

CUSTO E ECONOMIA NA DERIVAÇÃO SINTÁTICA E NA FORMULAÇÃO/PARSING INCREMENTAL DE SENTENÇAS – APROXIMAÇÕES E ESPECIFICIDADES

Erica dos Santos Rodrigues (LAPAL- PUC-Rio)

ericasr@puc-rio.br

1- Introdução

A Psicolinguística, ao longo de sua trajetória enquanto uma das áreas das chamadas Ciências Cognitivas, tem tido momentos de aproximação e de distanciamento em relação à Teoria Lingüística Gerativista. Nessa trajetória, destaca-se, em especial, o chamado período lingüístico da Psicolinguística, assim nomeado pela das idéias introduzidas por Chomsky no final da década de 50, no livro *Syntactic Structures*. Nesse período foi proposta a Teoria da Complexidade Derivacional, a qual procurava estabelecer uma relação de isomorfia entre o modo como a derivação sintática de uma sentença era caracterizada no âmbito do modelo formal de língua vigente naquele momento e as operações cognitivas associadas à compreensão de sentenças. No entanto, os resultados dessa teoria descritos por Chomsky como “a beautiful example of the way creative experimental studies can interweave with theoretical work in the study of language and of mental processes” (*apud* Smith, 2004), foram em pouco tempo revistos, dado não ser possível estabelecer correlação simples entre complexidade estrutural de determinadas sentenças e tipo de transformações estipuladas para a derivação sintática de tais estruturas.

Um distanciamento progressivo foi observado nos anos seguintes, motivado tanto pelas constantes mudanças na teoria lingüística chomskyana, que ocorriam em paralelo ao surgimento de outros modelos formais de gramática, como a *Lexical Functional Grammar* e a *Head Head-Driven Phrase Structure Grammar*, como pelo interesse por outros fatores de ordem cognitiva, que deflagraram estudos relativos a memória e compreensão de discurso, estudos sobre anáfora, processamento de inferências, trabalhos sobre reconhecimento e acesso lexical, modelos de *parsing*, etc. Esse segundo período, nomeado por alguns autores como período cognitivo, foi extremamente produtivo, e marca uma aproximação ainda mais estreita da Psicolinguística das chamadas ciências cognitivas, como a Psicologia Cognitiva, Filosofia da Mente e Inteligência Artificial. Nos anos 90, com o chamado Programa Minimalista da Teoria Gerativa (Chomsky, 1995), que se orienta por preocupações relativas à naturalidade conceitual de seus construtos, retomam-se discussões relativas à articulação entre processamento lingüístico e gramática (por exemplo, Phillips, 1996; Corrêa & Augusto, 2006, 2007; Lin, 2006).

Esse ‘reencontro’ engendra a retomada de antigas questões – como a da distinção entre competência e desempenho, a da realidade psicológica de construtos lingüísticos, da relação entre as operações computacionais implementadas no curso da derivação sintática e os processos envolvidos na produção e compreensão, etc. – e provoca a reflexão sobre em que termos essa reaproximação pode ser produtiva para a área da Psicolinguística. Neste trabalho, busca-se contribuir para essa reflexão, a partir de uma caracterização de como os conceitos de custo e de economia são tratados pela Lingüística Gerativista no Programa Minimalista e pela Psicolinguística no estudo do processamento. Considera-se que um esclarecimento sobre como cada área os emprega pode ser relevante para avaliar possibilidades de uma visão integrada processador/gramática, em que diferenças relativas ao curso incremental do processamento na produção e na compreensão sejam consideradas. Explora-se, de

forma inicial, como o conceito de fase, próprio da teoria lingüística, pode se apresentar como construto que viabiliza um tratamento integrado da derivação sintática e parsing/formulação de sentenças.

2- O Programa Minimalista e o conceito de Economia Derivacional

No Programa Minimalista (Chomsky 1995, 1999/2001), o sistema cognitivo da língua é concebido como um sistema cuja arquitetura seria perfeita, no sentido de o resultado das computações realizadas por esse sistema serem legíveis pelos demais componentes da mente humana (os chamados sistemas de desempenho) com os quais a língua faz interface. Essa concepção de que a língua é “projetada” para interagir com os sistemas de desempenho e que sofre restrições destes determina crucialmente o formato do modelo, permitindo a discussão acerca de quão psicologicamente “reais” seriam as operações implementadas pelo sistema computacional da língua.

Um conceito fundamental ao Modelo é o de Economia. No Minimalismo, o conceito de Economia tanto pode ser visto como remetendo a questões metodológicas, relacionadas a métricas de simplicidade e parcimônia do modelo, os chamados Princípios de Economia Metodológica, como remetendo a princípios gramaticais associados a uma minimização de esforços na computação sintática, os Princípios de Economia Substantiva (cf. Hornstein, Nunes & Grohmann, 2005).

