

Conferências do V Congresso Internacional da Associação Brasileira de Linguística



N.Cham. 410 C749c 2007

Autor: Congresso Internacional da Assoc

Título: Conferência do V Congresso Internacional



291730904

470362

Organizadoras

Roberto Silva

Heliana Mello

Conferência do V Congresso Internacional da Associação Brasileira de Lingüística

Diagramação Leonardo Araújo
Capa CEDECOM - UFMG

Associação Brasileira de Linguística. Congresso (5. : 2007 :
Belo Horizonte. MG)

A849c Conferências do V Congresso Internacional da As-
sociação Brasileira de Linguística / Thaís Cristóforo Silva e He-
liana Mello (organizadoras). - Belo Horizonte : FALE/UFMG,
2007. 280 p.

ISBN: 978-85-7758-018-7

Linguística - Congressos. I. Silva, Thaís Cristóforo. II. Mello,
Heliana. III. Título.

CDD: 410

Faculdade de Letras
BIBLIOTECA UNIVERSITARIA
13/09/09

2917309-04

FALE
BELO HORIZONTE

Belo Horizonte
Faculdade de Letras
Universidade Federal de Minas Gerais

Associação Brasileira de Linguística
www.abralin.org
Gestão UFMG: 2005-2007

Associação Brasileira de
Linguística
www.abralin.org

Gestão UFMG: 2005-2007

Diretoria

Presidente: Thaís Cristófaró Silva
Vice-Presidente: Heliana Ribeiro de Mello
1o Secretário: Luiz Francisco Dias
2o Secretário: Glaucia Lara Muniz
1o Tesoureiro: Ricardo de Souza
2o Tesoureiro: Ana Cristina Fricke Matte

Conselho

Profa. Esmeralda Negrão (USP)
Prof. Luiz Antônio Marcuschi (UFPE)
Profa. Maria Cecília Magalhães Mollica (UFRJ)
Profa. Maria do Socorro Silva de Aragão (UFCE)
Prof. Paulino Vandressen (UFSC-UCPEL)
Profa. Stela Maris Bortoni-Ricardo (UnB)

U.F.M.G. - BIBLIOTECA UNIVERSITÁRIA



291730904

Sumário

Apresentação	1
Visão Panorâmica da Associação Brasileira de Linguística em 2007 <i>Thais Cristófaru Silva</i>	5
Histórico da tabela de áreas do conhecimento na área de linguística e necessidades de atualização na tabela vigente <i>Rosalí Fernandez de Souza</i>	15
Por uma descrição gramatical mais concreta: as funções sintáticas <i>Mário A. Perini</i>	29
The bipartite structure of verbs cross-linguistically <i>Hcidi Harley</i>	45
Prosodic complexity and processing complexity: evidence from language impairment <i>John Harris, Nicola Gallon e Heater Van der Lely</i>	85
The role of production in the development of phonological representations <i>Marilyn May Vihman & Tamar Keren-Portnoy</i>	101
How might a rapid serial visual presentation of text affect the prosody projected implicitly during silent reading? <i>Eva M. Fernández</i>	117
Reflexos segmentais da organização rítmica do português do Brasil <i>Maria Bernadete Marques Abaurre</i>	135
L'enonciation Aphorisante <i>Dominique Mainqueneau</i>	155
Linguística comparativa e pré-história dos povos indígenas sul-americanos: a hipótese Tupí-Karíb e as relações genéticas entre Tupí, Karíb e Macro-Jê <i>Aryon Dall'Igna Rodrigues</i>	165
Language as a biocultural niche and social institution <i>Chris Sinha</i>	177
A Dança das linguagens na web: critérios para a definição de hipertexto <i>Antonio Carlos Xavier</i>	199
C-ORAL-ROM: Comparing romance languages in spontaneous speech corpora <i>Emanuela Cresti & Massimo Moneglia</i>	211

Apresentação

Durante o “V Congresso Internacional da ABRALIN”, realizado na Universidade Federal de Minas Gerais, Belo Horizonte, entre 28 de fevereiro e 03 de março de 2007, foram proferidas conferências, as quais perpassaram diversos campos temáticos dentro do escopo maior da Linguística contemporânea, assim como também foram apresentadas uma visão panorâmica da Linguística brasileira e um quadro geral da discussão em torno da especificação e rotulação das áreas do conhecimento na pesquisa brasileira.

