Síntese e considerações finais

A questão da autonomia do formulador sintático é um dos temas centrais para as teorias psicolinguísticas voltadas à investigação do processamento de sentenças. Nesse cenário, erros de concordância entre sujeito-verbo se apresentam como um caminho para se examinar em que medida processos sintáticos podem sofrer interferência de informação diferente da que pode ser codificada como traços formais do léxico em um modelo de língua. Nesta tese, buscou-se examinar qual a natureza dos erros de atração e o momento do processamento em que tais erros atuam. Também foi objetivo do estudo diferenciar erros de atração de casos de concordância licenciados pela gramática da língua. Assumiu-se como **hipótese de trabalho** a idéia de que o formulador sintático é autônomo e funciona de modo encapsulado, sem sofrer interferência de outras fontes de informação durante o processamento.

Buscou-se, nesta tese, uma **aproximação com a teoria gerativista** nos seus desenvolvimentos mais recentes – o Programa Minimalista, o qual assume que o sistema cognitivo da língua deve atender a condições de legibilidade impostas pelos sistemas de desempenho com os quais a língua faz interface. Pontos de compatibilidade entre o que seria a derivação de uma sentença nesse modelo de língua e a derivação de sentenças em tempo-real foram apresentados e questões relativas a uma integração entre o modelo de língua proposto no Minimalismo e modelos de processamento foram discutidas.

Com vistas a distinguir estruturas responsáveis por erros de atração de construções que admitem a concordância com um nome plural de um PP modificador, realizou-se uma **análise linguística das chamadas construções partitivas**, as quais permitem tanto a concordância com o termo que designa a parte quanto com o termo que designa o todo. Propôs-se que cada alternativa de concordância seria função da natureza “híbrida” (lexical e funcional) dos termos partitivos “maioria” e “parte”. Duas representações estruturais distintas para a construção partitiva foram, então, consideradas, associando-se cada estrutura a uma alternativa de concordância. Um experimento de julgamento de gramaticalidade indicou diferenças entre construções partitivas e DPs complexos,

tendo sido consideradas gramaticais as sentenças com construções partitivas, independentemente da concordância verbal. Também foi verificado que há menor rejeição à concordância plural com as partitivas.

A **revisão da literatura** acerca de erros de atração revelou que **fatores sintáticos, semânticos e morfofonológicos** podem atuar no processamento da concordância. Os fatores sintáticos que induziram erros foram a *distância linear entre o núcleo do sujeito e o verbo*; o *tipo de modificador* e a *posição hierárquica de um núcleo interveniente no DP sujeito*. Entre os fatores semânticos, destaca-se a *distributividade do DP sujeito*. Quanto a efeitos morfofonológicos, houve efeito de marcação associado ao *número* do núcleo do sujeito e número do nome local. Também foi verificado que as línguas podem exibir diferenças quanto a uma assimetria entre singular e plural na indução de erros de atração. Experimentos conduzidos em inglês mostraram ser possível dissociar um efeito morfofonológico de um efeito meramente fônico.

Foi visto que explicações fornecidas por **modelos interativos e não-interativos** não dão conta de efeitos de ordem linear entre o núcleo do sujeito e o verbo e de diferenças entre línguas quanto a efeitos morfofonológicos. O **modelo de produção monitorada por parser (PMP)** (Rodrigues & Corrêa, 2004; Corrêa & Rodrigues, 2005) apresenta-se como uma alternativa a esses modelos. As principais propriedades deste modelo são a **produção moderadamente incremental** e o **monitoração concomitante à fala por um parser, que atua de forma bottom-up**. A idéia de uma produção moderadamente incremental é compatível com a hipótese de que a concordância verbal (entendida como processo de valoração de traços formais da língua) seria computada antes de o DP sujeito ser enviado para a codificação morfofonológica com vistas a sua articulação. A proposta de um *parser* monitorador permite manter a autonomia do formulador sintático e explicar efeitos de distância linear e efeitos morfofonológicos no processamento da concordância. Neste modelo, os erros são atribuídos a uma interferência na codificação morfofonológica do verbo provocada por uma representação do DP sujeito gerada pelo *parser*.

