

Modelo de produção PMP revisto e ampliado

O Modelo **PMP** (produção monitorada por *parser*) foi concebido com vistas a explicar a interferência de fatores não sintáticos no processamento da concordância, nomeadamente efeitos de *distância linear* entre sujeito e verbo e efeitos de *marcação morfofonológica* dos núcleos nominais do DP sujeito.

Efeitos desse tipo permitem questionar a hipótese de um formulador sintático autônomo, não sujeito à interferência de outros componentes do sistema de produção, e parecem, em princípio, favorecer a idéia de interação entre níveis de processamento e de retroalimentação (*upward feedback*) no fluxo de informações.

Nesse contexto, o modelo apresenta-se como uma alternativa entre os modelos de natureza serial para explicar tanto interferências de fatores sintáticos como de fatores não-sintáticos. A incorporação de um *parser* monitorador à arquitetura do sistema de produção permite unificar a explicação para todos os efeitos em termos de acessibilidade de uma representação do DP sujeito gerada pelo *parser*. O erro deixa de ser atribuído a falhas no curso da formulação sintática do enunciado (a codificação gramatical) e passa a ser associado a limitações de processamento e armazenamento de informação pela memória de trabalho.

Nesta seção, apresenta-se uma versão revista e ampliada do modelo, na qual são consideradas três alternativas de explicação para os erros de concordância, a partir da hipótese de uma escala de acessibilidade das representações geradas pelo *parser*.

9.1

Erros de atração e escala de acessibilidade das representações geradas pelo *parser*

Em estudos relativo à acessibilidade de formas pronominais na produção e na compreensão de sentenças, Corrêa (1993; 2000) explica a possibilidade de

recuperação de um dado referente por uma forma anafórica a partir de uma **escala de acessibilidade** desse antecedente. Essa explicação tem como um de seus pressupostos teóricos a idéia de que o processamento da sentença no discurso é dependente de um sistema de memória de trabalho, o qual compreenderia pelos menos dois níveis distintos de processamento, que poderiam ser caracterizados em termos de subsistemas de memória. O **primeiro subsistema**, correspondente a uma **memória de curto-prazo**, manteria representações de natureza **literal** (envolvendo itens lexicais) ou **quase literal** (envolvendo informação de natureza gramatical depreendida dos itens lexicais). O **segundo sub-sistema** manteria de forma mais estável **representações conceptuais** abstraídas das primeiras, possivelmente já de natureza semântica ou proposicional.⁹⁴

Se considerarmos que as representações do DP sujeito geradas pelo *parser* estão sujeitas a essa escala de acessibilidade e que a acessibilidade de uma informação na memória é função tanto do próprio material lingüístico quanto de limitações individuais de memória, 3 possibilidades de explicação podem ser formuladas para dar conta dos erros de atração:

Explicação 1:

Esta explicação para os erros de atração corresponde exatamente à apresentada na primeira versão do modelo (cf. Rodrigues & Corrêa, 2004; Corrêa & Rodrigues, 2005). Nesta versão, assumia-se que o *parser* monitorador atuaria de forma *bottom-up*. Considerava-se que o primeiro DP emitido só seria analisado pelo *parser* como o sujeito da sentença quando o verbo fosse produzido. A memória manteria, num componente temporário, informação de **natureza morfofonológica** associada aos traços de número dos núcleos nominais que integram a representação do DP sujeito gerada pelo *parser*-monitorador. No caso de DPs com núcleo do sujeito singular, questões de acessibilidade associadas à saliência fônica da marca de número fariam com que um núcleo interveniente (plural) mais alto na estrutura hierárquica desse sintagma (ou num PP argumento)

⁹⁴Haveria ainda um nível de memória mais alto, relacionado ao processamento do discurso. Este nível manteria em ativação constante a representação dos elementos em torno dos quais o discurso como um todo se organiza bem como o planejamento de natureza pré-textual, o qual seria passível de alteração a partir da produção dos enunciados (van Dijk & Kintsch, 1983). Neste trabalho, como está sendo analisada a concordância sujeito-verbo em sentenças isoladas, não será discutido o processamento de informações nesse nível de memória.

fosse tomado como elemento definidor do número do DP sujeito como um todo. No momento em que fosse ocorrer a codificação morfofonológica do verbo, uma espécie de resíduo morfofonológico associado à representação do DP gerada pelo *parsing* poderia ainda estar presente na memória e vir a gerar interferência nessa codificação do verbo. No caso de DPs com N1 plural, não ocorreria erro porque a informação de número associada ao núcleo do sujeito seria menos suscetível a esvaecimento. Para uma representação do modelo, ver esquema apresentado ao final do capítulo 7.