Em termos de uma articulação com teorias de processamento, interessa-nos em especial o segundo conceito de Economia. Esse conceito é o que orienta, por exemplo, a comparação de derivações sintáticas quanto ao seu custo computacional. No Minimalismo, objetos sintáticos devem não só satisfazer condições impostas pelas interfaces como também ser construídos de forma ótima. Assim, considera-se que entre as derivações convergentes, isto é, cujo resultado é legível nas interfaces, são melhores aquelas que convergem de modo eficiente e econômico. Esta é a métrica usada para avaliar entre duas derivações qual deve ser assumida para caracterizar como foi gerado um dado objeto sintático. Entre os fatores que definem esse “custo” computacional estão a quantidade de operações realizadas, a natureza das operações e o momento em que tais operações ocorreram. Em relação ao primeiro fator, propõe-se que quanto menos operações estiverem envolvidas mais econômica será a derivação. Em relação à natureza das operações, assume-se que uma operação que envolva apenas a concatenação de elementos (a operação de *Merge*) seria menos custosa do que uma operação associada a movimento de constituinte (a operação *Move*). Isso porque esta operação compreenderia duas suboperações – concatenação e deslocamento de constituinte. Quanto ao momento em que as operações ocorrem, Chomsky (2000) observa que a legibilidade de um objeto sintático pode ser avaliada aos poucos, no sentido de que qualquer árvore parcialmente construída ao longo da computação pode ou não ser legível nas interfaces. Segundo Chomsky (1999/2001, 2000, 2004) as derivações procederiam por fases – entendendo-se por fase um objeto sintático cujas partes podem ser inspecionadas para verificar convergência. Logo, o conceito de convergência não se aplica apenas ao objeto sintático final que chega às interfaces. Isso significa que a cada fase da derivação podem-se comparar duas possíveis derivações convergentes para avaliar qual delas é mais econômica, o que implica que se pode abandonar uma delas tão logo se identifique que esta se apresenta como mais custosa.

Relacionando o conceito de Economia com a ideia de complexidade, pode-se considerar que, de acordo com a métrica adotada nesse modelo formal de língua, serão mais complexas computacionalmente as derivações que envolverem mais operações de movimento e um número maior de operações. Deve-se reforçar que essa métrica não é uma métrica psicológica, é uma métrica do modelo formal de língua, usada para

comparar, quanto à Economia, derivações que partam de uma mesma Numeração. Logo, em princípio, a métrica de custo derivacional adotada na proposta Minimalista parece se distanciar da que se faz necessária em modelos de processamento, em que se comparam sentenças com estruturas diferentes (por exemplo, declarativas e interrogativas; ativas e passivas, etc.), as quais partiriam de numerações distintas. Nesse sentido, os conceitos de complexidade e de custo computacional do modelo de língua não podem ser diretamente utilizados para se pensar questões de processamento. Note-se, contudo, que com uma nova concepção de numeração – entendida em termos de subarranjos (*subarrays*) e com o conceito de fase, talvez uma aproximação entre as duas métricas de complexidade possa ser estabelecida e seja possível comparar a computação de sentenças (com subarranjos parcialmente distintos) em determinados pontos da derivação sintática.

3- Incrementalidade na compreensão e produção de sentenças e o conceito de fase da teoria lingüística

Uma questão central com a qual modelos de produção e compreensão da linguagem precisam lidar é o da incrementalidade (Levelt, 1989). Por incrementalidade entende-se que, no processo de codificação/decodificação de informação, operações de um dado nível podem ter início antes de o processamento das unidades do nível anterior ter sido concluído. Quão incremental é esse processamento é ponto de discussão em um conjunto de trabalhos em particular no que tange à produção da linguagem, para a qual vem sendo considerada uma incrementalidade moderada (Ferreira, 2000; Rodrigues, 2006).

Dada a relevância desse conceito, ao se buscar estabelecer uma aproximação entre derivação sintática e parsing/formulação de sentenças é necessário buscar um construto, na teoria lingüística, que possa viabilizar uma caracterização incremental do processamento. Esse construto parece ser o de fase, como veremos em um dos itens desta questão. Antes, contudo, faz-se necessário uma breve caracterização do conceito de incrementalidade na compreensão e na produção.