Neste volume são reunidos os textos das conferências proferidas durante o “V Congresso Internacional da ABRALIN”, os quais foram submetidos por seus autores, com ajustes mínimos à adequação de formatação e padronização, para que se criasse uma homogeneidade e identidade para esta obra.

Os conferencistas apresentaram os seus resultados de pesquisas para um amplo público de lingüistas que participou do Congresso, bem como discutiram os seus trabalhos com tal público. Vale destacar que os conferencistas convidados foram indicados por pesquisadores afiliados à ABRALIN, em consulta democrática à comunidade científica, a qual, entendemos, reflete as linhas temáticas prioritárias dos estudos lingüísticos que foram trazidas à luz, durante o “V Congresso Internacional da ABRALIN”. A publicação deste volume tem como objetivo central registrar o foco de debate fomentado durante o Congresso por meio das conferências proferidas.

Este volume é organizado de acordo com a seguinte arquitetura: serão apresentados treze capítulos, sendo que inicialmente introduzem-se os capítulos panorâmicos sobre a Linguística brasileira e sobre o debate, para se classificarem as áreas de saber para fins de catalogação pelo CNPq. Seguem-se onze capítulos retratando palestras de cunho científico proferidas durante o evento.

O Capítulo de Thaís Cristófaró Silva (página 5) apresenta um retrato numérico da lingüística brasileira, enfocando as características principais dos associados da ABRALIN no cenário nacional. Avalia-se de maneira análoga a representatividade de lingüistas que são bolsistas em Produtividade em Pesquisa do CNPq. O mapeamento apresentado avalia a natureza da qualificação, do vínculo institucional, da distribuição geográfica e das áreas do conhecimento dos sócios da ABRALIN e dos bolsistas em Produtividade em Pesquisa do CNPq. Ao oferecer um mapeamento dos associados da ABRALIN, incluindo os bolsistas do CNPq, espera-se poder contribuir com o delineamento de ações futuras da ABRALIN, bem como com a instrumentalização de políticas de desenvolvimento científico e tecnológico no país.

O Capítulo de Rosali Fernandez (página 15) aborda a dificuldade de representação de áreas do conhecimento em tabelas ou esquemas de classificação para qualquer propósito de recuperação de informação. Segundo a autora, este processo tem como pontos referenciais a natureza da informação a ser retratada e os objetivos de uso da classificação.

O Capítulo de Mário Perini (página 29) discute uma proposta trazida por Culicover e Jackendoff (2005, p. 5), a Hipótese da Sintaxe Simples (HSS), que se

refere ao fato de que a teoria sintática mais explicativa é a que atribui o mínimo de estrutura necessário para exprimir a mediação entre fonologia e significado. Perini aplica a HSS a algumas relações abstratas comuns em gramática, a saber, as chamadas funções sintáticas, em especial a de sujeito.

O Capítulo de Heidi Harley (página 45) expõe a necessidade da proposta de uma estrutura complexa, bipartida para o sintagma verbal, que se aplique inclusive ao verbo. Tal estrutura configura-se por meio de um elemento *V'* e de outro *V*. A autora revisa algumas das motivações teóricas e empíricas para essa estrutura complexa. A seguir, apresenta fenômenos que essa abordagem estendida cobre, fenômenos esses que até então não haviam sido sintaticamente tratados. A autora demonstra também como essa abordagem permite uma nova e, talvez, mais abrangente explicação para uma peculiaridade da alternância dativa do inglês: isto é, o fato de essa alternância não se aplicar à porção de base latina do inventário verbal inglês.

O Capítulo de John Harris (página 85) discute se uma concepção de complexidade prosódica faz previsões válidas sobre o grau de complexidade ou dificuldade em processamento fonológico on-line. Por meio de métodos experimentais, o autor demonstra que a complexidade métrica e interna à sílaba pode afetar a acurácia na repetição, independentemente do número de segmentos ou sílabas em uma dada palavra inventada.

