

O TRATAMENTO DOS PAPÉIS TEMÁTICOS POR UM SISTEMA DE TRADUÇÃO AUTOMÁTICA: PARÂMETROS PARA A LÍNGUA PORTUGUESA

Mirna Fernanda de OLIVEIRA³¹

RESUMO: O presente trabalho parte do exame dos fundamentos linguísticos que motivaram a construção do UNITRAN, um sistema de tradução automática do tipo interlíngua que modela computacionalmente divergências sintático-semânticas entre pares de sentenças de línguas diferentes. Através do exame da parametrização dos papéis temáticos no sistema, demonstra-se como o sistema torna possível a implementação da correspondência sistemática entre a interlíngua e a estrutura léxico-semântica das línguas envolvidas, graças a uma base de conhecimento que codifica tipos específicos de divergências que o sistema deve manipular.

PALAVRAS-CHAVE: Tradução automática. Papéis temáticos. Princípios e parâmetros. Teoria de estruturas conceituais.

ABSTRACT: This paper examines the linguistic motivation behind the building process of UNITRAN, an interlingual machine translation system that formalizes syntactic and semantic divergencies between pairs of sentences of different languages. Through the study of the treatment given by the system to thematic roles, it is shown that it is possible to implement the systematic correspondence among the interlingua, syntactic and lexical-semantic structures of the languages involved, due to a knowledge base that codifies specific types of divergencies the system should be able to manipulate.

KEYWORDS: Machine translation. Thematic roles. Government and binding. Conceptual structures theory.

1. Introdução

A construção do UNITRAN (DORR, 1993), um sistema de tradução automática do tipo interlíngua, que traduz o inglês, o espanhol e o alemão, entre outros, entre si, parte da motivação de se modelar computacionalmente as divergências de tradução,³² que são definidas como um tipo específico de problemas de natureza sintática e léxico-semântica que devem ser enfrentados por um sistema de TA e que, segundo a abordagem de Dorr, podem ser solucionados, pelo menos em parte, com a especificação de informações sintáticas e léxico-conceituais sintaticamente parametrizadas³³ no léxico do sistema. Com essa motivação, o UNITRAN demonstra que é possível enfrentar problemas de divergências entre línguas sem recorrer a regras específicas e pontuais para cada uma delas. Em particular, o sistema

³¹ Doutora em Linguística e Língua Portuguesa. No momento é docente da UNIOESTE/Foz do Iguaçu, atuando na área de Língua Inglesa.

³² Dorr (1993) distingue sete tipos de divergências formais, que não incluem as divergências decorrentes do uso idiomático da língua e de conhecimentos aspectual, de discurso e de mundo. (DORR, 1993, *Preface*).

³³ É essencial ressaltar que esses parâmetros são os parâmetros do sistema UNITRAN e não os parâmetros que caracterizam a variação sintática entre as línguas, conforme a *Teoria dos Princípios e Parâmetros*. Dito de outra forma: trata-se de parâmetros implementados no sistema que são motivados por parâmetros linguísticos. Para evitar equívocos, vamos adotar dois rótulos: o rótulo simples “parâmetro”, para designar os parâmetros linguísticos, e o rótulo “parâmetros do sistema”, para designar os parâmetros implementados no sistema e que dão conta das divergências estruturais de tradução, quer de natureza sintática, quer de natureza léxico-conceitual.

demonstra que é possível a implementação da correspondência sistemática entre a interlíngua do sistema e a estrutura sintática e léxico-semântica das línguas graças à uma base de conhecimento que codifica tipos específicos de divergências que o sistema deve manipular.

O dois módulos do sistema UNITRAN que são relevantes para a discussão são o módulo sintático e o módulo léxico-conceitual. O primeiro é construído com a implementação computacional dos constructos da *Teoria dos Princípios e Parâmetros* (Chomsky, 1981) e parametriza a forma sintática das sentenças e dos seus constituintes; o segundo, construído com a implementação computacional dos constructos da *Teoria das Estruturas Conceituais* (Jackendoff, 1990), parametriza a estrutura conceitual das sentenças e as propriedades léxico-conceituais dos itens lexicais a partir dos quais ela é construída.

