

Divergencias en la traducción del mapudungün al castellano, una revisión de la propuesta de Bonnie Dorr

Andrés Chandía

andres@chandia.net

Màster Interuniversitari en ciència cognitiva i llenguatge

Asignatura de Aplicacions de Processament del Llenguatge Natural

Profesoras Irene Castellón Masalles, Mariona Taulé Delor y M. Antònia Martí Antonín

Universitat de Barcelona

Septiembre de 2007

Unde sit, ut Lingua Chilensis sit Lingua multo facillima; et tametsi Barbarorum, nom solum non barbara; sed aliis Linguis tanto melior; ut sicuti montes Andes alios montes; ita hæc alias Linguas usque superemineat; ut qui callet Chilense Idioma, alias Linguas velut e specula longe infra se despiciat, clare patenterque agnoscens quantum sit in illis superflui, quam multa quoque desint, et cetera, et quam jure merito unicuique non Chileno exprobrari posit: Si tua Lingua bona est, tamen est praestantior illa Lingua Chilena [...].¹
Bernhard Havestadt, 1777.

Introducción

El siguiente trabajo intenta llegar a algunas conclusiones en cuanto al problema de las divergencias lingüísticas producidas durante la traducción automática de una lengua a otra, en nuestro caso entre el mapudungün y el castellano, y después de aplicar los postulados que nos entrega Bonnie Dorr en su “Machine Translation Divergences: A Formal Description and Proposed Solution”, publicado en 1993 por la Association for Computational Linguistics².

1 La lengua chilena es una lengua facilísima; aun cuando es una lengua de bárbaros, no sólo no es una lengua bárbara, sino que es tanto mejor que otras lenguas; como los Andes superan a otras montañas, así sobresale esta lengua por sobre las otras. Quien conozca la lengua chilena verá otras lenguas muy por debajo de sí, como desde un atalaya, reconociendo de forma clara y patente cuánto hay de superfluo en ellas, como también que carecen de muchas cosas, etc., y a cualquier no chileno puede argüírsele justificadamente y con razón: si tu lengua es buena, la lengua chilena es superior. - Traducción de Fernando Zúñiga, 1995 -.

2 <http://www.aclweb.org/>

La investigadora argumenta sobre la utilidad de crear una interlingua basada en ELC - Estructuras Léxico-Conceptuales (LCS – Lexical Conceptual Structures, en inglés) y de la aplicación de ciertos parámetros, que iremos comentando a lo largo de este documento, para resolver las divergencias producidas durante la traducción automática sin perder la relación sistemática que existe entre interlingua y sintaxis. Es necesario comprender que Dorr desarrolla un sistema que trata las divergencias entre un lenguaje natural y la representación esquemática de un lenguaje, es decir, la interlingua, un lenguaje generado para interrelacionar las lenguas origen y destino, para la autora este concepto es clave ya que con el apropiado tratamiento de las divergencias entre lengua e interlingua es posible generar un sistema al que se puede ir sumando idiomas que potencialmente podrían traducirse a cualquiera de los introducidos con anterioridad al mecanismo de traducción.

Según la lingüista, mediante su investigación ha establecido una clasificación sistemática y formal de las divergencias en Traducción Automática, las que a su vez ha resuelto a través de relaciones o interconexiones (mapping) generalizadas y de algunos parámetros aplicados a los cruces lingüísticos. Especifica, eso sí, que las divergencias que se han clasificado corresponden a aquellas cuyas propiedades se asocian a entradas léxicas que no se basan exclusiva y estrictamente en información sintáctica, uso idiomático, conocimiento aspectual, conocimiento del discurso, del dominio, o del mundo; es decir, sólo a las divergencias léxico-semánticas, que sin embargo cubren en gran parte las posibles divergencias producidas durante la Traducción Automática. Explica cómo, mediante el mecanismo generado, es posible discernir si la traducción a la lengua destino cubre el concepto subyacente en la lengua origen y que incluso es posible generar una “salida” a partir de la aparición de dos o más divergencias simultáneas.

Justificación

La elección del mapudungün, por una parte, es sólo por una motivación personal, existen algunos vínculos entre quien escribe y el pueblo mapuche. Sin embargo, la elección del castellano como la lengua de interacción está motivada en la situación del mapudungün. El pueblo mapuche está ubicado al sur de Sudamérica, con una población repartida entre Chile y Argentina, países casi monolingües en castellano, lo que obliga a los mapuche a saber la lengua hispana para relacionarse con la

población criolla, por lo tanto, podemos estimar que por lo menos el 95% de las traducciones que se necesite hacer del mapudungün tendrán como lengua destino el castellano; a la inversa es, seguramente, otra la proporción que opera, ya que casi el total de la población mapuche conoce el castellano.

En su trabajo, Dorr toma como lenguas de interacción principalmente el castellano y el inglés, algunos de sus ejemplos consideran el alemán y también extrae algunas de sus referencias desde el francés, todas estas lenguas tienen muchos aspectos en común desde el punto de vista tipológico; el mapudungün, aunque en muchos aspectos puede tener equivalencias con estas lenguas, tipológicamente dista de ellas en general - más adelante daremos detalles de esta clasificación de la que hablamos -; esta consideración nos hace pensar que aplicar el marco de investigación propuesto y discutido en el trabajo de la lingüista puede arrojarnos resultados disímiles o confirmar sus afirmaciones, ciertamente esto también depende de nuestro dominio del tema, que puede ser insuficiente.

Definiciones

Para comenzar a adentrarnos en las divergencias y su problemática, en un conocimiento básico del mapudungün y en la resolución de las divergencias particulares³ que se presentan a la hora de la traducción automática desde el mapudungün al castellano haremos referencia a algunos conceptos básicos como el de divergencia, ELC - Estructura Léxico-Conceptual -, ELGR - Estructura Léxico-Conceptual Raíz -, ELCC - Estructura Léxico-Conceptual Compuesta - entre otros que irán apareciendo en este trabajo.

1. Probablemente una caracterización tipológica del mapudungün nos podrá dar una idea aproximada de algunos tipos de divergencias o dificultades en general con las que nos podemos encontrar al momento de ejecutar un sistema de traducción automática. Morfológicamente, en especial en las formas verbales del mapudungün, encontramos un comportamiento polisintético y aglutinante;

3 Las divergencias, tanto las expuestas por la autora como las que encontremos relacionadas al mapudungün, deben entenderse como las generadas dentro de un lenguaje estilísticamente neutro, es decir, tratando de no apartarse del orden canónico de los elementos correspondientes a cada lengua (origen/destino); en modo indicativo, con participantes explícitos, con un sujeto (S) definido, agentivo y humano, un complemento directo (O) paciente, definido, y un verbo que represente una acción en lugar de un estado o evento; claro está, dentro de lo posible.

veamos un ejemplo que lo justifica: *rüngkükonfemtuyami* – inmediatamente saltarás hacia adentro, de vuelta al punto de origen -, esta forma verbal puede ser dividida en constituyentes con significado individual, cuya concatenación da sentido a la forma compuesta.

<i>rüngkü</i>	saltar
<i>kon</i>	adentro
<i>fem</i>	inmediatamente
<i>rüngkükon</i>	saltar hacia adentro
<i>rüngkükonfem</i>	saltar hacia adentro inmediatamente
<i>tu</i>	hubo un salto previo hacia afuera
<i>rüngkükonfemt</i>	saltar hacia adentro inmediatamente regresando al punto de origen
<i>a</i>	futuro
<i>y</i>	enunciado de hecho (no hipótesis)
<i>m</i>	segunda persona
<i>i</i>	singular

Así justificamos la polisíntesis al ver que el verbo mapuche es de constitución compleja, formado por una serie de elementos significativos que se concatenan en un orden regulado. Y la aglutinación, al observar la identidad formal, propia, definida y constante de sus componentes, que los hace claramente segmentables. Además podemos observar incorporación en la estructura del mapudungün, dos o más categorías léxicas pueden aparecer en la estructura de las formas verbales:

<i>kutran</i>	enfermo, doliente, adolorido
<i>foro</i>	diente

se unen sin perder su identidad individual para generar

<i>kutranforo</i>	adolorido de los dientes
-------------------	--------------------------

que además puede incorporarse a una forma verbal como

<i>kutranfororkey</i>	dicen que él estaba adolorido de los dientes
-----------------------	--

en donde la encontramos aglutinada con los componentes de las formas verbales *-rke* (dicen que) e *-y* (enunciado de hecho), que carecen de autonomía.

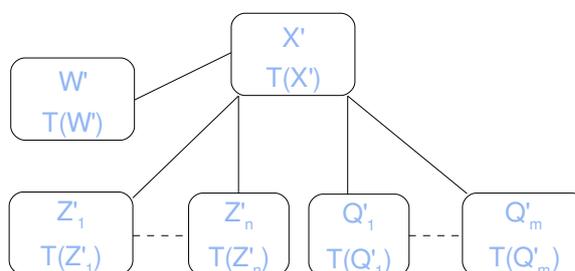
Así podemos generalizar diciendo que en mapudungün, especialmente en los verbos, hay expresiones léxicas que pueden contener un gran número de morfemas como en *di-tu-l-me-tu-a-fi-ñ* ('se lo alcanzaré' o 'se lo encontraré') y que estos morfemas generalmente aparecen con una forma estable, por ejemplo el sufijo de futuro *-a* en *ditulmetuafiñ* aparecerá con la misma forma *-a*

en otros compuestos verbales para cualquier persona.