O Capítulo de Marilyn May Vilman e Tamar Keren-Portnoy (página 101) investiga relações entre a produção e a percepção durante a aquisição da linguagem. As autoras discutem a relevância do léxico do adulto em ciclos de aprendizagem baseados em tipos (types). A seguir, discutem o modelo "Whole-word Phonology", avaliando dados do inglês e do português brasileiro. Ao discutirem dados sobre a percepção, relativos a reconhecimento e aprendizagem da linguagem, apresentam resultados sobre a produção e a memória fonológica.

O Capítulo de Eva Fernandez (página 117) revisita algumas investigações que testam como as prosódias explícita e implícita influenciam a resolução de ambigüidades. A autora descreve os resultados de cinco experimentos que examinam se a "prosódia visual" afeta a prosódia projetada durante a leitura silenciosa.

O Capítulo de Bernadete Abaurre (página 135) apresenta algumas considerações preliminares sobre a possível relação entre a organização rítmica do português do Brasil conforme refletida em dados de um corpus específico, a distribuição dos acentos secundários e a aplicação de processos de sândi vocálico de degeminação, elisão e ditongação no interior de um domínio prosódico específico, a palavra prosódica.

O Capítulo de Dominique Maingueneau (página 155) evoca a questão da enunciação aforisante, que se encontra no cruzamento entre a lingüística da enunciação e a análise do discurso, adotando uma perspectiva que o autor nomeia de "antropolingüística". Essa mesma questão já havia sido anteriormente mencionada pelo autor em um capítulo do livro *Cenas de enunciação*, sem que ela tivesse sido desenvolvida em profundidade como ele o faz neste artigo.

O Capítulo de Aryon Rodrigues (página 165) apresenta a Hipótese Tupi-Karib, argumentando em favor das relações genéticas entre as línguas Tupi-Karib e Macro-Jê, por meio do método comparativo da Lingüística Histórica.

O Capítulo de Chris Sinha (página 177) argumenta a favor da dupla natureza da linguagem como pertencente à cultura (segundo Condillac) ou como uma das faculdades superiores dos humanos (segundo Descartes). Segundo o autor,

as questões que se colocam nessa discussão são: em que se sentido podemos recolocar nossas idéias sobre a ontologia da linguagem, de forma a que ela seja vista tanto como uma forma sócio-cultural quanto algo intrínseco à biologia humana? Como essa nova visão "biocultural" da linguagem afeta nossas idéias sobre o seu processamento e sua aquisição?

O Capítulo de António Xavier (página 199) enfoca os seguintes objetivos: discutir um conjunto de parâmetros para definir o hipertexto e assim evitar confusões terminológicas, tais como: tomar hipertextualidade por intertextualidade e considerar a deslinearidade como característica exclusiva do hipertexto; e apontar a necessidade de os internautas compreenderem relativamente bem as regras dos jogos de linguagem que viabilizam sua incursão produtiva na rede.

O Capítulo de Emanuela Cresti e Massimo Moneglia (página 211) apresenta o corpus C-ORAL-ROM (Corpora de Referência Integrada para as Línguas Românicas Faladas) e sua arquitetura, que tenciona representar os principais fatores que afetam a variabilidade da língua falada. O artigo enfoca as condições que possibilitam a comparatividade de corpora para a análise lingüística, no nível da arquitetura de corpus e de sua anotação. O valor da anotação das quebras prosódicas para a determinação de enunciados e fronteiras é discutido em face de informação sintática.

Este volume registra as principais idéias veiculadas nas conferências proferidas no V Congresso Internacional da ABRALIN. A nossa expectativa é que através deste volume os leitores tenham a oportunidade de aprofundar a discussão das idéias inicialmente apresentadas nas conferências. Façam bom proveito!