Os resultados de cinco experimentos de produção induzida de erros conduzidos com falantes de português revelaram a interferência dos seguintes fatores:

- i) **Distância linear entre o núcleo do sujeito e verbo:** a incidência de erros foi maior nas condições em que a distância linear entre núcleo do sujeito era longa.
- ii) **Posição hierárquica do núcleo interveniente no DP sujeito:** em preâmbulos do tipo *A tinta do cartucho da impressora*, houve mais erros de concordância quando o núcleo interveniente era o N2 do que o 3, isto é, núcleos próximos do nó mais alto do DP sujeito induziram mais erros do que aqueles numa posição linear adjacente ao verbo.
- iii) **Status argumental do PP modificador:** núcleos intervenientes inseridos em PPs argumentos induziram mais erros do que aqueles inseridos em PPs adjuntos.
- iv) **Número do núcleo do sujeito:** a incidência de erros de atração foi maior quando o núcleo do sujeito era singular (não-marcado); não houve erros após núcleos do sujeito plural.
- v) **Número do DP do nome local:** houve mais erros quando o número do DP do nome local era plural do quando era singular. Esse efeito se observou independentemente da realização morfofonológica do nome local: se este era um substantivo flexionável (*os carros*) ou um substantivo de forma invariante (*os ônibus*).
- vi) **Tipo de DP distributivo:** DPs com núcleo interveniente plural (*a maçaneta das portas*) induziram mais erros de atração do que DPs distributivos com o operador “cada” (*a maçaneta de cada porta*).
- vii) **Distributividade:** foi obtido um efeito de distributividade, quando se ampliou o tamanho do DP distributivo com núcleo interveniente. O total de erros de atração foi muito maior após DPs distributivos (*a sola dos sapatos de couro*) do que após DPs não-distributivos (*o armário dos sapatos de couro*).

Efeitos de *distância linear entre o núcleo do sujeito e o verbo* podem ser explicados a partir do Modelo PMP que incorpora um *parser* monitorador. No processo de *parsing*, representações mantidas na memória de trabalho estariam sujeitas a esvaecimento progressivo. Na condição distância linear longa, haveria possível perda de informação de número codificada no núcleo do sujeito, o que favoreceria a interferência de informação de número de um núcleo interveniente.

Efeitos de *posição hierárquica* e de *número do DP do nome local* também são compatíveis com a proposta de um *parser* monitorador, pois uma representação gerada pelo *parser* conteria informação relativa à posição hierárquica dos elementos do léxico reconhecidos, bem como informação relativa aos traços morfofonológicos identificados. A acessibilidade do traço de número de um núcleo interveniente seria função de sua posição hierárquica e de marcação: quanto mais próximo um núcleo interveniente marcado (plural) estivesse do nó mais alto do DP sujeito mais acessível seria sua informação de número e maior a chance de esta vir a causar interferência.

Um efeito do *status argumental do PP* em que está inserido o núcleo interveniente também pode ser explicado em termos da manutenção, na memória, dos elementos que integram a representação gerada pelo *parser*. O núcleo interveniente que estivesse inserido em um PP argumento ficaria mais acessível do que aquele inserido em um PP adjunto, porque os argumentos estariam mais estreitamente associados à projeção do NP do que os PPs adjuntos. Estes seriam mais “marginais” na estrutura; logo, menos acessíveis. Daí o maior número de erros de atração provocados por núcleos intervenientes em PPs argumentos.

O modelo PMP também permite explicar o efeito de *número do núcleo do sujeito* e por que o português apresenta um comportamento semelhante ao do inglês (e diferente do francês) em relação a essa variável. De acordo com o modelo, a marcação de número do núcleo do sujeito afeta a manutenção, na memória, da representação do DP sujeito gerada pelo *parser*, a qual, por sua vez, pode interferir na codificação morfofonológica do verbo. Considerando-se que as línguas apresentariam diferenças quanto à **visibilidade** da informação de número do DP sujeito, haveria mais chance de se obter efeito da variável número do núcleo do sujeito naquelas línguas em que a informação de número do sujeito fosse mais visível. O português é bastante semelhante ao inglês em relação a essa propriedade; em ambas as línguas a informação de número é bastante visível (um afixo –s no final dos nomes, no inglês; um afixo –s no final dos determinantes e dos nomes, em português). No francês, por outro lado, a informação de número é menos visível, visto que na fala a informação de número não é marcada nos nomes e os determinantes não apresentam uma marca individualizada de número. Isso explica, portanto, por que em português e inglês apenas o singular induz erros de atração. Nessas línguas, nos preâmbulos com N1plural, a informação

morfofonológica de número estaria mais saliente e, por conseguinte, menos vulnerável a esvaecimento, o que dificultaria a ocorrência de erros nessa condição.