El hecho de que el mapudungün sea polisintético revela una gran capacidad para crear palabras nuevas, y que sea aglutinante la hace una lengua regular, por tanto más manejable desde determinados puntos de vista; por ejemplo, el mapudungün tiene muchas construcciones gramaticales basadas en la morfología más que en la sintaxis, por ejemplo, la voz pasiva está expresada mediante la partícula *-nge*, mientras que en castellano debemos utilizar un auxiliar que expresa el tiempo, debe cumplir con exigencias de concordancia y acompañar a un participio pasado. *pengen* se traduciría como “*fui visto*”, aunque aquí deberíamos entrar en una de las particularidades del mapudungün para ser más fieles al original, la inversión, de la cual hablaremos más adelante.

2. Las Divergencias no son otra cosa que los cruces lingüísticos que se generan entre la lengua origen y la lengua destino, cruces lógicos dadas las estructuras disimiles propias de cada lengua, sin embargo las que se producen en el proceso de la Traducción Automática, según comenta Dorr y como ya explicamos, son aquellas derivadas de la interconexión entre lengua origen e interlingua. En Traducción Automática, al tener las entradas léxico-semánticas de cada una de las lenguas en interacción, lo que correspondería a un modelo fijo de transferencia⁴, podemos observar y clasificar las divergencias que se producen.

3. La Estructura Léxico-Conceptual (ELC) es la representación léxico-semántica utilizada para exponer la estructura de la interlingua cuya representación formal se aprecia en el siguiente diagrama

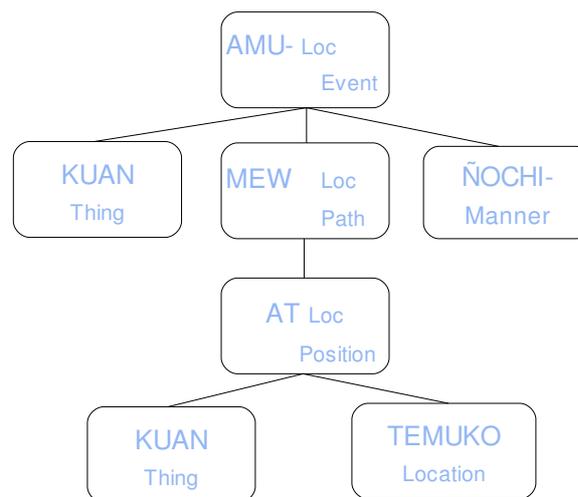


En este diagrama X' es el *núcleo lógico* de la frase; W' , el *sujeto lógico*; Z'_1, \dots, Z'_n son los *argumentos lógicos*; y Q'_1, \dots, Q'_m son los *modificadores lógicos*. $T(\Phi)$ es

4 Una entrada léxica de la lengua 1 tiene n correspondencias definidas en la lengua 2 y a la inversa, y ambas a su vez tienen una representación asignada en la interlingua

la *categoría conceptual*, el *tipo lógico* en palabras de Dorr (THING – EVENT – STATE – PLACE – PATH – PROPERTY – PURPOSE – MANNER – AMOUNT – TIME - “cosa – evento – estado – lugar – recorrido – propiedad – objetivo – manera – cantidad – tiempo), que corresponde a la primitiva conceptual Φ (BE – GO – STAY – CAUSE – INCH – EXT – EXCH – ORIENT - “estado – evento – BE + duración – expresión de causalidad – interpretación incoativa - extensión espacial - intercambio - orientación del objeto”, etc.), generalizadas en *campos semánticos* (localización “+loc” - tiempo “+time” - posesión “+poss” - características de la identidad “+char, +ident” - composición de la entidad “+comp, +ident”) . Las posiciones expresadas en el diagrama son vías de interconexión entre la representación en la interlingua y la representación sintáctica superficial del lenguaje natural.

La representación de la ELC para *kuan ñochika amuy temuko mew* (Juan va lentamente a Temuco) es:



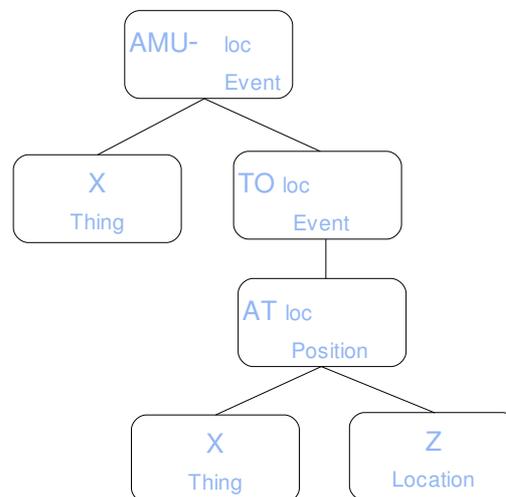
En este ejemplo el *núcleo lógico* es AMU-⁵_{Loc} (de la categoría Event); el *sujeto lógico* es KUAN (de la categoría Thing); el *argumento lógico* es MEW_{Loc} (de la categoría Path) formado a su vez por una ELC que contiene su propio *argumento lógico*, TEMUKO (de la categoría Location); el *modificador lógico* es ÑOCHI-⁶ (de la categoría Manner).

5 amu- corresponde a la raíz verbal de *ir* que se realiza junto a otros sufijos, sean estos obligatorios u opcionales, por lo tanto la forma que aparece en el esquema no es la que encontramos en el lenguaje natural en donde siempre la encontraremos con los sufijos correspondientes.

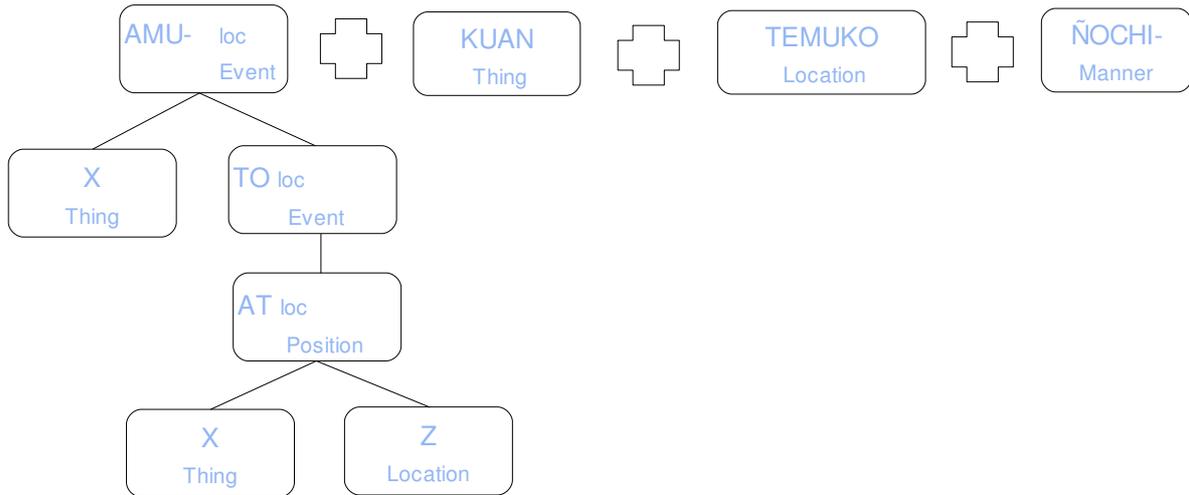
6 ñochi- corresponde a la raíz verbal de “ser o estar lento”, que en este caso viene complementada por el sufijo continuativo -ka, por lo tanto *ñochika* es algo así como “ser lento continuamente” que aplicado a otra forma verbal sería como el adverbio “lentamente”.

Tanto para la representación léxica como para la representación de la interlingua se utiliza el esquema de la ELC - Estructura Léxico-Conceptual -, en el caso de la representación léxica hablaremos de la ELCR - Estructura Léxico-Conceptual Raíz - y en el caso de la representación de la interlingua hablaremos de la ELCC - Estructura Léxico-Conceptual Compuesta, una de cuyas propiedades es la recursividad.

(a) ELCR - Estructura Léxico-Conceptual Raíz, esta es una ELC descontextualizada cuyas posiciones variables no están ocupadas, corresponde a una entrada léxica. Veamos el siguiente ejemplo de representación de la ELCR de amu- (ir)

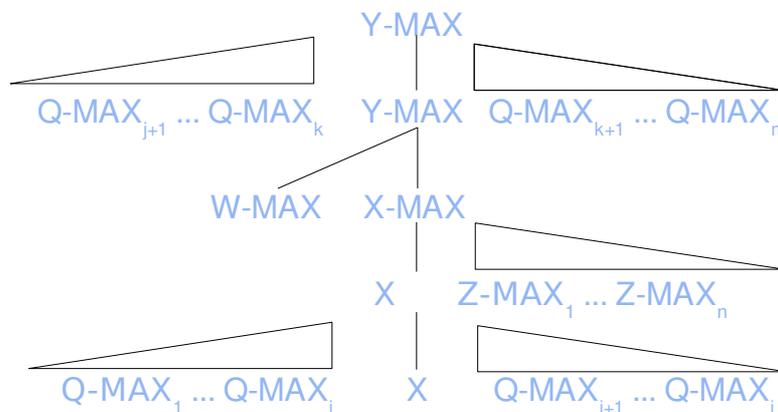


4. ELCC - Estructura Léxico-Conceptual Compuesta, esta es una ELC contextualizada producto de la unión de dos o más ELCR. De esta manera se estructura la interlingua que es una representación independiente del lenguaje natural y a través de la cual se relacionan los idiomas origen y destino que entran en juego durante la traducción. Si unimos la ELCR para AMU- con la de KUAN, la de TEMUKO y la de ÑOCHI-, obtenemos la ELCC correspondiente a *kuan ñochika amuy temuko*

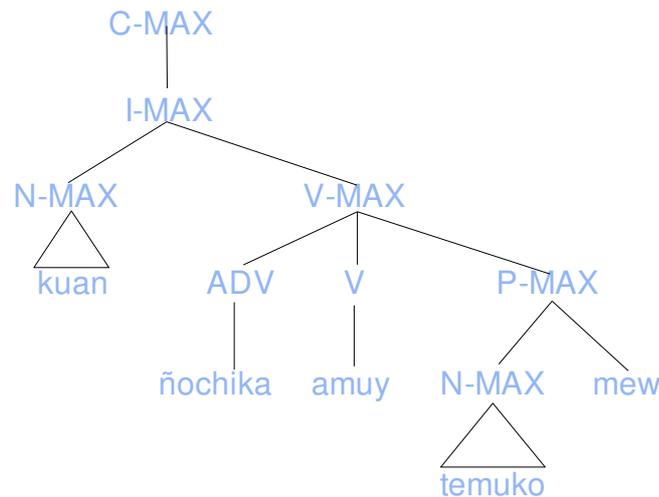


Cada categoría léxica tiene su ELCR cuyas posiciones variables presentan ciertas restricciones. La composición de ELC consiste en la unificación de las ELCR de las categorías léxicas con contenido que encontramos al interior de la *frase sintáctica* cuyas relaciones se establecen siguiendo las restricciones propias de cada pieza. Y es mediante este proceso que obtenemos la ELCC, es decir cada elemento se encaja en su posición deslizándose entre las restricciones que lo van acomodando en la representación final a la vez que se complementa con información léxica producto de la interconexión entre los componentes de la frase (mapping).