Thaïs Cristófaró Silva
Heliana Mello

Visão Panorâmica da Associação Brasileira de Lingüística em 2007

Thaís Cristófaru Silva
UFMG-CNPq

thaiscristofarosilva@ufmg.br

Abstract

This article presents the major characteristics of the members of the Associação Brasileira de Lingüística (ABRALIN). The data offer general information about the geographical distribution of members, their institutional affiliation, their qualification and also indicate to which area of knowledge they associate themselves to. It is expected that this article will contribute to a more comprehensive picture of ABRALIN members as well as it will offer support for future scientific policies in the area of linguistics.

1 Introdução

Este trabalho tem por objetivo apresentar as principais características dos associados da Associação Brasileira de Lingüística - ABRALIN no cenário nacional. O mapeamento apresentado avalia, em relação aos sócios da ABRALIN, o vínculo institucional, a natureza da qualificação, a distribuição geográfica e o vínculo com as áreas do conhecimento. Ao oferecer um mapeamento dos associados da ABRALIN, espera-se poder contribuir com o delineamento de ações futuras da ABRALIN, bem como com a instrumentalização de políticas de desenvolvimento científico e tecnológico no país. A primeira seção apresenta dados que contextualizam a Lingüística no cenário científico e tecnológico ao avaliar tal área no contexto da Coordenação do Programa de Pesquisa em Ciências Humanas e Sociais do CNPq - COCHS. A segunda seção apresenta um mapeamento dos sócios da ABRALIN em julho de 2007. A terceira seção finaliza o presente trabalho com uma avaliação geral dos dados apresentados.

2 A área de Lingüística no contexto de Ciências Humanas e Sociais

Esta seção tem por objetivos contextualizar a área de *Lingüística* na área de Ciências Humanas e Sociais. Os dados a serem apresentados foram recebidos pela Presidência da ABRALIN, em fevereiro de 2007, a partir de solicitação específica feita para a Coordenação do Programa de Pesquisa em Ciências Humanas e Sociais - COCHS do CNPq. O objetivo principal da solicitação dos dados ao CNPq foi o de buscar informações que permitissem contextualizar o fomento aos bolsistas de Produtividade em Pesquisa do CNPq na área de *Lingüística*. A ABRALIN recebeu um arquivo contendo dados referentes ao

número de pesquisadores de Produtividade em Pesquisa do CNPq, da área de ciências humanas em geral, cujos índices por área reproduzimos a seguir.¹

	Comitê	Área	N	%
1	PH	Psicologia	224	20,33
2	LL	Letras	179	16,24
3	HI	História	170	15,43
4	LL	Linguística	157	14,25
5	FI	Filosofia	102	9,26
6	AC	Comunicação	92	8,35
7	AC	Artes	74	6,72
8	PH	Serviço Social	58	5,26
9	AC	Ciência da Informação	34	3,08
10	CT	Divulgação Científica	9	0,81
11	FI	Teologia	2	0,18
12	CT	Museologia	1	0,09
			1102	

Tab. 1: Bolsistas em Produtividade em Pesquisa CNPq. Dados COCH - CNPq em fevereiro de 2007

A COCHS do CNPq engloba 12 áreas:² *Artes, Ciência da Informação, Comunicação, Divulgação Científica, Filosofia, História, Letras, Linguística, Museologia, Psicologia, Serviço Social e Teologia.* A tabela 1 lista as áreas do conhecimento em relação ao número de bolsistas em Produtividade de Pesquisa do CNPq. Pode-se observar que a área de Linguística representa o quarto lugar na área da COCHS, contando com 157 pesquisadores de Produtividade em Pesquisa, sendo precedido pelas áreas de Comunicação, Letras e História. Os dados da Tabela 1 sugerem que a Linguística tem uma atuação importante na área a qual se vincula no CNPq, sobretudo se consideramos a consolidação, relativamente recente, da Linguística como ramo independente da ciência. A distribuição dos 157 bolsistas de Produtividade em Pesquisa quanto à classificação do nível da bolsa é a seguinte:

¹Os Comitês de cada área são: AC (Artes, Comunicação e Ciência da Informação), HI (História), FI (Filosofia), LL (Letras e Linguística), PH (Psicologia e Serviço Social) e CT (Comitê Temático e Divulgação Científica).

