

A Aquisição da Morfologia de Verbos Regulares no Português Brasileiro: Uma Abordagem da Linguística Cognitiva

The Acquisition of Regular Verb Morphology in Brazilian Portuguese: A Cognitive Linguistics Approach

André Luiz Elias de Souza^{*,a} & Cláudia Cardoso-Martins^b

^aThe University of Texas at Austin & ^bUniversidade Federal de Minas Gerais

Resumo

A pesquisa em Linguística Cognitiva no Brasil tem crescido de maneira considerável nos últimos anos. No entanto, ainda são poucos os estudos sobre a aquisição da linguagem que adotam o aporte teórico da Linguística Cognitiva. Com o objetivo de preencher essa lacuna, o presente estudo investigou a produtividade de uma criança pequena no que tange o uso das diferentes flexões dos verbos regulares em português, em duas ocasiões diferentes: quando a criança tinha 1 ano e 10 meses de idade e quando ela tinha 2 anos e 2 meses de idade. Os dados foram retirados do banco de dados do projeto CHILDES. Apesar de ter havido um desenvolvimento significativo entre as duas ocasiões no uso das flexões verbais, os resultados sugerem que, em consonância com o paradigma da Linguística Cognitiva, o conhecimento da morfossintaxe de verbos se desenvolve de forma gradual e é inicialmente pouco abstrato. As implicações desses resultados para uma teoria sobre a aquisição da linguagem são discutidas.

Palavras-chave: Linguística cognitiva; Morfologia de verbos; Aquisição da linguagem.

Abstract

In recent years, there has been an increasing body of research on Cognitive Linguistics in Brazil. However, few studies have investigated first language acquisition from a Cognitive Linguistics viewpoint. The present study is an attempt to fill in this gap. The productivity of a young Brazilian Portuguese-speaking boy with regard to the use of regular verbs was investigated at two different times: when he was 1 year and 10 months old, and when he was 2 years and 2 months old. The data were taken from the CHILDES database. Although there was a significant progress between Times 1 and 2, results suggest that the knowledge of the morphosyntax of verbs develops gradually, being initially non-schematic, which is consistent with the Cognitive Linguistics framework. The implications of these results for a theory of language acquisition are discussed.

Keywords: Cognitive linguistics; Verb morphology; Language acquisition.

A investigação acerca da aquisição da morfologia verbal tem despertado o interesse de diversos pesquisadores e, por consequência, tem assumido um papel de destaque em disciplinas como a Psicologia Cognitiva, a Linguística e a Psicolinguística. Uma questão recorrente diz respeito à natureza do conhecimento linguístico subjacente ao uso produtivo das diferentes flexões verbais no início da aquisição da linguagem (Aguado-Orea, 2004; Rubino & Pine, 1998; Wittek & Tomasello, 2002, 2005). Em particular, vários pesquisadores têm procurado avaliar até que ponto o conhecimento inicial da criança em relação às marcações morfossintáticas de tempo e aspecto é baseado em regras morfossintáticas abstratas ou em itens linguísticos específicos (Akthar & Tomasello, 1997;

Olguin & Tomasello, 1993). Poucos estudos, no entanto, investigaram essa questão em línguas que apresentam um sistema morfossintático verbal complexo e denso, como é o caso da maioria das línguas românicas. O presente estudo propõe uma análise da fala espontânea de uma criança brasileira no que diz respeito à sua produção de verbos regulares, com vistas a fornecer evidência em favor da hipótese de que o conhecimento morfossintático inicial é pouco abstrato e ligado a itens verbais específicos presentes no *input* materno.

O Conhecimento Linguístico Inicial: Dois Posicionamentos Teóricos

Nas últimas quatro décadas, dois pontos de vista teóricos distintos têm procurado explicar o desenvolvimento da habilidade da criança de usar a gramática da sua língua materna de forma efetiva e produtiva. De acordo com um desses pontos de vista – de caráter claramente

* Endereço para correspondência: The University of Texas at Austin, Department of Psychology, 108, E. Dean Keeton St. 78712, Austin, TX, USA. E-mail: andreluiz@mail.utexas.edu

gerativista – a aquisição morfossintática é guiada por um conhecimento lingüístico inato (e.g., Clahsen, 1999; Hoekstra & Hyams, 1998; Poeppel & Wexler, 1993; Wexler, 1998). Esse ponto de vista assume que as crianças devem ser altamente produtivas desde os primeiros estágios do processo de aquisição da linguagem.

O ponto de vista alternativo, baseado nos pressupostos da Lingüística Cognitiva (Croft & Cruse, 2004), não assume qualquer tipo de conhecimento lingüístico inato. Pelo contrário, acredita-se que a competência morfossintática ‘emerge’ do uso efetivo que o falante faz da língua em eventos comunicativos reais. Em outras palavras, o conhecimento morfossintático emerge da relação entre a intenção da criança de se comunicar e sua capacidade de discernir padrões recorrentes de estruturas morfossintáticas nos enunciados concretos e funcionalmente situados que escuta ao seu redor. O processo de aquisição do conhecimento morfossintático pela criança deve ser, pelo menos inicialmente, restrito a itens lexicais e construcionais específicos da língua a que está exposta. Isso quer dizer que as crianças inicialmente aprendem as construções morfológicamente marcadas de sua língua por imitação e só posteriormente, com base em competências cognitivas gerais, passam a utilizá-las de forma criativa e produtiva (Tomasello, 2001).

A Linguagem Inicial da Criança: A Visão Gerativista

Pesquisadores simpáticos à visão da Lingüística Gerativista (e.g., Pinker, 1984) têm argumentado que o *input* a que a criança está exposta é altamente truncado e fragmentado e, como tal, não apresenta informações suficientes para a aquisição da linguagem. Esse argumento – conhecido como ‘pobreza do estímulo’ – tem um papel central na postulação de uma *Gramática Universal* geneticamente pré-determinada (Mioto, Silva, & Lopes, 2000). Em outras palavras, segundo a lingüística gerativista, somente um conhecimento inato da linguagem poderia explicar como as crianças são capazes de adquirir a sua língua de forma tão rápida e homogênea diante de um *input* tão pobre.

