# Proyecto evolutivo de traducción automática en un contexto controlado
# PETACC

Jocelyne Caron
Département de linguistique
Université du Québec à Montréal
e-mail:carmack@mlink.net

Dr Sylvie Ratté
École de technologie supérieure
http://www.etsmtl.ca/seg/sylvie/
e-mail: sratte@info.etsmtl.ca

## Introducción

El campo de la traducción automática (en lo sucesivo TA) se confronta a un escollo de envergadura: los sistemas no son suficientemente fiables y se reconoce que son muy poco eficaces en la comprensión de textos en una lengua terminal desconocida por el utilizador.

Respondiendo específicamente a las necesidades presentadas por el TLC, parece ser probable reducir el problema de la TA de dos maneras: limitándose en un primer lugar a tres idiomas muy bien documentados y en segundo lugar a un campo específico de la correspondencia comercial.

Además, pensamos como muchos que es menester, en un primer momento, plantear este amplio problema, considerando las trabas similarespresentadas en lenguas controladas, de las cuales nos demarcamos sin embargo, ubicando este proyecto en un contexto controlado. Ello significa, para nosotros, que debemos establecer, además de las limitaciones relativas al léxico y a la sintaxis, otras en el plano del discurso.

## 1. Marco metodológico

El marco metodológico que nos proponemos para el desarrollo de un sistema TA, además de ser elaborado en un contexto controlado, posee la particularidad de ser evolutivo. Explicamos a continuación estas dos particularidades.

### 1.1. Proyecto evolutivo

En un primer momento, nos centraremos a tratar un conjunto reducido de cartas que pertenecen a lo sumo a dos tipos, tales como las solicitudes de información o de pago. Este procedimiento nos conducirá a establecer las bases del sistema, en su conjunto.

En un segundo momento, se trata de ampliar las capacidades del sistema, aumentando progresivamente los modelos de cartas tratadas.

Finalmente, el marco estricto de los modelos de correspondencia,podría extenderse al tratamiento de otro tipo de documentos. Los imperativos del mercado podrían eventualmente guiar nuestras elecciones en el desarrollo del marco implantado.

Este marco metodológico constituye en sí mismo una parte del sistema, porque creemos que es la única manera de llegar a construir un sistema complejo que no se embrolle rápidamente.

### 1.2. Contexto controlado

Nuestro proyecto en un contexto controlado se desarrolla de acuerdo a tres grandes módulos que son:

*Módulo 1*

En el primer módulo de tratamiento, los textos originales son *limpiados*: a esto lo llamamos el pretratamiento.

*Módulo 2*
En este punto, la formalización de los datos linguísticos se realiza, a todos los niveles del análisis, sean discursivo, morfosintáctico y lexicoconceptual.

*Módulo 3*
Una vez que los datos linguísticos han sido formalizados, se trata de construir un apareamiento, a todos los niveles, para automatizar latransferencia, es decir la TA.

Para lograr el tratamiento de un conjunto importante de modelos de correspondencia, estas tres etapas (los tres modulos) se aplican cada vez que el sistema atravieza una fase más de su evolución.

## 2. Módulo I : pretratamiento de documentos originales

La correspondencia comercial es analizada y criticada a todos los niveles del análisis, estableciendo las trabas formales y conceptuales que deben encontrarse.

### 2.1. Nivel discursivo

En el nivel discursivo, el trabajo se efectuará en dos planos.

a) Identificación de modelos de correspondencia:

Ampliamos la noción de protocolo epistolar (que incluye comunmente los periféricos tales como la referencia, la fecha, etc.) para incluir además los marcadores exteriores a los nudos de la frase (i.e. los marcadores discursivos como por ejemplo los adverbios) y lo que llamaremos "fórmulas comerciales" (como por ejemplo las fórmulas de cortesía que se encuentran en este tipo de textos: le ruego que se digne...").