Explicação 2:

Diferentemente do que foi proposto na primeira versão do modelo, nesta segunda explicação considera-se que o *parser* poderia analisar o DP enviado da formulação como o sujeito da sentença e gerar uma árvore *top-down* para a sentença. Com base na informação referente ao número do DP sujeito, o *parser* já poderia antecipar (em termos de **expectativa**) qual seria a flexão do verbo. Vejamos, em detalhes, como isso se daria.

A representação do DP gerada pelo *parser* manteria ativa informação de **natureza morfológica**, associada à informação de número extraída dos núcleos nominais do DP sujeito. Nesse caso, a informação de número teria uma natureza menos “literal”, em termos de um morfema abstrato de número, o qual poderia vir a afetar a codificação morfofonológica do verbo da seguinte maneira:

O DP, tão logo é disponibilizado pelo sistema de produção, dentro de uma estrutura prosódica específica⁹⁵, é analisado pelo *parser*, que, após o acesso aos itens lexicais, começa a construir o esqueleto funcional da sentença. Partindo-se da idéia de que, na compreensão de sentenças, o número do verbo (que irá aparecer no afixo verbal) já pode ser **antecipado** com base em informação de número do DP sujeito, pode-se propor que esta informação ficaria ativa na memória de trabalho e poderia vir a afetar a codificação morfofonológica do verbo.

⁹⁵ Nessa explicação, considera-se que o *parser* poderia levar em conta informação prosódica ao segmentar o continuum da fala e que essa informação já daria indicações acerca da natureza do DP emitido, fazendo com que este fosse analisado como potencial sujeito da sentença. Assim, o esqueleto funcional começaria a ser projetado paralelamente ao reconhecimento de elementos do léxico em uma janela de processamento.

Paralelamente ao *parsing* do DP sujeito, o processo de codificação morfofonológica do restante da sentença estaria ocorrendo. Quando, no componente morfofonológico, o verbo fosse ser concatenado a um morfema abstrato de número resultante da computação da concordância na produção, **dois** morfemas de número estariam disponibilizados – aquele resultante da concordância na produção e aquele resultante da “**previsão**” estabelecida com base no número do DP sujeito. Haveria, então, uma espécie de **competição** entre dois possíveis morfemas. No caso de identidade entre os morfemas, o resultado seria uma forma verbal cujo afixo de número espelharia o número do DP sujeito. No caso de não-identidade, o erro poderia ocorrer.⁹⁶

Note-se que essa situação seria semelhante àquela que pode ocorrer quando há o processamento da fala de outra pessoa: o DP inicial seria analisado como o sujeito da sentença e uma previsão acerca do número e a pessoa do verbo seria estabelecida. Se um erro de concordância é produzido, o ouvinte pode imediatamente perceber o erro, visto que há uma quebra de expectativa em relação à informação inicial.

O erro, assim como na explicação 1, estaria associado a questões de acessibilidade e manutenção, na memória de trabalho, da informação de número da representação do DP sujeito gerada pelo *parser*. Quando o sujeito é um DP complexo, a informação de número do núcleo, em função de questões de distância linear e de marcação, pode ter sua acessibilidade afetada, levando a que se tome como informação de número aquela codificada em um segundo núcleo nominal mais acessível. A acessibilidade desse outro núcleo será função de questões de sua posição estrutural e de marcação morfofonológica. É, portanto, com base na informação do núcleo interveniente que o processador pode “antecipar” qual será o número do verbo que está prestes a ser produzido.

Um dado interessante que pode ser tomado como evidência a favor da idéia de competição entre os morfemas são as correções feitas pelo falante: é relativamente comum que o falante, tão logo produza o erro, corrija o que disse;

⁹⁶ Nessa proposta, assume-se que o radical verbal é representado separadamente de sua flexão. Essa separação entre morfema lexical e morfema flexional vem sendo atestada desde os primeiros trabalhos sobre erros produzidos na fala espontânea (Garrett, 1975, 1980). Erros como o da frase “A menina caixou a guarda” (no lugar de “A menina guardou a caixa”) são tomados como evidência de que os morfemas flexionais seriam representados separadamente dos radicais ou morfemas lexicais.