5. La *frase sintáctica* corresponde a la representación esquemática propuesta por Chomsky a través del diagrama de X' (X con barra) desarrollado en la Teoría del Régimen y Ligamen Gramatical, es la proyección sintáctica máxima de una frase y es un componente esencial para establecer las interconexiones entre la representación de la interlingua y la representación sintáctica superficial.



En este esquema X es el *núcleo sintáctico* (cuya categoría es V, N, A, P, I, o C); W-MAX es el *argumento externo*; Z-MAX₁, ..., Z-MAX_n son los *argumentos internos* y Q-MAX₁, ..., Q-MAX_m son los *adjuntos sintácticos*. Para nuestra frase la representación es la siguiente:



En donde el *núcleo sintáctico* es **V amuy**; el *argumento externo* es **N-MAX kuan**; el *argumento interno* es **P-MAX mew**, que a su vez es una *frase sintáctica* que contiene un *argumento interno*, **N-MAX temuko**, que confirma la recursividad de estas estructuras; el adjunto sintáctico es **ADV ñochika**.

6. Las interconexiones entre la representación de la interlingua y la representación sintáctica superficial de las que hablamos algunas líneas más arriba, se llevan a cabo mediante dos rutinas lingüísticamente definidas y formalizadas:

(a) Rutina de Enlace Generalizada (REG) – Generalized Linking Routine (GLR) en inglés – este proceso relaciona sistemáticamente las posiciones sintácticas de la ELC y las posiciones léxico-semánticas de la *frase sintáctica*, tal como se indica a continuación.

- i. $X' \Leftrightarrow X$
- ii. $W' \Leftrightarrow W$
- iii. $Z'_1 \dots Z'_n \Leftrightarrow Z_1 \dots Z_n$
- iv. $Q'_1 \dots Q'_m \Leftrightarrow Q_1 \dots Q_m$

Este concepto aplicado a nuestra frase nos generaría la siguiente representación:

- i. $X' = \text{AMU}_{\text{Loc}} \Leftrightarrow X = [\text{V amuy}]$
- ii. $W' = \text{KUAN} \Leftrightarrow W = [\text{N}_{\text{MAX}} \text{kuan}]$
- iii. $Z' = \text{MEW}_{\text{Loc}} \Leftrightarrow Z = [\text{PP mew}]$
- iv. $Q' = \text{ÑOCHI} \Leftrightarrow Q = [\text{ADV} \text{ño chika}]$

(b) La Realización Sintáctica Canónica (RSC) – Canonical Syntactic Realization (CSR) en inglés – es el proceso que interconecta la categoría conceptual o tipo léxico-semántico $T(\Phi')$ con una categoría sintáctica $CAT(\Phi)$, en donde Φ' es un constituyente de la ELCC relacionada con el constituyente sintáctico Φ por la REG - Rutina de Enlace Generalizada. Por ejemplo, la categoría conceptual de ELC *Thing* corresponde a la categoría sintáctica N, que finalmente se proyecta a su nivel máximo (N-MAX), de la misma manera podemos obtener las siguientes Realizaciones Sintácticas Canónicas (RSC) de la interrelación entre las Categorías conceptuales de las ELC y las Categorías Sintácticas.

Categoría conceptual de ELC	Categoría Sintáctica
EVENT	V
STATE	V
THING	N
PROPERTY	A
PATH	P
POSITION	P
LOCATION	ADV
TIME	ADV
MANNER	ADV
INTENSIFIER	ADV
PURPOUSE	ADV

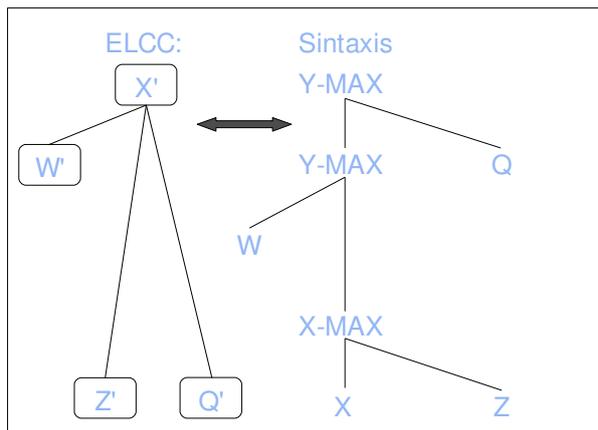
Ahora que ya tenemos nuestras definiciones de los elementos que participan del proceso de interconexión entre la lengua origen y la interlingua, y entre ésta y la lengua destino nos dirigiremos hacia las divergencias y a tratar de comprobar si la manera en que Bonnie Dorr resolvió esta problemática se aplica a una lengua como el mapudungün. Es preciso mencionar que la autora postula que las divergencias ocurren cuando se presenta una excepción en la REG - Rutina de Enlace Generalizada - o en la Realización Sintáctica Canónica (RSC), o en ambas a la vez en una de las lenguas, pero no en la otra, y es precisamente este hecho el que permite definir y clasificar formalmente las divergencias léxico-semánticas generadas durante el proceso de la Traducción Automática.

La lingüista dice que para resolver el problema debemos extraer y tratar la información generada por la REG y la RSC, y aplicarles unos cuantos mecanismos paramétricos para la codificación de la información léxica de una lengua en particular; al contrario que estos mecanismos la REG y la RSC se mantienen independientes de las lenguas. Los parámetros que suplen las excepciones a la REG y a la RSC son siete: :INT - :EXT - :PROMOTE - :DEMOTE - * - :CAT - :CONFLATED, que iremos definiendo a medida que avancemos.

Divergencias

1. Divergencias Temáticas

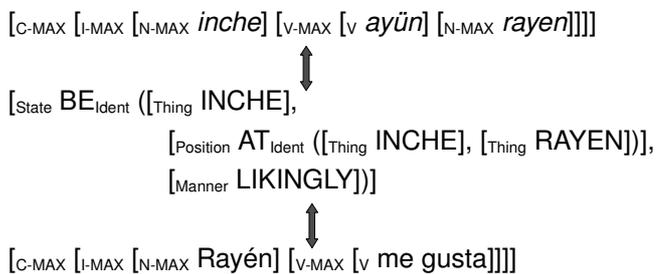
En este tipo de divergencias ocurre un reposicionamiento de los argumentos con respecto al núcleo de la frase, es decir, cuando en la REG no se siguen los pasos 2 y 3, y en cambio se generan las siguientes relaciones: $W' \Leftrightarrow Z$; $Z' \Leftrightarrow W$, representadas como sigue:



Este tipo de divergencia se presenta solamente cuando existe un sujeto lógico.

M: *ayün rayen* <=> C: Rayén me gusta

En este caso tenemos un intercambio de objeto y sujeto entre la lengua origen y la lengua destino, en mapudungün *rayen* es el tema expresado como objeto verbal, mientras que en castellano es el sujeto del verbo (literalmente, la frase mapuche se traduciría como “gusto de Rayén”), por su parte el sujeto tácito *inche*⁷ toma la posición del objeto “me” en castellano. Las correspondientes estructuras sintácticas y ELCC son estas:



La forma en que la autora resuelve esta divergencia es mediante un par de los parámetros antes mencionados, :INT y :EXT, éstos evitan que se produzca la interrelación generada por la REG con respecto al posicionamiento del sujeto y argumentos lógicos en castellano, para ilustrar el uso de estos parámetros recurriremos a las entradas léxicas para *ayün* y *gustar*.

7 Si el sujeto hubiese estado explícito en la frase se tomaría como un intensificador o un desambiguador, por ejemplo, a la pregunta “¿le gustas a Rayén?”, la respuesta “Rayén me gusta a mí”, se obtendría con

<i>inche</i>	<i>ayü</i>	-n	<i>rayen</i>
yo	gustar.de	-indicativo-1ª pers.-singular	Rayén

Como se ve, el pronombre personal de primera persona *inche* está reforzando al sufijo -n de la forma verbal.