De acordo com a Teoria de Princípios e Parâmetros (Chomsky, 1995), a Gramática Universal pode ser descrita em termos de uma série de princípios e parâmetros: os princípios correspondem ao que é invariável entre as línguas do mundo (por exemplo, é um princípio a noção de que as frases têm ‘sujeito’), enquanto os parâmetros correspondem a formulações mais detalhadas das propriedades gramaticais de cada língua específica (por exemplo, conforme a língua, o ‘sujeito’ pode ou não ser oculto). Como esse exemplo demonstra, os parâmetros são entendidos como assumindo um valor positivo ou negativo (Mioto et al., 2000). Uma vez que eles constituem os únicos elementos que diferem entre as línguas, tudo o que uma criança precisa fazer no processo de aquisição de linguagem é marcar o valor (positivo ou negativo) que corresponde à sua língua. Pesquisadores simpáticos à Teoria de Princípios e Parâmetros têm

procurado definir quais são os aspectos da gramática que precisam assumir valores paramétricos diferentes para que a variação existente entre as línguas naturais seja satisfatoriamente explicada (Roeper & Williams, 1987).

Segundo Hyams (1986), a idéia de marcação de parâmetros pode também ser útil para explicar a diferença entre a linguagem do adulto e a linguagem inicial da criança. Os erros apresentados por uma criança podem ser indícios dos valores paramétricos que são considerados padrões. Por exemplo, como observamos anteriormente, algumas línguas (como o português) permitem que o ‘sujeito’ seja opcionalmente ocultado (um fenômeno que sob a perspectiva da Teoria da Vinculação e Regência é conhecido como *pro-drop*). Línguas como o inglês e o francês não permitem tal opcionalidade (veja, por exemplo, o uso obrigatório da partícula expletiva *there* em *There seems to be true* [=Parece ser verdade]). No entanto, Brown (1973) observou que crianças falantes do inglês, em fase inicial de aquisição de linguagem, freqüentemente produziam frases sem sujeitos. Hyams (1986) interpretou esse fenômeno como evidência de que, para esse aspecto da gramática, o valor padrão é o valor *pro-drop*. Segundo ela, as crianças observadas por Brown apresentaram esse padrão, pois não haviam ainda marcado o valor paramétrico correto para esse aspecto da gramática do inglês.

O processo de aquisição de formas verbais finitas, ou seja, de verbos que apresentam marcação flexional de tempo, aspecto e concordância tem sido também estudado de acordo com a Teoria de Princípios e Parâmetros. De acordo com Wexler (1998), por exemplo, os parâmetros básicos relacionados ao uso dos verbos de uma língua são marcados muito cedo no processo de desenvolvimento da linguagem. Em particular, Wexler (1998) argumenta que as flexões mais funcionalmente centrais e básicas (as do tempo presente, por exemplo) são aprendidas muito rapidamente pelas crianças.

Wexler (1998) propõe que o uso dessas flexões deve ser produtivo desde os estágios iniciais do processo de desenvolvimento da linguagem. No caso da língua portuguesa, essa afirmação sugere que as crianças deveriam ser capazes de anexar todas as flexões já conhecidas de uma determinada conjugação verbal a todos os radicais da mesma conjugação que elas já conhecem. Dessa maneira, uma criança brasileira que já tenha aprendido os radicais *am-*, *guard-* e *sopr-* e as flexões *-o*, *-a*, *-amos* e *-ei*, deveria ser capaz de produzir todas as combinações possíveis (*am-o*, *am-a*, *am-amos*, *am-ei*, *guard-o*, *guard-a*, *guard-amos*, *guard-ei*, *sopr-o*, *sopr-a*, *sopr-amos* e *sopr-ei*).

A Linguagem Inicial da Criança: A Visão da Lingüística Cognitiva

Uma prática comum entre os pesquisadores da escola gerativista interessados no desenvolvimento inicial da linguagem é observar a fala da criança e assumir que ela

baseia-se nas mesmas regras e categorias gramaticais abstratas que geram a linguagem dos adultos. Em outras palavras, a linguagem da criança é analisada em termos das categorias que fazem parte da competência linguística de falantes adultos (Hoekstra & Hyams, 1998). Por outro lado, Tomasello (2007) tem argumentado que um procedimento mais adequado para a análise da linguagem infantil consiste em observar, de forma sistemática, os usos que a criança faz da linguagem e, a partir dessa observação, levantar hipóteses sobre a estrutura subjacente que permite tal produção. Tal estrutura subjacente não precisa ser, necessariamente, a mesma que sanciona a produção linguística de um falante adulto.

Uma série de estudos que utilizaram esse procedimento mais sistemático de observação dos usos (e, principalmente, dos não-usos) da linguagem pela criança tem sugerido que, inicialmente, a linguagem organiza-se em torno de itens lexicais e construcionais específicos. Em relação à morfologia de verbos, esses estudos têm mostrado a existência de um longo período no desenvolvimento da linguagem da criança em que ela não utiliza os morfemas verbais em contextos em que seu uso é obrigatório (Brown, 1973). O uso que as crianças fazem dos morfemas verbais é, inicialmente, restrito a apenas alguns verbos específicos (Pine, Lieven, & Rowland, 1998).

Tomasello (1992) realizou uma análise sistemática da linguagem da sua filha dos 15 aos 24 meses de idade. Ele observou que, apesar de grande parte da sua linguagem ser gramatical do ponto de vista da linguagem do adulto, era também significativamente restrita a itens linguísticos específicos. Para Tomasello (1992), isso ficou bastante evidente no uso que sua filha fez dos verbos. Durante exatamente o mesmo período de desenvolvimento, alguns verbos foram utilizados em apenas um tipo de enquadramento (*frame*) sentencial enquanto outros verbos, semântica e argumentativamente semelhantes, foram utilizados em enquadramentos sentenciais distintos e mais complexos. Aos 23 meses de idade, por exemplo, o verbo *to cut* ‘cortar’ foi usado em apenas um tipo de enquadramento sentencial (*cut X*) enquanto o verbo *to draw* ‘desenhar’ foi utilizado em uma variedade de construções (*draw X*; *draw X on X*; *draw X for X*; *X draw on X*). Tomasello (1992) observou, ainda, que (a) quase metade dos 162 verbos e predicados produzidos pela sua filha foram utilizados em apenas um tipo de configuração argumentativa (por exemplo, *mommy break*; *dad break*; *baby break*); (b) dois terços de todos os verbos produzidos pela sua filha não apresentavam nenhuma marcação morfológica de tempo e aspecto; (c) um sexto dos verbos foram utilizados apenas com a marcação de passado (*-ed*) e um sexto apenas com a marcação de progressivo (*-ing*). Somente 2% dos verbos produzidos pela criança foram utilizados com ambas essas marcações. Com base nesses resultados, Tomasello (1992) concluiu que, inicialmente, o uso de verbos pela

sua filha parecia estar organizado não em termos de uma categoria abstrata de verbos, mas sim em torno de verbos e itens lexicais específicos.