²Outras áreas relacionadas com as Ciências Humanas encontram-se sob a Coordenação do Programa de Pesquisa em Ciências Humanas e Sociais Aplicadas - CGCHS do CNPq e Coordenação do Programa de Pesquisa em Ciências Sociais Aplicadas e Educação - COSAE. Estas e outras coordenações, incluindo a COCHS que encampa a área de Linguística, encontram-se vinculadas à Diretoria de Programas Horizontais e Instrumentais - DPH do CNPq.

Nível PQ	N
1A	24
1B	16
1C	26
1D	24
2	67
Total	157

Tab. 2: Bolsistas de Produtividade em Pesquisa da Área de Linguística

Nesta seção observamos que a área de Linguística conta com 157 bolsistas de Produtividade em Pesquisa, representando 14,25% das bolsas sob a Coordenação do Programa de Pesquisa em Ciências Humanas e Sociais - COCHS do CNPq. A distribuição dessas bolsas é equilibrada contando com 57,4% de bolsistas de nível 1 e com 42,6% de bolsistas de nível 2. Na próxima seção apresentamos alguns dados referentes aos associados da ABRALIN.

3 Visão Panorâmica dos sócios da ABRALIN em julho 2007

De acordo com o estatuto vigente, a ABRALIN apresenta as seguintes categorias de sócios:

Sócios efetivos: são aqueles que se dedicam direta ou indiretamente à pesquisa linguística ou exercem o ensino de Linguística ou de línguas em nível universitário.

Sócios colaboradores: são aqueles interessados nos objetivos da Associação, mas que não preenchem as condições de sócios efetivos. Os sócios colaboradores dividem-se em graduados ou com formação acadêmica superior e alunos com a graduação em curso.

Os sócios podem indicar o tipo de associação desejado no ato da inscrição, e a proposta é posteriormente aceita ou declinada pelo Conselho da ABRALIN. Os dados da Tabela 3 caracterizam a distribuição dos sócios da ABRALIN em julho de 2007.

	N	%
Efetivos	581	55,70
Colaboradores	385	36,91
Alunos de Graduação	77	7,39
TOTAL	1043	

Tab. 3: Sócios da ABRALIN em julho de 2007

Foram considerados nesta análise os 1043 sócios que preencheram os dados cadastrais no banco de dados implementado em 2006, no sítio da ABRALIN. Em tal banco de dados, os sócios deveriam informar dados pessoais e acadêmicos, de

maneira que o gerenciamento de um conjunto significativo de dados foi possível. Os dados foram preenchidos diretamente pelo associado na expectativa de se caracterizar com fidelidade o perfil de cada um dos sócios.

Os dados apresentados na Tabela 3 indicam que membros efetivos constituem a maioria dos associados da ABRALIN. De acordo com as categorias previstas no estatuto da ABRALIN, podemos afirmar que o perfil majoritário dos sócios da ABRALIN, i.e. sócios efetivos, é de profissionais que se dedicam direta ou indiretamente à pesquisa *Linguística* ou exercem o ensino de *Linguística* ou de línguas em nível universitário.

Vale ressaltar que em conjunto os sócios efetivos e colaboradores totalizam 92,61% (966 membros). Com relação ao tipo de qualificação dos sócios da ABRALIN, observamos os seguintes dados:

	N	%
Doutores	527	50,53
Mestres	348	33,36
Graduação completa	91	8,73
Graduação em curso	77	7,39
TOTAL	1043	

Tab. 4: Qualificação dos Sócios da ABRALIN

A maioria dos associados da ABRALIN tem Doutorado (50,53%) ou Mestrado (33,36%). Estes dados oferecem evidências adicionais de que os sócios da ABRALIN são bastante qualificados, contribuindo para a caracterização da ABRALIN como Associação de destaque no cenário nacional com relação ao desenvolvimento científico e tecnológico da área. Os sócios com graduação completa totalizam 8,73% e graduandos totalizam 7,39%. Em conjunto, os Doutores, Mestres e Graduados totalizam 966 sócios (92,61%), número correspondente aos sócios efetivos e colaboradores. A Tabela 5 sistematiza a relação entre sócios efetivos e colaboradores e a qualificação deles.