(a) Entrada léxica de *ayün*:

[State BE_{Ident} ([Thing W],
 [Position AT_{Ident} ([Thing W], [Thing Z])],
 [Manner LIKINGLY])]

(a) Entrada léxica de *gustar*:

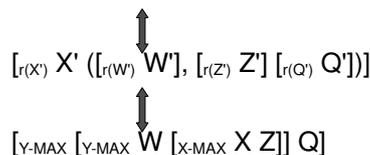
[State BE_{Ident} ([Thing :INT W],
 [Position AT_{Ident} ([Thing W], [Thing :EXT Z])],
 [Manner LIKINGLY])]

Como vemos, para el mapudungün no necesitamos la especificación de parámetros adicionales, por lo tanto los constituyentes de la estructura ocupan la posición especificada por la REG; en castellano, por el contrario, los elementos marcados por los parámetros :INT y :EXT deben cambiar de posición al componer la estructura sintáctica, operando de la siguiente manera:

ELCR 1: [_{r(X)} X' ([_{r(W)} W'], [_{r(Z)} Z'] [_{r(Q)} Q'])]⁸

ELCR 2: [_{r(X)} X' ([_{r(W)} :INT W'], [_{r(Z)} :EXT Z'] [_{r(Q)} Q'])]

Traducción: [_{Y-MAX} [_{Y-MAX} W [_{X-MAX} X Z]] Q]



Explicaremos aquí una de las particularidades del mapudungün para probar si es posible resolver las divergencias en estos casos mediante el mismo proceso antes explicado.

Las formas verbales transitivas del mapudungün se expresan de manera directa o inversa, la primera es aquella que expresa una acción ejecutada por una primera o segunda personas y que recae sobre una tercera

*ayü-y-u*⁹ *rayen*
 gustar.de-indicativo-1ª pers.¹⁰-dual Rayén
 "a nosotros dos nos gusta Rayén"

Si la acción recae sobre la primera o segunda personas y es ejecutada por una

8 La notación _{r(Z)} es una representación simplificada para denotar un elemento del conjunto {Z₁ ... Z_n}, lo mismo para Q.

9 Por simplicidad no hemos colocado la forma (probablemente más correcta) *ayüfiyu*; el sufijo *-fi* indica que existe un receptor concreto de la acción del verbo, que en el caso de no ser explícito puede subentenderse.

10 La forma debería ser *ayü-y-i-u*, la *i* en negrita corresponde a la primera persona, pero en este caso el morfema no aparece porque se produce asimilación fonológica.

tercera, en vez de alterar el orden de los componentes de la forma verbal, se agrega un par de sufijos que indican que la acción es inversa

*ayü-e-y-u-mew*¹¹ *rayen*
gustar.de-inversión-indicativo-1ª p.-dual-3ª p. agente.forma.inversa Rayén
“nosotros dos le gustamos a Rayén” (Rayén gusta de nosotros dos).

Para dos terceras personas también opera este sistema, sin embargo aquí entran en juego los conceptos de persona focal y satélite; en la frase

achawall pe-fi *kodkod*
gallina ver-3ª pers.determinada.paciente gato.montés, o
gallina ver-lo gato.montés
“la gallina vio al gato montés

achawall es la **Persona Focal** y *kodkod* la **Persona Satélite**, en esta frase la persona focal se comporta como sujeto y la persona satélite como complemento directo, sin embargo en una frase como

achawall pe-e-y-ew *kodkod*
gallina ver-inversión-indicativo-3ª p.-sing.-3ª p. en 2ª posición.agente gato.montés

tendríamos que jugar con las alternativas de traducción al castellano, por ejemplo podríamos generar una forma pasiva “la gallina fue vista por el gato montés”, en donde “la gallina” es sujeto y “por el gato montés” pasa de ser complemento directo en la frase mapuche a complemento circunstancial o adjunto en la frase castellana. Otra posible traducción sería “a la gallina la vio el gato montés” con la cual conservaríamos las funciones de la frase original en mapudungün. Incluso podríamos hacer que operara la inversión sobre la estructura superficial y obtener la frase “el gato montés vio a la gallina”; sin embargo en esta última traducción no se refleja que “la gallina” es la persona focal y “el gato montés” es la satélite. Entre dos terceras personas es focal aquella introducida con anterioridad en el discurso, y satélite aquella introducida posteriormente para interactuar con la primera, por ejemplo, en un cuento sobre una puma encontramos

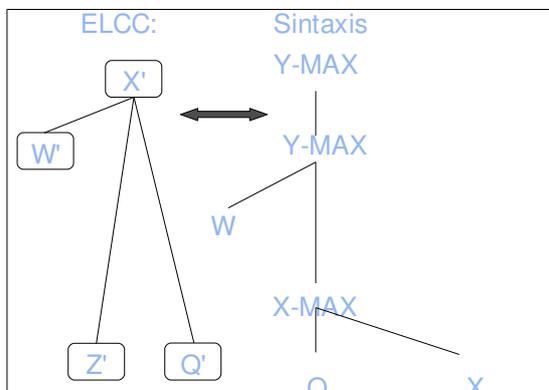
11 A este sufijo le podríamos considerar doble porque tiene una primera realización antes del sufijo que indica modo verbal y una segunda, que expresa focalización del segundo participante cuando es tercera persona, al final de la forma verbal. Salas lo identificó como un sufijo discontinuo, porque para él el sufijo *-e* de *leli-e-n* “me miraste” es distinto al que aparece en *leli-e-n-ew* “me miró” que viene acompañado del sufijo *-(m)ew*, Salas identifica al de *leli-e-n* como indicador de una persona satélite singular segunda; sin embargo este sufijo aparece en todas las formas inversas, si la segunda forma *-(m)ew* está presente siempre está la primera forma *-e* antes del sufijo de modo verbal, pero no a la inversa; esto es así porque *-e* es la marca de formas inversas finitas y no finitas, mientras que *-(m)ew*, que sólo aparece cuando hay inversión, está presente cuando el agente corresponde al participante pospuesto al verbo en la estructura superficial, y es una tercera persona como en *leli-e-n-ew*, por lo tanto, estamos ante dos sufijos uno de los cuales es complementario del otro, *-(m)ew* de *-e*.

<p>2 → 3</p> <p>ayü- fi- -mi te gusta</p> <p>leli- fi- -mu lo miran (dual)</p> <p>leli- fi- -mün lo miran (plural)</p> <hr/> <p>3' → 3''</p> <p>leli- fi- lo mira</p> <p>etc...</p>	<p>3 → 2</p> <p>ayü- e- y- -mew le gustas</p> <p>leli- e- ymu- -mew los (dual) mira</p> <p>leli- e- ymün- -mew los (plural) mira</p> <hr/> <p>3'' → 3'</p> <p>leli- e- y- -ew lo mira</p> <p>etc...</p>
---	---

Si aplicamos la forma correspondiente del verbo a nivel morfológico, como lo que tenemos que hacer coincidir son las representaciones superficiales, sólo nos tendremos que preocupar por la generación del pronombre adecuado en los casos en que este no venga explícito, como en *ayüeyumew rayen*, para obtener “A nosotros dos nos gusta Rayén”, pero esto ya no es un problema de divergencias. Por el contrario, será cuestión de divergencias cada vez que un verbo como *ayün* se presente en su forma directa, en cuyo caso procederemos como nos detalla Dorr en su trabajo.

2. Divergencias promocionales

Dorr llama promocionales a aquellas divergencias en que un modificador lógico en la forma sintáctica de la lengua origen aparece en la posición del verbo en la lengua destino, es decir, durante la REG se cambian los pasos 1 y 4 por las siguientes relaciones: $X' \Leftrightarrow Z$ y $Q' \Leftrightarrow X$, respectivamente, representadas en el cuadro siguiente:



En la representación sintáctica del esquema no tenemos Z porque no ocurre un

intercambio con X, sino que X se desplaza a la posición de Z que continua en relación de dependencia con X, que el esquema no representa por simplicidad. Esta divergencia, cuando se presenta, se soluciona mediante el parámetro ':PROMOTE especificado en la entrada léxica necesaria para forzar a intercambiar posiciones entre el núcleo y el adjunto (el ejemplo del trabajo de referencia es [_{ADV}usually] <=> [_Vsoler]) en la representación sintáctica superficial de la lengua destino.

En el caso del mapudungün todo dependerá de la estructura que elijamos como equivalente en castellano, veamos un ejemplo que lo explica mejor. Si en mapudungün tenemos la frase

i-ke-la-yimi

comer-**habitual**-negación-indicativo 2ª p. sing.

Podemos generar las siguientes traducciones, todas ellas correctas ya que expresan la idea que subyace en la lengua origen:

“no comes habitualmente”, “siempre comes poco”, “sueles no comer”.

Si elegimos aquellas en que se mantiene como verbo principal la raíz verbal presente en la frase mapuche, obviamente no se generará la divergencia ya que los modificadores verbales serán adverbios o frases adverbiales que conservarán su estatus dentro de la frase, sin embargo si elegimos la representación que contiene dos formas verbales, tendremos que la partícula que modificaba a la raíz verbal de la frase original, ahora es el verbo principal. Podríamos, incluso, formular una frase que creara una disyuntiva aún más complicada, por ejemplo si tuviéramos en nuestra frase otra partícula modificadora como en

i-ufaluw-ke-la-yimi

comer-**simulado-habitual**-negación-indicativo 2ª p. sing.

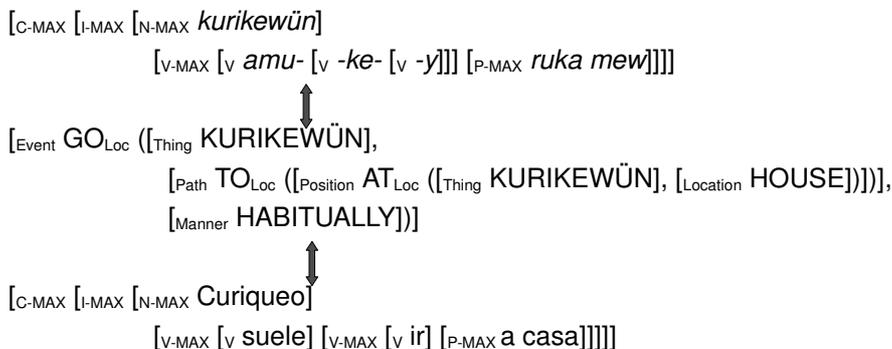
Para la cual, tal vez, tendríamos que exponer una traducción agramatical en primera instancia, pero capaz de expresar la idea subyacente, y posteriormente generar la representación en la lengua destino más ajustada a la estructura canónica del castellano en nuestro caso, por ejemplo,

“*simuladamente no comes habitualmente”, “siempre simulas comer poco”, “sueles simular no comer”.