O sistema UNITRAN é uma alternativa para sistemas de TA dos tipos direto e de transferência. O que o torna do tipo interlíngua é o fato de utilizar uma forma subjacente de representação da informação linguística independente de língua e estabelecer a correspondência entre essa representação e a estrutura sintática das línguas envolvidas. Por exemplo, uma diferença gramatical entre sentenças do alemão e do inglês pode ser descrita em termos de um parâmetro sintático que descreve a ordem dos constituintes nas duas línguas. Em particular, esse parâmetro, denominado parâmetro de “núcleo de sintagma verbal”, tem o valores “*inicial*” para os verbos do inglês (Ex. *I have seen him*) e “*final*” para os verbos do alemão (Ex. *Ich habe ihn gesehen*). Assim, para o inglês, o parâmetro de núcleo de sintagma verbal colocaria, durante a geração da sentença do inglês, o SN objeto logo depois do verbo principal; para o alemão, esse parâmetro colocaria o verbo principal no final da sentença.

Desta forma, a fundamentação teórica do UNITRAN (DORR, 1993) é constituída pela *Teoria dos Princípios e Parâmetros* (CHOMSKY, 1981) e pela *Teoria das Estruturas Conceituais* (JACKENDOFF, 1990). A primeira subsidia a construção do seu módulo sintático e a segunda do seu módulo léxico-semântico. Esse sistema, que traduz do inglês para o alemão e o espanhol bidirecionalmente, é parametrizado nos dois níveis, a saber:

No nível sintático: que diz respeito à estrutura da sentença;

No nível léxico-semântico: que diz respeito às propriedades semântico-conceituais dos itens lexicais.

Para este artigo selecionamos dois aspectos dessa aplicação das teorias tendo em vista o tratamento do problema da representação dos papéis temáticos para a língua portuguesa.

2. Fundamentos teóricos

2.1 Estrutura sintática: a Teoria- θ

A Teoria- θ , parte integrante da *Teoria de Princípios e Parâmetros* (CHOMSKY, 1981), descreve a projeção da estrutura argumental das categorias lexicais; ocupa-se, assim, da atribuição de papéis temáticos (θ) como Agente, Paciente, etc. O Critério θ estipula que cada SN argumento realiza um único papel temático, e cada papel temático, por sua vez, é atribuído a um único SN argumento. O papel de Agente é aquele atribuído ao argumento que realiza a ação denotada pelo predicador; o papel Paciente é aquele atribuído ao argumento que é afetado pela ação denotada pelo predicador; o papel Meta é aquele atribuído ao argumento que é o destinatário da ação denotada pelo predicador (COOK, 1989, p.117).

A teoria diz respeito a três parâmetros: os Clíticos, Duplicação de Clíticos e *Nom-Drop*.

O parâmetro Clíticos é utilizado para especificar a categoria elementos foneticamente fracos. Esse sub-módulo dá conta dos clíticos *him*, *le* e *lhe*, que ocorrem, por exemplo, em sentenças como *I gave him the book*, *Yo le di el libro*, *Eu lhe dei o livro*, respectivamente, do inglês, do espanhol e do português.

O parâmetro Duplicação de Clíticos dá conta da atribuição de papel temático a sentenças do espanhol do tipo *Lo vimos a Guille*. Nessa sentença, esse parâmetro assegura que o clítico *lo*, co-referente ao N^o *Guille* (que não tem papel temático), lhe transfira o seu papel temático.

Por fim, o parâmetro *Nom-Drop* efetua a atribuição de elementos pleonásticos. As sentenças 1, 2 e 3 abaixo ilustram o processo, com a aplicação de duas condições: a primeira condição estipula que o parâmetro seja marcado *negativo* e o elemento seja um sujeito explícito, isto é, o pleonástico *it* do inglês; a segunda condição estipula que o parâmetro seja marcado *positivo* e a língua seja *pro-drop*, como o espanhol e o português.

1. *It is raining.*
2. *Está lloviendo*
3. *Está chovendo*

2.2 A Teoria das Estruturas Conceituais

A *Teoria das Estruturas Conceituais* (TEC) de Jackendoff (1990 e ss.), que subsidia a construção da interlíngua do sistema de tradução automática UNITRAN de Dorr (1992, 1993), fornece a metalinguagem para a representação semântica do significado sentencial independente de língua e se mostra apropriada para tratar do problema das divergências.