Tomaremos como referencia los estudios de Ratté (1988) para realizar las propriedades esenciales que permitirán identificar el tipo de correspondencia. Especialmente, los aspectos que se relacionan muy en particular con el protocolo epistolar (mencionado arriba) que pueden servir como cimiento para la formalización de correspondencia según el tipo al que pertenezcan.

b) Reestructuración de las cartas:

Siempre tomando como referencia los estudios de Ratté y procediendo al análisis del contenido de la correspondencia, realizaremos las partes escenciales para reagruparlas, entre otras, en párrafos. Esto nos permitirá formalizar la carta de tal modo que a un tipo dado de carta, corresponda una macroestructura.

Aún no disponemos de una muestra suficiente de cartas para poder evaluar hasta que punto deberíamos depurar las cartas, o sea simplificarlas. En ciertos casos por ejemplo, deberíamos reescribir la mención de objeto y reestructurar los párrafos. De esta manera, podríamos decir que en el caso preciso de una carta de solicitud de información, deberíamos esperarnos a encontrar tres párrafos muy esquematizados, dentro de los cuales el contenido del primero y del último serán previsibles.

### 2.2. Reestructuración de frases - nivel morfosintáctico

Sin lugar a dudas esta etapa es la que más se aproxima a las ideas avanzadas dentro de los lenguajes controlados. En efecto, se procederá a una depuración específica de cada uno de los idiomas. Por ejemplo en francés, podremos decidir que las frases no conlleven mas de dos verbos conjugados, etc.

Existen muchos documentos en los cuales se definen estándares muy precisos (en particular en la redacción técnica) en los cuales podemos sustentar nuestra definición de las reglas sintácticas del lenguaje comercial.[1]

Por lo tanto, es una gran empresa de reducción de la complejidad de las estructuras sintácticas que tiene por objetivo facilitar lasetapas posteriores.

### 2.3. Anotación de las informaciones lexicoconceptuales

La anotación de las informaciones lexicoconceptuales se organiza de la siguiente manera: primeramente se procede a la identificación de las categorías sintácticas (nombres, verbos, preposiciones, etc.); luego se identifica las locuciones verbales, nominales y locuciones preposicionales.

Además, preveemos construir dos tipos de léxicos de especialidad. El primero reagrupará las locuciones comerciales, que pertenecen estrictamente a este campo (cf. Clas & Horguelin). Incluimos aquí las *fórmulas comerciales* mencionadas en el 2.1. El segundo léxico reagrupará los diccionarios que pertenecen específicamente a campos especializados.

### 2.4. El resultado: recolección de las informaciones linguísticas en los tres idiomas

En esta etapa, creamos un muestrario representativo de modelos de cartas comerciales, frecuentes en los tres idiomas que han pasado por un proceso de normalización. El resultado está constituido de un *corpus anotado* de informaciones linguísticas específicas a estos tres niveles:

- discursivo
- morfosintáctico
- lexicoconceptual

## 3. Módulo II : formalización, codificación, uniformización

Esta etapa constituye la de la formalización del conjunto de los modelos de carta elaborados en los tres idiomas y a todos los niveles.

### 3.1. Nivel discursivo

Dentro de los elementos del protocolo epistolar que se encuentran al exterior del cuerpo de la carta, la mención de objeto está formulada de una manera muy estricta, lo que nos permite anticipar el tratamiento de otros elementos de la carta.