Tipo de sócio e qualificação		N	%
Efetivo	Doutores	456	47,20
	Mestres	125	12,93
Colaborador	Doutores	71	7,34
	Mestres	223	23,08
	Graduado	91	9,42
TOTAL		966	

Tab. 5: Tipo e Qualificação dos Sócios da ABRALIN

A maioria dos associados efetivos são Doutores, atuantes nas diversas áreas da *Linguística*. A distribuição geográfica dos sócios da ABRALIN é apresentada na Tabela 6.

Estado	N	%
1.Minas Gerais	269	25,81
2.São Paulo	214	20,52
3.Rio de Janeiro	143	13,72
4.Bahia	67	6,42
5.Rio Grande do Sul	49	4,69
6.Distrito Federal	47	4,51
7.Paraíba	40	3,84
8.Paraná	34	3,26
9.Pernambuco	32	3,06
10.Santa Catarina	30	2,88
11.Ceará	20	1,91
12.Pará	15	1,44
13.Goiás	14	1,34
14.Mato Grosso do Sul	12	1,15
15.Alagoas	10	0,96
16.Espírito Santo	10	0,96
17.Mato Grosso	7	0,67
18.Rio Grande do Norte	7	0,67
19.Acre	5	0,48
20.Rondônia	5	0,49
21.Piauí	4	0,38
22.Sergipe	4	0,38
23.Amazonas	2	0,19
24.Amapá	1	0,09
25.Maranhão	1	0,09
26.Roraima	1	0,09
27.Tocantins	0	0,00

Tab. 6: Distribuição de Sócios da ABRALIN por Estado

Os dados da Tabela 6 indicam o estado brasileiro de residência dos sócios da ABRALIN. Os dados foram listados em ordem decrescente. Minas Gerais conta com um grande número de sócios, possivelmente, devido ao fato de o "V Congresso Internacional da ABRALIN" ter sido realizado em Minas Gerais em 2007. Podemos observar que os estados de Mato Grosso, Rio Grande do Norte, Acre, Rondônia, Piauí, Sergipe, Amazonas, Amapá, Maranhão, Roraima apresentam um baixo número de associados. Nenhum associado da ABRALIN reside no estado do Tocantins. Estes números expressam que tais estados brasileiros possivelmente não contam com desenvolvimento de pesquisa na área de Lingüística e seria importante captar associados nestes estados.

Finalmente, apresentamos a distribuição dos sócios da ABRALIN em relação à área do conhecimento à qual se vinculam. Considerando-se as 14 áreas do conhecimento selecionadas para inscrição no "V Congresso Internacional da ABRALIN" em 2007, temos a seguinte distribuição:

	Área do Conhecimento	N	%
1	Análise do texto e do discurso	261	25,02
2	Sociolinguística e dialetologia	123	11,79
3	Ensino de língua(s)	90	8,62
4	Morfologia e sintaxe	84	8,05
5	Semântica e pragmática	76	7,28
6	Psicolinguística e aquisição da linguagem	69	6,61
7	Fonética e fonologia	63	6,04
8	Linguística histórica	61	5,85
9	Práticas Sociais da Linguagem	58	5,56
10	Línguas Indígenas	42	4,02
11	Linguística de corpus e computacional	36	3,45
12	Alfabetização e Letramento	35	3,36
13	Tradução	26	2,49
14	Terminologia	19	1,82
	TOTAL	1043	

Tab. 7: Área de atuação dos Sócios da ABRALIN

Os números da Tabela 7 indicam a área do conhecimento que foi selecionada pelo associado quando houve o preenchimento do cadastro no banco de dados da ABRALIN. Os índices diferenciados entre as áreas listadas na Tabela 7 expressam a organização atual dos pesquisadores e a dinâmica de organização do conhecimento neste momento.