A TEC investiga o nível conceitual da linguagem, buscando um modo de especificar a estrutura conceitual nos níveis lexical e sentencial e as relações que se estabelecem entre os planos linguístico e conceitual. A TEC focaliza a semântica lexical de maneira sintaticamente relevante, ao se preocupar com a construção da estrutura conceitual da sentença em função das estruturas conceituais parciais dos seus componentes. Por exemplo: a estrutura conceitual nuclear (abstraindo-se do tempo) da sentença de *João colocou o carro na garagem*, é dada na figura abaixo.

[Evento CAUSAR
 ([Coisa JOÃO],[Evento IR ([Coisa CARRO],[Trajetória PARA ([Lugar EM([Coisa GARAGEM])])])])])]

Pelo expediente da **decomposição semântica**, a TEC aponta que é possível se concebermos que os significados das unidades lexicais não são atômicos. Assim, a decomposição do significado em elementos menores permite a especificação de um sistema composicional que impõe restrições à combinação desses elementos e, portanto, impõe também restrições nas configurações sentenciais.

Da mesma forma, a proposição de uma **ontologia** permite a construção de blocos de representação conceitual; as estruturas conceituais sancionadas pela TEC são construídas por meio de categorias da ontologia. Cada constituinte é decomposto numa estrutura funcional do tipo [Categoria FUNÇÃO ([Categoria Argumento 1], [Categoria Argumento 2],..., [Categoria Argumento n])]. As funções codificam as relações entre os argumentos e são munidas de restrições conceituais, rotuladas pelas categorias ontológicas.

Em outras palavras, Jackendoff procura identificar funções que expliquem padrões gramaticais de combinação, usando a ontologia como um ponto de referência. A proposição das categorias ontológicas é o ponto de partida para a identificação de generalizações de como as palavras podem se combinar para formar constituintes maiores. Graças à capacidade criativa da linguagem, fica claro que não se aprende simplesmente maneiras específicas de relacionar as palavras, existem padrões de combinação regulares e produtivos que se aplicam a classes de palavras e sentenças. Ao permitir o mapeamento de palavras e sintagmas em

classes mais gerais, a ontologia dá suporte para a identificação de tais padrões de combinação (VERSPOOR, 1997).

Vale ressaltar ainda que, com o crescente interesse no estudo da representação e organização lexical³⁴, torna-se cada vez mais possível isolar os componentes de significado comuns a verbos participantes de uma mesma classe. Tais componentes podem então ser usados para determinar a representação de verbos através de línguas distintas. A abordagem da TEC, ao considerar a representação semântica como um subconjunto da estrutura conceitual, ou seja, da linguagem de representação mental. A figura abaixo mostra as principais funções conceituais da TEC (JACKENDOFF, 1990, p. 43)³⁵.

- (a) [Lugar FUNÇÃO-LUGAR ([Coisa])]
- (b) [Trajetória PARA ou DE ou EM_DIREÇÃO_A ou DISTANTE_DE ou VIA ([Coisa] ou [FUNÇÃO-LUGAR])]
- (c) [Evento IR ([Coisa], [FUNÇÃO-TRAJETÓRIA])]
- (d) [Estado ESTENDER-SE ([Coisa], [FUNÇÃO-TRAJETÓRIA])]
- (e) [Evento CAUSAR ([Coisa] ou [Evento IR ([Coisa], [FUNÇÃO-TRAJETÓRIA])]) ou [Evento PERMANECER ([Coisa], [FUNÇÃO-LUGAR]), [Evento IR ([COISA], [FUNÇÃO-TRAJETÓRIA])] ou [Evento PERMANECER ([Coisa], [FUNÇÃO-LUGAR])]]
- (f) [Evento PERMANECER ([Coisa], [FUNÇÃO-LUGAR])]
- (g) [Estado ESTAR ([Coisa], [FUNÇÃO-LUGAR])]
- (h) [Estado ORIENTAR-SE ([Coisa], [FUNÇÃO-TRAJETÓRIA])]

2.3 Os papéis temáticos na teoria de estruturas conceituais

Na TEC, os papéis temáticos são considerados posições na estrutura conceitual. Por exemplo, os papéis temáticos ATOR e PACIENTE são considerados entidades representadas em suas relações, movimentações, lugares, etc. Uma série de ações e uma série temática são introduzidas na representação dos EVENTOS. Num exemplo de Jackendoff (1990, p.130):

- a) *What Bill did to the books was to load them on the truck.*
O que Bill fez aos livros foi carregá-los no caminhão.
- b) *? What Bill did to the truck was load the books onto it.*
O que Bill fez ao caminhão foi carregar os livros sobre ele.