La primera constituyente de la mención de objeto corresponde exactamente al tipo de carta, que vemos por ahora como exclusivamente un nombre de acción tal como *solicitud, oferta, respuesta, envío*, etc. Este primer nombre está seguido de uno de los dos elementos siguientes, sea este un nombre o un término. A este primer nombre le sigue a continuación uno o dos de los elementos siguientes, sea un nombre, sea un

---

[1] Pensamos que muchos de los cambios realizados por los traductores humanos, frecuentemente, tienen una finalidad puramente estética. Hasta la fecha, la posición que adoptamos consiste en preservar prioritariamente las construcciones sintácticas equivalentes en dos idiomas dados. Esto no quiere decir necesariamente que por ejemplo, los gerundios (empleados especialmente en inglés) deberían traducirse de la misma manera en español y francés, idiomas que los usan con menos frecuencia.

término. Este nombre pertenece al mismo conjunto de la primera constituyente, por lo tanto es un nombre de acción que corresponde también a un tipo de carta como *pago* o *información*. Bajo la designación de "término" caben varios nombres perteneciendo estrictamente al campo comercial, como *precio* o *fecha*, tanto como términos técnicos que formarían parte del sujeto de la carta.

Además, la codificación y la ubicación de los marcadores exteriores a los nudos de la frase y de las *fórmulascomerciales* constituyen referencias que permiten alinear las frases a procesarse. En resumen, sirven como cimiento para la transferencia, simplificándola a su vez.

Realizando un inventario de las cartas según sus tipos, se crea un tipo de carta para cada protocolo epistolar particular, lo que permite definir la macroestructura de una carta. Estos tres grandes conjuntos de constituyentes serán excluidos del análisis sintáctico propiamente dicho.

**3.2. Codificación de informaciones morfosintácticas**

Al igual que para las informaciones morfológicas, se busca codificar las descripciones relativas a las informaciones sintácticas de manera a permitir representaciones epuradas, o sea recurriendo a las constantes y las variables, a la subespecificación.

a) Informaciones morfológicas

Las informaciones morfológicas son anotadas de manera tal que no se tenga que utilizar informaciones linguísticas exclusivas a un idioma. Se las codificarán como conjuntos de rasgos permitiendo así la subespecificación.

Deseamos poder utilizar esta particularidad a dos niveles: 1) en el léxico 2) en el apareamiento de las estructuras morfosintácticas.

Los ejemplos (1) y (2), retomados de Abeillé & al. (1990),ilustran correctamente la codificación que se busca realizar. Por lo tanto, no se ha especificado el número de la palabra *lettre*, mientras que *bagages* conlleva ya el rasgo plural.

(1) N [lettre, {<num, *a*>}]
(2) N[bagages, {<num, *Pl*>}]

La subespecificación (1) cubre así las dos instanciaciones posibles de la palabra *lettre* (singular o plural), mientras que la palabra *bagages* no puede ser más que plural. Esto permite una reducción de las reglas de transferencia, como lo mencionaremos más abajo.

b) Informaciones sintácticas

Veamos ahora como las selecciones metodológicas anteriormente implicadas, trascienden a un nivel sintáctico.

Primeramente, al nivel de las representaciones sintácticas, la experiencia adquirida en el tratamiento automático de las lenguas naturales nos demuestra que hace falta recurrir arepresentaciones arborescentes similares a las que se encuentran en las gramáticas de dependencia (Badia, 1993). Las árboles tienen entonces un nivel de profundidad reducido, lo cual facilita el apareamiento de las estructuras.

Luego, los estudios de Rosetta nos permiten ver hasta donde es importante proceder a una identificación de las propriedades abstractas de los nudos maximales de manera a preparar el terreno para el apareamiento de idioma a idoma.

Es preciso notar que estas propiedades deben ser suficientemente generales para poder aplicárselas a todos los idiomas tratados.

Por ejemplo: en francés se producen las preguntas de muchas maneras. Entre otras, existe una manera idiosincrática de hacerlo incluyendo a continuación de la palabra interrogativa, el segmento*est-ce-que*. Sabemos perfectamente que este segmento nunca tendrá lugar a ser traducido en otra lengua: su desaparición del árbol final no presentaningún problema si se conserva la propiedad abstracta deinterrogación. Esta información escencial nos sirve para orientar las construcciones interrogativas enla lengua terminal.