3.1 Incrementalidade na compreensão

Há um conjunto expressivo de evidências de que os ouvintes reconhecem palavras e analisam e compreendem sentenças de forma incremental. No caso do processamento sintático, que nos interessa mais diretamente neste projeto, há uma vasta literatura sobre efeitos de *garden-path*. Pesquisas sobre esses efeitos indicam que o mecanismo de *parsing* realiza uma análise automática do material lingüístico à medida que este vai sendo disponibilizado. Isso explica por que frases com ambiguidades temporárias como *The horse raced past the barn fell* (Bever, 1970) trazem dificuldades para o processamento. O mecanismo de *parsing* concatena, de modo incremental, imediato, a forma *raced* ao SN *The horse*, assumindo que se trata do verbo da oração principal, a qual se completaria em *The horse raced past the barn*. No entanto, ao acessar a forma *fell*, o *parsing* verifica que tomou uma decisão incorreta/escolheu o caminho errado no labirinto e que será necessário refazer a análise. Efeitos similares ocorrem com outros tipos de sentença, apontando para uma atuação muito rápida do *parser* (Altmann & Steedman, 1988; Frazier & Rayner, 1982; Rayner, Carlson & Frazier, 1983; Trueswell, Tanenhaus & Garnsey, 1994). Nesses trabalhos em geral se utiliza a técnica de leitura auto-monitorada e se assume como variável dependente o tempo que se leva para ler os segmentos da frase. Cumpre notar que, embora os trabalhos apontem para a incrementalidade do processamento, há muita controvérsia quanto aos tipos de informação e momento em que estas são utilizadas pelo

processador, sendo possível delimitar duas grandes abordagens sobre o processamento de sentenças – uma abordagem modular, que trabalha com a ideia de encapsulamento informacional e uma separação entre informação sintática e semântica/discursiva, e uma abordagem interativa, em que várias fontes de informação (sintática, semântica, discursiva, relativa a frequência) estariam disponíveis simultaneamente para o processador (ver van Gompel & Pickering, 2007).

Mais recentemente, com o emprego de um paradigma conhecido como “paradigma do mundo visual”- em que se pode verificar, por meio de um equipamento de *eye-tracking* (ou recurso semelhante), o movimento do olhar de um indivíduo enquanto este está ouvindo uma sentença – um outro aspecto tem sido considerado ao se avaliar o processamento incremental, qual seja, o do papel do contexto e de processos antecipatórios.

Spivey, Tanenhaus, Eberhard, & Sedivy (2002), em experimento no qual manipularam contexto visual prévio no processamento de sentenças ambíguas do tipo "Put the apple on the towel in the box", verificaram que a interpretação inicial atribuída à sentença foi consistente com o contexto visual em que esta foi ouvida. Assim, no contexto de 1-referente (uma única maçã), a primeira análise foi de “on the towel” como locativo, modificador do verbo; no contexto de 2-referentes (duas maçãs – uma sobre uma toalha e outra sobre um guardanapo), a primeira análise foi de “on the towel” como modificador do nome. Esses resultados foram tomados como evidência de que na compreensão em tempo-real o processador leva em conta de forma imediata informação não linguística disponível no contexto.

Altmann & Kamide (1999), por sua vez, analisaram o movimento do olhar em situação de inspeção de uma cena visual semi-realista, com vistas a investigar se uma dada informação poderia ser antecipada com base em informação linguística. No experimento realizado, os participantes ouviam uma frase como “The boy will move the cake” or “The boy will eat the cake” enquanto viam uma cena em que havia um menino, um bolo e vários objetos distratores (uma bola, um carrinho, um trenzinho). Os resultados foram bastante robustos – o *onset* do movimento sacádico para o objeto alvo (bolo) ocorreu significativamente mais tarde na condição com o verbo “move” do que na condição com o verbo “eat” e as sacadas para o alvo foram disparadas depois do *onset* da palavra “cake” na condição com o verbo “move”, mas antes do *onset* da palavra “cake” na condição com “eat”. Com base nesses dados, conclui-se que informação expressa no verbo seria usada para restringir de algum modo o contexto de busca do referente subsequente, no caso o objeto pós-verbal, o que sinaliza a atuação de processos antecipatórios.

Em outro estudo (Altmann e Kamide, 2007), os autores confrontaram sentenças como “The man has drunk all of the wine” e “The man will drink all of the beer” apresentadas juntamente com figuras de dois tipos: figura contendo um copo vazio (congruente com verbo no passado) e um copo cheio (congruente com verbo no futuro). Tanto na condição de futuro, quanto na de passado, os participantes tornaram-se propensos a olhar mais para os alvos “congruentes” no início da expressão referencial. Esses resultados, juntamente com os do trabalho anterior, apontam para o papel de processos antecipatórios no processamento da sentença e para o caráter altamente incremental da compreensão da linguagem.