³⁴ Grimshaw (1990), Jackendoff (1983, 1990), Levin & Rappaport (1993), Pustejovsky (1995) e Hale & Keyser (2002).

³⁵ Como se trata de uma metalinguagem formal para a representação de estruturas conceituais e, portanto, independentes de língua, optamos, por julgar mais conveniente e inteligível, por transpor as categorias ontológicas e as funções conceituais para o português.

- c) * *What Bill did to the books was load the truck with them.*
O que Bill fez aos livros foi carregar o caminhão com eles.
- d) *What Bill did to the truck was load it with books.*
O que Bill fez ao caminhão foi carregá-lo com livros.

Em cada caso, o evento expresso é o mesmo: os livros vão para cima do caminhão. O que muda, como se reflete na forma sintática, é a entidade que é vista como PACIENTE, a que é mais afetada pela ação de Bill. O papel PACIENTE, em (a), é desempenhado por “os livros”, em (b), por “o caminhão”. A razão da baixa aceitabilidade de (c), por exemplo, é que há um conflito entre a entidade (X) explicitamente identificada como recebendo a ação de Bill; *Bill fez alguma coisa para X*, e a entidade sugerida pela sintaxe da sentença clivada é a de PACIENTE.

O exemplo sugere que a noção semântica de PACIENTE é relevante para o mapeamento entre a sintaxe e a semântica, e que essa relação temática deve ser representada para dar conta da distinção entre as sentenças paralelas acima (VERSPOOR, 1997).

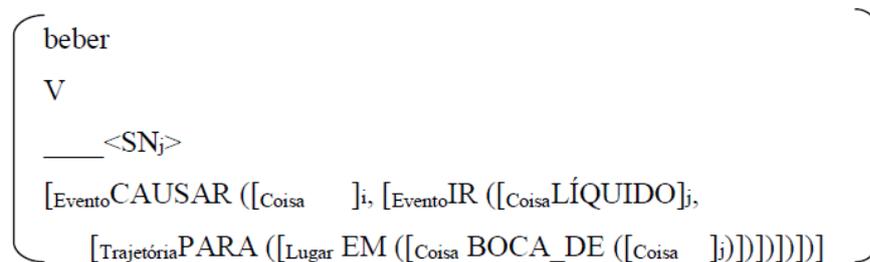
Na visão de Jackendoff as relações temáticas não estão associadas a posições sintáticas, mas antes correspondem a configurações particulares na estrutura conceitual (1990, p. 46). Assim, por exemplo, o papel temático Origem (*Source*), ou seja, “o objeto do qual parte o movimento”, é o argumento da FUNÇÃO-TRAJETÓRIA FROM; o papel temático Meta (*Goal*), ou seja, “o objeto em direção ao qual o movimento é direcionado”, é o argumento da FUNÇÃO-TRAJETÓRIA PARA; o papel temático Agente (*Agent*) é o primeiro argumento do FUNÇÃO-CAUSAR, e assim por diante.

Essa estratégia de não se considerar os papéis temáticos como elementos primitivos da teoria permite torna-la parcimoniosa. Por exemplo, em uma sentença como *João escalou a montanha*, a Meta do verbo *escalar* não é a montanha, mas o seu topo. Ou seja, a estrutura conceitual da sentença *João escalou a montanha* é dada por [Evento IR ([Coisa JOÃO], [Trajetória PARA (TOPO-DE ([Coisa MONTANHA]))]]. Assim, não é preciso postular esse novo papel temático.

Segundo Higginbotham (1985, p. 555, *apud* JACKENDOFF, 1990, p. 48), deve-se, no entanto, considerar as grades temáticas das unidades lexicais com um mecanismo de indexação sintática. Por exemplo, o verbo *ver* possui a grade temática <1, 2, E>, ou seja, um evento constituído de dois argumentos. Já Jackendoff (1990, p. 48) propõe apresentar o que subjaz esse e outros modelos de estrutura de argumentos. Usando a as estruturas conceituais como suporte, ele propõe que se postule uma lista dos argumentos do verbo ao se extrair os índices da sua estrutura conceitual; o papel temático de cada constituinte da sentença é determinado por meio da conexão que se estabelece entre esse constituinte e a posição

