

**TransType2: une nouvelle approche pour un système d'aide à la traduction**

Guy Lapalme

Elliott Macklovitch

### Résumé

Nous présentons le projet *TransType* développé depuis plusieurs années dans notre laboratoire en collaboration avec des chercheurs français, allemands et espagnols dans le cadre d'un projet de la communauté économique européenne. *TransType* est une approche hautement interactive à la traduction où le traducteur humain débute une traduction pour laquelle le système propose continuellement des suites possibles. Si le traducteur juge une suggestion appropriée, il peut facilement l'accepter et ainsi éviter d'avoir à taper cette partie de la traduction. S'il ignore cette suggestion en continuant à taper sa propre traduction, le système trouvera une autre complétion appropriée au nouveau texte. Cette approche permet d'améliorer la productivité des traducteurs en termes de mots traduits à l'heure. Nous présenterons une démonstration de cette approche innovatrice, les modèles de langues et de traduction utilisés ainsi que les résultats d'évaluations tenues auprès de traducteurs professionnels.

***TransType: une nouvelle approche pour un système d'aide à la traduction***

Le laboratoire de Recherche Appliquée en Linguistique Informatique (RALI)<sup>1</sup> de l'Université de Montréal développe depuis plusieurs années *TransType*, un outil inédit d'aide à la traduction qui propose de compléter des traductions partielles fournies par un traducteur humain.

La figure 1 illustre l'utilisation de *TransType* pour traduire un texte de l'anglais vers le français. Le traducteur choisit une phrase à traduire dans la fenêtre de gauche et commence à entrer sa traduction dans la fenêtre de droite.

Insérer la figure 1 ici.

Après chaque caractère tapé par le traducteur, le système affiche une liste de propositions de complétion que l'utilisateur peut soit accepter (à l'aide d'une touche ou de la souris) soit ignorer en continuant à taper. Dans ce dernier cas, le système recalcule aussitôt de nouvelles suggestions de complétion en fonction de la nouvelle entrée du traducteur. Cette interface, développée suite à plusieurs alternatives testées auprès de traducteurs professionnels, est facile à utiliser. On peut en mesurer la performance en calculant la proportion de frappes épargnées en tapant une traduction, mais nous avons développé des mesures plus appropriées de la productivité des traducteurs. Contrairement à l'approche classique à la traduction automatique où la machine traduit une première ébauche que l'homme doit ensuite vérifier et corriger, dans *TransType* le traducteur garde le contrôle du processus de traduction et c'est la machine qui doit s'adapter continuellement aux choix faits par l'humain. Même s'il est raisonnable de penser que cette forme de complétion sera utile aux traducteurs, il est important de valider cette conjecture dans un contexte opérationnel. Cet article décrit certains résultats d'évaluations effectuées auprès de traducteurs professionnels.

### Origines du projet

Le projet **TT2** est le dernier d'une série de projets de recherche portant sur de nouveaux types de systèmes d'aide à la traduction qui intègrent le traducteur humain pour en tirer le meilleur parti. *TransTalk* (Brousseau et coll., 1995) est un système de dictée de traduction qui prend en compte l'information tirée du texte source de la traduction dictée afin d'améliorer les performances du système de reconnaissance de la parole.

L'adaptation des techniques utilisées dans *TransTalk* dans le but d'accélérer la frappe de traduction a donné lieu à une technologie appelée *Target-Text Mediated Interactive Machine Translation* (Foster et coll. , 1997). Cette approche a par la suite été développée par le RALI entre 1997 et 2000 dans le cadre d'une *subvention stratégique CRSNG*. Ce projet a permis de montrer la faisabilité de produire des suggestions utiles en temps réel dans le cadre de la traduction entre l'anglais et le français et de faire une première série d'évaluations contrôlées avec des traducteurs humains (Langlais et coll. 2002).

Les résultats très encourageants de ces premières expériences ont permis de réunir autour de ce thème une équipe internationale de recherche comprenant des industriels et des universitaires de plusieurs pays dans le *Projet européen du 5e programme cadre TransType2* (2002-2005)<sup>2</sup>. Plusieurs versions du système y sont développées permettant de traduire de l'anglais vers le français, l'espagnol ou l'allemand (et inversement) et ce, sur deux types de textes très différents: des débats parlementaires (canadiens et européens) et des manuels techniques de Xerox. Afin de s'assurer que le prototype répond aux besoins des traducteurs professionnels, des traducteurs de deux des plus grands bureaux de traduction du Canada et d'Espagne sont impliqués dans différentes évaluations *in-situ* des prototypes. Il est également

prévu d'ajouter une composante de reconnaissance de la parole pour la sélection des suggestions du système (Esteban et coll. 2004).