Ainda em relação a fatores morfofonológicos, também é possível explicar a partir do modelo PMP por que não há diferenças em termos de erros de atração entre preâmbulos cujo nome local é um substantivo flexionável (*os carros*) ou um substantivo de forma invariante (*os ônibus*). Em português brasileiro, a informação de número relevante é aquela codificada no determinante do DP do nome local; logo, o *parser*, ao extrair a informação de número desse DP, leva em consideração o traço de número do determinante, independentemente de informação expressa no nome local. Quando o número do determinante é plural (marcada), fica mais acessível na memória e tem maiores chances de prevalecer sobre a informação do núcleo do sujeito.

O modelo PMP, em sua primeira versão, foi concebido com vistas a explicar particularmente efeitos sintáticos e morfofonológicos no processamento da concordância. Para explicar efeitos semânticos, como o de distributividade, bem como para tentar distinguir interferências provocadas por informação morfológica de número daquelas provocadas por informação morfofonológica, uma **versão revista e ampliada do modelo** foi proposta.

Nessa versão, buscou-se integrar **3 possibilidades de explicação** para o erro tendo em vista uma **escala de acessibilidade**, na memória de trabalho, da representação do DP sujeito gerada pelo *parser*.

Uma das explicações do modelo ampliado corresponde exatamente ao proposto na primeira versão do modelo PMP. Após o *parsing* do DP, informação de natureza literal (envolvendo itens lexicais) ou quase literal (envolvendo informação de natureza gramatical depreendida dos itens lexicais) seria mantida numa memória temporária para que pudesse ser integrada ao restante da sentença que está sendo produzida em paralelo. Quando o verbo fosse ser codificado morfofonologicamente, uma espécie de **resíduo morfofonológico** resultante do *parsing* do DP produzido, ainda mantido na memória, poderia interferir no processo de codificação do verbo, com reflexos na sua representação fonética. Importante notar que nesta versão o *parser* atua de forma *bottom-up* e o DP inicialmente produzido só é integrado em uma estrutura sintática, na função de sujeito, quando o verbo é efetivamente emitido.

Outra possibilidade, no entanto, precisa ser considerada, a de que o *parser* comece a construir o esqueleto da sentença de modo *top-down* e que o DP inicial seja identificado como o sujeito da sentença. Essa possibilidade é que levou a **ampliação do modelo PMP**.

Nessa segunda explicação para os erros, considera-se que informação de número de natureza menos literal, associada a um morfema abstrato de número do DP sujeito seria mantida na memória. Com base nessa informação, o falante poderia **antecipar**, em termos de expectativa, qual seria o morfema abstrato de número do verbo, que seria expresso por meio de um sufixo no caso do português. Paralelamente ao *parsing* do DP sujeito, o processo de codificação morfofonológica do restante da sentença estaria ocorrendo. Assim, 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 – o morfema de número resultante da concordância na produção e o **morfema “antecipado”** a partir de uma expectativa do número do verbo. Quando por questões de acessibilidade houvesse perda de informação relativa ao traço de número do núcleo do sujeito, a “previsão” feita com base na representação do DP sujeito gerada pelo *parser* seria incongruente em relação ao resultado da computação sintática na produção. Nesse caso, dois morfemas abstratos de número não idênticos estariam competindo no momento da codificação morfofonológica do verbo e poderia ocorrer uma associação do verbo ao morfema resultante da “expectativa” gerada no processo de *parsing*, fazendo com que parecesse ter ocorrido um erro de concordância.