estrutural correspondente na estrutura conceitual. Ou seja, a estrutura de argumentos de um verbo pode ser vista como a forma abreviada da parte da estrutura conceitual que é visível para a sintaxe. As restrições referentes ao número e ao tipo de relações temáticas advêm de quaisquer restrições existentes no escopo das funções conceituais necessárias para expressar os significados de verbos e preposições. Surgem portanto três conseqüências desse ponto de vista: (i) não há papel temático para o sujeito; essa é uma relação sintática, não conceitual, e sujeitos sintáticos podem possuir uma variedade de papéis temáticos; (ii) não apenas os SNs recebem papéis temáticos. Por exemplo, Em *A luz mudou de vermelho para verde*, *verde* é Meta; (iii) não deve existir um papel temático *default* como o papel temático Objetivo (“neutro”) de Fillmore (1968).

Na proposta de Jackendoff, as **restrições seletivas** são parte do significado do verbo e devem estar integradas à estrutura de argumentos. No verbo *beber*, assumindo seu significado como “fazer com que um líquido entre na boca de alguém”, é dada na figura abaixo uma possível representação:



A restrição seletiva do objeto direto (o SN_j) aparece como a informação conceitual LÍQUIDO, com ele coindexada. Para que se faça uso dessa informação, não se deve simplesmente substituir a leitura do objeto direto pelo constituinte *j*; ao contrário, é necessário considerar a leitura do argumento SN de alguma maneira fundida (fusão, *merge*) com a informação semântica já contida no constituinte. Assim, propõe-se que se troque a substituição do argumento com uma regra para fusão de argumento.

Em *Pedro bebeu o vinho*, a fusão de argumentos combina a leitura de *vinho* com o constituinte [_{Coisa} LÍQUIDO]_j; o traço redundante LÍQUIDO é apagado. A leitura resultante é que em *Pedro bebeu o vinho*, lê-se “o líquido específico no contexto”. A fusão do argumento lida com fenômenos observáveis a respeito das restrições seletivas, incluindo o uso dessas para se ter certeza de que a maior categoria conceitual de um argumento combine com aquela estipulada pelo verbo. Portanto, as restrições seletivas dos argumentos funcionam da mesma forma na estipulação de argumentos totalmente incorporados como Tema, como no

caso dos verbos (a) *amanteigar* (*Maria amanteigou o bolo*) e (b) *embolsar* (*Pedro embolsou o dinheiro*):

a) [Evento CAUSAR ([Coisa]i, [Events IR ([Coisa MANTEIGA],
[Trajetória PARA ([Lugar SOBRE ([Coisa]j))]]))]]]

b) [Evento CAUSAR ([Coisa]i, [Evento IR ([Coisa]j,
[Trajetória PARA ([Lugar EM ([Coisa BOLSO]))]]))]]]

Comparando as entradas de *beber* e *amanteigar*, por exemplo, percebe-se que não há necessidade de se postular um nível na estrutura de argumentos para codificá-los. Ambos diferem semanticamente a respeito da informação sobre TEMA e TRAJETÓRIA; eles também diferem sintaticamente em que *amanteigar* é obrigatoriamente transitivo, e *beber*, é opcionalmente transitivo. (JACKENDOFF, 1990, p. 53-54).

Assim, se o constituinte do verbo está indexado, seus traços semânticos são codificados como restrições seletivas; se o constituinte não está indexado, seus traços semânticos são codificados na estrutura conceitual e ele se torna implícito. Portanto, é possível estabelecermos como relacionar argumentos da estrutura conceitual a constituintes sintáticos: cada item lexical na sentença especifica como seus argumentos conceituais estão relacionados com posições sintáticas no sintagma de que é núcleo. A fusão de argumentos usa a informação lexical para relacionar as leituras dos complementos e sujeitos sintáticos com as posições argumentais indexadas na estrutura conceitual do núcleo sintagmático. Estabelecemos também que *papel temático* é um termo que designa uma posição estrutural na estrutura conceitual, que consiste de um conjunto de índices que relacionam os argumentos sintáticos e conceituais. Essa integração entre estrutura de argumentos e estrutura conceitual é possível porque a estrutura de argumentos é concebida como uma representação que é parte da estrutura conceitual, e não um conjunto de diacríticos que etiquetam argumentos sintáticos. Para Jackendoff, essa é uma grande virtude da TEC. (1990, p. 55).