### 3.3. Construcción de léxicos de unidades simples y complejas

Retomamos de Schenk la idea de un léxico con CBE (complex basic expressions- expresiones básicas complejas) con el propósito de codificar más particularmente las locuciones que pertenecen a categorías y formas múltiples (ver sección 2.3.). Las CBE están representadas como árboles de superficie,permitiendo la subespecificación de todos los tipos de información.

Dentro del léxico, la expresión idiomática inglesa *lend a hand* se la representa por un árbol (ver a continuación) dentro del cual las variables permitirán el análisis de todas las variantes de esta expresión. (p.ej. *lend often a hand* o también *lend Mary a hand*). La expresión idiomática: *kicked the bucket* está también representada por un árbol (ver 2.) pero esta vez las variantes son más limitadas.

1. [VP[x] [V'[V lend] [y] [NP [det a [N hand] ] ] ]

2. [VP[x] [V'[V kick] [NP [det the [N buckett] ] ] ]

Finalmente, dentro de un marco evolutivo, el sistema exijirá refinar las entradas relativas al léxico, de tal manera que puedan recibir las informaciones lexicoconceptuales.

### 3.4. Resultado: construcción de las estructuras de datos en tres idiomas

Puesto que a todos los niveles de representación las informaciones linguísticas están codificadas de manera uniforme, las estructuras de datos ( a todos los niveles) así obtenidas en los tres idiomas muestran suficientemente similitudes formales para facilitar inmediatamente los diferentes tipos de apareamiento.

## 4. Módulo III : construcción del módulo de transferencia

Teniendo una colección de cartas pertenecientes al mismo tipo en los tres idiomas, procedemos acontinuación a establecer la correspondencia, a todos los niveles de análisis, entre estas informaciones, lo que nos conlleva a transformarlas (volver acodificar), y relacionarlas automáticamente mediante la transferencia.

Cabe destacar que los modelos con *interlingua* nos parecen por el momento difíciles de realizarlos aunque el trabajo ya realizado en el marco de las descripciones mediante las primitivas conceptuales, resulta muy prometedor(Jakkendoff, Pustejovsky). Sin embargo, es importante dirigirnos hacia este ideal, organizando los módulos de la transferencia de manera tal que tomen en cuenta lo mejor posible las informaciones lexicoconceptuales.

Pasemos revista ahora cómo se efectúa el proceso de apareamiento a todos los niveles.

**4.1. Apareamiento sobre una estructura de discurso similar en los tres idiomas**

Las informaciones anotadas al nivel discursivo contribuyen a orientar nuestro análisis. Les recordamos que hemos privilegiado el protocolo epistolar en el sentido definido en el punto 1.1 y en la estructura misma del documento.

Por ejemplo, una carta puede identificarse como el resultado de un contacto anterior (teléfono, carta, cita) si su introducción comienzapor *Following* + *NP*. Podemos notar que el proceso de traducción puede simplificarse puesto que los datos que quedan por ser traducidos están reducidos considerablemente (limitándose posiblemente a un NP en este caso).

En consecuencia, el análisis discursivo simplifica las etapas posteriores del tratamiento, reduciendo la complejidad estructural, puesto que las partes identificadas como siendo parte del protocolo no necesitan el llamado a un proceso de traducción total (análisis, transferencia + generación), estas expresiones forman parte intrínseca de la concepción del sistema.[2]

**4.2. Apareamiento de las estructuras morfológicas y sintácticas de idioma a idioma**

a) Estructuras morfológicas

Al nivel morfológico, favorecemos un mecanismo de apareamiento por unificación en el sentido que la codificación de las informaciones morfológicas (entrevista en 3.2.) nos permiten simplificar la transferencia.

Por ejemplo, el apareamiento en (1) permite englobar las dos instancias (Sg ou P1) de *lettre*, pero tambien permite transferir esta información morfológica hacia la lengua terminal. Esto evita recurrir a mecanismos externos, de tal manera que estas informaciones sean mantenidas hasta el final del tratamiento.