Pine e Lieven (1997) encontraram resultados semelhantes em uma amostra de 12 crianças entre 2 e 3 anos de idade no que tange o uso dos determinantes da língua inglesa. O uso que as crianças faziam dos artigos *a* [= um/uma] e *the* [= o/os/a/as], por exemplo, parecia estar organizado em torno de itens nominais específicos. Em outras palavras, os substantivos usados com o artigo definido diferiam daqueles utilizados com o artigo indefinido, sugerindo que o uso dessas partículas se organizava em torno de itens lexicais concretos.

O mesmo padrão de aquisição tem sido observado em línguas diferentes da língua inglesa. Em um estudo com três crianças falantes do italiano, Pizzuto e Caselli (1992) observaram que, apesar de haver seis terminações distintas para cada verbo da língua italiana (primeira, segunda e terceira pessoas do singular e do plural), 47% de todos os verbos produzidos pelas crianças que participaram do seu estudo foram utilizados com uma única terminação e 40% foram utilizados com duas ou três terminações distintas. Entre os verbos utilizados com mais de quatro terminações, praticamente a metade consistia de verbos irregulares e muito frequentes e, dessa maneira, é possível que tenham sido aprendidos por imitação.

Rubino e Pine (1998) encontraram resultados semelhantes em um estudo sobre a aquisição da concordância verbal no português brasileiro. Esses autores coletaram 300 minutos da fala espontânea de uma criança de 3 anos de idade. A imensa maioria (97%) das concordâncias verbais foi produzida corretamente. No entanto, apenas 50 das 1.464 concordâncias produzidas pela criança (i.e., menos de 4%) envolviam verbos no plural. Além disso, a proporção de erros em contextos de concordância verbal no plural foi significativamente mais alta do que a proporção de erros em contextos envolvendo concordância no singular. Enquanto a proporção de erros foi de apenas 2,1% nesses contextos, a proporção de erros em contextos de concordância no plural foi de 28% (14 das 50 ocorrências). Uma vez que, na linguagem dos adultos, as formas verbais no singular ocorrem com uma frequência alta em comparação às formas plurais, Rubino e Pine (1998) concluíram que a baixa taxa de erro na fala da criança não deve ser considerada indício de conhecimento abstrato acerca da morfossintaxe de verbos.

Mais recentemente, Aguado-Orea (2004) investigou a produtividade morfossintática de verbos de duas crianças (um menino e uma menina) falantes do espanhol. Os dados foram coletados longitudinalmente. No início da coleta, o menino tinha 1 ano e 10 meses e a menina 2 anos e 2 meses. Ao final da coleta, o menino tinha 2 anos e 5 meses e a menina 2 anos e 7 meses. Com um corpus relativamente denso (correspondente a mais de 50 horas para cada participante), Aguado-Orea (2004) observou que as crianças foram significativamente menos produ-

tivas do que seus pais na produção de diferentes flexões verbais. Isso ocorreu mesmo após o controle do efeito de diferenças entre os pais e as crianças no que diz respeito ao número de diferentes verbos e flexões produzidos e no que diz respeito ao número total de verbos produzidos.

O presente estudo é uma tentativa de replicar, no português do Brasil, os resultados obtidos por Aguado-Orea (2004) em espanhol. Em comparação à língua inglesa, as línguas românicas possuem uma morfologia verbal muito rica. Por exemplo, enquanto o verbo inglês *to sew* pode ser produzido de cinco formas diferentes (*sew*, *sews*, *sewing*, *sewed*, *sewn*), o mesmo verbo em português, “costurar”, admite mais de 30 flexões distintas (*costuro*, *costura*, *costuras*, *costuramos*, *costuravam*, *costuráveis*, *costurará*, *costurarão*, etc.). Em princípio, essa característica deveria facilitar a detecção (ou não) de produtividade. Além disso, ao contrário do que ocorre em inglês, os erros apresentados pelas crianças falantes do português no início da aquisição da linguagem não podem ser facilmente interpretados em termos do chamado estágio do *Infinitivo Opcional* (Wexler, 1998).

Esse estágio foi formulado por teóricos gerativistas para explicar a frequência elevada com que formas verbais não-finitas ocorrem na fala inicial de crianças falantes do inglês, mesmo em contextos que exigem uma forma verbal finita (e.g., a oração **I sew yesterday* ao invés da oração *I sewED yesterday*). Uma vez que o uso de *Infinitivos Opcionais* não é considerado um erro morfológico (ver Deen & Hyams, 2006; Hoekstra & Hyams, 1998; e Wexler, 1998, para uma explicação desse argumento), a frequência relativamente alta de formas verbais não-finitas em contextos que pressupõem formas verbais finitas não é considerada evidência da inexistência de um mecanismo morfossintático abstrato. Em português, por outro lado, os radicais verbais não podem ocorrer como *morfemas livres*, ou seja, só podem ser utilizados com uma das várias flexões verbais possíveis. Até mesmo as formas não-finitas – incluindo o infinitivo – apresentam um morfema específico para sua marcação (e.g., *am-AR*, *am-ANDO*, *am-ADO*). Do ponto de vista gerativista, portanto, não parece haver uma razão óbvia que explique a pouca produtividade de uma criança falante do português brasileiro no que tange o uso de flexões verbais. É possível, portanto, que uma investigação acerca da produtividade verbal no início da aquisição do português possa contribuir para esclarecer a natureza do conhecimento inicial da criança a respeito da morfologia verbal.