**TransType2** regroupe les partenaires européens suivants :

- Atos Origin (Espagne) : gestionnaire du projet et coordonnateur technique.
- RWTH de l'Université d'Aachen (Allemagne) : moteur de traduction statistique anglais- {allemand, espagnol, français} et reconnaissance de la parole.
- ITI de l'Université de Valence (Espagne) : moteur de traduction anglais- {allemand, espagnol, français} exploitant les techniques d'états finis et reconnaissance de la parole.
- Xerox Research Centre Europe (France) : fournisseur de corpus et modélisation statistique de la traduction
- Celer Soluciones (Cabinet de traduction) : évaluation des moteurs de traduction dans le contexte de travail.

Les partenaires canadiens de ce projet sont :

- RALI où sont impliqués les chercheurs suivants:  
Elliott Macklovitch: responsable des évaluations par les traducteurs humains et liaison avec les chercheurs européens  
Guy Lapalme: professeur administratif des demandes de fonds, direction des étudiants dans les domaines de l'adaptation des modèles de langues  
Philippe Langlais: professeur responsable des modèles statistiques de traduction et direction d'étudiants dans le domaine de la reconnaissance de la parole  
Simona Gandrabur: responsable des mesures d'estimation de confiance et responsable du moteur de traduction anglais-français

Ngoc Tran Nguyen: responsable de l'implantation de l'interface et des outils de trace pour l'évaluation de la productivité des traducteurs

Les étudiants gradués suivants sont aussi impliqués dans le projet: Christian Jauvin (PhD), Laurent Nepveu (MSc terminée), Thomas Le Plus (MSc).

Il faut aussi souligner la contribution de George Foster qui pendant plusieurs années a été le chercheur principal du projet ayant complété en 2002 sa thèse sur le sujet et qui y a continué à y travailler au RALI pendant 2 ans par la suite.

- Société Gamma (Cabinet de traduction): évaluation des moteurs de traduction dans le contexte de production.

### **Modèle Probabiliste**

La linguistique cherche à caractériser et à expliquer la multitude d'observations sur la langue utilisée dans les conversations, les écrits et les médias. Cette discipline porte à la fois sur le côté cognitif de l'acquisition, de la production et de la compréhension du langage ainsi que sur la compréhension de la relation entre l'expression linguistique et le monde réel.

En linguistique, on s'intéresse également à la compréhension des structures utilisées dans le langage. Pour y arriver, on peut proposer des « règles » qui définissent la bonne formation des énoncés. Toutefois, ces règles évoluent dans le temps; les locuteurs les violent parfois par ignorance, par paresse ou pour être créatifs sans pour autant empêcher le message de passer. C'est en partie cette flexibilité qui rend la formalisation de la langue naturelle si difficile. On peut distinguer deux grandes approches en linguistique: rationaliste et empirique. La première suppose une connaissance à priori de règles linguistiques qui encodent la connaissance sur la langue alors que la seconde s'appuie sur une capacité d'association, de reconnaissance de patrons et de généralisation à partir d'exemples. Nous allons nous concentrer ici sur cette seconde approche qui revient en force depuis le milieu des années 80. Les méthodes empiriques ont débuté dans les années vingt pour faire place, dans les années soixante, à l'approche rationaliste promue entre autres par Chomsky. La disponibilité grandissante de textes sous format électronique ainsi que des moyens de stockage et de traitement de plus en plus performants ont permis un développement accéléré de l'approche empirique au cours des dernières années. Cette approche suppose qu'on peut apprendre (voire expliquer) la structure du langage en spécifiant un modèle général dont on peut déterminer la valeur de ses paramètres à l'aide de statistiques et d'inférences de schémas sur de grandes quantités de textes.

Les modèles statistiques de langue ont l'avantage d'être robustes, c-à-d de pouvoir traiter de « vrais » textes et non pas seulement des idéalizations théoriques. Ils pourront de plus fournir toujours une réponse (parfois même plusieurs) et ils sont en général plus rapides à appliquer que les méthodes symboliques. Il faut toutefois convenir de quelques inconvénients : ils sont difficiles à mettre au point et à corriger; il est très difficile d'en expliquer les résultats en termes de principes de plus haut niveau plus facilement compréhensibles par l'utilisateur.