Dois trabalhos recentemente conduzidos no LAPAL, com falantes de português (no primeiro caso, crianças; no segundo, adultos) também podem ser tomados como evidência da relevância do contexto e de processos antecipatórios.

Ribeiro (2012) observou que o tipo de material visual utilizado e a manipulação do contexto visual em tarefa de seleção de imagem podem facilitar a identificação da

referência e reduzir a carga de processamento linguístico de interrogativas e relativas de objeto (ex.: *Que gato que o cachorro molhou?/ O gato que o cachorro molhou tremeu de frio*). Esse trabalho corrobora pesquisa anterior realizada no LAPAL com adultos, em que se verificou que informação visual prévia relativa a atores e ações, passíveis de serem instanciados em uma relativa de objeto, tem papel relevante no processamento desse tipo de estrutura, podendo vir a contribuir para antecipação de informação linguística durante o processamento (Forster et al. 2010).

O estudo de Foster et al. focalizou o mapeamento *on-line* entre um DP-sujeito contendo uma oração relativa (*a garota que o bombeiro levantou vai comprar um brinquedo*) e um referente apresentado visualmente. Considerando trabalhos como o de Altmann & Kamide (1999; 2007), a pesquisa teve por objetivo verificar em que medida o referente de um DP-sujeito complexo, contendo uma relativa de objeto, poderia ser identificado antes que sua posição de origem (o *gap* da relativa) fosse detectada. No estudo de Forster et. al (2010), foi empregada a técnica de rastreamento do olhar, que permite registrar movimentos do olhar durante a visualização de cena/leitura de um texto. No experimento 2 desse estudo, os participantes viam uma sequência de figuras – uma primeira figura, com duas meninas, posicionadas ao lado da outra (menina ruiva e menina loira); uma segunda figura apresentada antes da relativa e uma terceira figura, idêntica à primeira. Na figura intermediária (antes da relativa), apareciam representadas simultaneamente, uma ao lado da outra, duas ações envolvendo, cada qual, um personagem agente e um personagem paciente. Foram manipuladas as seguintes informações visuais nessa figura intermediária: os personagens agentes, os personagens pacientes e as ações, dando origem a quatro condições distintas, quais sejam: condição A: figura em que as duas personagens pacientes eram diferentes mas sofriam a mesma ação executada pelo mesmo tipo de agente (um **menino** sendo levantado por um bombeiro e uma **menina** sendo levantada por outro bombeiro); condição B: figura em que a diferença estava apenas no agente (um **bombeiro** levantando uma menina e um **palhaço** levantando outra menina); condição C: figura em que a diferença estava na ação (um bombeiro **levantando** uma menina e um bombeiro **puxando** uma menina); condição D: figura em que a diferença estava no agente e na ação (um **bombeiro levantando** uma menina e um **palhaço puxando** uma menina). As variáveis dependentes consideradas foram o número de primeiras fixações do olhar e a duração total de fixações na figura apresentada ao final do experimento. Foi verificado o número de fixações e o tempo de fixação em três segmentos diferentes da sentença: Segmento 1 (do núcleo da relativa ao determinante do sujeito da relativa), Segmento 2 (o nome sujeito da relativa) e Segmento 3 (do verbo da relativa em diante). Os dados foram submetidos a três ANOVAS, um para cada segmento analisado. Os resultados sugerem que **o DP complexo é mapeado em um referente assim que há informação distintiva disponível** e a despeito do preenchimento do *gap* da relativa, como mostram as fixações na condição (B) para o Segmento 2 da sentença. Para essa condição (B), cuja diferença crucial está no agente (um **bombeiro** levantando uma menina e um **palhaço** levantando outra menina), assim que o participante ouvia a palavra bombeiro na relativa (a menina que o bombeiro...), já direcionava o olhar para o personagem que havia sofrido a ação. Não precisava, portanto, aguardar a produção do verbo, pois a informação dada pela palavra “bombeiro” já lhe permitia mapear o DP.