A terceira explicação considera que uma **representação semântica** do DP sujeito, tomado como tópico da sentença, poderia ser retomada na forma de um **elemento pronominal nulo** e que seria com base na informação de número desse elemento pronominal que a concordância sujeito-verbo seria computada. O elemento pronominal codificaria informação de número associada à representação semântica do referente do DP sujeito, o que explicaria efeitos de distributividade no processamento da concordância. Note-se que, de acordo com essa explicação, a valoração do traço de número do verbo seria feita com base no traço de número do pronome e, nesse sentido, não haveria uma falha ou um erro de concordância.

Com base, pois, nas explicações do modelo PMP ampliado é possível sustentar a idéia de que a computação sintática da concordância na produção não

seria em si afetada por informação de natureza semântica ou morfofonológica. Os erros seriam decorrentes de questões de acessibilidade, na memória, da representação do DP sujeito gerada pelo *parser* e, nesse sentido, o erro seria posterior ao *parsing* do DP sujeito. Logo, pode-se dizer que os resultados da presente pesquisa dão suporte à hipótese da **autonomia** do formulador sintático.

Vários tópicos de pesquisa podem ser desenvolvidos como desdobramentos do presente trabalho. Em termos mais amplos, tem-se a articulação entre modelos de processamento e modelos de língua, em particular o modelo desenvolvido no âmbito do Programa Minimalista. Foi visto, ao longo da tese, que a relação gramática e processador não pode ser estabelecida em termos de uma relação de identidade plena e que especificidades do processamento *on-line* de sentenças precisam ser consideradas ao se procurar estabelecer uma aproximação entre o que ocorre na produção/compreensão de sentenças e o modo como se dá a derivação formal destas em modelos de língua. Nesse cenário, questões ligadas a movimento sintático e direcionalidade da derivação se apresentam como pontos de investigação (cf. Corrêa 2005a, 2006a). Em termos do tratamento da concordância sujeito-verbo na produção, seria interessante explorar em que medida os mecanismos formais de implementação da concordância no modelo de língua (checagem, valoração no sistema de *Agree*) estariam de acordo com a implementação da concordância em um modelo de processamento da produção.

Em termos mais específicos, a tese abre para uma investigação de tópicos tanto em Teoria Lingüística quanto em Psicolingüística. Em Lingüística, a análise da estrutura das construções partitivas indicou que ainda há muito o que se explorar quanto à questão da natureza semi-lexical dos itens que entram na derivação sintática da sentença. A computação da concordância de número interna ao DP no caso de DPs complexos também é uma área bastante promissora em termos de pesquisa. Questões de representação de adjuntos e de argumentos e seu impacto para a computação sintática é outro ponto que merece ser investigado.

Em Psicolingüística, há vários aspectos do processamento da concordância a explorar. Seria interessante, por exemplo, comparar resultados de compreensão com os de produção. Não são muitos os trabalhos em compreensão nesta área e certamente há bastante o que ser examinado, em particular quanto à hipótese de que o *parser* atuaria baseado em “expectativas” geradas a partir da análise do

material lingüístico em processamento. No estudo específico da produção, a distinção entre concordância de gênero e número, passível de ser observada em estruturas com adjetivos ou com participios, também merece ser investigada. Nesta tese, foram identificados erros em que uma forma de participio concordava em número com o núcleo do sujeito e em gênero com o nome local (*O depósito das armas de fogo foi desativada; A prateleira dos livros de história estava quebrado*), o que sugere que haveria diferenças em termos de como o processador lida com informação de número e informação de gênero no DP sujeito. Quanto às estruturas responsáveis por erros de atração, foram examinados em detalhes DPs modificados por PPs. Cumpre ainda verificar o que ocorre com modificadores oracionais em termos de indução de erros de atração. Em que medida, como propõe Bock & Nicol (1993), a oração de fato insularia informação de um núcleo interveniente nela inserido é um ponto a ser explorado. Também em termos metodológicos, há bastante o que se desenvolver. É preciso buscar metodologias experimentais alternativas que permitam induzir erros de atração, sem que seja necessário o falante partir de um preâmbulo apresentado pelo experimentador.

Em suma, há muito trabalho a ser feito e, nesse cenário, a presente tese pretende ter contribuído com elementos que possam servir como ponto de partida para pesquisas futuras sobre o processo de produção de sentenças, área em que ainda há poucos trabalhos em português.