Un modèle de langue probabiliste peut être présenté comme une fonction qui donne la probabilité de rencontrer un mot en fonction des  $n$  mots précédents. Par souci de simplification, on ne considère que le contexte immédiat comme pouvant influencer le prochain mot, car des raisons pratiques imposent des restrictions sur l'empan du texte tenu en ligne de compte. En principe, on aimerait que  $n$  soit grand mais alors on aurait trop de paramètres à estimer comme le montre le tableau 2 pour un vocabulaire de 20 000 mots.

#### Insérer le tableau 2 ici

Différentes méthodes sont appliquées pour restreindre l'espace des probabilités, chercher des estimateurs à vraisemblance maximum et surtout traiter l'éparpillement des données. Cette approche a déjà montré son utilité dans plusieurs applications dont la reconnaissance de la parole et de caractères, la correction de fautes d'orthographe et les systèmes de traduction automatique.

Manning et Schütze (Manning99) présentent une excellente introduction au domaine.

#### *Modèles de traduction*

L'approche statistique aux modèles de traduction estime des probabilités de rencontrer des mots de la langue cible étant donné l'historique du texte cible et du texte source. On utilise souvent l'analogie du canal bruité dans lequel entre une phrase  $t$  d'une langue  $T$  ayant  $T$  comme vocabulaire. Le canal est si bruité que la phrase  $s$  qui en sort est la traduction de  $t$  dans une autre



langue  $S$  ayant  $\mathcal{S}$  comme vocabulaire. Le processus de traduction consiste donc, en présence d'une phrase de sortie  $s$  de la langue  $S$ , à trouver  $t$  dans la langue  $T$  qui l'a produite. L'analogie du canal bruité représente donc le processus inverse de la traduction: la phrase cible est entrée dans le canal et la phrase source en ressort!

La modélisation statistique de la traduction est définie avec l'équation suivante simplifiée à l'aide la loi de Bayes pour séparer les processus de traduction et de génération de la phrase cible.

$$\hat{t} = \arg \max_{t \in T} p(t | s)$$

$$\hat{t} = \arg \max_{t \in T} \frac{p(s | t) \times p(t)}{p(s)}$$

$$\hat{t} = \arg \max_{t \in T} \underset{\text{traduction}}{p(s | t)} \times \underset{\text{langue}}{p(t)}$$

Pour ce processus, il faut d'abord estimer  $p(t)$  la distribution de la langue cible et  $p(s | t)$  la distribution de la traduction pour ensuite rechercher la phrase  $\hat{t}$  qui maximise le produit de ces probabilités. Les distributions des modèles de langue et de traduction sont estimées à partir d'observations sur des traductions déjà faites.

À cause du caractère interactif de *TransType* et du fait qu'il doive fonctionner en temps réel, il a fallu adapter ce modèle de base en ajoutant une notion de vocabulaire actif pour restreindre la recherche des mots candidats au sous-ensemble des mots les plus susceptibles d'apparaître dans la traduction de la phrase source courante. Il est également possible de moduler les contributions relatives des modèles de traduction et de langue à l'aide d'un facteur  $\lambda$  compris entre 0 et 1. Ce processus peut être défini par l'équation suivante :

$$p(t | t', s) = \frac{p(t | t')}{\text{langage}} \lambda + \frac{p(t | s)}{\text{traduction}} (1 - \lambda)$$

et illustré dans la figure 2.

Insérer la figure 2 ici

## Évaluation

En plus du défi technique visant à développer les bases mathématiques et l'infrastructure informatique nécessaire pour implanter les modèles et développer une interface conviviale pour les traducteurs, on a voulu mesurer l'impact de cette nouvelle approche sur la productivité des traducteurs.

Il faut d'abord définir la notion de productivité: le fait d'avoir moins de caractères à taper ou d'actions de souris à effectuer devrait, en principe, accélérer le travail du traducteur mais la vraie mesure reste le temps global nécessaire pour traduire des textes. Pour le développement informatique des prototypes, nous utilisons cette mesure, mais il faut la corrélérer avec de mesures plus réalistes.