2.2 Incrementalidade na produção

Enquanto não parece haver dúvida quanto à natureza incremental da compreensão, no caso da produção o cenário não é tão claro. Discute-se, por exemplo,

em que medida a formulação linguística do enunciado ocorre concomitantemente à conceptualização da mensagem, guiada por elementos individuais (proeminentes no discurso ou no contexto visual), ou tem início após a construção de uma estrutura conceptual de natureza proposicional, em que os elementos da mensagem se articulam em torno de um dado evento ou estado de coisas, conforme proposto originalmente por Wundt (cf. Blumenthal, 1970). Esse problema, que remete ao que Bock et al. (2004) se referem como hipótese do ponto inicial (*starting point hypothesis*), tem sido abordado mais recentemente a partir de dados de rastreamento ocular registrados durante tarefa de descrição de cenas representadas em figuras envolvendo, em geral, dois participantes. A lógica por trás dos experimentos é buscar verificar se escolhas linguísticas – relativas, por exemplo, ao tipo de estrutura linguística (exemplo: ativas ou passivas) são determinadas pela ordem de fixação em elementos isolados da cena (no caso, personagem agente ou paciente) ou por uma apreensão global da cena (*the gist of the scene*), que ocorreria nos 300ms de análise inicial de uma imagem. Os resultados obtidos até o momento são conflitantes (cf.. Griffin & Bock, 2000 vs. Gleitman et al. 2007).

No que tange especificamente ao processo de codificação gramatical (durante o qual ocorre a estruturação sintática da sentença) e o estágio seguinte de codificação fonológica (assumindo-se o modelo de Levelt, 1989; Bock & Levelt, 1994), também há posições divergentes em relação ao grau de incrementalidade do processamento.

Levelt (1989) incorpora em seu modelo de produção o conceito de incrementalidade, conforme proposto por Kempen & Hoenkamp (1987) no âmbito de uma teoria da produção sintática de sentenças que se pretendia fosse dotada de plausibilidade linguística e estrutural – a Gramática Procedural Incremental (IPG – *Incremental Procedural Grammar*).

Na IPG, a formulação de sentenças se dá paulatinamente (*piecemeal*), isto é, a sentença é vista como compreendendo uma série de unidades e, enquanto se está articulando uma dessas unidades, a unidade seguinte já está sendo planejada. Tão logo um fragmento de conteúdo conceptual esteja disponível, é transmitido ao formulador sintático que procura traduzi-lo em um fragmento de sentença, o qual é então articulado. Enquanto isso, o processamento de outros fragmentos estaria sendo realizado.

Essa idéia de incrementalidade foi incorporada por Levelt (1989) a um modelo psicolinguístico de produção da linguagem, de natureza serial. Em termos da estruturação sintática da sentença, Levelt propõe que, assim que um elemento nominal inicial é ativado, ele projeta não apenas nós do tipo N e NP, mas também o nó da própria sentença. Logo, antes mesmo de informação gramatical relacionada ao lema do verbo estar disponível, um nó IP já seria projetado. Mais do que isso, consoante com uma visão altamente incremental, o NP (correspondente ao sujeito) seria enviado para a codificação fonológica antes mesmo de o verbo ser parte da representação sintática da sentença.

Ferreira (1999), com base em uma proposta de formulação sintática ancorada no modelo da *Tree-Adjoining Grammar* - TAG (Joshi, 1985; Joshi, Levy & Takahashi, 1975; Kroch & Joshi, 1985), discorda de Levelt no que tange à estrutura linguística que seria construída antes de ter início a codificação fonológica. O verbo, na visão de Ferreira, tem um papel sintático fundamental e um NP só pode ser enviado para a codificação fonológica após uma árvore elementar nucleada pelo verbo ter sido ativada.¹ A ideia básica dessa proposta é que a codificação de um conceito como sujeito depende fundamentalmente das propriedades do verbo. Ferreira reporta resultados experimentais

¹ Para uma visão detalhada do formalismo e da implementação da proposta das TAGs, ver Ferreira (1999)

que apontam nessa direção (cf. Lindsley, 1975; Kempen & Huijbers, 1983; Ferreira, 1994; Meyer, 1996).

No estudo do processamento da concordância de número entre sujeito e verbo que conduzimos no doutorado (Rodrigues, 2006), adotamos uma visão moderadamente incremental ao caracterizarmos o fluxo de informação entre os níveis da codificação gramatical e da codificação (morfo)fonológica. Propusemos que a concordância entre sujeito e o “verbo” (no caso, um núcleo funcional com traço não-interpretável de número) seria computada antes de o DP sujeito ser enviado para a codificação morfofonológica com vistas à sua posterior articulação. Na tese, interessava-nos investigar que fatores poderiam induzir os chamados erros de atração – do tipo “O estudo dos erros de concordância foram conduzidos...” e em que momento tais erros ocorreriam. Assumimos que os erros não seriam propriamente erros de produção; na verdade, assumiu-se que o formulador sintático seria bastante robusto e funcionaria de forma automática na produção, não sendo sujeito a erros ele próprio. Como será visto a seguir, os erros ocorreriam em uma etapa posterior à computação sintática da concordância na produção.