3. Compatibilização de teorias

3.1 Dois tipos de coindexação

Para relativizar a Teoria θ e compatibilizá-la com a TEC, Jackendoff (1990, p.59-68) postula que no mapeamento entre as estruturas sintática e conceitual haja dois tipos de coindexação; esse ajuste da teoria sintática de papéis temáticos é necessária, dado que, na TEC, a um mesmo argumento sintático pode corresponder mais de uma posição na estrutura conceitual, o que significa dizer que a um mesmo SN podem estar associados mais de papel

temático, contrariando, assim, a Teoria θ . Os exemplos em que essa fato ocorre são com predicadores que denotam algum tipo de transação, como aqueles lexicalizados pelos verbos *comprar*, *vender* e *trocar*. Tomemos, então, a sentença (a) com o verbo cuja estrutura conceitual é dada a seguir (o verbo *comprar*):

a) *Maria comprou o livro de João.*

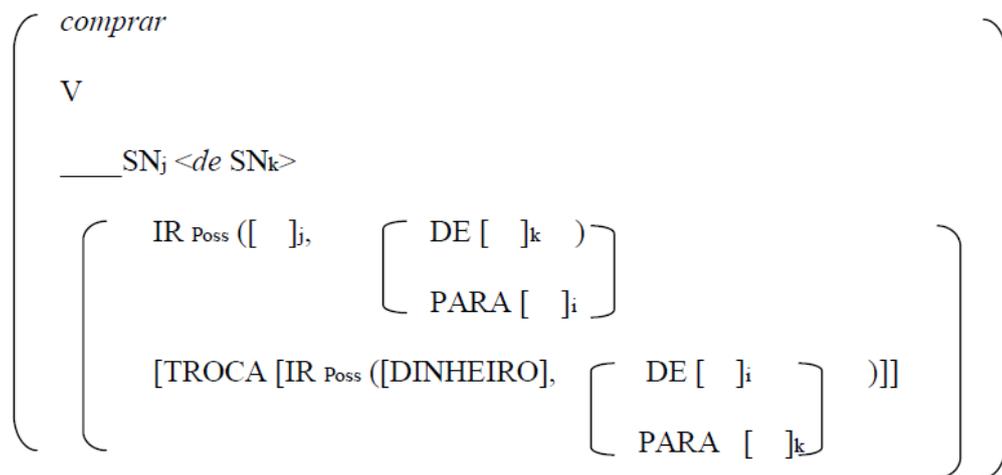
O significado dessa sentença envolve, pelo menos, dois componentes conceituais, duas transferências de posse:

i. a transferência de posse do livro, que passa de João para Maria (a transferência principal)

ii. a transferência de posse do dinheiro envolvido na transação, que passa de Maria para João.

(a contra-transferência)³⁶

Essa análise mostra que os SNs *Maria* e *João* recebem dois papéis temáticos³⁷.



Esse fato é problema para a Teoria θ , porque, para essa teoria, os papéis temáticos são etiquetas (*tags*) colocadas na estrutura sintática. Já na TEC, como já foi discutido anteriormente, os papéis temáticos são definidos em termos de posições na estrutura conceitual. Em outras palavras, na TEC, duas (ou mais) posições na estrutura conceitual podem ter o mesmo índice, o que significa dizer que um mesmo constituinte da estrutura sintática pode preencher duas (ou mais) posições na estrutura conceitual.

Na figura abaixo, papéis temáticos múltiplos são representados porque, na sentença (a), cada um destes SNs, *Maria* e *João* (objeto sintático da preposição *de*), está coindexado a duas posições, a saber: o SN *Maria* está coindexado simultaneamente à posição indexada pelo

³⁶ Embora o Tema da contra-transferência, o dinheiro, não esteja sintaticamente representado, ele está presente na estrutura conceitual como o primeiro argumento da função IR, que, por sua vez, é o argumento da função TROCA.

³⁷ Observe-se que não se poderia dizer que os papéis temáticos da contra-transferência são irrelevantes, dado que é a contra-transferência que permite a distinção entre *comprar* e *obter*.