A maioria dos estudos sobre aquisição da morfologia parte do pressuposto de que os falantes adultos fazem uso exaustivo das possibilidades que o sistema morfossintático da sua língua oferece. Por outro lado, Cameron-Faulkner, Lieven e Tomasello (2003) argumentam que, sob uma análise mais sistemática, a linguagem do adulto mostra-se mais lexicalmente específica do que se tende a

acreditar (ver Aguado-Orea, 2004 e Pizzuto & Caselli, 1992, para um argumento semelhante no tocante ao uso das flexões verbais). Assim, para afirmar que a criança é menos produtiva do que o adulto no que diz respeito à produção de diferentes flexões verbais é necessário que haja uma diferença significativa entre a produtividade real (observada) do adulto e a produtividade da criança. Aguado-Orea (2004) aponta ainda que a diferença entre a produtividade da criança e do adulto pode, na realidade, ser influenciada por diferenças na frequência de ocorrência e de tipo de verbos produzidos em uma dada amostra. A frequência de ocorrência (do inglês *token frequency*) é definida como a frequência com que unidades lingüísticas específicas (e.g., verbos, substantivos ou até mesmo um item lexical específico) ocorrem em um determinado *corpus*. A frequência de tipo (do inglês *type frequency*), por outro lado, diz respeito ao número de instâncias diferentes de uma determinada unidade lingüística (e.g., o número de verbos ou substantivos diferentes) em uma dada amostra. Como Aguado-Orea (2004) observa, a probabilidade de se encontrar usos contrastivos de flexões verbais está diretamente relacionada à frequência de tipo e de ocorrência na amostra analisada. Uma vez que o adulto supostamente produz mais tipos e mais ocorrências de verbos do que a criança, é bem provável que a fala da criança pareça, à primeira vista, menos produtiva do que a fala do adulto.

A Produtividade Morfológica e o Papel das Frequências de Tipo e de Ocorrência

Ao contrário dos modelos de aquisição da linguagem de base gerativista, os modelos baseados no uso (*usage-based*) não postulam a existência de regras abstratas inatas. Segundo a Linguística Cognitiva, por exemplo, o que existe são esquemas mentais abstraídos das instâncias reais de uso da linguagem. São esses esquemas os responsáveis pela sanção de novas produções, ou seja, a base do conhecimento lingüístico. Conseqüentemente, a noção de esquema está intimamente ligada à noção de produtividade e à noção de uso criativo da linguagem.

Um esquema lingüístico – assim como qualquer outro esquema – é uma abstração. Nesse caso é uma abstração das similaridades existentes entre diversas estruturas lingüísticas (Langacker, 1987). Por exemplo, as formas *amar*, *cantar*, *andar* e *guardar*, por apresentarem uma mesma terminação (*-ar*), possibilitam a formação de um esquema lingüístico do tipo [X-AR]. Esse esquema, por sua vez, sanciona novas produções tais como *deletar* e *coisar*. Não obstante, os esquemas variam quanto ao seu grau de produtividade ou, nos termos da Linguística Cognitiva, quanto ao seu grau de entrincheiramento. Um esquema se torna entrincheirado (ou produtivo) na medida em que cresce o número de instâncias que o elabora. Contrariamente, um esquema baseado em poucas instâncias tende a ser fraco ou pouco-produtivo (Taylor, 2002).

As palavras *casas*, *mesas*, *cadeiras*, *janelas* e *portas*, por exemplo, são elaborações (exemplos) do esquema da formação do plural regular no português do Brasil (ver Figura 1). Dizemos que esse é um esquema entrincheirado (forte) e, conseqüentemente, altamente produtivo, uma vez que existem muitas instâncias que o elaboram. Esse esquema é tão entrincheirado que é considerado o esquema prototípico de formação do plural pelos falantes do português (ver Lakoff, 1987, para uma discussão acerca

de prototypicalidade em categorias lingüísticas). Em contrapartida, as palavras *cães*, *pães* e *capitães* são instâncias do esquema de formação do plural de algumas palavras terminadas em *-ão*. Uma vez que são poucas instâncias que elaboram esse esquema (e também pelo fato de que existem outras formas de plural de palavras com essa mesma terminação, por exemplo, *cidadãos*, *mãos*, etc.), ele é considerado um esquema pouco entrincheirado e conseqüentemente, pouco produtivo.

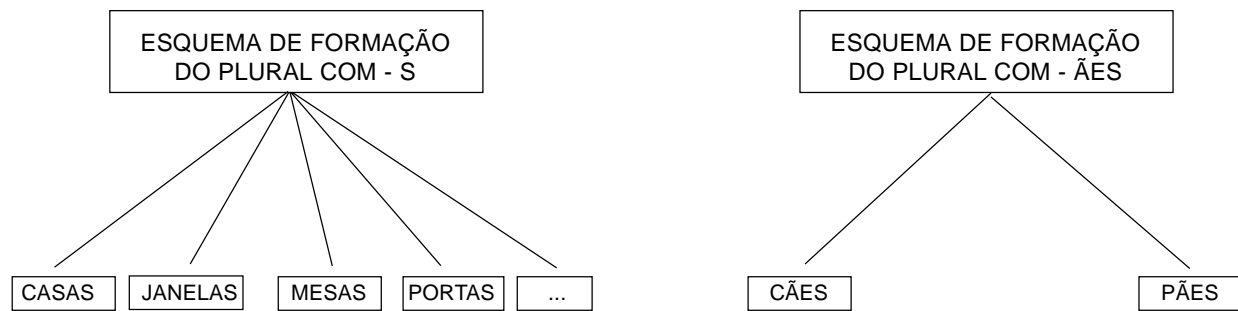


Figura 1. Relação esquema-instância na formação do plural em português

Intimamente coadunadas à noção de entrincheiramento de um esquema, estão as noções de freqüência de tipo e freqüência de ocorrência (Barlow & Kemmer, 2002; Bybee, 1995). Em outras palavras, podemos dizer que o entrincheiramento de uma estrutura lingüística (incluindo os esquemas) é o resultado da freqüência com que esse esquema ou estrutura é utilizado em eventos comunicativos (Croft & Cruse, 2004). Segundo Taylor (2002), por exemplo, a forma *never* ocorre muitas vezes em *corpora* de linguagem oral em língua inglesa. Dizemos, então, que essa palavra apresenta uma alta freqüência de ocorrência na língua inglesa. Já o esquema de formação do plural regular em português é instanciado por diversas palavras distintas, de forma que a freqüência desse 'tipo' de plural (o regular) é alta. O grau de produtividade de um esquema está, segundo Bybee (1995), diretamente ligado à sua freqüência de tipo. Em outras palavras, quanto mais instâncias um esquema sancionar, maior será sua freqüência de tipo e maior será sua produtividade. Por outro lado, uma estrutura lingüística com alta freqüência de ocorrência (como no caso da palavra inglesa *never*), tende a ficar entrincheirada como unidade lingüística específica e, conseqüentemente, enfraquecer o esquema que a sanciona.