En vista de esta situación, si nos decidimos por la traducción que genera la divergencia promocional, obtendremos las siguientes representaciones de las

estructuras sintácticas y sus correspondientes ELCC para una frase como:

kurikewün¹⁴ amu-ke-y ruka mew
 Curiqueo ir-habitual-indicativo casa a
 “Curiqueo suele ir a casa”



En efecto, y como adelantamos hace algunas líneas, la partícula (adverbial) *-ke-* que modifica a la raíz verbal *amu-* pasa, en castellano, a ser el verbo principal “soler” cuyo argumento interno es el evento descrito como principal en la frase mapuche, “ir a casa”. El parámetro *:PROMOTE*, que ya hemos introducido, es el encargado de ajustar esta divergencia obligando a que la interconexión (mapping) generada en la REG sea anulada en los pasos que posicionan al núcleo y modificador lógicos, tal como demanda la entrada léxica correspondiente a ambas formas verbales:

(a) Entrada léxica de *-ke-*

[_{Manner} *HABITUALLY*]

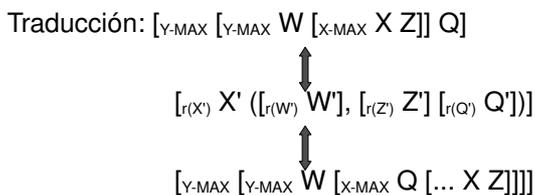
(b) Entrada léxica de *soler*

[_{Manner} *:PROMOTE HABITUALLY*]

Como se aprecia en el siguiente diagrama el paso del mapudungün a la estructura sintáctica está controlado por la REG, mientras que desde el castellano y haciendo uso del parámetro *:PROMOTE* establecido, observamos un intercambio de posiciones entre el núcleo y su modificador:

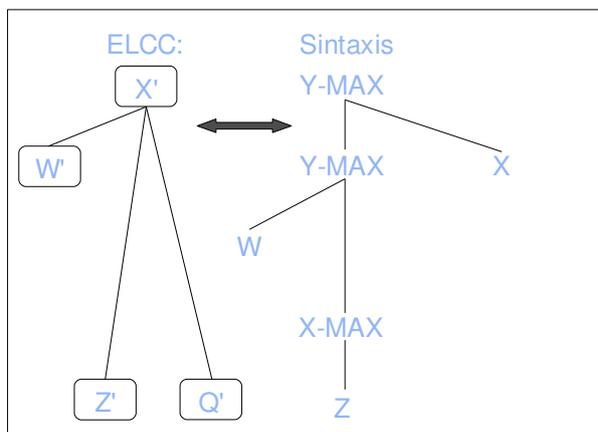
14 *kurikewün* “Palabra oscura” de *kuri* “negro” *-ke* habitual, *kurike* ennegrecido y *wün* que probablemente es una abreviación de *dungün* “palabra, asunto, mensaje, etc.”, nótese el uso del sufijo de habitualidad que en este caso funciona como adjetivo.

ELCR 1: $[_{r(Q)} Q']$
 ELCR 2: $[_{r(Q)} :PROMOTE Q']$



5. Divergencias Descendientes (Demotional)

Este tipo de divergencia describe el cambio de posición que afecta a un núcleo lógico de la frase origen transformándolo en un argumento interno del núcleo en la frase destino, de esta manera el núcleo original se asocia con la posición de adjunto y su argumento lógico con la del núcleo. Es decir, en la REG se produce un bloqueo de los pasos 1 y 3 de esta rutina y se realizan de esta manera: $X' \Leftrightarrow Q$ y $Z' \Leftrightarrow X$, respectivamente; en el primer caso no existe un intercambio de posiciones, la relación continúa siendo la misma, sólo que ahora X' pasa a ser un modificador de un núcleo superior, mientras que Q sigue siendo su modificador (si está presente).



Desde el mapudungün al castellano podríamos argumentar que este tipo de divergencia se da a través de la verbalización de raíces o formas nominales, adjetivales e, incluso, adverbiales¹⁵. Por ejemplo, si a la forma nominal *pun'*

15 Al ser raíces, no son exclusivas de una categoría, pero su comportamiento en general apunta a una determinada categoría, los sufijos que suelen acompañar a estas raíces son típicos de aquellas categorías a las que nos referimos, a diferencia de las formas completas, es decir, nombres, adjetivos o adverbios, cuya pertenencia a la categoría gramatical correspondiente, es obvia.

“noche” le añadimos el sufijo de modo indicativo *-i*, obtenemos *pun'i* “anocheció”, si agregamos el mismo sufijo al adjetivo *kelü* “rojo”, obtenemos “rojear” - *kelüy ñi mollfüñ* “le rojeaba su sangre” -; otras partículas o sufijos pueden causar el mismo efecto: *teltong* “trote” más el durativo *-küle* producen “ir al trote”, etc. En el mismo contexto tenemos formas nominales que se verbalizan mediante un sufijo temático, así por ejemplo, si tenemos el sustantivo *wen'i* “amigo” junto al sufijo continuativo *-ka* quiere decir “convertir en amigo” como en

wen'ikafi	t'a	chi	pu	mapuche
los.hizo.amigos.él.a.ellos		los	plural	mapuche

“(el diablo) convirtió en sus amigos a los mapuche”

En fin, variadas son las posibilidades de verbalizar las diferentes formas léxicas del mapudungün, todas ellas mediante partículas o sufijos cuya función consiste en modificar la raíz verbal, por esta razón exponemos como forma de inducción de divergencias este fenómeno. Sin embargo, esta divergencia ocurre a nivel morfológico, por lo tanto lo que posteriormente tenemos en la representación sintáctica es una forma verbal en plena función como tal, en consecuencia, y como propusimos en el punto de divergencias temáticas, la solución de este tipo de problemática pasa por una descripción y clasificación tanto de las raíces verbales como de los sufijos verbalizadores en este caso. Sólo a modo de ejemplo veamos la siguiente frase.

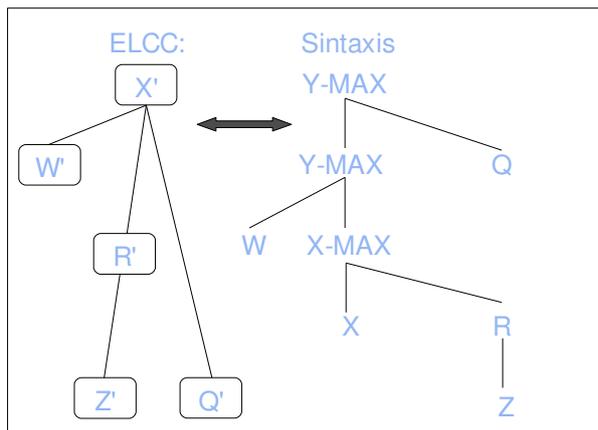
fey	chi	epeko	trükür-uw-i
ese	el	espejo	niebla-transformación-indicativo
ese	el	espejo	convertirse.en.niebla-indicativo

“el espejo se convirtió en niebla”

En el caso de esta frase tenemos una forma verbal que proviene de una nominal *trükür* “niebla”, que para los efectos de nuestra explicación, y al ser la raíz corresponde al núcleo lógico, y la partícula *-uw-*, que indica transformación en el sustantivo al que acompaña, es el modificador. En la forma castellana, “niebla” también es una forma nominal, sin embargo en este caso pertenece a un sintagma preposicional que ocupa el lugar de argumento lógico, pero repetimos, es un proceso morfológico.

6. Divergencias estructurales

Por divergencia estructural entendemos aquella que, a diferencia de las anteriores, no produce cambios en la REG, sino que establece una correspondencia diferente entre las distintas posiciones, como muestra el diagrama



que describe cómo la función de complemento se despliega bajo un tipo de categoría sintáctica en la lengua origen y bajo otro distinto en la lengua destino.

Desde el mapudungün al castellano nos podemos encontrar con esta divergencia al traducir algunas oraciones que en ciertos contextos la producen y en otros no, expondremos aquí una serie de frases para constatarlo y una posible¹⁶ explicación sobre la naturaleza de los contextos:

(a) *ñi chaw engu amuan temuko*
 mi padre ellos.dos iré Temuco
 “iré a Temuco con mi padre”

(b) *ñi chaw iñchiw amuayu temuko*
 mi padre nosotros.dos iremos.los.dos Temuco
 “iré a Temuco con mi padre”

¹⁶ Sólo es una posible explicación porque los datos de los que disponemos son muy pocos, haría falta un corpus más grande y conocer más situaciones que nos brindaran un contraste entre la distintas frases, como también mayor contacto con informantes que corroboraran las distintas utilizaciones contextuales, sin embargo creemos que no estamos muy alejados con nuestra explicación.

- (c) *eymi amuaymi temuko mi chaw eymu*
tú irás Temuco tu padre ustedes.dos
“irás a Temuco con tu padre”

En la versión castellana de estas tres frases el verbo “ir” necesita estar complementado por un sintagma preposicional encabezado por la preposición “a” para indicar dirección o destino, su equivalente mapuche, por el contrario, realiza este complemento con un nombre, o un sintagma nominal si se prefiere. Con respecto al adjunto comitativo no debemos preocuparnos porque aunque todos ellos se expresan en castellano en forma de sintagma preposicional encabezado por la preposición “con”, su presencia puede resolverse mediante una estrategia diferente de la que resuelve las divergencias. Para ser más concretos, cuando en mapudungün tenemos un segundo participante de compañía se utiliza el pronombre personal correspondiente que equivale a la preposición “con”, por ejemplo, en la frase *ñi chaw engu amuan temuko*, el pronombre de primera persona dual *engu* equivale a “con”. Con estrategia diferente queremos decir un análisis que determine la presencia de este fenómeno y lo resuelva mediante un módulo específico en el proceso de la traducción.