muitas vezes, a correção tem início antes mesmo de o verbo ter sido concluído (ex.: *O tecido das cortinas do teatro rasga(ram)/rasgou*). Também ocorre de o falante começar com a forma correta e substituir pela forma errada (ex. *O tecido das cortinas do teatro rasgou/rasgaram*). O quadro a seguir ilustra como se daria o processamento da concordância e o ponto em que ocorreria o erro:

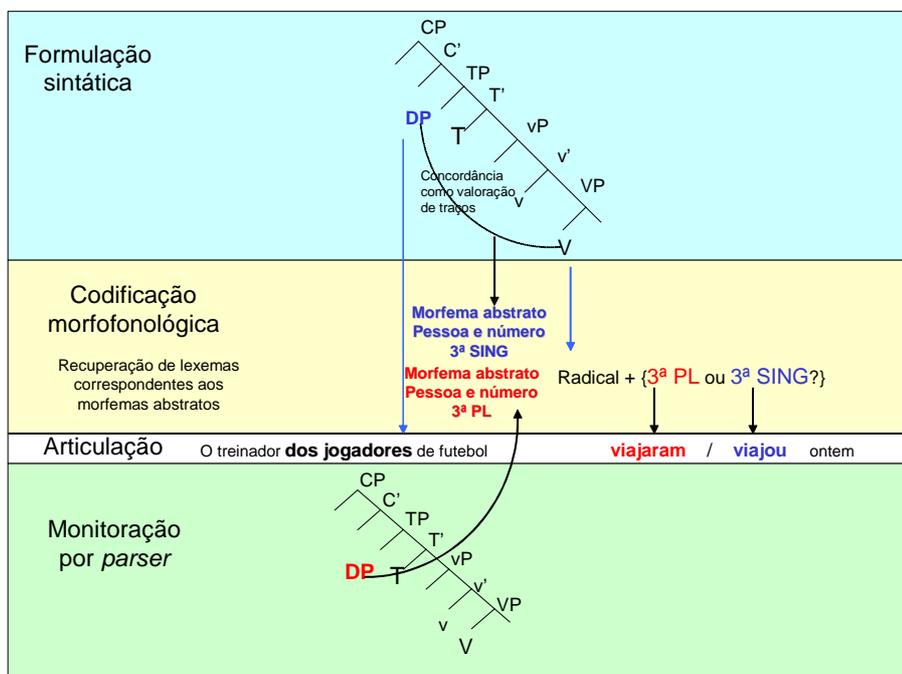


Fig. 9: Modelo PMP revisado e ampliado (explicação 2) – Em vermelho, morfema abstrato de número gerado pelo *parser* a partir de previsão feita com base em informação de número do DP analisado como sujeito da sentença.

Explicação 3:

A representação do DP gerada pelo *parser*, uma vez interpretada e integrada com informação prévia, daria origem a uma **representação semântica/conceptual**, a qual seria retomada na forma de um **elemento pronominal nulo**. Esse elemento pronominal, por sua vez, teria seu traço de número especificado a partir dessa representação semântica do DP.

A idéia de atribuir a interferência de fatores semânticos/contextuais na concordância à retomada do DP por um elemento pronominal nulo foi originalmente proposta por Corrêa (2005b; 2006) para explicar efeitos semânticos na concordância de gênero sujeito-predicativo. Nesses trabalhos, Corrêa investiga a possibilidade de o DP inicial, processado como tópico, ser fechado como uma unidade semanticamente independente antes de o verbo ser introduzido na

formulação, o que iria requerer a introdução de um elemento pronominal que retomasse seu referente como sujeito do verbo. Esse elemento pronominal nulo recuperaria, então, os traços semânticos do referente do DP inicial fechado como uma unidade semanticamente autônoma e a concordância procederia, sem problemas, com o predicativo concordando com o sujeito nulo recém incorporado à formulação do enunciado. Uma vez que o predicativo concordaria com o sujeito nulo, o resultado soaria como um erro de concordância entre este e o DP processado como tópico.

Assim, por exemplo, no caso de erros de concordância em sentenças com DPs com leitura distributiva, o que seria retomado pelo elemento pronominal nulo sujeito do verbo seria a idéia de pluralidade do referente do DP inicial, levando a um aparente erro de concordância.

É possível que também com DPs **não**-distributivos a concordância verbal se estabeleça com um pronome nulo. Isso ocorreria toda vez que o DP inicial fosse tomado como um tópico e que a representação mantida na memória fosse de natureza conceptual, correspondente ao DP “fechado”. Como nos DPs não-distributivos (*o engradado das garrafas*), o que seria retomado seria um referente singular, esse processo não seria evidenciado na forma verbal, que seria singular.

