

2. Pressupostos teóricos: um modelo de computação on-line integrado a uma concepção minimalista de língua

2.1 Concepção Minimalista de Língua

A Teoria Linguística Gerativista concebe a linguagem como uma faculdade mental inerente ao ser humano. Conforme argumentado por Chomsky (1997) as crianças desenvolvem, em curto período, um sistema de conhecimentos cuja riqueza não está diretamente relacionada com a quantidade de informações presentes no meio em que estão inseridas. A não correspondência entre a relativa pobreza de estímulos ante a complexidade da gramática desenvolvida na aquisição da linguagem coloca um problema a que Chomsky refere como *problema de Platão*. O problema traz a seguinte questão: como é possível que as crianças desenvolvam uma gramática não sendo explicitamente instruídas para tal? Além disso, como é possível que crianças produzam palavras e estruturas frasais complexas nunca antes ouvidas? Diante desse problema, a Teoria Linguística parte do pressuposto de que tem de haver restrições para que a gramática de uma língua seja passível de ser aprendida em período tão curto.

As diretrizes do Gerativismo (Chomsky, 1981, 1986, 1995) no que concerne as metas de uma teoria linguística são satisfazer aos critérios de adequação empírica, em caráter descritivo e explanatório. O primeiro visa a garantir que um modelo de língua seja capaz de descrever sua gramática de forma que esta gere as sentenças da língua e somente estas. O segundo, por sua vez, visa a garantir que a teoria linguística apresente as propriedades que gramáticas de línguas naturais devem ter para que a criança convirja para a gramática da língua a que é exposta. Em outras palavras, visa a caracterizar o estado inicial das gramáticas das línguas humanas, em termos de uma Gramática Universal (GU)². Nesse sentido, abordagens psicolinguísticas dirigidas à aquisição da linguagem

²Gramática Universal: teoria das propriedades da forma das gramáticas das línguas naturais e/ou dos princípios que regem a constituição das mesmas, concebida como um modelo formal do estado inicial da aquisição da linguagem entendido como um programa biológico específico da espécie humana que garante a criação/aquisição de línguas naturais. (cf. Corrêa, 2006)

pressupõem uma gama de hipóteses acerca do que é extraído pela criança, nesse período, como relevante na língua em que a mesma está imersa. A hipótese de *bootstrapping*³ fonológico (Morgan & Demuth, 1996; Christophe et al., 1994) baseia-se, especificamente, na questão da segmentação do sinal da fala pela criança. Segundo essa proposta, a análise fonológica feita pela criança do sinal acústico da fala forneceria material necessário sobre a estrutura gramatical da língua a que ela está exposta, permitindo, assim, a aquisição do léxico, em geral, e, particularmente, a aquisição da sintaxe da língua. Em outras palavras, a partir de enunciados linguísticos recortados do ambiente linguístico, a criança mapearia o léxico, a morfofonologia, a morfossintaxe e outras informações gramaticais, em sua inter-relação.

Cabe ressaltar que, ainda que o estímulo linguístico contenha uma série de pistas que possibilitam a identificação de uma gramática, é necessário supor que a criança tem de ser predisposta a reconhecer tais “pistas”, ou seja, a reconhecer informação que seja linguisticamente relevante nos dados da fala. Essa visão possibilita que se compreenda GU como um estado inicial “rico”, que viabiliza o reconhecimento do material linguístico pelo criança.

De acordo com a proposta do Programa Minimalista, a forma assumida pelas gramáticas das línguas humanas atende às pressões das interfaces da língua interna com os sistemas que atuam no desempenho linguístico, assim satisfazendo as condições de legibilidade impostas pela arquitetura da mente/cérebro na qual a língua se insere. Cito: Chomsky (1997)

A faculdade da linguagem se encaixa dentro da arquitetura maior da mente/cérebro. Ela interage com outros sistemas, que impõem condições que a linguagem deve satisfazer se for para ser de todo usável. Estas poderiam ser pensadas como ‘condições de legibilidade’, no sentido que outros sistemas precisam ser capazes de ‘ler’ as expressões da língua e delas fazer uso para o pensamento e a ação. (p. 46)

No Programa Minimalista (PM) (1995-2007) é concebida uma língua-I, língua interna, constituída por dois componentes, um léxico e um sistema computacional, e a derivação linguística tem como resultado a representação da forma fonética (*PF*) e da forma lógica (*LF*). Estas constituem interfaces com os

³ O termo *Bootstrapping* refere-se ao desencadeamento ou “inicialização” do processamento de acesso e consiste num modo de teorizar a aquisição da linguagem que privilegie a identificação por parte da criança de padrões (sintáticos, morfológicos, fonológicos ou semânticos) que possibilitem o desencadeamento da sintaxe ou de aspectos do significado lexical.

sistemas articulatório-perceptual (A-P) e o sistema conceitual-intencional (C-I), respectivamente, os chamados sistemas de desempenho.