(1) N[lettre, {<num,*a*>]} <==> N[letter, {<num,*a*>}]
(2) N[bagages, {<num,*Pl.*>]} <==> N[luggage,{<num,*Sg*.>}]

El apareamiento es diferente para *luggages*, (2), en la medida en que la información morfológica esta modificada en el momento de la transferencia.

b) Estructuras sintácticas

Las estructuras sintácticas del idioma de origen están sometidas a un mecanismo llamado de recodificación, el cual nos permite efectuar el apareamiento de las estructuras casi isomorfas.

Suscribimos de esta manera al concepto de isomorfia propuesto en los estudios de Rosetta.

La recodificación puede definirse como la transferencia dentro de un idioma dado, de manera que resulte una estructura que sea lo más general posible. Consiste, por una parte, a simplificar las estructuras y por otra parte a eliminar las informaciones idiosincráticas y reeemplazarlas por propiedades más generales.

---

[2] Esta reducción podría también aplicarse al nivel léxicoconceptual restringiendo, por ejemplo, elconjunto de los sentidos propios a un verbo.

Los rasgos sintácticos seleccionados para describir las propiedades abstractas deseadas deben determinarse de manera que se apoyen sobre los análsis refinados que solo las teorías linguísticas pueden ofrecernos.

De esta manera, conocemos que existe en francés una expresiónparticular que nos permite construir preguntas incluyendo *est-ce-que* a continuación de un pronombre interrogativo, como en el caso *qu'est-ce que le produit contient ?* (¿Qué contiene el producto?). Todo lo que quedará en el árbol sintáctico, después de la recodificación, es la información de una construcción interrogativa determinada.

La recodificación linguística permite entoncés subsumir todo lo que es particular al idioma de origen de manera que se pueda asentar la transferencia sobre los rasgos linguísticos comunes a todos los idiomas. En otras palabras, gracias a la recodificación, obtendremos una estructura adecuada para realizar la traducción.

**4.3. Apareamiento de las estructuras lexicoconceptuales.**

Efectivamente se trata de establecer una correspondencia entre las unidades lexicales de la lengua de origen, sean éstas simples o complejas ( ver 3.3.), y los de la lengua terminal. Aquí, nos distanciamos de Schenk, quien para esquivar toda forma de transferencia, introduce en el léxico *falsos idiomas.*

En efecto, codificando las lenguas como *Complex BasicExpressions* (CBE), se ve obligado a codificar en un árbol una expresión completamente literal de la lengua de origen, de tal manera que pueda responder de su traducción en la lengua terminal, sea idiomática o que corresponda simplemente a una palabra.

De esta manera, la expresión literal inglesa *get up early* debe ser codificada como una CBE en el léxico inglés, de tal manera que en el español se obtenga, a la traducción, la palabra *madrugar*.

Decidimos más bien operar este apareamiento durante la transferencia propiamente dicha.

Definimos así un léxico de transferencia propio a cada par de idiomas. Este léxico podrá contener apareamientos tales como :

1- unidad simple <==> unidad simple
1- unidad simple <==> unidad compleja
1- unidad compleja <==> unidad compleja

Al igual que los apareamientos sintácticos, los apareamientos relativos al léxico pueden ser subespecificados.

Recordamos que los léxicos son estructurados de tal manera que permitan su evolución progresiva. Además, a medida que surgen necesidades como resultado del desarrollo progresivo del sistema, integraremos las informaciones lexicoconceptuales. Los estudios de Pustejovski aparecen los más adaptados a la computación ya las descripciones flexibles.

Por el momento, no hemos retenido el marco de las estructuras conceptuales (SC, Dorr, Jackendoff), porque constatamos que éstas nos hacen perder informaciones en el curso del tratamiento.[3]

---

[3] Por otra parte, el estudio detallado de los trabajos de Dorr (1994) nos indica que sus primitivas conceptuales son tomadas delos conocimientos del mundo. Así, para emparejar el verbo *stab* con una traducciónespañola, o sea *dar puñaladas*, se aumentan las representaciones a la Jackendoff de una primitiva conceptual *sharp-objet*. p.476.