índice i da função conceitual IR (Meta do livro) e à posição k da função conceitual TROCA (Origem do dinheiro); o SN *João* está coindexado simultaneamente à posição indexada pelo índice i da função conceitual TROCA (Origem do livro) e à posição k da função conceitual IR (Meta do dinheiro). A figura abaixo ilustra o resultado da operação de fusão de argumentos, necessária para construção da estrutura conceitual da sentença (a), que “preencheu”, com as duas interpretações (MARIA e JOÃO), as posições previstas na estrutura conceitual do verbo *comprar*.

$$\left(\begin{array}{c} \text{IR}_{\text{Poss}} ([\text{LIVRO}]_j, \left[\begin{array}{c} \text{DE } [\text{JOÃO}]_k \\ \text{PARA } [\text{MARIA}]_i \end{array} \right]) \\ \\ \left[\text{TROCA } [\text{IR}_{\text{Poss}} ([\text{DINHEIRO}], \left[\begin{array}{c} \text{DE } [\text{MARIA}]_i \\ \text{PARA } [\text{JOÃO}]_k \end{array} \right]) \right] \end{array} \right)$$

Essa representação, entretanto, ainda exige mais refinamento, porque em sentenças passivas como (b), por exemplo, o argumento sintático sujeito (externo ao SV) torna-se implícito. Nessa sentença, os papéis temáticos Meta da transferência principal e Origem da contra-transferência não estão sintaticamente realizados. Porém, na estrutura conceitual, as duas posições que correspondem a esses papéis precisam ser preenchidas pelo mesmo indivíduo, porque essa configuração de participantes é parte integrante da transação COMPRAR.

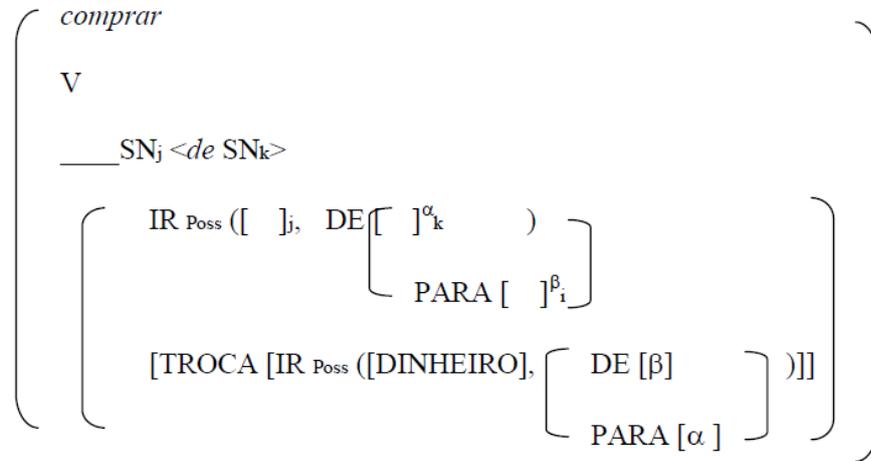
(b) *O livro foi comprado de Maria.*

Já em sentenças como (c), em que não se explicitam a Origem da transferência principal e a Meta da contra-transferência, não se pode especificar que ambos sejam desempenhados pelo mesmo indivíduo (Maria pode muito bem ter recebido o livro de um indivíduo e ter dado o dinheiro para outro).

(c) *Maria comprou o livro.*

Para solucionar as lacunas na TEC ilustradas pelos exemplos (b) e (c), Jackendoff postula dois tipos de coindexação: (1) uma coindexação que conecta um constituinte sintático a um constituinte conceitual, os índices i e k da figura anterior; e (2), uma coindexação que

estipula que dois ou mais papéis temáticos são realizados pelo mesmo indivíduo, conforme ilustra a figura seguinte.³⁸



Na figura acima, são explicitados os dois tipos de coindexação: as letras do alfabeto latino sinalizam a relação entre posições sintáticas e conceituais e as letras do alfabeto grego sinalizam a ligação entre posições na estrutura conceitual.

A Teoria θ relativizada passa a ter a seguinte forma:

O novo critério- θ

Cada índice que liga uma estrutura sintática a uma estrutura conceitual numa entrada lexical deve estar especificado uma única vez na estrutura conceitual. Todos os demais papéis temáticos que o SN por ventura desempenhar deverão estar expressos por argumentos ligados ao constituinte conceitual indexado. (JACKENDOFF, 1990, p. 64).

Assim temos o efeito formal desejado para a *Teoria de Princípios e Parâmetros*: cada argumento sintático está ligado a um único argumento conceitual.