O léxico é adquirido no decorrer do desenvolvimento linguístico e é constituído por elementos, caracterizados em termos de traços semânticos, formais (ou gramaticais) e fonológicos. Traços semânticos se referem às propriedades relativas ao significado e vinculam-se aos sistemas intencionais e conceituais da cognição não linguística e são restringidos por esta. Os traços fonológicos representam as propriedades dos fonemas que definem a forma fônica das palavras/morfemas e são restringidos pelas propriedades do sistema perceptual e articulatório do ser humano. Os traços formais dizem respeito às propriedades relevantes para a sintaxe da língua, tais como gênero, número, pessoa (traços- ϕ), caso, QU etc., uma vez que viabilizam a computação sintática, informando ao sistema computacional de que maneira as estruturas devem ser construídas. Os elementos do léxico são ainda divididos em duas categorias, *funcionais* e *lexicais*, em função de suas propriedades semânticas e sintáticas, o que tem certo efeito morfofonológico. Categorias lexicais agrupam palavras ou morfemas dotados de conteúdo semântico, como verbos, nomes, adjetivos e algumas preposições, como *de em o direito de todos*, que expressa posse. Elementos lexicais são de classe aberta e, portanto, são constantemente adquiridos, dada a riqueza do vocabulário das línguas. Categorias funcionais, por sua vez, são de classes fechadas, restritas a poucos elementos, como determinantes, conectivos, verbos auxiliares e morfemas flexionais, basicamente compostos de traços formais que operam de acordo com a parametrização de uma língua em particular. Do ponto de vista fonológico, os elementos funcionais tendem a ser reconhecidos no material linguístico a partir das palavras de conteúdo (cf. Morgan, Shi & Allopena, 1996). Assim, pode-se concluir que apenas a partir de informação fonológica a criança seria capaz de determinar se palavra pertence a categorias lexicais ou funcionais. No experimento conduzido em Morgan, Shi & Allopena (1996), verificou-se, por meio de combinação de diferentes pistas fonológicas, se a criança seria capaz de predizer se um determinado elemento seria pertencente a categorias lexicais ou funcionais. Os resultados apontam que em mais de 80% das situações de teste foi possível predizer se os itens eram funcionais ou lexicais. Estes resultados indicam que, mesmo em línguas com parâmetros distintos, os dados de input percebidos

pelas crianças contêm pistas fonológicas suficientes que viabilizam a categorização das palavras.

Isto posto, elementos de categorias funcionais parecem imprescindíveis para o provimento de posições estruturais relevantes para o comportamento sintático de elementos lexicais nos domínios nominal, verbal e oracional. No caso das orações interrogativas QU e relativas, aqui investigadas, propriedades que permitem que um elemento seja movido de uma dada posição sintática para a periferia esquerda da sentença são representadas na língua em termos de traços formais de categorias funcionais.

O sistema computacional, concebido como universal, ou seja, comum a todas as línguas, opera sobre os traços formais do léxico, disparando as operações: *select*, *merge*, *agree* e *move* que atuam sobre uma Numeração⁴, um conjunto de itens lexicais selecionados. A derivação linguística resultante das operações do sistema computacional é, então, disponibilizada para a representação da Forma Fonética (Phonetic Form) e da Forma Lógica (Logical Form), e deve ser passível de interpretação nesses níveis – *PF* só interpreta traços fonológicos e *LF* só interpreta traços semânticos e traços formais interpretáveis, conforme ilustrado no esquema de derivação linguística abaixo:

⁴ Uma Numeração é dada a partir de um arranjo de itens lexicais (LI) e seus índices (i) que corresponde ao número de vezes em que LI foi selecionado do léxico e ao número de vezes em que o elemento poderá ser acessado na sintaxe. A derivação linguística conclui-se apenas quando todos os índices estiverem zerados.