- (d) *amuy temuko küdaw-a-lu*
fue Temuco trabajar-futuro-finalidad
“fue a Temuco a trabajar”

En esta frase nos encontramos con una situación análoga a la anterior, la única diferencia es que aquí el adjunto, que también es un sintagma preposicional se resuelve a nivel morfológico, ya que la raíz verbal *küdaw-* está modificada por las partículas *-a* de futuro y *-lu* que junto a la anterior indica finalidad, así tendríamos por ejemplo *amu-a-lu* “para ir”, *n'o-a-lu* “a cruzar las aguas”, etc.

- (e) *mülekan ruka mew*
sigo.estando casa en
“sigo estando en casa”

- (f) *amutun waria mew*
me.fui ciudad a
“me fui a la ciudad”

- (g) *waria mew küpan*
ciudad de he.venido
“he venido de la ciudad”

Con estas tres frases ejemplificamos el uso de la posposición mapuche *mew* cuyo significado depende del contexto en que se encuentre, de la información que se desprenda de los verbos y sustantivos que la acompañen. En este caso, claro está, no se generaran divergencias entre las frases originales en mapudungün y sus correspondientes traducciones al castellano.

- (h) *amuay temuko waria mew*
irá Temuco ciudad a
“irá a Temuco”

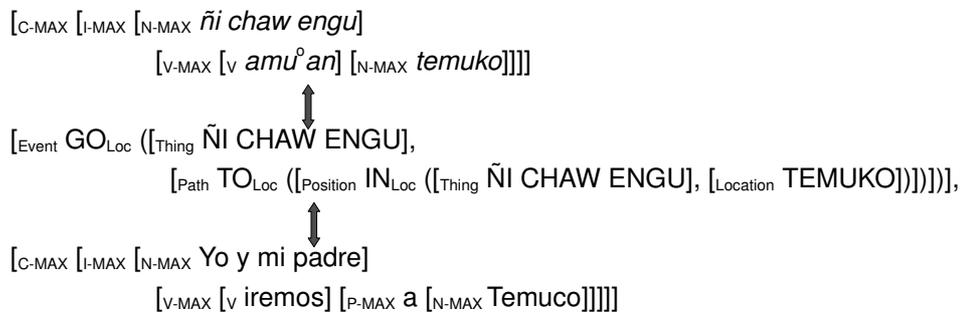
La posposición *mew* es un relacionador entre el verbo y el nombre que establece pertenencia, locación, direccionalidad u otras situaciones contextuales dependiendo de los sentidos del verbo y nombre en conjunto.

Esta posposición también toma otros valores, pero este no es el lugar para explayarnos con respecto a ella; con la información que ya hemos mencionado podemos generalizar la situación que nos llevará de vuelta a las divergencias. En los primeros ejemplos notamos que el mapudungün no hecha mano de la posposición que nos ha ocupado en los párrafos anteriores, mientras que en los últimos la encontramos; tenemos el verbo *amu-* y el nombre propio *temuko* implicado en ambas situaciones, lo que marca la diferencia, según observamos, es que en las primeras frases la locación es secundaria, mientras que en las últimas es preponderante; así, en las primeras frases es más relevante el hecho de “ir con el padre” o de “ir a trabajar” que el lugar relacionado con estos acontecimientos; en las últimas frases la acción y el lugar tienen relación directa, se predica la acción de “ir a la casa”, no la de “ir con alguien a la casa” por ello se utiliza la posposición relacionante *mew* que contextualiza la situación.

Bien, retomemos ahora nuestro propósito inicial y veamos cómo se resuelve la divergencia cuando esta se produce.

De la frase *ñi chaw engu amuan temuko*, sólo para hacer un poco más simple nuestro trabajo, no analizaremos el sintagma nominal *ñi chaw engu*; también

supondremos que mediante el proceso de diferenciación contextual del verbo, establecido en un módulo previo del sistema de traducción, obtendremos los verbos *amun^p* (con posposición) y *amun^o* (sin posposición) por ejemplo, y entonces nuestra frase la representaremos de la siguiente manera:



Este diagrama muestra que el objeto verbal en mapudungün es el sintagma nominal *temuko*, mientras que en castellano es el sintagma preposicional “a Temuco”. Para resolver esta divergencia, como dice Dorr, tenemos que aplicar el marcador * que provoca una realización composicional de los constituyentes lógicos en distintos niveles. Principalmente, mediante el marcador * se indica la posición de una ELCR que debe combinarse con otra ELCR para conformar una ELCC o parte de una. A través de las entradas léxicas de *amun^o* e “ir” ilustraremos como opera este parámetro.

(a) Entrada léxica de *amun^o*



(a) Entrada léxica de *ir*



Al estar señalada la posición $[Location Z]$ con el marcador * en la entrada del mapudungün, este constituyente se realizará como complemento del verbo principal bajo la forma *temuko* en la estructura superficial. En la estructura del castellano, como la aparición del marcador * se da en una posición más elevada $[Path TO_{Loc} \dots]$, el componente al cual se asocia se realiza de manera más compleja en la estructura sintáctica, representada por el SP “a Temuco”. Por lo tanto, la solución de la divergencia estructural se genera de la siguiente manera:

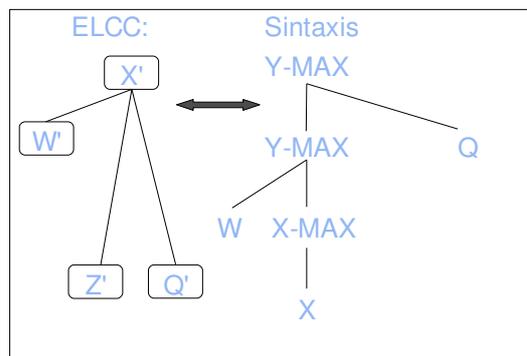
ELCR 1: $[_{r(X)} X' ([_{r(W)} W'), [_{r(R)} R' ([_{r(Z)} * Z')]] [_{r(Q)} Q']]]$
 ELCR 2: $[_{r(X)} X' ([_{r(W)} W'), [_{r(R)} * R' ([_{r(Z)} Z')]]] [_{r(Q)} Q']]$

Traducción: $[_{Y-MAX} [_{Y-MAX} W [_{X-MAX} X Z]] Q]$
 \updownarrow
 $[_{r(X)} X' ([_{r(W)} W'), [_{r(Z)} Z'] [_{r(Q)} Q']]]$
 \updownarrow
 $[_{Y-MAX} [_{Y-MAX} W [_{X-MAX} X [... R Z]]] Q]$

Lo que provoca que, en castellano en nuestro caso, se realice un sintagma R que domina a Z, es el hecho de tener el parámetro * asociado a R' en la ELCR de la lengua destino, y no tenerlo de esta manera en la lengua origen, en la cual no se produce esta relación de dominancia.

5. Divergencias combinatorias¹⁷ (Conflational)

La equivalencia de correspondencias cambia en este tipo de divergencias, el proceso específico es la supresión o aparición (dependiendo de la lengua desde la cual se mire) de uno de los componentes de la ELCC, realizándose generalmente en la posición de argumento lógico o de modificador lógico, por ello la correspondencia de equivalencia “<=>” se cambia, dependiendo de la posición que genere un resultado expresado por combinación de elementos, ya sea en el paso 3 o en 4 de la GER. Tenemos un esquema de este proceso en donde además podemos notar que la posición Z' en la ELCC no tiene una realización correspondiente en la sintaxis.

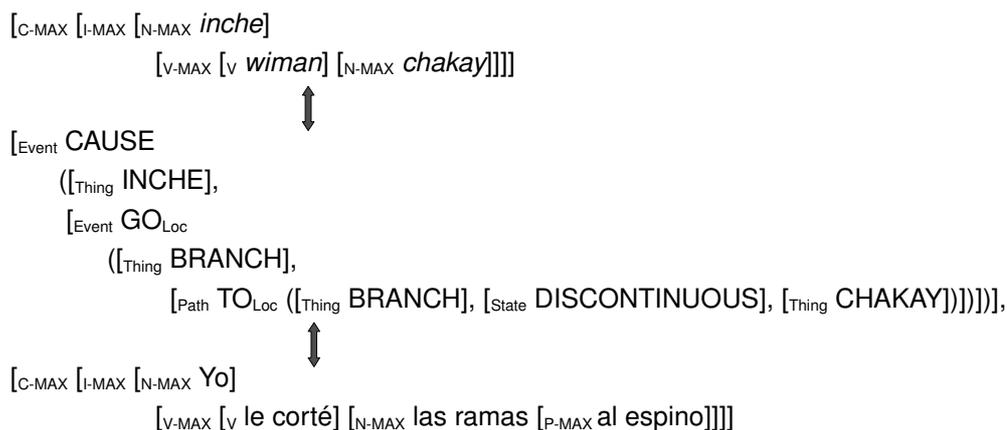


17 Pensamos que es necesario aclarar a qué se refiere el término combinatoria; “conflate” en inglés es un tecnicismo que quiere decir combinar, en el caso de las divergencias no es que se combinen los tipos de divergencia, sino que el nombre puntualiza el hecho de tener en una lengua expresiones realizadas a través de una forma léxica simple, mientras que en otras el mismo concepto se expresa mediante una combinación de elementos o formas léxicas, el ejemplo que utiliza Dorr es “stab” equivalente a “dar puñaladas”.