Objetivos

Com base nos resultados dos estudos descritos anteriormente e das colocações teóricas advindas da Linguística Cognitiva, o presente estudo busca evidências de que o conhecimento lingüístico inicial dos seres humanos é baseado em itens lexicais e construcionais específicos e não em torno de regras esquemáticas e abstratas. O estu-

do busca também evidência de que a formação de esquemas morfossintáticos está ligada à freqüência com que itens lingüísticos ocorrem no *input* dirigido à criança. Vale a pena ressaltar que a escolha da Linguística Cognitiva como sustentáculo teórico-metodológico para a presente pesquisa não pretende desmerecer o importante trabalho e as pesquisas ancoradas dentro do escopo da Linguística Gerativa. Como bem aponta Oliva (2001), o processo de aquisição da linguagem deve ser entendido como um fenômeno altamente complexo e que não pode sujeitar-se a reducionismos teóricos nem posicionamentos extremistas. A descrição ora proposta pretende tão-somente suscitar reflexões sobre o tema e contribuir para a ampliação do escopo explicativo de modelos teóricos distintos, de forma a aperfeiçoar o entendimento geral do processo de aquisição e desenvolvimento da linguagem. A seguir, descrevemos os procedimentos utilizados na coleta e análise dos dados.

Método

Os Dados da Pesquisa

Os dados para essa pesquisa foram retirados do *corpus* de Florianópolis (Scliar-Cabral & MacWhinney, 2005), disponível para o uso da comunidade científica no banco de dados do Projeto CHILDES – *Child Language Data Exchange System* (MacWhinney & Snow, 1990). O *corpus* foi incorporado ao banco de dados do CHILDES em 1993 e consiste de 5.530 falas de uma criança falante nativo do português brasileiro. O *corpus* também inclui as falas maternas dirigidas à criança. Os dados foram transcritos de acordo com o formato CHAT (MacWhinney,

2000), que permite algumas análises lingüísticas automáticas a partir do software CLAN (*Computerized Language Analysis*).

Os dados foram coletados em três ocasiões distintas. Os dados da primeira ocasião consistem de 1.320 falas da criança e foram coletados quando a criança tinha 1 ano, 8 meses e 21 dias. Os dados da segunda ocasião equivalem a 2.245 falas da criança e foram coletados quando ela tinha 1 ano, 10 meses e 20 dias. Os dados da última ocasião foram coletados quando a criança tinha 2 anos, 2 meses e 8 dias e correspondem a 1.966 falas da criança. Como a ocorrência de verbos na primeira ocasião foi baixa, optou-se por agrupar os dados das duas primeiras ocasiões e considerá-los conjuntamente. Assim, daqui para frente, a primeira ocasião refere-se às falas correspondentes às duas primeiras coletas (idade média da criança = 1 ano e 10 meses) e a segunda ocasião refere-se às falas correspondentes à terceira coleta.

Análises dos Dados

Apesar de transcritos no formato CHAT, os dados do *corpus* de Florianópolis ainda não apresentam a linha %mor (Hausser, 1990), a qual permite a indexação das falas de acordo com sua composição morfossintática. Assim, as contagens e análises dos verbos foram feitas manualmente. A partir do comando FREQ, acessamos a frequência de todas as palavras enunciadas pela criança e pela mãe, separadamente para as duas ocasiões. Com essa lista de frequência, criamos um arquivo do tipo CUT com palavras que correspondiam a verbos (arquivos do tipo CUT, são arquivos, criados no programa CLAN e que tem extensão .cut). Esses arquivos funcionam como fonte de onde o programa busca as palavras de interesse para um determinado estudo. Por exemplo, suponhamos que um pesquisador esteja interessado em acessar o contexto de ocorrência de todos os pronomes possessivos da língua portuguesa. O arquivo de extensão .cut iria então conter uma lista com todos os pronomes possessivos do português. Através do comando KWAL e do arquivo do tipo CUT, tivemos acesso ao contexto de ocorrência de cada verbo produzido pela criança (ver MacWhinney, 2000 para definições técnicas acerca dos comandos utilizados pelo programa CLAN).

Com os arquivos em mãos, procedemos à análise de todos os contextos de ocorrência dos verbos presentes no *corpus*. Apenas os verbos regulares produzidos pela criança foram incluídos nas análises. Não foram incluídos nas análises os verbos produzidos pela criança imediatamente após a produção dos mesmos pela mãe. Foram também excluídos os verbos produzidos em canções e músicas infantis e a forma verbal 'olha'. Por ser muito frequente no português brasileiro, essa forma é por vezes usada apenas como uma ferramenta pragmática para direcionar a atenção do interlocutor. Além disso, como consequência da sua alta frequência de uso, sua realização fonética se reduziu ao ponto de muitas vezes ser pronunciada ape-

nas como [ó]. Uma vez que a presente pesquisa teve como objetivo analisar o uso de diferentes flexões, a não realização fonética da suposta flexão dessa forma verbal poderia desviar os resultados das análises. Ademais, como afirma Bybee (1995), a alta frequência de uso de uma estrutura lingüística acaba causando uma espécie de reformulação semântico-funcional (*semantic bleaching*), de maneira que essa estrutura passa a incorporar outros valores semânticos e pragmáticos. É possível, portanto, que a criança não reconheça a forma [ó] como um verbo, mas sim como uma ferramenta cujo objetivo é chamar a atenção do interlocutor para algum evento externo. Da fala da mãe, foram incluídos somente os verbos dirigidos à criança, visto que algumas vezes a mãe dirigia sua fala ao pai da criança ou à pesquisadora, ambos presentes durante a observação.

Após a análise dos contextos de ocorrência, e a exclusão dos casos citados anteriormente, listamos um a um (na ordem de ocorrência no *corpus*) os verbos que seriam incluídos nas análises. Esses verbos foram classificados em termos de sua conjugação (primeira, segunda e terceira), tipo (regular e irregular), pessoa/número (primeira, segunda e terceira pessoa do singular e plural) e tempo/modo. Finalmente, para cada um dos verbos regulares, foram listadas as flexões e seus respectivos radicais.