Como o falante é também ouvinte de seus enunciados, um mecanismo de *parser* monitorador 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. 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 fosse disponibilizado pelo sistema de produção, dentro de uma estrutura prosódica específica, seria analisado pelo *parser*, que, após o acesso aos itens lexicais, começaria 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.

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.²

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

² 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.

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.

Logo, nessa proposta, o erro **não** é resultante de uma falha no estabelecimento da concordância em si; o custo de processamento relacionado à computação da concordância é função de custos de memória, relativos à acessibilidade e manutenção 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 marcação³, 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 morfológica⁴. 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.

3.3 O conceito de fase

Como vimos na seção 2, no Programa Minimalista assume-se uma visão derivacional da computação sintática, sendo os objetos sintáticos construídos passo a passo, a partir da aplicação de operações sintáticas. Essa visão é altamente compatível com a ideia de incrementalidade, em especial quando conjugada ao conceito de fase, que corresponde a subpartes dos arranjos de itens lexicais, ponto de partida da derivação sintática.

De forma pouco técnica, poderíamos caracterizar as fases como pequenos ‘chunks’, que seriam manipulados pela memória de trabalho, ao longo de uma derivação sintática. Nesse sentido, o conceito de fase poderia ser tomado como equivalente a uma unidade de processamento e passaria a ter realidade psicológica em modelos psicolinguísticos (cf. Corrêa & Augusto, 2006, 2007; Lin, 2006).

Como aponta Smith (2004), o conceito de fase é compatível com a ideia de minimização de custos no âmbito do processamento. Pode-se pensar que uma derivação em fases reduz a carga de processamento do sistema, o que é totalmente congruente tanto com a visão de Economia relevante para a computação sintática, no modelo formal de língua, como para a visão de Economia em um modelo de processamento, em que questões relativas à manutenção de informação na memória de trabalho nos processos de codificação e decodificação de enunciados linguísticos precisam necessariamente ser avaliadas.

Um ponto crítico a ser explorado é se o que se está considerando como correspondente a uma fase no modelo linguístico seria exatamente equivalente aos chunks com os quais o processador trabalharia e se haveria diferenças em relação ao parsing e à formulação. Nessa direção, pode-se discutir, por exemplo, se o DP poderia

³ Foi verificado que quando o núcleo do sujeito é singular, não-marcado, ele está mais sujeito a esvaecimento e, portanto, ocorrem mais erros de concordância nessa condição do que quando ele é plural, marcado. Também se observou que quanto mais distante o núcleo do sujeito está do verbo mais chances há de ocorrerem erros de concordância.

⁴ As condições em que o núcleo interveniente era plural induziram mais erros de concordância do que aquelas em que o núcleo interveniente era singular, não-marcado. Também foi verificado que núcleos intervenientes mais próximos hierarquicamente do núcleo do sujeito (“A tinta dos cartuchos da impressora”) tendem a induzir mais erros de concordância do que os mais distantes (“A tinta do cartucho das impressoras”), mesmo que estes estejam linearmente próximos ao verbo. Também houve efeito relacionado ao status argumental do PP modificador, com mais erros após núcleos intervenientes inseridos em PPs argumentos do que em PPs adjuntos, o que foi relacionado a questões de acessibilidade influenciada por fatores estruturais.

vir a constituir uma fase, se a produção, tida como moderadamente incremental, operaria com fases de mesmo tipo que a compreensão, se informação semântica codificada em uma fase, ao ser enviada para a forma lógica, poderia ser processada em paralelo à computação sintática em curso, etc.

Estes são pontos a serem explorados em estudos futuros e que merecem ser vistos de forma articulada a discussões relativas à direcionalidade da derivação sintática, que na proposta do Programa Minimalista é bottom-up, concepção que, segundo Corrêa e Augusto (2006 e 2007), se apresenta como pouco compatível com uma concepção incremental de processamento, da esquerda para a direita. Nosso objetivo neste trabalho foi o de apontar o que nos parece ser um caminho interessante de pesquisa no estudo da relação processador/gramática.

4- Referências bibliográficas

ALTMANN, G. T. M.; KAMIDE, Y. Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition*, v. 73, p. 247–264, 1999.

ALTMANN, G. T. M.; KAMIDE, Y. The real-time mediation of visual attention by language and world knowledge: linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, v. 57, p. 502–518, 2007.

ALTMANN, G.; STEEDMAN, M. Interaction with context during human sentence processing. *Cognition*, v. 30, p. 191-238, 1988.

BEVER, T. G. The cognitive basis for linguistic structure. In: HAYES, J. R. (Ed.) *Cognition and the development of language*. New York: Wiley, 1970.

BLUMENTHAL, A. L. *Language and psychology: historical aspects of psycholinguistics*. New York: Wiley, 1970.