En mapudungün, como en todas las lenguas, tenemos verbos que se expresan con una combinación de formas léxicas en otras lenguas, no nos referimos a verbos compuestos, sino que a formas verbales cuyas raíces ya es necesario traducirlas utilizando dos o más formas léxicas, del castellano en nuestro caso. Por ejemplo tenemos *n'on* “cruzas las aguas”, *külün* “estar (alguien) inclinado”, *ilon* “comer carne”, *wiman* “cortar ramas”. Recordemos que el mapudungün tiene como estrategia verbalizar formas típicamente nominales, como ya explicamos, y entre estas formas tenemos muchas que generan divergencias del tipo combinatoria. Ejemplos: *kofketun* “comer pan” de *kofke* “pan”, *tokitun* “golpear con el hacha” de *toki* “hacha”, *rüpütun* “hacer el camino” de *rüpü* “camino”, *mumülltun* “recoger leña” de *mumüll* “leña”, en estos casos el sufijo *-tu* transforma al nombre en algo así como “lo que más comúnmente se hace implicando al sustantivo que lo origina”.

Para determinar si este tipo de formas se debe tratar a nivel morfológico o sintáctico, es decir considerarlos generadores de divergencias combinatorias, tendríamos que hacer un estudio al respecto y sopesar las estrategias para resolver este asunto, por lo tanto, para efectos del presente trabajo continuaremos con las raíces consideradas un poco más arriba y trabajaremos con la siguiente frase y su respectiva esquematización:

inche wiman chakay
 yo corté.ramas espino
 “le corté las ramas al espino”



Como ya hemos explicado, en mapudungün sólo utilizamos una forma léxica para lo que en español tenemos que utilizar dos, *wiman* <=> “cortar ramas”, debido a que el mapudungün incorpora el objeto afectado “rama”, dentro de la forma verbal. El parámetro que introduciremos para resolver este tipo de

divergencia será el marcador :CONFLATED que impide la utilización de una posición determinada. Las entradas léxicas para *wiman* y “cortar” nos aclaran la acción de este parámetro:

(a) Entrada léxica de *wiman*

[Event CAUSE
 ([Thing W],
 [Event GO_{Loc}
 ([Thing :CONFLATED BRANCH],
 [Path TO_{Loc} ([Thing BRANCH], [State DISCONTINUOUS], [Thing TREE])])])])]

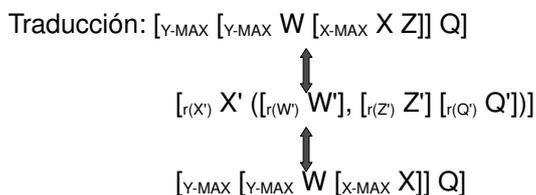
(b) Entrada léxica de *cortar*

[Event CAUSE
 ([Thing W],
 [Event GO_{Loc}
 ([Thing Y],
 [Path TO_{Loc} ([Thing Y], [State DISCONTINUOUS])])])])]

El constituyente que contiene la posición lógica que corresponde a BRANCH no se representa en la estructura sintáctica del mapudungün debido a que contiene el marcador :CONFLATED. Pasa lo contrario en la estructura sintáctica del castellano, el constituyente que corresponde a la posición del marcador ':CONFLATED, que no aparece en este caso, se debe realizar para poder dar sentido a la frase.

El diagrama de la solución para las divergencias combinatorias es el que sigue:

ELCR 1: [_{r(X)} X' ([_{r(W)} W'], [_{r(Z)} * Z'] [_{r(Q)} Q'])]
 ELCR 2: [_{r(X)} X' ([_{r(W)} W'], [_{r(Z)} :CONFLATED Z'] [_{r(Q)} Q'])]



Se puede notar además, que el argumento Z' se asocia con un marcador ':CONFLATED en la ELCR del castellano, pero no en la ELCR del mapudungün, lo que hace que la realización sintáctica del castellano suprima este constituyente.

6. Divergencias Catoriales

Ahora describiremos una divergencia que se produce en la RSC - Realización Sintáctica Canónica – cuya principal característica es que la $CAT(\Phi)$ se ve forzada a tener un valor diferente del que normalmente se asigna a $T(\Phi)$ de acuerdo a lo expresado en el cuadro que repetimos aquí para mayor claridad:

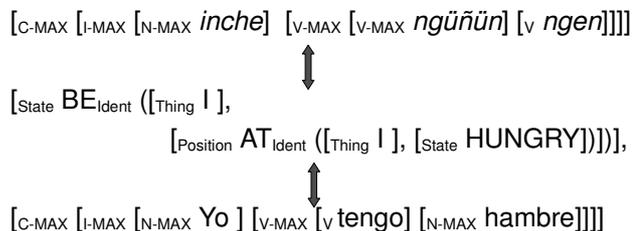
Categoría conceptual de ELC	Categoría Sintáctica
EVENT	V
STATE	V
THING	N
PROPERTY	A
PATH	P
POSITION	P
LOCATION	ADV
TIME	ADV
MANNER	ADV
INTENSIFIER	ADV
PURPOUSE	ADV

Formalmente tenemos la siguiente descripción: una categoría conceptual - tipo lógico – $T(\Phi)$ se relaciona con una categoría sintáctica $CAT(\Phi)$, siendo $CAT(\Phi) \neq RSC(T(\Phi))$. En estos casos la $CAT(\Phi)$ debe ser especificada a través de una parametrización léxica.

Es la dinámica normal del mapudungün expresar los estados experimentados por el hablante como agente de la acción, por ejemplo, “tengo mucha sed” es *fücha wiwün* cuya traducción literal sería algo así como “sedeo mucho”, *fücha ngüñün* es “hambre mucho” (tengo mucha hambre) y lo mismo con *füta wütren* “tengo mucho frío”; sin embargo hay otro tipo de construcción sintáctica que no es muy común, aunque válida en determinados contextos, como el que encontramos en la siguiente frase, *küdawuma ngen* “soy uno que ha trabajado”, que para nuestra frase toma la siguiente forma:

inche	ngüñün	ngen	
yo	hambre	soy	(“yo soy uno que hambrea”, “yo estoy hambreado”)
“yo tengo hambre”			

De esta manera encontramos nuestro caso para la divergencia categorial, estas son las representaciones formales y los pasos a seguir para resolver la divergencia, la estructura sintáctica y su correspondiente ELCC:

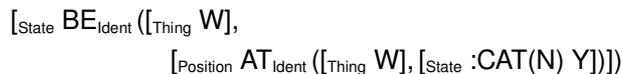


El predicado es verbal en mapudungün *ngũñün* “hambrear”, pero en castellano es nominal “hambre”. Para resolver esta divergencia utilizaremos el parámetro :CAT, que veremos reflejado en las siguientes entradas léxicas:

(a) Entrada léxica de *ngen*

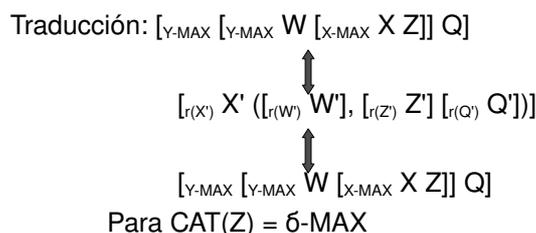
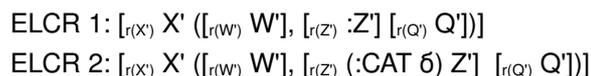


(b) Entrada léxica de *tener*



Debido a que la entrada mapuche no contiene el marcador :CAT en la posición que corresponde a [State Y], este constituyente se realiza en la estructura sintáctica como RSC(Estado) → Verbo (*ngũñün*). En castellano este tipo de realización es imposible, de hecho sólo se puede parafrasear una traducción aproximativa o se debe recurrir a la correspondencia idiomática que genera divergencia, por ello es necesario el marcador :CAT(N) que evita la realización de la categoría por defecto de la RSC, pero que permite la realización del argumento de Estado como una forma nominal, es decir como “hambre”.

La solución se representa en el siguiente diagrama:



7. Divergencias léxicas

Nos encontraremos frente a una divergencia de este tipo cuando entre lengua origen y lengua destino hay elementos que expresan un mismo sentido a través de formas léxicas cuyas propiedades semánticas no tienen relación, la autora pone como ejemplo la expresión “brake into” y “forzar” como su equivalente, en este caso la semántica de la expresión inglesa se acerca a la idea de “forzar”, que podría haber sido traducida por “irrumpir” por ejemplo, sin embargo los elementos que forman la expresión, “brake” e “into” por sí solos no dicen relación con sus equivalencias castellanas. Este fenómeno suele ser provocado por los tipos de divergencias que hemos revisado hasta ahora debido a que éstas modifican las propiedades de composición y realización de las entidades léxicas que participan en la representación sintáctica; por esto se considera como una combinación de los elementos de las divergencias la formalización de la divergencia léxica. Por otra parte esta divergencia se resuelve durante el proceso de selección léxica, en el que no hay marcadores o parámetros restrictivos específicos.

Bajo estas consideraciones y conociendo la forma compositiva de las raíces verbales del mapudungün, podemos llegar a concluir que este tipo de divergencia se dará casi en cada frase:

Partícula *-lu*, este sufijo combinado con el sufijo de futuro *-a*, expresa finalidad, en castellano se expresa mediante una preposición, así, dependiendo del contexto, la combinación *-a-lu* seleccionará “a”, “para” o “por”.