Análises de Produtividade

Um dos principais objetivos do presente estudo foi replicar os resultados de Aguado-Orea (2004) de que, inicialmente, a criança é menos produtiva do que a mãe no que diz respeito ao uso das diferentes flexões verbais. Para proceder a essa análise, calculamos a frequência de ocorrência (número de verbos no *corpus*) e a frequência de tipo (número de radicais verbais distintos no *corpus*), separadamente para a criança e para a mãe. Finalmente, calculamos o número de flexões anexadas a cada radical distinto (ou, coincidentemente, o número de formas verbais distintas). Para cada participante, a razão entre o número de formas verbais distintas e a frequência de tipo resultou no número médio de flexões por radical verbal. Por exemplo: suponha que a criança tenha produzido as seguintes formas: *amar, amou, partiu, partir, andei, comi, comi*. Para essa amostra, a frequência de ocorrência é igual a 7, a frequência de tipo é igual a 4 (*am-*, *part-*, *and-* e *com-*), o número de flexões para cada radical é: *am*-=2, *part*-=2, *and*-=1 e *com*-=1 e, conseqüentemente, o número de diferentes formas verbais é igual a 6. O número médio de flexões por radical é igual a 1,5.

Sempre que possível, tomamos o cuidado de controlar o efeito de diferenças entre a criança e a mãe no que diz respeito às frequências de ocorrência e de tipo. A razão para isso é que o número médio de flexões por radical está diretamente relacionado à frequência de ocorrência e de tipo de verbos (Aguado-Orea, 2004). Diferenças entre a mãe e a criança poderiam, portanto, resultar de dife-

renças na frequência de ocorrência e de tipo e não, necessariamente, de diferenças na produtividade. Para controlar diferenças na frequência de tipo, foram excluídos das análises os radicais verbais produzidos somente por um dos participantes, ou seja, verbos produzidos somente pela mãe ou somente pela criança. Após essa exclusão, procedemos ao controle do tamanho da amostra: excluímos aleatoriamente ocorrências de verbos da maior amostra até que seu número de ocorrência se igualasse ao número de ocorrência da menor amostra.

Resultados

Primeira Ocasão

Na primeira ocasião, a díade produziu um total de 3.167 verbos (1.483 verbos regulares e 1.684 verbos irregulares). Apesar do número relativamente elevado de verbos regulares, a frequência de tipo (número de radicais distintos) foi relativamente pequena, tanto para a mãe quanto para a criança (Tabela 1).

Tabela 1

Frequência de Ocorrência, Frequência de Tipo e Número Médio de Flexões por Radical em função da Ocasião, para a Mãe e para a Criança (Verbos Regulares)

| Ocasião | Mãe | Criança |
|--------------------------|------|---------|
| 1ª. Ocasião | | |
| Frequência de Ocorrência | 824 | 659 |
| Frequência de Tipo | 106 | 82 |
| Formas Verbais Distintas | 200 | 145 |
| Média de Flexões/Radical | 1,90 | 1,77 |
| 2ª. Ocasião | | |
| Frequência de Ocorrência | 254 | 502 |
| Frequência de Tipo | 70 | 96 |
| Formas Verbais Distintas | 109 | 181 |
| Média de Flexões/Radical | 1,56 | 1,88 |

Virtualmente 60% dos verbos regulares foram produzidos na forma finita (530 para a mãe e 359 para a criança). Desse percentual, quase 65% foi produzido no tempo presente (419 para a mãe e 153 para a criança). Ainda entre os verbos finitos produzidos pela díade, a maioria foi produzida na terceira pessoa (94,2% - 502 para a mãe e 335 para a criança) e no singular (99,5% - 528 para a mãe e 359 para a criança).

A flexão mais frequentemente produzida pela mãe foi a flexão *-a* (364 vezes), seguida da flexão *-ar* (204) e da flexão *-ou* (63). Essas também foram as flexões mais frequentemente produzidas pela criança, embora a ordem tenha sido diferente (*-ar*, 211 vezes, seguida da flexão *-ou*, 121 vezes, e da flexão *-a*, 96 vezes). Todas essas flexões são próprias de verbos da primeira conjugação. Logo, esses números não surpreendem, uma vez que a

frequência de ocorrência de verbos da primeira conjugação foi a mais alta para ambos os participantes (85% e 67% de todos os verbos para a mãe e para a criança, respectivamente). Quarenta e dois por cento dos radicais distintos foram produzidos com uma e apenas uma flexão verbal. Apenas 2,4% dos verbos foram produzidos com mais de quatro flexões distintas.

Como pode ser visto na Tabela 1, tanto a mãe quanto a criança produziram um número relativamente pequeno de flexões por verbo na primeira ocasião. A criança produziu, em média, 1,77 flexões por radical verbal ($DP = 0,97$) e a mãe, 1,90 ($DP = 1,06$). Essa diferença não foi estatisticamente significativa ($t(181) = 0,85$, n.s.). No entanto, após excluirmos da análise os verbos produzidos somente pela mãe ou somente pela criança, o número médio de flexões por radical foi significativamente menor para a criança ($M = 1,83$; $DP = 0,99$) do que para a mãe ($M = 2,15$; $DP = 1,15$), $t(125) = 1,71$, $p < 0,05$. Os mesmos resultados foram encontrados quando, além de reduzir a análise aos verbos produzidos pela mãe e pela criança, também controlamos o efeito de diferenças entre a mãe e a criança na frequência de ocorrência desses verbos, $t(123) = 1,71$, $p < 0,05$.

Conforme observamos anteriormente, os verbos pertencentes à segunda e à terceira conjugação foram muito pouco frequentes tanto na fala da mãe quanto na fala da criança. Em vista disso, realizamos uma análise de produtividade incluindo apenas os radicais pertencentes à primeira conjugação. Nessa análise, não controlamos o efeito do conhecimento de radicais distintos e nem o efeito do tamanho da amostra, uma vez que esse procedimento reduziria drasticamente o número de itens para a análise estatística. Nessa análise, a criança produziu um número significativamente menor de flexões por radical da primeira conjugação do que a mãe ($t(83) = 2,33$, $p < 0,02$).

Segunda Ocasião

Como pode ser visto na Tabela 1, a frequência de ocorrência e a frequência de tipo foram menores na segunda ocasião do que na primeira, tanto para a mãe quanto para a criança.

As formas finitas corresponderam a 60% dos verbos produzidos pela díade (167 para a mãe e 287 para a criança). Desse total, 82% dos verbos foram produzidos no presente (152 para a mãe e 223 para a criança) e 90% na terceira pessoa (159 para a mãe e 251 para a criança). Nenhum verbo foi produzido no plural.