Aunque no se ha tocado el tema en este trabajo, creemos que los conocimientos del mundo no pertenecen al sistema de TA y que solamente el utilizador puede decidir cuando surge un problema.

### 4.4. Resultado: automatización de todos los niveles de la transferencia

En este punto, llegamos a la automatización de todos los niveles de la transferencia, es decir, la orientación hacia una estructura de discurso idéntica en los tres idiomas, la recodificación y la transferencia, tanto morfosintáctica como lexicoconceptual.

Sabemos que en este punto, el proceso de traducción está constituido de tres etapas precisas:

- la aplicación de reglas de recodificación;
- la aplicación de reglas de la transferencia relativa al léxico y a la morfosintaxis en la lengua terminal;
- generación de la frase en la lengua terminal.

Nos queda por comprobar hasta que punto al nivel discursivo tendremos que producir un módulo de transferencia o bien si se tratará únicamente de una forma de apareamiento que permitirá solamente guiar al resto del proceso de traducción.

Un sistema de traducción concebido de esta manera permitirá que los mecanismos de transferencia se apliquen sobre estructuras cuyas representaciones serán libres de toda condición o de toda limitación, que en caso contrario hubieran debido ser introducidas en diferentes momentos para guiar el pasaje de un idioma a otro.

## 5. Mecanismo de transferencia

El mecanismo de transferencia se aplica a diferentes niveles, o sea en la transferencia relativa al léxico (de palabra a palabra), en latransferencia categorial (de categoría a categoría) y en la transferencia estructural.

De manera general, está constituido de un conjunto de reglas quevan desde las reglas estructurales hasta las reglas relativas al léxico. El enfoque que privilegiamos consiste en intentar conservar la rupturaentre los niveles (es decir, por ejemplo al no encontrar la codificación de una transferencia relativa al léxico en las reglas de transferencia estructural) dentro de lo posible. Sin embargo, y aunque las reglas de transferencia se encuentren en módulos separados, los niveles comparten el mismo formalismo.

Para ilustrar lo planteado, adjuntamos un extracto reducido de lo que podría constituir las reglas de la transferencia estructural y a la vez de la transferencia relativa al léxico.

Todas estas reglas respetan el mismo tipo de sintaxis: un lado izquierdo y un lado derecho, los dos representados por un árbol dedependencia. La sintaxis de las reglas es la siguiente (el símbolo´´|´´ identifica a un O lógico, la Δ identifica las variables y “ ”encuadran a los elementos literales):

**Nombre de categoría** nombre de la categoría, tal cual aparece en un árbol

**Propiedad** nombre y valor de las propiedades relativas al léxico, a la sintaxis o a la semántica

**Palabra con sus flexiones** palabra que conlleva todas las concordancias (ej.: venden)

**Palabra básica** palabra que no conlleva ninguna información morfológica (ej.: vender)

**Palabra de relación** nombre de relación tal cual aparece en un árbol

**Secuencia** toda secuencia de caracteres (ej. : Δx es una variable)

1. Regla -> Árbol <==> Árbol
2. Árbol ->(Nudo Dependiente* )
3. Nudo ->(Categoría, Elemento, Conjunto_de_rasgos)
4. Dependiente ->(Relación, Árbol)
5. Categoría -> NombreDeCategoría | Variable
6. Elemento -> "Palabra con sus flexiones" |PalabraBásica | Variable
7. Conjunto_de_rasgos -> { <rasgo, valor>* }
8. Relación-> NombreDeRelación | Variable
9. Variable ->Δ Secuencia

La organización de las informaciones bajo la forma de conjunto de rasgos nos permite recurrir a la subespecificación de los elementosrelativo al léxico y de esta manera introducir en el formalismo variables y constantes a todos los niveles.