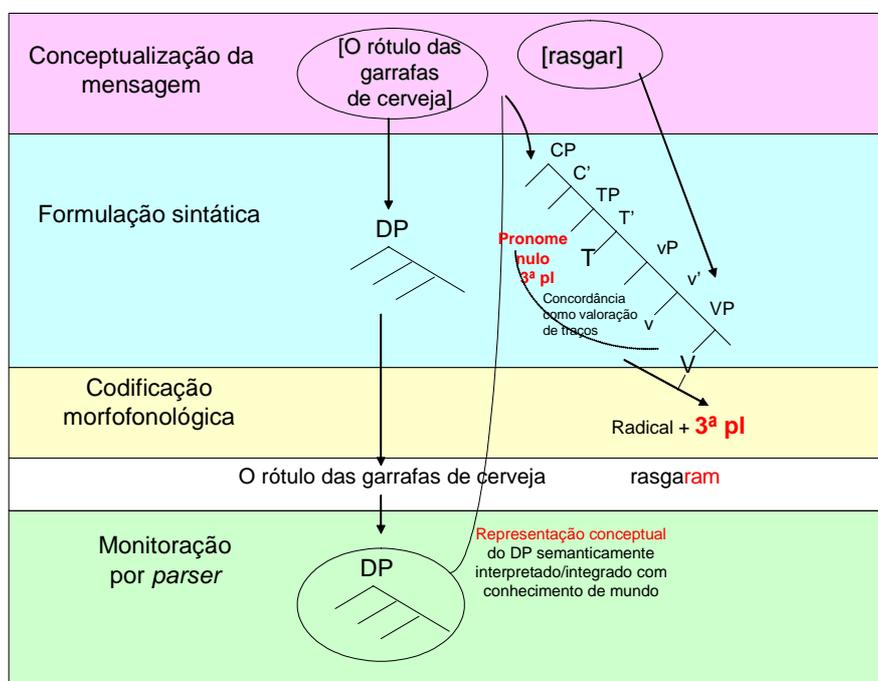


Fig. 10: Modelo PMP revisto e ampliado (explicação 3) – Retomada da representação conceptual do DP resultante do *parsing* através de pronome resumptivo e computação da concordância. Os colchetes indicam representação de natureza conceptual.

9.2

Modelo PMP e Programa Minimalista: considerações acerca de um tratamento unificado para a computação sintática da concordância

Para fechar a seção, cumpre fazer algumas observações acerca da compatibilidade entre o tratamento da concordância sujeito-verbo no modelo PMP e a computação sintática da concordância no modelo de língua proposto no Programa Minimalista. Em primeiro lugar, cabe notar que, embora no modelo PMP, assumam-se a idéia de valoração de traços formais para caracterizar a computação da concordância, não há um comprometimento com uma implementação específica da computação da concordância desenvolvida no âmbito do Programa Minimalista. Isso significaria tomar a derivação sintática no modelo formal de língua como idêntica à derivação *on-line* da sentença, o que pode ser problemático em termos de dois aspectos básicos: i) a direcionalidade da derivação sintática x formulação sintática; ii) questões de ordem associadas a movimento sintático.

Em relação ao primeiro ponto, tem-se apontado como dificuldade para um tratamento unificado da derivação/formulação de sentenças, o fato de a derivação sintática no Programa Minimalista ser *bottom-up* (ascendente), da direita para a esquerda, enquanto a formulação de sentenças seria *top-down* (descendente), da esquerda para a direita (cf. Phillips, 1996, 2003). Corrêa (2005a, 2006a) observa que, em termos estritamente formais, no momento em que o Minimalismo passa a assumir uma abordagem relacional para a caracterização de nós sintáticos (*bare phrase*), não é mais tão claro o quão crucial é a direção da derivação sintática. Nesse sentido, talvez seja possível buscar uma compatibilidade entre derivação em termos formais e derivação em tempo real. Corrêa propõe que a derivação em tempo real se daria tanto numa direção *top-down* quanto numa direção *bottom-up*. O esqueleto funcional, que permite a explicitação da força ilocucionária da sentença e a indicação de informação temporal acerca do evento descrito, seria derivado *top-down*, e a concatenação de categorias lexicais, que permitem expressar o conteúdo propriamente dito da mensagem (a partir das relações argumentais), seria feita *bottom-up*. Um modelo de processamento que incorpore uma derivação *on-line* do formulador sintático nos termos propostos por Corrêa

precisará, portanto, acoplar uma derivação *top-down* com uma derivação *bottom-up*.⁹⁷