Não houve diferença significativa entre a mãe e a criança em relação ao número médio de flexões por verbos na segunda ocasião (ver Tabela 1). Isso ocorreu tanto na análise que incluiu todos os verbos quanto naquela que incluiu apenas os verbos da primeira conjugação, os quais continuaram sendo muito mais frequentes do que os verbos da segunda e da terceira conjugação (85% e 74% de todos os verbos produzidos pela mãe e pela

criança, respectivamente). Além disso, ao contrário do que ocorreu nas análises da primeira ocasião, o controle do efeito de diferenças entre a mãe e a criança no que diz respeito às frequências de tipo e de ocorrência não alterou os resultados.

Apesar de ter havido um aumento no número médio de flexões por verbo produzidas pela criança entre a primeira e a segunda ocasião, o número de radicais produzidos com apenas uma flexão verbal continuou relativamente elevado. O mesmo também ocorreu para a mãe.

Análises de Desenvolvimento: Comparações Entre as Duas Ocasões

A Figura 2 ilustra o número médio de flexões verbais produzidas pela criança e pela mãe na primeira e na segunda ocasião. Esse cálculo incluiu apenas os verbos produzidos pela criança (ou pela mãe) tanto na primeira quanto na segunda ocasião. Além disso, tomamos o cuidado de igualar o tamanho das amostras nas duas ocasiões, para a criança e para a mãe. Enquanto o número médio de flexões por verbo não diferiu entre a primeira e a segunda ocasião para a mãe ($t(181) = 0,85$, n.s.), a criança foi significativamente mais produtiva na segunda do que na primeira ocasião ($t(105) = 2,13$, $p < 0,05$).

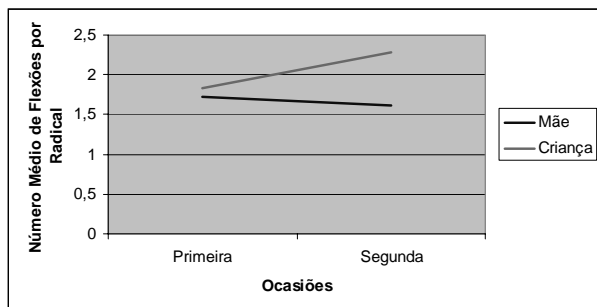


Figura 2. Número médio de flexões por radical em função do participante e da ocasião.

Discussão

O presente estudo pretendeu verificar em que medida uma criança falante do português do Brasil pode ser considerada produtiva no uso das diferentes flexões verbais no início da aquisição da linguagem. Comparamos o número médio de flexões por verbo produzidas pela criança e pela mãe, em duas ocasiões diferentes: quando a criança tinha 1 ano e 10 meses e quando ela tinha 2 anos e 2 meses. Na primeira ocasião, a criança produziu um número significativamente menor de flexões por verbo do que a mãe. De fato, a imensa maioria dos verbos finitos produzidos pela criança na primeira ocasião ocorreu com apenas uma única flexão. O mesmo ocorreu na segunda ocasião, embora não tenhamos encontrado uma diferença significativa entre a mãe e a criança nessa ocasião no que tange o número médio de flexões por verbo.

É pouco provável que a diferença entre a mãe e a criança na primeira ocasião possa ser explicada em função do número maior de verbos produzidos pela mãe. Com efeito, a diferença entre a mãe e a criança ficou ainda mais evidente após igualarmos os dois *corpora* em termos da frequência de ocorrência e da frequência de tipo. Além disso, conforme observamos, a mãe produziu um número menor de verbos na segunda do que na primeira ocasião. No entanto, ela não apresentou um número menor de flexões verbais na segunda ocasião do que na primeira.

A grande maioria dos estudos que tratam da produtividade de crianças acerca da gramática de sua língua parte do pressuposto de que um falante competente da língua faz uso extensivo das possibilidades que a língua oferece. Como os resultados de Cameron-Faulkner et al. (2003), nossos resultados sugerem que a linguagem espontânea dos adultos subestima as possibilidades produtivas que o sistema lingüístico oferece. É possível que o número pequeno de flexões por verbo encontrado para a mãe (e também para a criança) possa ser explicado em função do contexto de interação entre uma criança pequena e sua mãe, o qual, em geral, é restrito ao aqui-e- agora. De qualquer maneira, essas considerações não parecem suficientemente capazes de explicar os dados sobre a produção da criança. Por exemplo, mais da metade dos verbos que ela produziu com uma única flexão, foram produzidos pela mãe com a mesma flexão pelo menos uma vez. Esses resultados sugerem que, pelo menos inicialmente, a criança pode estar aprendendo os verbos por imitação. Com efeito, seis das 10 flexões distintas produzidas pela criança na primeira ocasião apareceram em verbos utilizados com apenas uma flexão.

Esse padrão fica evidente na segunda ocasião. Semelhantemente às observações de Tomasello (1992), os verbos *gostar* e *ganhar*, apesar de pertencerem à mesma conjugação, foram utilizados pela criança da forma como foram produzidos pela mãe, mesmo que essa forma resultasse em erro de concordância verbal. Por exemplo, o verbo *gostar* foi produzido pela mãe apenas nas formas *gosto* e *gosta*. Já o verbo *ganhar* foi produzido pela mãe na forma *ganhei*. Com efeito, a criança produziu a frase *eu ganhei bá(bala)*, o que sugere que ela conhecia a flexão *-ei*. No entanto, para o verbo *gostar*, ela produziu a construção **eu gosta* para se referir a um evento ocorrido no passado. Do ponto de vista gerativista, ela já deveria ser capaz de produzir a forma *gostei*, uma vez que ela já possuía o conhecimento da flexão *-ei*. No entanto, ela parece ser sensível ao efeito da frequência de ocorrência e de tipo na fala da mãe. O mesmo padrão foi observado para os verbos *amassou*, *gravou* e *achou* que foram produzidos dessa forma pela mãe e somente dessa forma pela criança, mesmo que isso resultasse em erros do tipo **eu amassou*, **eu gravou* e **eu achou*.

Todos os verbos produzidos somente pela criança foram produzidos com pelo menos uma das três flexões

mais frequentemente utilizadas pela mãe. Esse padrão de uso sugere que a criança é sensível ao efeito da frequência de tipo desde o início da aprendizagem da linguagem. Como é sugerido por Bybee (1995), a alta frequência de tipos verbais que envolvem essas três desinências (-a, -ar e -ou) fortalece e contribui, ainda que moderadamente, para formação dos primeiros esquemas morfossintáticos.