BOCK, J. K.; LEVELT, W. J. M. Language production: grammatical encoding. In: GERNSBACHER, M. A. (Ed.) *Handbook of Psycholinguistics*. San Diego, CA: Academic Press, 1994, p. 945-984.

BOCK, J. K.; IRWIN, D. E.; DAVIDSON, D. J. J. Putting first things first. In: FERREIRA, F.; HENDERSON, M. (Eds.) *The integration of language, vision and action: eye movements and the visual world*. New York: Psychology Press, p. 249-278, 2004.

CHOMSKY, N. *The minimalist program*. Cambridge, MA: MIT Press, 1995.

CHOMSKY, N. Derivation by Phase. *MIT Occasional Papers in Linguistics* 18, 1999. [Posteriormente publicado In: Kenstowiz, M. (Ed.) *Ken Hale: A Life in Language*. Cambridge, MA: MIT Press, 2001. p. 89-155]

CHOMSKY, N. Minimalist inquiries: the framework. In: Martin, R., Michaels, D., Uriagereka, J. (Eds.). *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, MA: MIT Press, 2000. p. 89-155.

CHOMSKY, N. Beyond explanatory adequacy. In: Belletti, A. (Ed.) *Structures and Beyond: The Cartography of Syntactic Structures*, vol. III. Oxford: Oxford University Press, 2004.

- CHRISTIANSON, K.; FERREIRA, F. Planning in sentence production: evidence from a free word-order language (Odawa). *Cognition*, v.98, p. 105-135, 2005.
- CORRÊA, L. M. S. Relação processador linguístico-gramática em perspectiva: problema de unificação em contexto minimalista. DELTA. Documentação de Estudos em Linguística Teórica e Aplicada (PUC-SP. Impresso), v. 24, p. 231-282, 2008.
- CORRÊA, L. M. S.; AUGUSTO, M. R. A. Computação linguística no processamento on-line: em que medida uma derivação minimalista pode ser incorporada em modelos de processamento? Texto apresentado na sessão Inter-GTs da ANPOLL, 2006.
- CORRÊA, L. M. S.; AUGUSTO, M. R. A. Computação linguística no processamento on-line: soluções formais para a incorporação de uma derivação minimalista em modelos de processamento. *Cadernos de Estudos Linguísticos*, v. 49, 167-183, 2007.
- FERREIRA, F. Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language*, 33, 715-736, 1994.
- FERREIRA, F. Syntax in language production: an approach using tree-adjoining grammars. In: WHEELDON, L. (org.) *Aspects of language production*. East Sussex: Psychology Press, 1999.
- FORSTER, R.; CORRÊA, L. M. S.; ALMEIDA, M. J. A.; RODRIGUES, E. dos S. Integrating information: the incremental processing of restrictive object relative clauses in Brazilian Portuguese In: *First International Psycholinguistics Congress; 22nd Meeting of ANPOLL's Psycholinguistics Workgroup*, 2010, Rio de Janeiro.
- FRAZIER, L.; RAYNER, K. Making and correcting errors during sentence comprehension: eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, v.14, p.178-210, 1982.
- GIBSON, E. Linguistic complexity: locality of syntactic dependencies. *Cognition*, v. 68, p. 1-76, 1998.
- GIBSON, E. The dependency locality theory: a distance-based theory of linguistic complexity. In: MYASHITA, Y.; MARANTZ, A.; O'NEIL, W. (Eds.) *Image, language, brain*. Cambridge, MA: MIT Press, 2000.
- GLEITMAN, L. R.; JANUARY, D.; NAPPA, R.; TRUESWELL, J. C. On the give and take between event apprehension and utterance formulation. *Journal of Memory and Language*, v. 57, p. 544-569, 2007.
- GORDON, P.C.; HENDRICK, R.; JOHNSON, M. Memory interference during language processing. *Journal of experimental psychology: learning, memory and cognition*, v. 27, p.1411-1423, 2001.
- GORDON, P. C.; HENDRICK, R.; JOHNSON, M.; LEE, Y. Similarity based interference during language comprehension: evidence from eye tracking during reading. *Journal of experimental psychology: learning, memory and cognition*, v. 32, p. 1304-1321, 2006.
- GORDON, P.C.; HENDRICK, R.; JOHNSON, M. Effects of noun phrase type on sentence complexity. *Journal of memory and language*, v. 51, p. 97-114, 2004.