Em relação ao segundo ponto – a questão do movimento sintático, Corrêa (2005a, 2006a) argumenta que uma teoria que proponha identificação entre derivação lingüística e processamento precisaria distinguir dois tipos de movimento sintático: movimentos que descrevem processos pertinentes ao posicionamento de elementos na ordem canônica com que se apresentam na língua (ex. movimento de núcleo e movimento de DP para posição A) e movimentos que descrevem processos que alteram essa ordem canônica, com vistas a adequar a fórmula do enunciado lingüístico a um particular contexto discursivo (ex. Movimento WH, Topicalização, e Focalização). Segundo Corrêa, o primeiro tipo de movimento estaria associado ao processo de aquisição da língua e poderia ser visto como uma descrição do tipo de operação envolvida quando da fixação de parâmetros de ordem. Esse movimento corresponderia, pois, a um processo bem inicial na aquisição da língua e possivelmente estaria “gravado” no cérebro, o que tornaria sua condução automática quando os parâmetros correspondentes estivessem fixados. Já o segundo tipo de movimento seria determinado por ajustes do enunciado a propósitos comunicativos específicos e, nesse sentido, seria decorrente de cada emissão/compreensão e ocorreria durante o processamento dos enunciados. Essas diferenças se refletiriam em termos de custos computacionais: enquanto o primeiro tipo deveria ser isento de custo, dada sua natureza automática, o segundo deveria ter um custo determinado por demandas específicas de processamento.

Logo, uma aproximação entre derivação lingüística e processamento precisaria levar em conta essa distinção, o que possivelmente teria impactos no

⁹⁷ “ [...] o acesso a elementos de categorias funcionais do léxico na produção partiria de relações de interface entre o sistema da língua e “sistemas intencionais”, viabilizadas pelos traços semânticos/pragmáticos dos mesmos e os elementos recuperados seriam tomados como projeção máxima. O acesso a elementos de categorias lexicais, por sua vez, partiria de relações de interface entre o sistema da língua e “sistemas conceptuais”, viabilizadas pelos traços semânticos desses elementos. Torna-se necessário, pois, dissociar os chamados sistemas C-I (Conceptuais - intencionais), quando se considera computação em tempo real. Os sistemas conceptuais interagiriam com elementos de categorias lexicais, enquanto que os intencionais com elementos de categorias funcionais, cada um dos quais dando início, possivelmente em paralelo, a derivações/computações com direcionalidades distintas, as quais se acoplam, uma vez que verbo e tempo se relacionam numa única estrutura hierárquica.” (Corrêa, 2005a)

que tange aos mecanismos formais usados para a caracterização da derivação sintática da sentença.

Não é claro, por exemplo, em que medida a computação sintática da concordância na produção/compreensão poderia ser caracterizada em termos de um sistema de *Probe-Goal*, implementado pela operação *Agree*. Nesse sistema, a valoração de traços se dá a partir de uma relação de c-comando, que se estabelece entre um núcleo mais alto (*probe*- sonda) cujos traços são não-interpretáveis e precisam ser valorados e um núcleo mais baixo (*goal* – alvo), com traços interpretáveis. Se, como propõe Corrêa, o DP inicial já se encontra na sua posição estrutural final, definida no processo de fixação de parâmetros, não haveria como o traço não-interpretável de número de T ser valorado. Isso só seria possível se imaginássemos que uma cópia do DP estaria nessa posição mais baixa, como uma forma de estabelecer uma relação temática entre o DP e o verbo da sentença.

Em **suma** um tratamento unificado da computação sintática concordância ainda carece de investigação mais aprofundada. De todo modo, a despeito de o modelo PMP não se comprometer com um sistema específico da implementação da concordância, pode-se dizer que é compatível com a idéia de que a derivação sintática transcorreria automaticamente. Isso porque o modelo explica os “erros de atração” **não** como falhas no processo de formulação sintática, mas sim em termos de uma interferência de uma representação gerada pelo *parser*, a qual estaria sujeita a limitações da memória de processamento. Outro ponto interessante do modelo é que ele permite explicar diferenças entre línguas em termos da visibilidade de como informação de número é expressa, o que também é compatível com a idéia de que as línguas se diferenciam em relação a seus traços formais e a como estes são realizados.