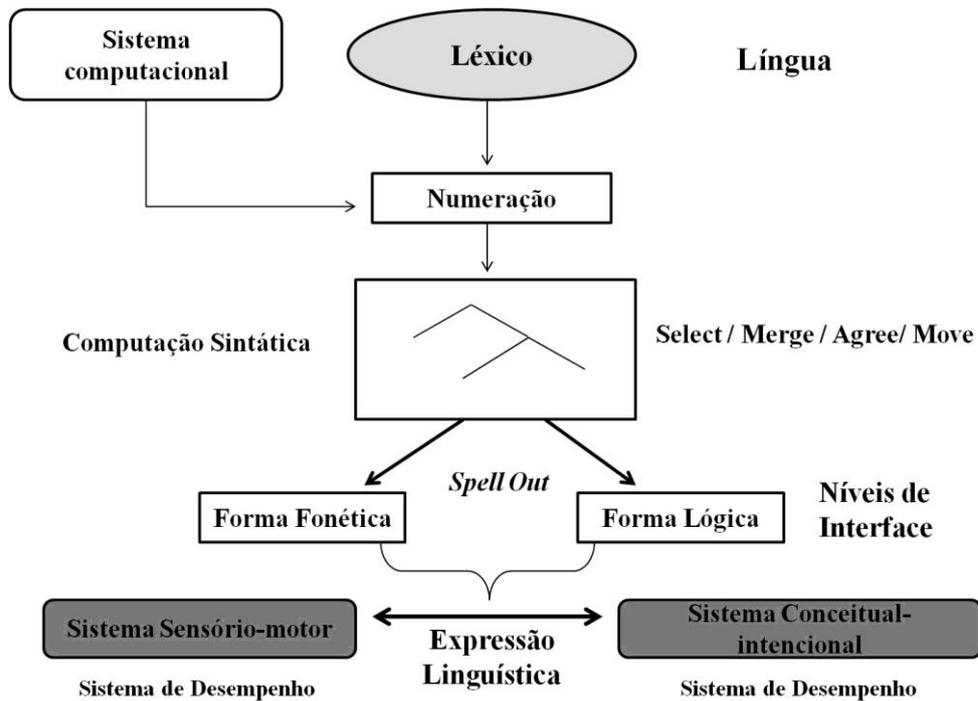


Figura 1: Esquema de uma derivação linguística no Programa Minimalista. (Corrêa, 2008, p. 196)

O sistema linguístico constitui-se, assim, de módulos específicos relativos ao léxico, à sintaxe, à semântica e à fonologia. Tal concepção revela uma especificidade não somente no que concerne ao sistema cognitivo da língua em suas relações com os demais sistemas cognitivos da mente, mas, ainda, de uma especialização de subsistemas dentro do próprio módulo linguístico. Admitindo-se tal pressuposto, pode-se considerar que as manifestações decorrentes de patologias de linguagem seriam específicas de tais subdomínios ou mesmo de suas intersecções, como no caso do chamado DEL pragmático.

As operações de que se vale o sistema computacional atuam de forma recursiva, isto é, o sistema é constituído, estritamente, como um gerador de estruturas linguísticas dotadas de som e significado – digo estritamente já que o mesmo apenas “enxerga” os traços formais selecionados dos elementos advindos do léxico. A operação *select*, seleciona os elementos da Numeração. Em seguida, a operação recursiva de concatenação *Merge* agrupa os elementos selecionados em uma estrutura sintática de forma hierárquica. A checagem (Chomsky, 1995) ou valoração (Chomsky, 1999) dos traços, na operação *Agree*, garante que a derivação linguística seja visível nas interfaces, eliminando ou valorando os traços formais não interpretáveis no curso da derivação. A operação *Agree* é regida por

uma das condições que atuam na derivação linguística, o Princípio de Interpretabilidade Plena (PIP). Este princípio garante que os dados de uma computação linguística passíveis de interpretação semântica e fonológica sejam visíveis nas interfaces, possibilitando que a expressão linguística se realize em sons (forma fônica) e significado (forma lógica) pelo falante ou analisado na codificação de uma mensagem pelo ouvinte.