- GRIFFIN, Z. M.; BOCK, K. What the eyes say about speaking. *Psychological Science*, v. 11, p. 274-279, 2000.
- HÄUSSLER, J.; BADER, M. Agreement checking and number attraction in sentence comprehension: insights from germana relative clauses. *Travaux du cercle linguistique de prague* 7. 2009.
- HORNSTEIN, N.; NUNES, J.; GROHMANN, K. K. *Understanding Minimalism*. Cambridge: Cambridge University Press, 2005.
- JACKENDOFF, R. S. *Semantic interpretation in generative grammar*. Cambridge, Mass: MIT Press, 1972.
- JACKENDOFF, R. S. The status of thematic relations in linguistic theory. *Linguistic Inquiry*, v.18, p.369-412, 1987.
- JACKENDOFF, R. S. *Semantic structures*. Cambridge, Mass: MIT Press, 1990.
- JOSHI, A.K.; LEVY, L; TAKAHASHI, M. Tree Adjunct Grammars. *Journal of the Computer and System Sciencies*, v. 10, p. 136-163, 1975.
- JOSHI, A.K. How much context-sensitivity is required to provide reasonable structural descriptions: tree adjoining grammars. IN: DOWTY, D; KARTUNNEN, L; ZWICKY, A (Eds.). *Natural language parsing: psychological, computational, and theoretical perspectives*. Cambridge: Cambridge University Press, 1985.
- JOSHI, A.K.; KROCH, A. *The linguistic relevance of tree adjoining grammar*. Technical Report MS-CS-85-16, Dept. of Computer and Information Sciences. 1985.
- KEMPEN, G. Human grammatical coding: Shared structure formation resources for grammatical encoding and decoding. Trabalho apresentado em *The 17th Annual CUNY Conference on Human Sentence Processing*, University of Maryland, p. 66, 2004.
- KEMPEN, G.; HOENKAMP, E. An incremental procedural grammar for sentence formulation. *Cognitive Science*, v. 11, p. 201-258, 1987.
- KEMPEN, G.; HUIJBERS, P. The lexicalization process in sentence production and naming: indirect election of words. *Cognition*, v.14, p.185-209, 1983.
- KEMPEN, G.; OLSTHOORN, N.; SPRENGER, S. Grammatical workspace sharing during language production and language comprehension: evidence from grammatical multitasking. *Language and Cognitive Processes*, v. 27, p. 345-380, 2012.
- LIN, Charles Chien-Jer Grammar and and parsing: a typological investigation of relative-clause processing. PhD thesis, University of Arizona.2006.
- LEVELT, W. J. M. *Speaking: from intention to articulation*. Cambridge, Mass: MIT Press, 1989.
- LINDSLEY, J.R. Producing simple utterances: how far ahead do we plan? *Cognitive Psychology*, v.7, p. 1-19, 1975.

- MEYER, A.S. Lexical access in phrase and sentence production: results from picture-word interference experiments. *Journal of Memory & Language*, v.35, p. 477-496, 1996.
- NICOL, J.; FORSTER, K.; VERES, C. Subject-verb agreement processes in comprehension. *Journal of Memory and Language*, v.36, p. 569-587, 1997.
- RAYNER, K.; CARLSON, M.; FRAZIER, L.. The interaction of syntax and semantics during sentence processing: eye movements in the analysis of semantically biased sentences. *Journal of verbal learning and verbal behavior*, v. 22, p. 358–374, 1983.
- RIBEIRO, Vanessa Gouveia. O que torna orações relativas e interrogativas de objeto de difícil processamento para crianças? Um estudo experimental com foco no DEL (Déficit Específico da Linguagem). Dissertação de Mestrado. Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Letras, 2012. (Orientadora: Letícia Maria Sicuro Corrêa).
- RODRIGUES, E. dos S. *Processamento da concordância de número entre sujeito e verbo na produção de sentenças*. Tese de Doutorado, Departamento de Letras, Pontifícia Universidade Católica do Rio de Janeiro. Tese de doutorado. PUC-Rio, 2006.
- SMITH, N. *Ideas and Ideals*. 2 ed. Cambridge University Press, 2004.
- SPIVEY, M. J.; TANENHAUS, M.K.; EBERHARD, K.M.; SEVIDY, J.C. Eye movements and spoken language comprehension: effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology*, v. 45, p. 447-481, 2002.
- TRUESWELL, J.C.; TANENHAUS, M.K.; GARNSEY, S.M. Semantic influences on parsing: use of thematic role information in syntactic disambiguation. *Journal of memory and language*, v. 33, p. 285-318, 1994.
- van GOMPEL, R. P. G.; PICKERING, M. J. Syntactic parsing. In: GASKELL, Gareth. (Ed.) *The Oxford handbook of psycholinguistics*. Oxford, UK: Oxford University Press, 2007, p. 289–307.