilingue à un site Web ...
 - sans toucher au contenu de ce site
 - *à la Google Translate*
- **Originalité**
 - ... avec amélioration incrémentale et contrôlée de la qualité des traductions ...
 - ... via une passerelle iMAG dédiée au site Web « élu »



Passerelles coopératives d'accès multilingue

- Mise en œuvre (1/2)
 - Le site d'origine n'est pas affecté
 - Les pages traduites sont construites à la volée
 - grâce à une **mémoire de traductions** qui contient
 - des prétraductions obtenues par TA (externe, interne)
 - des traductions obtenues manuellement ou par post-édition
 - Les améliorations sont apportées
 - **par les visiteurs** dans le contexte de navigation
 - **par un locuteur natif** servant de **modérateur** pour chaque langue

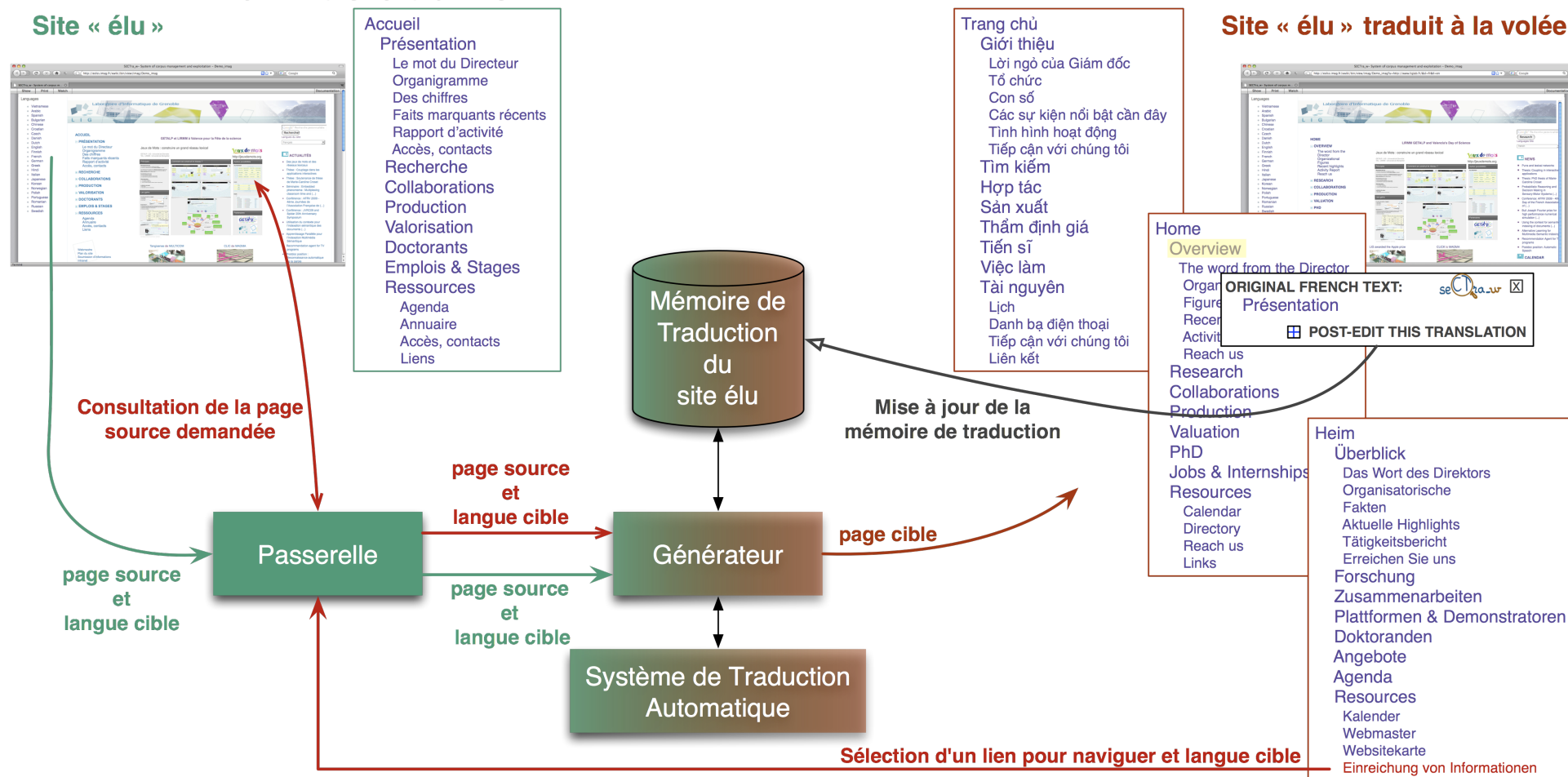


Passerelles coopératives d'accès multilingue

- Mise en œuvre (2/2)
 - le processus fonctionne en **boucle infinie**
 - les **prétraductions sont régulièrement recalculées**
 - suivi des améliorations des systèmes de TA externes
 - un système de TA dédié (interne) peut être mis en place et continûment mis à jour
 - recalcul de ses données à partir des bonnes traductions

Passerelles coopératives d'accès multilingue

• Architecture





Passerelles coopératives d'accès multilingue

- Page Web source

CALL FOR PAPERS

The Conference

The first Cranfield International Conference on Machine Translation, jointly organised by the BCS-NLTSG and Cranfield University, was held in 1984. A second Cranfield conference was held in 1994. Machine Translation is now some 60 years old and in view of the success of the previous Cranfield conferences, we are marking the twenty-fifth anniversary of the first conference at Cranfield by holding another this November.