A operação *Move* permite que os constituintes sejam movidos de sua posição original para um outro nível da estrutura. Conforme exemplificado por Chomsky (1997), tal operação toma um objeto X e um objeto Y (parte de X) e junta Y a X. A expressão então gerada inclui o que se chama de *cadeia não trivial*, que consiste de duas ocorrências distintas de Y; a ocorrência na posição original inicialmente chamada de *vestígio* – na teoria de regência e ligação (Chomsky, 1982) – é hoje caracterizada como *cópia* (usualmente sem expressão fonológica). Os elementos concatenados (*Merge*) ou movidos (*Move*) são alocados de forma hierárquica ocupando uma posição correspondente àquela em que ocorrem na ordenação linear da fala – imposta pelos parâmetros de ordem da língua e, no caso dos elementos movidos, por motivações discursivas. A concatenação por *Merge* obedece a parâmetros de ordem e a por *Move* é motivada por questões de interface com a sintaxe, FF e LF.

O sistema computacional obedece ainda às condições de Economia, relacionadas ao mecanismo de derivação sintática que implica que as operações sejam minimamente complexas⁵. A proposta do PM assume que se em determinado passo derivacional houver a possibilidade de *Merge* e *Move* o sistema computacional optará por *Merge* (Chomsky, 1995). Isso é dado por conta de as estruturas concatenadas por *Merge* não acarretarem custo computacional, dado que não alteram a configuração da língua. Por outro lado *Move* é uma operação complexa que engloba *Merge+Cópia*, modificando a ordem dos constituintes de sua posição canônica. Portanto, o sistema computacional opta por fazer uso da operação mais simples.

Diante disto, operações derivadas de movimento sintático de constituintes – estruturas passivas, relativas, topicalizações e interrogativas QU – têm custo computacional alto previsto para o processamento linguístico (Friedmann &

⁵ Em sua vertente metodológica, o Programa Minimalista procura simplificar análises, eliminar estipulações descritivas e outras soluções de “engenharia linguística” (Raposo, 1999).

Novogrodsky, 2004; Corrêa & Augusto, 2011). Tal complexidade está vinculada a alteração da ordem canônica da língua em questão, motivada por demanda discursiva, atribuindo-se, assim, maior carga à memória de trabalho, que precisa manter ativo um elemento que posteriormente deverá ser reativado para que a computação sintática seja bem-sucedida. Este ponto será investigado, na próxima seção, por meio de uma perspectiva de processamento.

2.2 Uma abordagem psicolinguística nos estudos da linguagem

A Psicolinguística se dedica tanto ao estudo do modo como o conhecimento linguístico é adquirido e representado quanto dos modos como se realizam os processos de produção e de compreensão de enunciados verbais. Tal abordagem tem como objeto de estudo o processamento linguístico ocorrido na mente, com vistas à sua explicitação. Acredita-se que essa explicitação auxiliaria no tratamento de problemas relativos ao processamento linguístico por parte de alguns falantes, em que se observa problemas específicos na linguagem.

Tendo essa pretensão, a Psicolinguística lida com os seguintes direcionamentos: investigar como se realiza a produção de enunciados, a compreensão e a aquisição da linguagem de um ponto de vista funcional; prover modelos teóricos das entidades e processos mentais pertinentes a cada um desses três processos. No entanto, a Psicolinguística investiga a identificação dos fatores que atuam nesses mesmos processos, de modo que seja possível prover uma explicação efetiva sobre a natureza de diferentes manifestações do desempenho linguístico humano.

A linha investigativa em que este trabalho se baseia busca uma articulação entre os campos da Psicolinguística e da Linguística, mais especificamente da Teoria Linguística Gerativista, em sua vertente minimalista (Corrêa, 2002; 2005; 2006; 2008; Corrêa & Augusto, 2007; 2011). Tal articulação mostra-se profícua nos estudos relacionados à aquisição e processamento da linguagem, uma vez que a abordagem linguística se propõe a prover um modelo formal de língua que se adéqua à arquitetura da mente. De acordo com a proposta minimalista todos os princípios de GU podem ser vistos como decorrentes do Princípio de Interpretabilidade Plena nas interfaces. Isso significa que tudo o que é

gramaticalmente relevante se faz visível nas interfaces da língua com sistemas de desempenho. Nesse sentido, a incorporação das diretrizes do Programa Minimalista a uma abordagem psicolinguística se faz interessante pelas condições de Economia e Legibilidade impostas pela faculdade da linguagem em suas interfaces e na interação com os demais sistemas cognitivos imprescindíveis para a linguagem.