Interessantemente, não observamos nenhum erro de regularização (i.e., erros em que a criança generaliza flexões regulares para verbos irregulares). Esse resultado evidencia mais uma vez a sensibilidade da criança à frequência de ocorrência. Uma vez que os verbos irregulares são muito frequentes no *input* lingüístico, é possível que sejam fortalecidos como itens específicos e, dessa forma, ‘barrem’ a aplicação de esquemas regulares. Com efeito, esse tipo de erro geralmente só é observado mais tarde, por volta dos 2,5-3 anos de idade. É possível que com o aumento da experiência lingüística, os esquemas regulares se entrincheirem de tal forma que passem a sancionar a aplicação de flexões regulares a verbos irregulares.

O presente estudo apresenta uma série de limitações que impossibilitam conclusões mais pontuais. A primeira delas diz respeito ao *corpus* utilizado nas análises. Ainda é necessário em português, um *corpus* mais denso do que o utilizado no presente estudo. Como apontam Tomasello e Stahl (2004), a probabilidade de se observar fenômenos infreqüentes na fala espontânea da criança é fortemente influenciada pela densidade temporal com que o *corpus* é coletado. Um *corpus* mais denso certamente possibilitaria análises mais confiáveis e estatisticamente mais bem fundamentadas. Outra limitação do presente estudo diz respeito ao fato de seus resultados serem baseados em observações da fala de uma única criança.

Apesar dessas limitações, nossos resultados parecem indicar que, consoante com diversas pesquisas no escopo da Linguística Cognitiva (ver Tomasello, 2000, para uma revisão desses estudos), o processo de aquisição da morfossintaxe verbal é gradual e restrito a itens lexicais específicos, e que as frequências de tipo e de ocorrência no *input* lingüístico da criança parecem ter papel cogente no processo de aquisição da linguagem.

Referências

- Aguado-Orea, J. (2004). *The acquisition of morpho-syntax in Spanish: Implications for current theories of development*. Unpublished doctoral dissertation, University of Nottingham, UK.
- Akthar, N., & Tomasello, M. (1997). Young children's productivity with word order and verb morphology. *Developmental Psychology, 33*, 952-965.
- Barlow, M., & Kemmer, S. (Eds.). (2002). *Usage-based models of language*. Stanford, CA: CSLI.
- Brown, R. (1973). *A first language: The early stages*. London: George Allen.
- Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes, 10*, 425-455.
- Cameron-Faulkner, T., Lieven, E., & Tomasello, M. (2003). A construction-based analysis of child directed speech. *Cognitive Science, 27*, 843-873.
- Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.
- Clahsen, H. (1999). Lexical entries and rules of language: A multidisciplinary study of German inflection. *Behavioral and Brain Sciences, 22*(6), 991-1060.
- Croft, W., & Cruse, A. (2004). *Cognitive linguistics*. Cambridge, MA: Cambridge University Press.
- Deen, K., & Hyams, N. (2006). The morphosyntax of mood in early grammar with special reference to Swahili. *First Language, 26*, 67-102.
- Hausser, R. (1990). Principles of computational morphology. *Computational Linguistics, 47*.
- Hoekstra, T., & Hyams, N. (1998). Aspects of root infinitives. *Lingua, 106*, 81-112.
- Hyams, N. (1986). *Language acquisition and the theory of parameters*. Norwell, MA: Reidel.
- Lakoff, G. (1987). *Women, fire and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.
- Langacker, R. (1987). *Foundations of cognitive grammar: Vol. 1*. Stanford, CA: Stanford University Press.
- MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk*. Mahwah, NJ: Lawrence Erlbaum.
- MacWhinney, B., & Snow, C. (1990). The child language data exchange system: An update. *Journal of Child Language, 17*, 457-472.
- Miotto, C., Silva, M., & Lopes, R. (2000). *Manual de sintaxe*. Florianópolis, SC: Insular.
- Olguin, R., & Tomasello, M. (1993). Twenty-five-month-old children do not have a grammatical category of verb. *Cognitive Development, 8*, 245-272.
- Oliva, A. (2001). *Papel do input e da interação na aquisição de linguagem: Proposta de um modelo explicativo*. Tese de Doutorado não-publicada, Universidade de São Paulo, SP.
- Pine, J., & Lieven, E. (1997). Slot and frame patterns and the development of the determiner category. *Applied Psycholinguistics, 18*, 123-138.
- Pine, J., Lieven, E., & Rowland, G. (1998). Comparing different models of the development of the English verb category. *Linguistics, 36*, 4-40.
- Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.
- Pizzuto, E., & Caselli, M. (1992). The acquisition of Italian morphology: Implications for models of language development. *Journal of Child Language, 19*, 491-557.
- Poepfel, D., & Wexler, K. (1993). The full competence hypothesis of clause structure in Early German. *Language, 69*, 1-33.
- Roeper, T., & Williams, E. (1987). *Parameter setting and language acquisition*. Norwell, MA: Kluwer Academic.
- Rubino, R., & Pine, J. (1998). Subject-verb agreement in Brazilian Portuguese: What low error rates hide? *Journal of Child Language, 25*, 35-60.
- Scliar-Cabral, L., & MacWhinney, B. (2005). *A aquisição da morfologia verbal em português brasileiro*. Paper presented at the meeting of the XIV Congreso Internacional de la Asociación de Lingüística y Filología de América Latina, Monterrey, México.

- Taylor, J. (2002). *Cognitive grammar*. Oxford, UK: Oxford University Press.
- Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. Cambridge, MA: Cambridge University Press.
- Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition*, 74, 209-253.
- Tomasello, M. (2001). The item-based nature of children's early syntactic development. In M. Tomasello & E. Bates (Eds.), *Language development. The essential readings* (chap. 11). Oxford, UK: Blackwell.
- Tomasello, M. (2007). Cognitive linguistics and first language acquisition. In D. Geeraerts & H. Cuyckens (Eds.), *The Oxford handbook of cognitive linguistics* (pp. 1092-1112). Oxford, UK: Oxford University Press.
- Tomasello, M., & Stahl, D. (2004). Sampling children's spontaneous speech: How much is enough? *Journal of Child Language*, 31, 101-121.
- Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of optional infinitive stage. *Lingua*, 106, 23-79.
- Wittek, A., & Tomasello, M. (2002). German children's productivity with tense morphology: The perfekt. *Journal of Child Language*, 29, 567-590.
- Wittek, A., & Tomasello, M. (2005). German-speaking children's productivity with syntactic constructions and case morphology: Local cues help locally. *First Language*, 25, 103-125.