4. Considerações finais

Desta forma, o sub-módulo Papéis Temáticos pode comportar três parâmetros e operar entre o *Componente Sintático* e o *Componente Léxico-Semântico* do UNITRAN: Clíticos, Duplicação de Clíticos e *Nom-Drop*.

Cumprê destacar que a atribuição dos papéis temáticos depois da análise gramatical leva à construção da representação do significado da sentença de entrada em termos da interlíngua do UNITRAN. A atribuição dos papéis temáticos é consolidada depois da geração

³⁸ Observe-se que nesta figura não há mais duplicação de índices que especificam a correlação entre as estruturas sintática e conceitual.

da estrutura sintática, a fim de assegurar que os argumentos gerados combinem com os papéis temáticos que lhes são atribuídos pelo predicador que os seleciona. A atribuição dos papéis temáticos, seguindo a TEC, não é especificada com a aposição de símbolos como *Agente* e *Tema* aos N^o, mas é automaticamente calculável a partir das posições, na Estrutura Conceitual da sentença, a que correspondem os constituintes sintáticos. Quanto ao seu grau de dependência, o sub-módulo Papéis Temáticos depende de informações contidas nas entradas lexicais dos léxicos das línguas envolvidas no processo de tradução.

O parâmetro *Clíticos*, acessado durante a precompilação, é utilizado para especificar a categoria dos clíticos (elementos foneticamente fracos).

Clíticos	
Inglês	N ^o
Espanhol	CL - dativo CL - acusativo
Português	CL - dativo CL - acusativo

O parâmetro *Duplicação de Clíticos* dá conta da atribuição de papel temático a sentenças do espanhol do tipo *Lo vimos a Guille*. Nessa sentença, esse parâmetro assegura que o clítico *lo*, co-referente ao N^o *Guille* (que não tem papel temático), lhe transfira o seu papel temático.

Duplicação de Clíticos	
Inglês	Não
Espanhol	Sim
Português	Não

Por fim, o parâmetro *Nom-Drop* efetua o a atribuição de elementos pleonásticos.

Nom-drop	
Inglês	Não
Espanhol	Sim
Português	Sim

Com a aplicação das restrições sintáticas implementadas nos dezenove parâmetros distribuídos nos sete sub-módulos previstos pelo sistema, o Componente Sintático do UNITRAN dá conta de divergências sintáticas significativas entre o inglês, o espanhol e o português e constrói tanto a estrutura sintática da sentença da língua de partida quanto a estrutura sintática da sentença da língua de chegada, a partir da representação em interlíngua

do significado da sentença da língua de partida. Desta forma, tem-se a teoria de princípios e parâmetros juntamente com a TEC formando o arcabouço teórico para a representação dos papéis temáticos por sistemas de tradução automática.

Referências

- CHOMSKY, N. (1981). **Lectures on Government and Binding: The Pisa Lectures**. Holland: Foris Publications. Reprint. 7th Edition. Berlin and New York: Mouton de Gruyter, 1993.
- COOK, V. J. **Chomsky's Universal Grammar: an Introduction**. Blackwell, 1989.
- DORR, B. J., "The Use of Lexical Semantics in Interlingual Machine Translation," *Machine Translation*, 7:3, pp. 135--193, 1992.
- _____. **Machine Translation: A View from the Lexicon**, MIT Press, Cambridge, MA, 1993.
- FILLMORE, C. "The Case for Case". In: Bach and Harms (Ed.): *Universals in Linguistic Theory*. New York: Holt, Rinehart, and Winston, 1-88. 1968.
- GRIMSHAW, J. *Argument Structure*. The MIT Press, 1990.
- HALE, K; KEYSER, S. J. **Prolegomenon to a Theory of Argument Structure**. MIT Press, 2002.
- HIGGINBOTHAM, J. **On Semantics**. *Linguistic Inquiry* 16:547--931.
- JACKENDOFF, R. **Semantic Structures**. Cambridge, MA: MIT Press, 1990.
- LEVIN, B. **English Verb Classes and Alternations: A Preliminary Investigation**, University of Chicago Press, Chicago, IL. 1993.
- PUSTEJOVSKY, J. **The Generative Lexicon**. MIT Press, Cambridge. 1995.
- VERSPoor, C. **Contextually-Dependent Lexical Semantics**. Edimburgh: University of Edimburgh. College of Science and Engineering. 1997.