2.3. Modelo Integrado de Computação On-line (MINC)

O Modelo Integrado de Computação *online* (Corrêa & Augusto, 2006-11) foi concebido de modo a caracterizar a computação sintática em tempo real. Conforme mencionado acima, o modelo incorpora os pressupostos minimalistas relativos à concepção de língua. Portanto, é previsto, em tal modelo de processamento, um mecanismo que faz uso das operações de um sistema computacional linguístico que extrai do léxico elementos ricos de informações pertencentes a uma dada língua. Em outras palavras, a caracterização formal de língua, proposta no Programa Minimalista, facilita a articulação com uma abordagem psicolinguística na formulação de modelos implementacionais. Modelos de processamento tornam-se importantes na medida em que explicitam o funcionamento do sistema cognitivo da língua na compreensão e produção de enunciados linguísticos, ou no processo de aquisição de uma língua.

Sob os pressupostos do MINC, em relação à computação sintática, esta é iniciada pela composição de um arranjo de itens do léxico (correspondente à Numeração, em derivação minimalista), dada uma intenção de fala (produção) ou na/pela segmentação e reconhecimento lexical na codificação de uma mensagem (compreensão). Enquanto em uma derivação linguística considera-se a computação a partir de um arranjo de elementos do léxico (em princípio obtidos aleatoriamente), na computação on-line os elementos recuperados do léxico mental tanto na produção da fala quando no reconhecimento lexical na compreensão partem de uma intenção de fala e de uma mensagem. Logo a constituição do “arranjo” inicial é motivada.

No que concerne à direcionalidade da computação on-line, esta é caracterizada no modelo em duas vias, que se relacionam com os sistemas

conceptuais-intencionais. Assim, a derivação se dá de cima para baixo (*top-down*), de modo que elementos de categorias funcionais associam-se aos sistemas intencionais motivados pela força ilocucionária do enunciado e pela definição da referência a determinadas entidades e eventos; de baixo para cima (*bottom-up*), a partir da interação de elementos lexicais com sistemas conceptuais motivados pelo conteúdo (proposicional) da mensagem. O diagrama, abaixo, ilustra sinteticamente como se daria a direcionalidade da derivação linguística adotada pelo MINC.

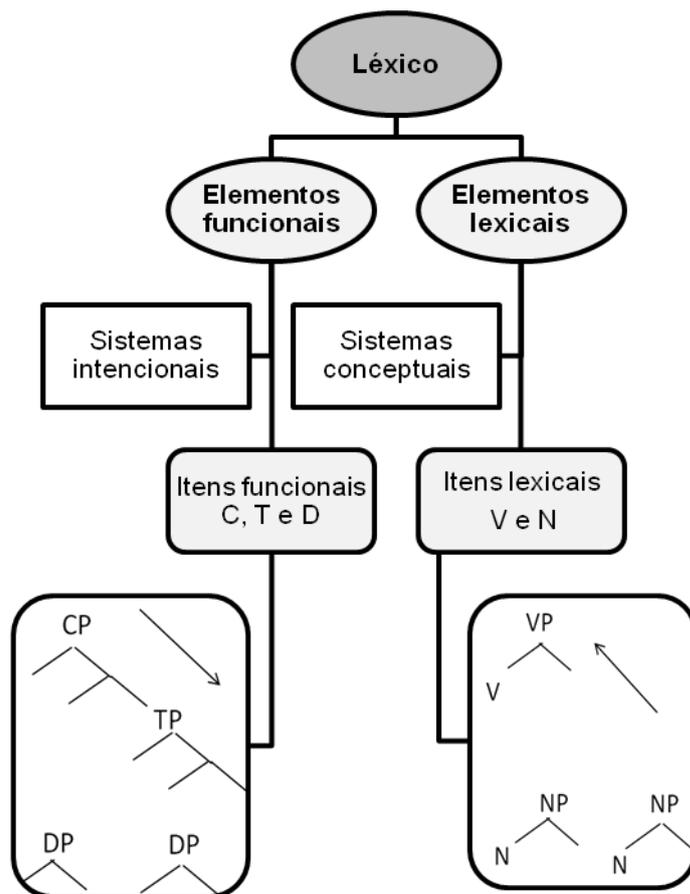


Figura 2: Bidirecionalidade na computação on-line

Conforme ilustrado na Figura 2, o léxico é concebido como um domínio de interface entre a língua e os outros sistemas incluídos no conceito de faculdade da linguagem em sentido amplo (Chomsky, Hauser & Fitch, 2002), como os sistemas conceptual e intencional. Assim, uma estrutura sintática para ser composta deve incluir itens lexicais denotantes (entidades e eventos).

