

2

Arquitetura do sistema de produção e o processamento da concordância: a autonomia do formulador sintático em modelos interativos e não-interativos

A produção da linguagem é concebida, na maior parte dos modelos de produção, como uma atividade complexa que envolve processamento de informação em diferentes níveis ou estágios (cf. Fromkin, 1971, 1973; Garrett 1975, 1980, 1988; Levelt, 1989; Bock & Levelt, 1994). Considera-se que a produção teria início com a conceptualização da mensagem que se deseja transmitir. Seguir-se-ia a esta etapa um processo de busca de elementos do léxico, os quais portariam traços semânticos compatíveis com os conceitos envolvidos na mensagem assim como com aspectos pertinentes à adequação de registro e a outros fatores de ordem pragmática. O acesso lexical envolveria também a recuperação de informação léxico-sintática do lema³, um tipo de representação específica que agregaria informação gramatical acerca de um dado item lexical. De posse dessas representações, um formulador conduziria a codificação gramatical da sentença. Nessa fase de codificação gramatical, ocorreria a organização hierárquica da sentença e a ordenação dos constituintes de acordo com a ordem linear dos constituintes na língua. A essa fase seguir-se-ia a etapa de codificação morfofonológica da informação proveniente do componente sintático e a codificação fonológica, que possibilitaria a articulação da sentença.

É compatível com esses modelos a idéia de um *parser* monitorador, nos moldes do que propõe Levelt (1989). De acordo com Levelt, o falante usaria o próprio sistema de compreensão para monitorar os enunciados à medida que estes fossem sendo produzidos. O autor propõe que o *parser* teria acesso tanto às representações fonéticas pré-articulatórias, decorrentes do processo de codificação

³ O termo *lema*, introduzido por Kempen & Hoenkamp (1987; porém já citado em Kempen & Huijbers, 1983), foi empregado originalmente para fazer referência à palavra como uma entidade semântico-sintática em oposição ao conceito de *lexema*, usado para designar os traços fonológicos da palavra (cf. Levelt, 1989). Com os desenvolvimentos da teoria de acesso lexical, o termo ganhou um uso mais restrito, passando a designar apenas a informação sintática acerca de uma dada palavra. Segundo Corrêa (2005a, 2006a), as propriedades que definem os lemas podem ser relacionadas a traços formais do Léxico de um modelo formal de língua enquanto que as propriedades que definem lexemas podem ser relacionadas com traços fonológicos.

fonológica da sentença, quanto à fala já articulada. A vantagem dessa proposta é que não são necessários recursos adicionais de monitoramento para a detecção de erros: o falante, por ser ouvinte de sua própria fala, também seria capaz de verificar falhas em sua produção (cf. Postma, 2000). A seguir apresenta-se uma representação esquemática do processo de produção em modelos seriais.