Le RALI a déjà mené une série d'évaluations préliminaires avec des traducteur humains sur une version précédente de *TransType* (Langlais et coll. 2002) ce qui a permis d'obtenir un premier feedback et de démontrer l'intérêt de l'approche. Globalement, tous les traducteurs ont beaucoup apprécié les suggestions interactives toujours bien orthographiées. Les suggestions fournies par *TransType* sont beaucoup moins intrusives et lourdes à manipuler que les systèmes de mémoire de traduction actuels. Les traducteurs avaient l'impression d'avoir augmenté leur productivité; malheureusement les mesures réelles en termes de frappes sauvées et de temps de traduction ne corroboraient pas cette impression positive.

En moyenne, le temps nécessaire pour lire, évaluer et accepter une suggestion correspond au temps de taper trois caractères, il est donc inutile de suggérer de très courtes complétions. Nous avons aussi remarqué que la lecture des suggestions amenait une charge cognitive importante et c'est pourquoi il est quelques fois préférable de n'afficher qu'une seule longue suggestion plutôt que plusieurs courtes. Afin de limiter le nombre de choix et leur fréquence

d'apparition, il est important d'obtenir des mesures de confiance (Gandrabor 2003) sur les suggestions de façon à ne pas les afficher si les scores sont trop bas, même s'ils sont parmi les cinq ou sept meilleurs.

Dans le cadre du projet TT2, nous avons prévu des évaluations trimestrielles chez les deux cabinets de traduction qui faisaient partie de notre consortium, soit la Société Gamma à Ottawa et Celer Soluciones à Madrid. Le but de ces évaluations était double : d'une part, en mettant la dernière version du système entre les mains de vrais traducteurs, ces essais nous permettaient de mesurer de façon concrète et fiable son impact sur la productivité; d'autre part, les essais fournissaient un canal de communication par lequel les utilisateurs pouvaient fournir des suggestions et de la rétroaction aux développeurs. Deux types de textes très différents ont été retenus pour ces évaluations : des manuels d'utilisateur fournis par la société Xerox et des débats parlementaires, canadiens et européens. En fait, il fallait trouver des corpus qui offrent à la fois des textes en grande quantité, afin de permettre l'entraînement de nos systèmes statistiques, et des traductions de haute qualité vers plusieurs langues cibles. Chaque cabinet de traduction a sélectionné trois traducteurs chevronnés pour participer aux essais. Chez Gamma, les participants traduisaient de l'anglais vers le français, alors que chez Celer les participants traduisaient vers l'espagnol.

Afin d'analyser les résultats des séances de traduction, le RALI a développé un programme utilitaire qui s'appelle *TT-Player*. Ce programme prend en entrée un fichier trace très détaillé dans lequel est enregistrée chaque interaction entre l'utilisateur et le système *TransType*. En lisant cette trace, *TT-Player* peut non seulement rejouer la séance de traduction, à la manière d'un magnéscope, mais il peut aussi calculer des statistiques sur toutes sortes de paramètres qui caractérisent la façon dont le traducteur s'est servi du système. Ainsi, *TT-Player* nous renseigne

sur le nombre de prédictions que l'utilisateur a acceptées dans une séance, la longueur moyenne de ces prédictions, de quelle façon l'utilisateur a accepté les complétions (soit par le clavier soit par la souris), le nombre de caractères que l'utilisateur a lui-même tapés pour obtenir son texte cible, le nombre de caractères qu'il a effacés, etc. Et il nous dit, bien sûr, combien de temps le traducteur a mis pour produire son texte, ce qui nous permet de calculer sa productivité, c'est-à-dire, le nombre de mots qu'il traduit à la minute ou à l'heure.