Vale ressaltar que o léxico, tal como é tomado numa abordagem psicolinguística, pode ser concebido como um sistema no qual todo o

conhecimento sobre o que é específico de uma dada língua é organizado e mantido numa memória de longo prazo, em contraste com o léxico em modelos formais de língua, nos quais essa dimensão relativa à manutenção e acesso não é pertinente. Esse modo de conceber o léxico na Psicolinguística possibilita explicar o acesso lexical na construção e na compreensão de enunciados linguísticos.

Constata-se, assim, que o léxico em um modelo de língua e em um modelo de computação on-line são objetos em princípio equivalentes, mas caracterizados de forma distinta, dado os interesses de cada abordagem teórica de pesquisa.

Levando-se em conta custo computacional, relevante para este trabalho, o modelo assume dois tipos de movimento sintático: movimento com custo computacional e movimento sem custo computacional. O primeiro identifica estruturas com demandas discursivas específicas e se utiliza de cópias sequenciais, como um recurso formal para explicar esse custo. O segundo diz respeito ao resultado da fixação de parâmetros de ordem estrutural da língua no início da aquisição da linguagem, e faz uso de cópias simultâneas cuja implementação se daria instantaneamente, como um recurso formal para caracterizar a ausência de curso na computação on-line.

A identificação de estruturas com custo computacional elevado, como orações relativas e interrogativas QU e QU+N de objeto, contribui para a investigação dos fatores que implicam dificuldade na compreensão e na produção da linguagem. A assunção de que tais estruturas fazem uso de cópias sequenciais para a sua formação permite prever que, na compreensão, o elemento alvo em uma interrogativa QU+N como em *Que gato; o cachorro mordeu__i?* tem de ser mantido numa “caixa de memória” o elemento movido para posterior reativação numa segunda posição estrutural (indicada na sentença como um segundo índice (*i*)) no decorrer do processamento. Essa distinção teórica sobre os tipos de movimento permite esclarecer o fato de crianças DEL apresentarem comprometimento apenas em estruturas que envolvem movimento custoso. Traços presentes em função de condições discursivas específicas informariam ao sistema computacional da necessidade deste posicionar elementos em posições distintas das posições canônicas, acarretando custo computacional mensurável.

As ilustrações abaixo, apresentadas por Corrêa & Augusto (2011), demonstram a computação conduzida em espaços paralelos na produção de uma sentença interrogativa QU de objeto (Figura 3). Tal derivação seria decorrente de

uma intenção de fala em que núcleos funcionais (C, T e D) entram na computação, direcionada de cima para baixo, como projeções máximas (CP, TP e DP). Por outro lado, os elementos lexicais entram na computação (de baixo para cima) como projeções mínimas (N, V e P (algumas preposições)), a partir do conteúdo da mensagem, incluindo o elemento QU (de natureza pronominal). Cabe dizer que, neste caso, tem-se um DP ramificado por se tratar do operador “Que+N”, como na frase, *Que gato o cachorro molhou?*. No caso de interrogativas QU (“quem” ou “o que”), o DP não é ramificado, portanto, tem-se uma projeção mínima e máxima do determinante (D = “quem” ou “o que”).

Na figura 4, ilustra-se o posicionamento do DP-QU no local onde recebe papel temático (objeto do verbo) e na posição onde ocorre, em decorrência de movimento sintático, com o uso de cópias sequenciais.