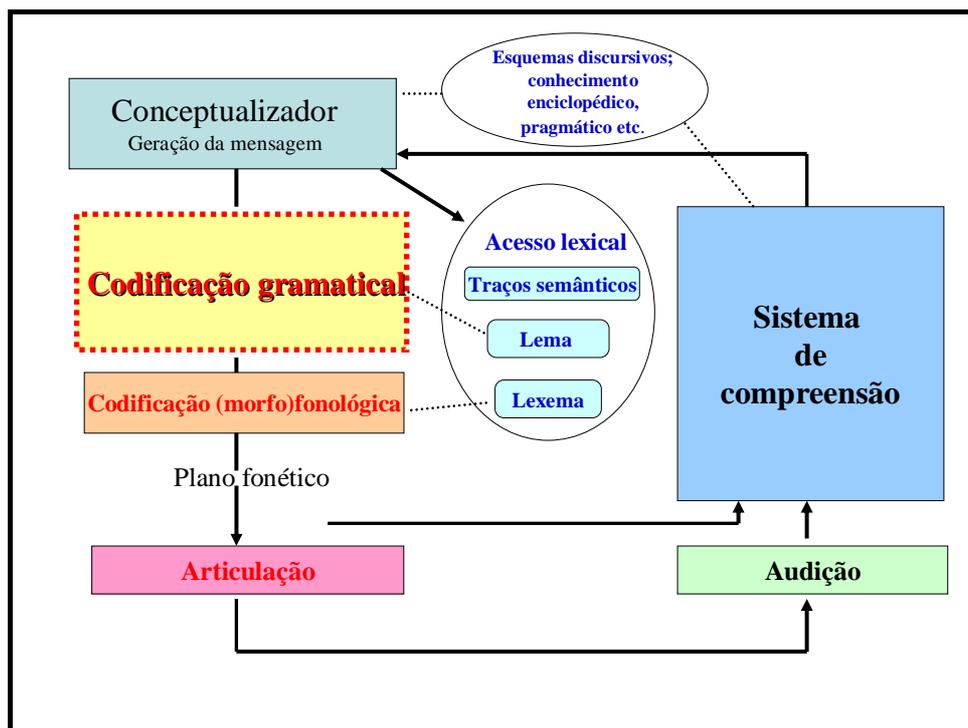


Fig. 1: Arquitetura do sistema de produção da linguagem e monitoração da fala. Adaptação da representação de Levelt (1989) para o processamento de informações pelo falante. As caixas representam componentes de processamento e o círculo representa conhecimento armazenado.

Modelos interativos x não interativos e o processamento da concordância

Uma questão acerca da qual os modelos de produção de sentenças divergem diz respeito a como se dá o **fluxo de informação** ao longo do processamento, mais especificamente, se poderia haver ou não interação entre informações de natureza distinta num dado estágio do processamento. Dependendo da resposta dada a essa questão, os modelos podem ser definidos como não-interativos ou interativos (cf. Vigliocco & Harstuiker, 2002).

Modelos não-interativos, como o caracterizado na seção anterior, apresentam, em geral, uma visão serial e unidirecional da produção da linguagem, em que não há possibilidade de interação entre os níveis de processamento (*feedforward models*). O *input* de um nível para o outro é mínimo, isto é, apenas informação relevante e necessária é encaminhada de um estágio para outro.

Os modelos interativos, por sua vez, assumem uma concepção interativa do processamento com possibilidade de convergência de informações de fontes distintas ao longo do processo de formulação de sentenças (Dell, 1986; Stemberger, 1985). Sob esse rótulo, agrupam-se desde modelos de orientação probabilística, baseados em satisfação de restrições, originalmente desenvolvidos no âmbito do estudo da compreensão (cf. Haskell & MacDonald, 2003), até modelos alinhados com a idéia de níveis de processamento (Vigliocco & Hartsuiker, 2002). No caso dos modelos interativos de níveis, considera-se a possibilidade de retro-alimentação de informações provenientes de um dado nível no nível anterior (*feedbackward models*).

Tanto em modelos seriais não-interativos quanto em modelos interativos de níveis, a concordância é computada no estágio de **codificação gramatical**. Nos modelos não-interativos, o formulador sintático estabelece a concordância exclusivamente com base em informação de natureza léxico-sintática, não havendo interferência de fatores semânticos ou morfofonológicos no processamento. Já nos modelos interativos de níveis, considera-se que a concordância, a despeito de ser um processo sintático, poderia mobilizar informação de natureza não-sintática. Nesse caso, o formulador sintático não atuaria de forma autônoma, encapsulada.

Quanto aos mecanismos sintáticos de implementação da concordância entre sujeito-verbo, duas alternativas são consideradas na literatura voltada à investigação de erros de atração: um **mecanismo de cópia de traços** (Kempen & Hoenkamp, 1987) e um **mecanismo de unificação de traços** (Kempen & Vosse, 1989; De Smedt, 1990).⁴

No mecanismo de **cópia de traços** (Kempen & Hoenkamp, 1987), uma fonte ou controlador transmite seus traços sintáticos (por exemplo, traços de

⁴ Pode-se estabelecer uma certa correspondência entre os mecanismos de cópia e de unificação de traços e os procedimentos de valoração (Chomsky, 1999) e de checagem de traços (Chomsky, 1995), propostos no âmbito do Programa Minimalista. O tratamento da concordância no Minimalismo será apresentado no capítulo 3.