Après cinq rondes d'essais de *TransType* chez les deux cabinets de traduction, les résultats ont été très encourageants. Pour avoir un étalon de comparaison, nous avons demandé aux participants de traduire une section du manuel Xerox en se servant de l'éditeur de *TransType*, mais sans l'aide des prédictions fournies par le système. Ensuite, nous leur avons demandé de traduire huit autres sections du même manuel en utilisant *TransType* en mode normal, c'est-à-dire avec son moteur de prédictions activé. En comparant les taux de productivité des participants avec et sans l'aide des prédictions, nous avons constaté que *TransType* a permis des gains de productivité assez importants chez cinq des six participants, des gains qui varient entre 30 à 55%. D'autres statistiques fournies par *TT-Player* tendent à renforcer ces résultats. Par exemple, nous avons trouvé que le nombre moyen d'actions par mot cible – le nombre de caractères tapés ou de cliques de la souris – diminuent grandement lorsque les traducteurs se prévalent des complétions offertes par *TransType*. Autrement dit, le système permet aux traducteurs non seulement de produire leur traductions plus rapidement mais aussi avec moins d'effort. Lors de cette dernière ronde d'essais, nous avons également introduit une étape de contrôle de la qualité, afin de s'assurer que ces gains de productivité n'ont pas été réalisés au détriment de la qualité. Or, les réviseurs ont trouvé que les textes produits à l'aide des prédictions de *TransType* ne contenaient pas plus d'erreurs que ceux de l'essai à blanc; en fait, le taux

d'erreurs était très faible dans tous les textes produits lors de l'essai, que ce soit avec ou sans les prédictions du système.

### **Conclusion**

*TransType* est une approche innovatrice à la traduction interactive qui combine les forces respectives de la machine et de l'humain dans une interface conviviale. Elle a déjà montré son efficacité dans le cadre d'évaluations avec des traducteurs professionnels. Il reste maintenant à l'intégrer dans un éditeur commercial (p.e. Microsoft Word) afin que *TransType* puisse plus facilement se combiner aux outils courants de traducteurs. Il serait aussi intéressant de voir jusqu'à quel point cette approche pourrait aussi servir à des utilisateurs qui doivent souvent rédiger des textes dans une langue qui n'est pas leur langue maternelle ou même servir à l'apprentissage de la rédaction bilingue.

### Bibliographie

- Brousseau J., Drouin C., Foster G., Isabelle P., Kuhn R., Normandin Y. et Plamondon P. (1995) French Speech Recognition in an Automatic Dictation System for Translators: the TransTalk Project. *Proceedings of Eurospeech*, p. 193-196, Madrid, Spain.
- Esteban, J. , Lorenzo, J., Valderrábanos, A. S. et Lapalme, G. (2004) TransType2 - An Innovative Computer-Assisted Translation System. *The Companion Volume to the Proceedings of 42st Annual Meeting of the Association for Computational Linguistics*, p. 94-97, Barcelona, Spain.
- Foster G., Isabelle P. et Plamondon P. (1997) Target-Text Mediated Interactive Machine Translation. *Machine Translation*, p. 175-194, 1997.
- Simona Gandrabur et George Foster. (2003) Confidence Estimation for Text Prediction. *Proceedings of the Conference on Natural Language Learning (CoNLL 2003)*, Edmonton, Mai.
- Langlais P., G. Lapalme et M. Loranger. Transtype: Development-evaluation cycles to boost translator's productivity. *Machine Translation (Special Issue on Embedded Machine Translation Systems)*, 17(2):77-98, 2002.
- Manning C., Schütze, H. (1999) Foundations of Statistical Natural Language Processing, MIT Press, 1999.



Author Note

Guy Lapalme est professeur titulaire au département d'informatique et de recherche opérationnelle de l'Université de Montréal. Depuis plus de 20 ans, il fait des recherches dans le domaine du traitement de la langue et depuis 1997 dans le cadre du laboratoire RALI.

Elliott Macklovitch est le coordonnateur du laboratoire RALI et est le responsable des évaluations de TransType avec les traducteurs professionnels

Des renseignements supplémentaires peuvent être obtenus auprès de Guy Lapalme, RALI-DIRO, Université de Montréal, CP. 6128, Succ Centre-Ville, Montréal, H3C 3J7, Québec, Canada, courriel: [lapalme@iro.umontreal.ca](mailto:lapalme@iro.umontreal.ca)

Footnotes

<sup>1</sup><http://rali.iro.umontreal.ca>

<sup>2</sup> No de projet: IST-2001-32091, <http://tt2.sema.es>

Pour des informations plus spécifiques au RALI voir

<http://rali.iro.umontreal.ca/Traduction/index.fr.html>

Tableau 1

*Nombre de paramètres à estimer pour des n-gram.*

	n	# de paramètres	
bigramme	2	$20\,000^2$	400 millions
trigramme	3	$20\,000^3$	8 trillions
4-gramme	4	$20\,000^4$	$1.6 \cdot 10^{17}$

## Figure Captions

*Figure 1.* Interface de TransType telle que vue par l'utilisateur qui traduit le texte à gauche en tapant sa traduction dans le panneau de droite où le système propose des suggestions de complétions de texte. Dans la version en couleur, les caractères en rouge ont été entrés par le système suite à des acceptations de l'utilisateur. On peut avoir une idée rapide de la proportion des caractères acceptés par l'utilisateur sur l'ensemble du texte.