La partícula *rume* tiene varios significados, cuando no acompaña a una raíz significa “al menos”, “siquiera” o “sumamente”

<i>rume</i>	<i>fücha</i>	<i>illkuyngün</i>
sumamente	mucho	se.enojaron.ellos
“se enojaron muchísimo”		

si en la frase aparece *kiñe* “uno”, su significado es “simple”, “puro”, “directo”

<i>kiñe</i>	<i>rume</i>	<i>amunge</i>	<i>ruka</i>	<i>mew</i>
uno	directo	ir.imperativo	casa	a
“vete directamente a casa”				

Al formar parte de una raíz verbal la modifica indicando que la acción se realiza repentinamente

trapial rüngkü-rume-y

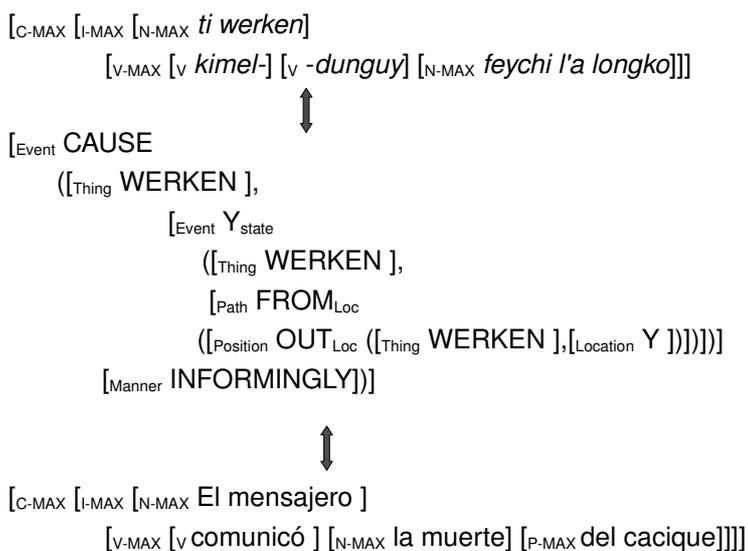
puma saltar-repentino-indicativo

“el puma saltó repentinamente”, “el puma saltó de pronto”, “el puma saltó de manera imprevista”, etc.

En fin, en mapudungün muchas son las partículas, de ellas podemos definir su significado general, pero el contexto determinará cuál es la forma léxica más adecuada en la lengua destino, para su traducción. Pero no sólo esto ocurre, también tenemos términos compuestos equivalentes a términos simples en castellano, *longkomoyo* es “pezón” de *longko* “cabeza” y *moyo* “seno”, *adkintun* es “otear (mirar a lo lejos)” de *ad* “imagen” y *kintun* “buscar”, etc. En la siguiente frase tenemos una divergencia léxica provocada por una divergencia combinatoria,

ti werken kim-el-dungü-y feychi l'a longko
 el mensajero hacer.saber-sustantivador¹⁸-mensaje-indicativo esa muerte cacique
 “el mensajero comunicó la muerte del jefe”

veamos ahora su estructura sintáctica y la correspondiente ELCC.



Hemos elegido separar la forma verbal en esta representación porque no tenemos claro cómo debería ser la estrategia previa a la de la fase sintáctica, nos referimos a la etapa de análisis morfológico, probablemente la solución en este nivel al presentarse raíces complejas carezca de expectativas concretas y la forma verbal deba descomponerse como si fueran distintos elementos de una misma frase, tal como lo hemos hecho en nuestro ejemplo; a partir de esta resolución podemos observar que los componentes de la raíz verbal mapuche

18 El sufijo *-el-*, entre otros efectos, hace que la forma verbal tome la función de sustantivo verbal, en el ejemplo tenemos por lo tanto “su hacer saber el mensaje (del mensajero)”.

corresponden a “hacer saber” y “el mensaje”, que probablemente constituirán, sino una solución definitiva cuando sea necesario traducir raíces complejas, al menos una primera propuesta, ya que pasar a “comunicar” requiere una abstracción mayor y por ende más compleja.

Una vez presentados todos los tipos de divergencias y las soluciones propuestas por Dorr, sólo nos queda mencionar cuáles han sido las singularidades con que nos hemos encontrado a través del análisis aplicado al mapudungün. En primer lugar la rica variedad de partículas y sus diversas aplicaciones y sentidos. Este tipo de elementos del discurso ha planteado discusiones y generado material de estudio en varias lenguas, y probablemente es un área de la gramática con una dificultad generalizada desde el punto de vista del análisis lingüístico tradicional. Lo que más influye a este respecto es el poco interés que estas partes de la oración suscitaron a los estudiosos de la antigüedad clásica, a pesar del extendido uso que de ellas hacían los hablantes; también es determinante el hecho de no tener suficientes datos confiables como resultado del uso mayoritariamente oral de estas partículas, a todo esto se suma la dificultad - quizás por ello todo lo anterior - de encontrar los parámetros de utilización sujetos a la interacción entre pragmática, semántica, prosodia y sintaxis. Pensamos que el estudio minucioso de las partículas del discurso mapuche puede arrojar muchas luces en el camino a establecer un sistema de Traducción Automática eficiente, al menos para lenguas con un uso extendido o medianamente extendido de este recurso; una de ellas, que por el contrario cuenta con investigación más profunda, es el alemán, en donde se denominan *Modalpartiken* “partículas modales” o *Abtönungspartiken*, y de la cual, los interesados en el mapudungün, podemos extraer muchos conocimientos.

Otra cuestión que produjo dudas, titubeos y errores fueron los **temas verbales complejos**. Con respecto a lo primero sabemos que en castellano una raíz verbal puede transformarse en un tema más complejo a través de sufijos como re- (revolver), de(s)- (desarmar), etc. Pero en mapudungün nos encontramos con varias dificultades añadidas, estos prefijos no existen como tales; las raíces se pueden combinar (Concatenación Radical), y éstas, que ya forman un tema complejo, a su vez pueden combinarse con raíces nominales (Incorporación Nominal).

A la hora de buscar modificadores nominales o verbales también nos encontramos con fenómenos a los cuales no estamos acostumbrados en castellano. Los adjetivos en lenguas romance, son de naturaleza principalmente nominal, modifican a un nombre o funcionan como tal “es un chico bueno”, “lo bueno es ser niño”. En la lengua mapuche la naturaleza del adjetivo es fundamentalmente verbal y

por ello podemos tener verbos como *lefün* “velohear”, que traducimos como “correr”, del adjetivo *lef* “veloz”; la equivalente a la frase castellana “el otro” es *chi kangelu* “el ser otro”.

Bien, para dar una visión global y finalizar podemos decir que en general la propuesta de Bonnie Dorr es acertada y la solución de las divergencias es viable, sin embargo en lo que concierne a las Estructuras Léxico Conceptuales, pensamos que se debería generar una representación más rígida, en el sentido formal, queremos decir que, mientras más dominio se tiene de este tipo de representación más fácil es generar esquemas ad hoc comprometiendo la generalización de un proceso que debería mantenerse homogéneo a través de los diversos casos.

Creemos que mucha de la arbitrariedad presente en los procesos de selección léxica es perfectamente evitable, se debe definir y delimitar el proceso de generación dentro de un marco más bien restrictivo, condicionado si se quiere. Lo que pretendemos decir es que un buen resultado no pasa por ampliar por ejemplo las equivalencias léxicas haciendo crecer el sistema como propone la lingüista, sino más bien definiendo módulos específicos desarrollados para producir resultados controlados, quizás el sistema podría crecer en cantidad de módulos y en potencia y versatilidad de un motor central capaz de decidir los módulos a utilizar y hacerlos interactuar en función de una situación parametrizada, hacer crecer los módulos de entradas léxicas sólo entorpecería el proceso de generación. La interconexión (mapping) entre las formas léxicas y la representación sintáctica debe estar definida en función de un resultado preconcebido, de esta manera se mantiene la coherencia entre lengua origen y lengua destino.

Con respecto a la aplicación de esta metodología a lenguas como el mapudungün, vemos una posibilidad concreta de resolución específica de las divergencias, eso sí, como ya planteamos en muchos de los puntos que fuimos comentando, un estudio exhaustivo del comportamiento morfosintáctico y su relación con la semántica oracional, probablemente nos brindaría nuevas soluciones para este tipo de desajustes en traducción, con lo cual podríamos generar resultados aceptables a partir de las reglas de combinatoria de sufijos y raíces, por ejemplo.

Referencias

Dorr, Bonnie J. (1993) "Machine Translation Divergences: A Formal Description and Proposed Solution". Association for Computational Linguistics.

Zúñiga, Fernando (2006) "MAPUDUNGUN El habla mapuche". Centro de Estudios Públicos, Santiago de Chile.

Salas, Adalberto (2006) "El Mapuche o Araucano (1992)" 2ª Edición revisada por Fernando Zúñiga. Centro de Estudios Públicos, Santiago de Chile.

De Augusta, Felix José (1991) "Diccionario Español – Mapuche. Tomo 2 (1916)". Ediciones Séneca, Santiago de Chile.

Espinal, M. Teresa (coord.); Macia, Josep; Mateu, Jaume; Quer, Josep; Ynglès, M. Teresa (2002) "Semàntica. Del significat del mot al significat de l'oració". En Ariel Lingüística, Editorial Ariel, Barcelona, España.

Martí, M. Antonia (2003) "Consideraciones sobre la polisemia". Artículo realizado en el marco del proyecto X-TRACT (PN98-1226) y X-TRACT2 (BFF2002-04226-C03-03).

Saint-Dizier, Patrick (1999) "An Introduction to the lexical semantics of predicative forms". En "Predicative Forms in Natural Language and in Lexical Knowledge Bases". Kluwer Academic Publishers, Dordrecht, Holanda.

Pustejovsky, James; Boguraev, Branimir (1993) "Lexical knowledge representation and natural language processing", En "Artificial Intelligence 63, pp. 193-223". Elsevier Science Publishers, USA.

Moreno, Juan Carlos (2000) "Curso Universitario de Lingüística General". Editorial Síntesis, Madrid, España.