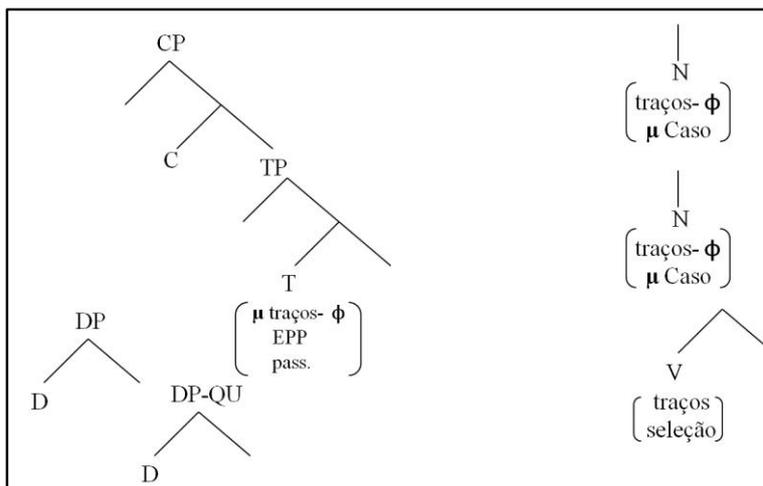


Figura 3: Derivação em espaços paralelos (Corrêa & Augusto, 2011, p. 5)

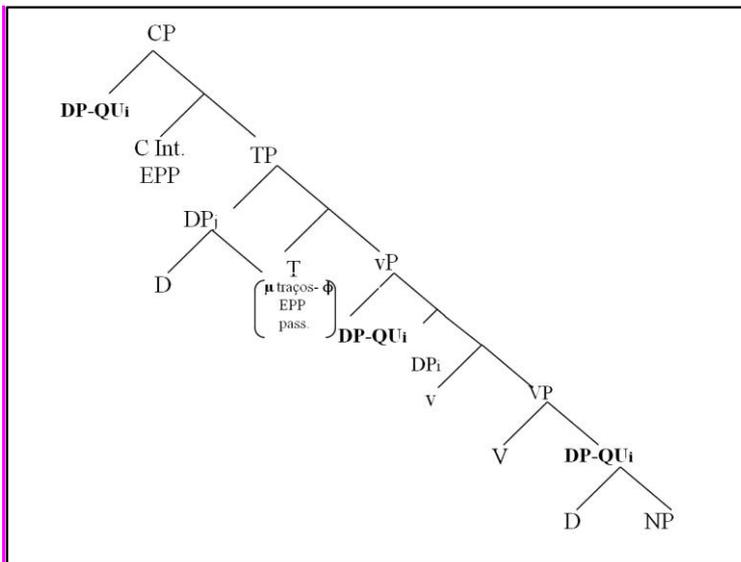


Figura 4: Cópias sequenciais de elementos QU (Corrêa & Augusto, 2011, p. 5)

As autoras explicam que as demandas impostas pelo movimento de um DP-QU para a posição de objeto do verbo impõe a construção de uma cópia extra em Spec, vP a ser recuperado posteriormente. Dessa forma, o DP sujeito da oração principal estaria associado a duas posições ao mesmo tempo: a de sujeito sintático, que controla a concordância verbal em Spec, TP; e a posição de Spec, vP, em função das exigências semânticas do verbo. O acoplamento das derivações em espaços paralelos ocorre em função da relação estabelecida entre T e vP. Dessa forma, a exigência do elemento QU (C) passa a ser suprida pelo movimento do DP-QU para o seu especificador.

Nota-se na Figura 4 que a criação de cópias sequenciais adiciona um custo à memória, uma vez que esta tem de manter por longo período o DP sujeito movido somando-se, ainda, a complexidade computacional imposta por um elemento interveniente até que o primeiro DP seja reativado na posição de objeto. Sendo assim, pode-se prever que orações interrogativas QU de objeto são mais custosas do que as de sujeito, e o mesmo acontece em orações relativas. Tal diferença é decorrente da posição de origem do elemento movido. Nas estruturas de sujeito, a posição de objeto é preenchida logo que o verbo é processado, demandando menor carga de processamento à memória de trabalho. Estruturas de objeto, por sua vez, requerem que o elemento movido seja mantido por mais tempo na memória, já que este tem de se manter ativo até que a posição de objeto seja preenchida.

As considerações do MINC no que se refere à complexidade imposta no processamento de relativas e interrogativas QU de objeto permite uma melhor compreensão acerca dos fatores que motivam o custo. Deste modo, assumiram-se, aqui, as diretrizes do modelo com vistas a verificar em que medida tal custo pode ser acentuado ou minimizado, em função de demandas diferenciadas quanto ao tipo de elementos computados no processamento de tais sentenças. Buscou-se verificar também se o custo pode ser diferenciado em situações discursivas em que um referente prévio do DP interveniente seja imediatamente integrado no processamento das construções de objeto investigadas aqui, antecipando, dessa forma, a atribuição de seu papel temático.