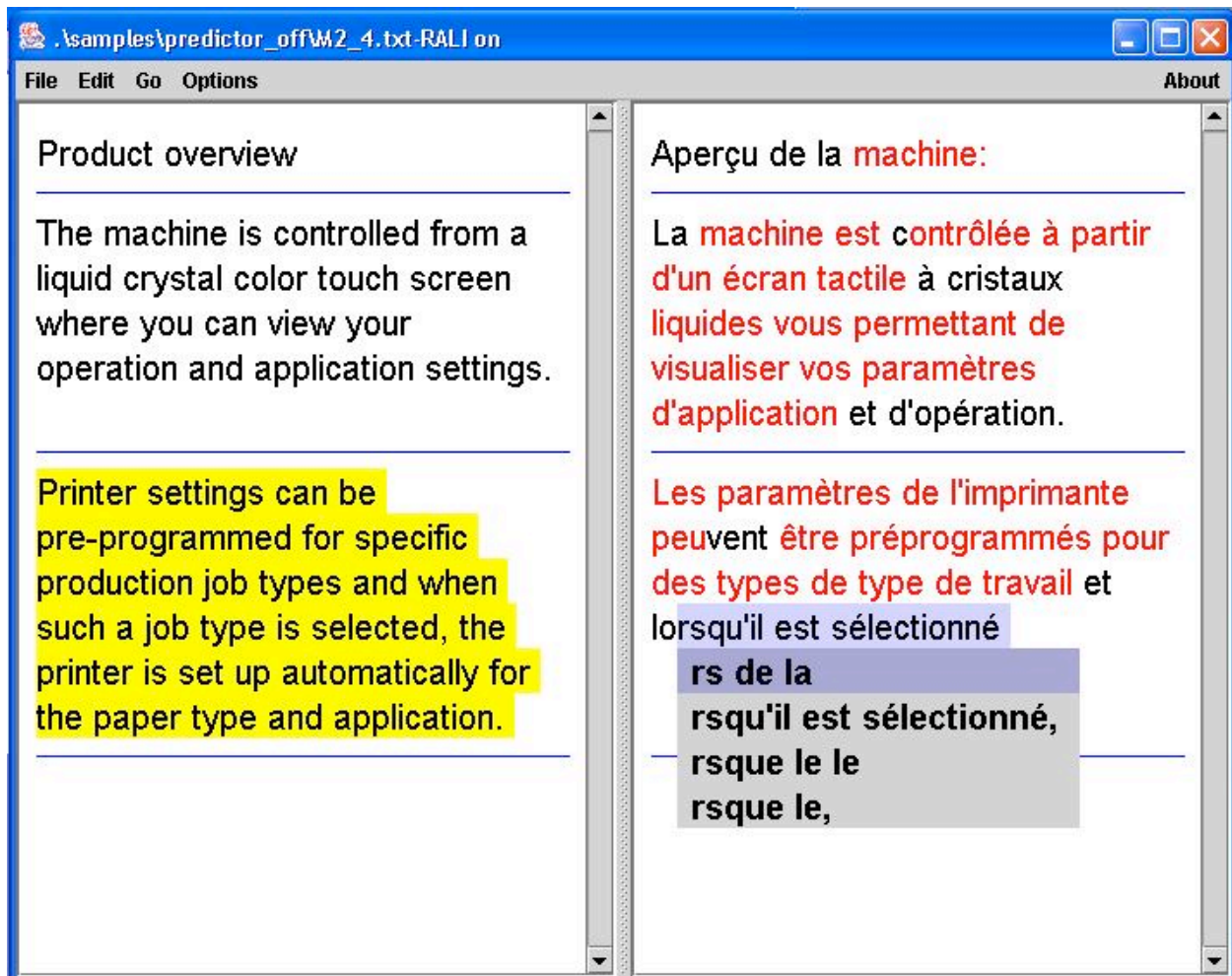


Figure 2 : Combinaison de modèle de langue et de traduction utilisé dans TransType