pessoa, número, gênero) a um alvo. No caso da concordância sujeito-verbo, assume-se que o sujeito é o elemento controlador e que o verbo herda informação de número e pessoa do sujeito. Os valores dos traços de número e pessoa do sujeito são especificados a partir de um mapeamento de informação de natureza conceptual (no nível da mensagem) para uma representação léxico-gramatical. A transferência dos traços de número-pessoa do sujeito para o verbo se dá através da aplicação sucessiva de operações de cópia.

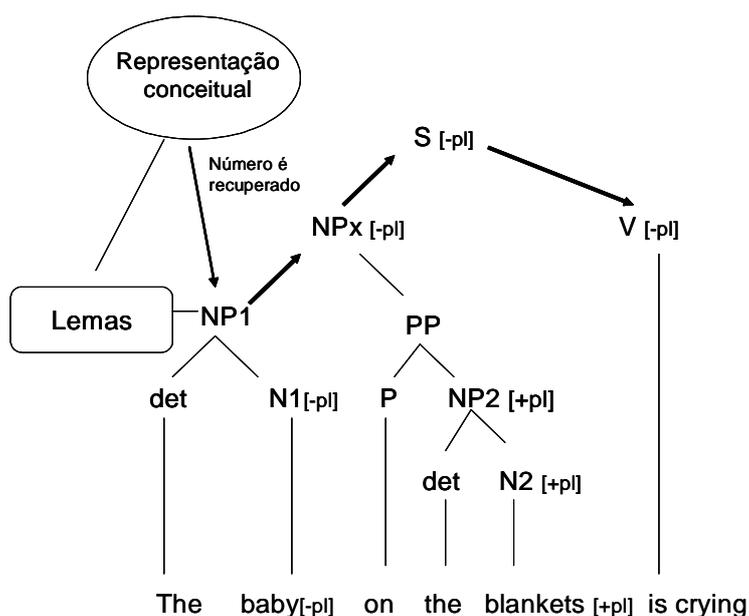


Fig. 2: Concordância como operação de cópia de traços (Vigliocco, Butterworth & Garrett, 1996).⁵

No mecanismo de **unificação de traços** (Schieber, 1986; De Smedt, 1990; Kempen & Vosse, 1989), os elementos envolvidos na relação de concordância apresentam informação relativa a número já especificada. Esses elementos funcionam como ‘parceiros’ de mesmo *status* no estabelecimento da concordância

⁵Na representação da concordância como um processo de cópia de traços, Vigliocco, Butterworth & Garrett (1996) assumem o modelo de gramática procedural incremental (*Incremental Procedural Grammar*), desenvolvido por Kempen & Hoemkamp (1987), no contexto de teorias computacionais de processamento de linguagem natural. Nesse modelo de gramática, os nós categoriais não seriam “elementos estruturais passivos”, mas “procedimentos ativos”, responsáveis pela construção de um tipo de constituinte sintático. Como em qualquer linguagem de programação, é permitida aos procedimentos sintáticos a chamada de subprocedimentos, dando origem a uma hierarquia de procedimentos representada em uma árvore sintática. Os nós terminais são procedimentos lexicais que não evocam nenhum subprocedimento. Os números subscritos servem para distinguir várias instanciações de um mesmo procedimento sintático na hierarquia.

de número. Nessa proposta, o valor do número tanto do sujeito quanto do verbo é resultante de informação proveniente de representações conceituais. Para que os elementos envolvidos na concordância terminem com um mesmo valor, é proposto um procedimento que examina compatibilidades, verifica a interseção dos valores e realiza a unificação dos traços de ambos os nós. Veja-se, a seguir, figura ilustrativa do funcionamento desse processo:

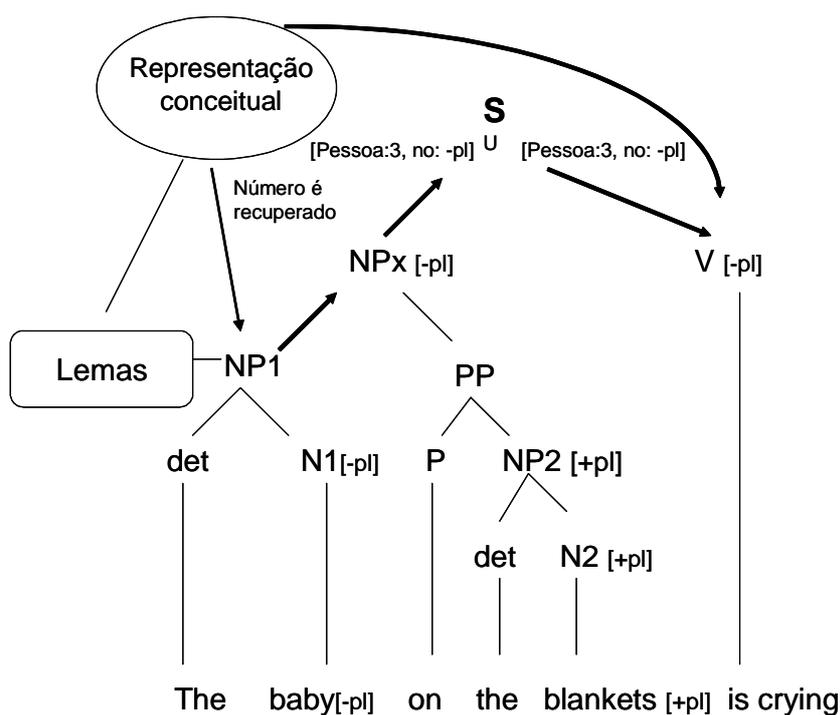


Fig. 3: Concordância como operação de unificação de traços (Vigliocco, Butterworth & Garrett, 1996).

É importante notar que os mecanismos de cópia e de unificação de traços adotados nos modelos de produção para caracterizar a concordância integram modelos computacionais de gramática voltados à geração de linguagem natural. Embora alguns desses modelos busquem atingir um certo grau de plausibilidade lingüística e psicológica (cf. Kempen & Hoenkamp, 1987), eles não estão preocupados com questões de aprendibilidade e tampouco levam em conta limitações próprias de um processador humano ao processamento lingüístico de

sentenças.⁶ Enquanto nos modelos psicolinguísticos, a questão da precisão e da rapidez de como o sistema de produção opera são fundamentais, nos modelos computacionais, a facilidade de implementação do sistema e de gerenciamento das tarefas é crucial. Assim, embora seja possível uma aproximação entre modelos psicolinguísticos e modelos computacionais, visto que ambos buscam caracterizar os procedimentos que permitem a geração de sentenças e os tipos de representação utilizados, é preciso ter em mente as diferenças quanto aos critérios que orientam a construção desses modelos.

Em relação a essa questão dos critérios, tem se revelado como promissora uma aproximação entre modelos de processamento e o modelo de língua desenvolvido no Programa Minimalista da teoria linguística gerativista (Chomsky, 1995; 1998, 1999; 2001). Isso porque nessa proposta o que orienta a organização da arquitetura do sistema linguístico não são apenas critérios de adequação descritiva e explanatória, mas também critérios de economia e de naturalidade conceptual, relativos a como o sistema da língua interage com os demais sistemas cognitivos. No próximo capítulo, será apresentado, em linhas gerais, o modelo de língua proposto no Minimalismo e serão discutidas as compatibilidades entre a derivação sintática de uma sentença no Minimalismo e o processo de formulação em tempo real de uma sentença.

⁶ “[...] *Humans and computers are subject to different constraints, in particular with respect to attention span, memory size, and processing speed. There are also differences concerning the granularity of processes, the communication cost between processors, etc. With respect to representation, computer systems are often made efficient and manageable by a homogeneity at a high level (the representation formalism), whereas humans probably depend on homogeneity at a much lower level (the neuron).* [...]” De Smedt, Horacek & Zock (1996, p.2).