Passerelles coopératives d'accès multilingue

- Page Web de traduction brute obtenue via la passerelle

APPEL DE COMMUNICATIONS

La Conférence

Le premier Cranfield Conférence internationale sur la traduction automatique, organisée conjointement par le BCS-NLTSG et l'Université de Cranfield, a eu lieu en 1984. Une deuxième conférence de Cranfield a eu lieu en 1994. La traduction automatique est maintenant âgé de 60 ans et compte tenu du succès des précédentes conférences Cranfield, nous célébrons le vingt-cinquième anniversaire de la première conférence de Cranfield en tenant un autre, cela Novembre.



Passerelles coopératives d'accès multilingue

- Fenêtre flottante d'accès au segment original

Organisé par: Le Natural Language Translation Specialist Group de la
British Computer Society (BCS-NLTSG)

Original English text: 
CALL FOR PAPERS
 Post-edit this translation...

APPEL DE COMMUNICATIONS

Le premier Cranfield Conférence internationale sur la traduction automatique, organisée conjointement par la



Passerelles coopératives d'accès multilingue

- Accès à la post-édition contributive



The screenshot shows a web page titled "APPEL DE COMMUNICATIONS" on a light green background. A text box contains the following text: "ence internationale sur la traduction automatique, organisée conjointement par le e Cranfield, a eu lieu en 1984. Une deuxième conférence de Cranfield a eu lieu en 1994. La traduction automatique est maintenant âgé de 60 ans et compte tenu du succès des précédentes conférences". A browser window is overlaid on the left, showing "Original English text: seCTra_w x" and "CALL FOR PAPERS". Below this, a text area contains "APPEL DE COMMUNICATIONS". At the bottom of the browser window are two buttons: "Contribute" and "Postedit in SECTra_w".



Passerelles coopératives d'accès multilingue

- Validation de la post-édition contributive

APPEL À COMMUNICATIONS

La Conférence

Le premier Cranfield Conférence internationale sur la traduction automatique, organisée conjointement par le BCS-NLTSG et l'Université de Cranfield, a eu lieu en 1984. Une deuxième conférence de Cranfield a eu lieu en 1994. La traduction automatique est maintenant âgé de 60 ans et compte tenu du succès des précédentes conférences Cranfield, nous célébrons le vingt-cinquième anniversaire de la première conférence de Cranfield en tenant un autre, cela Novembre.



Perspectives de recherche

- Traduction Interactive Fondée sur le Dialogue et Documents Auto-Explicatifs
 - DI pour l'auteur
 - Certification du sens
- Désambiguïsation Interactive en langue cible
 - DI pour le lecteur
- TA fondée sur la connaissance
 - probable retour en force de cette idée grâce à l'émergence du **Web sémantique** (Web 3.0)
 - existence d'ontologies *métier* utilisables pour la TA **mais** développées et maintenues pour d'autres usages



Conclusion

- La fonction traduisante n'est automatisable que pour la **suivi** et la **diffusion**
- La TAO quasi-parfaite est possible dans des cadres restreints
 - Technique non extensible pour l'instant
 - Exemples isolés, peu de situations réelles
- La TAO adaptée aux documents à traduire permet d'obtenir des résultats de qualité suffisante pour être post-édités en assurant un gain de productivité
 - Nombreux systèmes commerciaux opérationnels



Références

- Petit guide de l'acheteur de traduction
 - <http://www.itl.org.uk/pdfs/trans/trans03.pdf>
- The State of Machine Translation in Europe and Future Prospects (2002)
 - Accès aux systèmes en ligne et commerciaux
 - http://ccl.pku.edu.cn/doubtfire/NLP/Machine_Translation/Overview/Article%20-%20MT_John_Hutchins.htm
- Machine translation: An Introductory Guide (1994)
 - <http://www.essex.ac.uk/linguistics/clmt/MTBook/>
- Survey of the State of the Art in Human Language Technology (1996)
 - <http://cslu.cse.ogi.edu/HLTsurvey/HLTsurvey.html>
 - http://www.lt-world.org/HLT_Survey/master.pdf
- La traduction automatique (2001)
 - <http://www.etudes-francaises.net/entretiens/15traduction.htm>
- Machine translation: Publications by John Hutchins
 - <http://www.hutchinsweb.me.uk/>
- TA statistique
 - <http://www.statmt.org/>



Merci de votre attention