El mecanismo que nos proponemos utilizar y que hemos tenido la oportunidad de construir y comprobar, es el recubrimiento de los árboles sintácticos por las reglas de la transferencia, mecanismo elaborado por Watanabe (1986). La estructura de salida está construida combinando, de manera no destructiva, las reglas de la transferencia que pueden ser las más utilizadas de manera bidireccional.

Como lo describen las reglas presentadas arriba, conservaremos e inclusive aumentaremos el formalismo de las CBE de Rosetta, adaptándolo a las necesidades de la transferencia.

En cuanto al problema de los *falsos idiomas,* se lo trata más bien en nuestro módulo de transferencia con las reglas correspondientes. Por ejemplo, la asociación entre *get upearly* y *madrugar* se efectúa al nivel de la transferencia lexicoconceptual.

Por lo tanto, el principio que privilegiamos es el de la transferencia pura y simple.

**6. Conclusión**

El paso de los ciclos sucesivos constituidos por los tres grandes módulos nos permitirá desarrollar una experiencia sobre el procedimiento a seguir (metodología) para continuar en la TA de otros tipos de cartas y por consiguiente de documentos más diversificados.

Nos preguntamos si no sería mas oportuno comenzar por la traducción inglés-español, para responder a una necesidad del TLC.

Necesariamente, el hecho de ir acumulando experiencia y competencianos conducirá a cambiar y ajustar el rumbo.

Finalmente, no hemos abordado el último módulo, el de la generación. Pensamos que solamente como resultado de la experiencia en la aplicación integral del modelo, podremos resolver adecuadamente este problema.

**Bibliografía**


ABEILLÉ A., SCHABES, Y. AND JOSHI, A. K., *Usinglexicalized tags for machine translation*. In **InternationalConference on Computational Linguistics COLING** (1990).

BADIA, T., *Dependency and Machine Translation*. In **Linguistic Issues in Machine Translation**. Frank Van Eynde(ed), Pinter (1993), pág. 105-137.

CLAS, A., HORGUELIN, P.A., **Le Français, langue des affaires**. Mc Graw-Hill (1984).

DORR, B. J., *Interlingual machine translation: a parametrized approach*. In **Natural language processing**, F. C. Pereira and B. J. Grosz, editors, MIT Press (1994), pág. 429-492.

DYNDA, A. ET Zavada, D., **Commercial correspondence, GuÌa de correspondencia commercial,** Paraninfo, 1988.

HUTCHINS, W. J., AND SOMERS, H. L., **An Introduction to Machine Translation**. Academic Press (1992).

JACKENDOFF, R., **Semantic Structures**, MIT Press (1990).

MARUYAMA, H. ET WATANABE, H., *Tree cover search algorithm for example-based translation*. In **4th International Conference on Theoretical and Methodological Issues in Machine Translation** (1992),pág. 173-184.

PUSTEJOVSKY, J., **The Generative Lexicon**. MIT Press (1995).

RATTÉ, S., *Classification automatique de la correspondance commerciale*. Mémoire de maîtrise (M. Sc.), Université du Québec à MontrÈal (1988)

SCHENK, A., *Idioms in the Rosetta machine translation system*. In **International Conference on Computational Linguistics COLING** (1986), pág. 319-324.

SEARLE, J. R., *A Taxonomy of Illocutionary Acts*. In J.R. Searle, **Expression and Meaning**, Cambridge University Press(1979), pág. 1-29.

TOUSIGNANT, É., **PrÉsentation de la correspondance d'affaires**, Éditions de la Liberté, 1995.

TSUJI, J. AND FUJITA, K., *Lexical transfer based on bilingual signs : Toward interaction during transfer*. In **Seoul International Conference on NLP** (1990).

WATANABE, H., *A Similarity-driven Transfer System*. In **International Conference on Computational Linguistics COLING**(1986), pág. 770-776.