As áreas do conhecimento listadas na Tabela 7 foram propostas pela comissão organizadora do “V Congresso Internacional da ABRALIN”. O agrupamento em áreas do conhecimento sugerido para o Congresso buscou expressar uma proposta de áreas do conhecimento aprovada em reunião entre o CNPq e Associações da área de Letras e Linguística. Tal proposta é reproduzida a seguir.³

A Tabela 8 reproduz a lista apresentada no documento final de uma reunião que se realizou no Rio de Janeiro, no dia 20 de dezembro de 2005, organizada pelo Professor Manoel Domingos, Vice-presidente do CNPq, e por alguns membros da Comissão de reorganização da Tabela de Áreas do Conhecimento (TAC), com as seguintes associações: ABRALIN, ABRALIC e ANPOLL (ALAB já havia se reunido em Brasília com o referido professor).

De acordo com a Tabela 8, as seis subáreas do conhecimento – Fonética e Fonologia, Morfologia e Sintaxe, Semântica e Pragmática, Sociolinguística e Dialetologia, Psicolinguística e Aquisição da linguagem, e Análise do texto e do discurso – figuram como parte da área de Linguística. As quatro subáreas do conhecimento – Ensino de língua(s), Práticas Sociais da Linguagem, Tradução e Terminologia – figuram como parte da área de Linguística Aplicada. As áreas Linguística histórica, Línguas Indígenas, Linguística de corpus e computacional,

³ A proposta apresentada na Tabela 8 contém, adicionalmente, uma lista de *Especialidades*. A sugestão dessa lista foi discutida na referida reunião, sendo que tal lista é reproduzida no Apêndice para efeito de documentação.

<p>ÁREA: Lingüística</p> <p>1.Fonética e fonologia</p> <p>2.Morfologia e sintaxe</p> <p>3.Semântica e pragmática</p> <p>4.Sociolingüística e Lingüística Histórica</p> <p>5.Psicolingüística e Aquisição da Linguagem</p> <p>6.Teoria e análise do discurso e do texto</p>
<p>ÁREA: Lingüística Aplicada</p> <p>1.Ensino de Línguas</p> <p>2.Práticas sociais da linguagem</p> <p>3.Tradução</p> <p>4.Terminologia</p>

Tab. 8: Documento de área das Associações

Alfabetização e Letramento não haviam sido contempladas na proposta apresentada na Tabela 8, mas foram apresentadas aos associados como opção de cadastro no banco de dados da ABRALIN.

Levando-se em conta o vínculo dos sócios da ABRALIN com as áreas da *Lingüística* e *Lingüística Aplicada* e *Outras* subáreas que não apareceram na proposta apresentada na Tabela 8, podemos apresentar a seguinte sistematização:

	Área do Conhecimento	N	%
1	Lingüística	676	64,81%
2	Lingüística Aplicada	193	18,51%
3	Outras	174	16,68%
	TOTAL	1043	

Tab. 9: Área Lingüística, Lingüística Aplicada e Outras

Os dados da Tabela 9 indicam que a maioria dos associados da ABRALIN atua na área denominada *Lingüística* (64,81%) apresentada na Tabela 8. Um total de 35,19% dos associados atua na área denominada *Lingüística Aplicada* (18,51%) e em *Outras* áreas temos (16,68%). A próxima seção avalia, de maneira geral, os dados apresentados.

4 Conclusão

Este artigo teve por objetivo documentar as principais características dos sócios da Associação Brasileira de Lingüística - ABRALIN no cenário nacional, com dados obtidos a partir de consulta ao banco de dados da Associação em julho de 2007. Os dados apresentados refletem informações fornecidas pelos associados individualmente. Associados podem alterar seus dados cadastrais, de maneira que a qualquer momento é possível que a Associação proceda avaliações semelhantes a esta apresentada neste artigo.

Os dados apresentados avaliaram, em relação aos sócios da ABRALIN, o vínculo institucional, a natureza da qualificação, a distribuição geográfica e o

vínculo com as áreas do conhecimento. Podemos afirmar que a ABRALIN constitui um fórum importante de debate e circulação de informação de ponta em Lingüística, ao congregar 1043 sócios (ver Tabelas 3-5). Os associados, em sua maioria são sócios efetivos, que "se dedicam direta ou indiretamente à pesquisa lingüística ou exercem o ensino de Lingüística ou de línguas em nível universitário" (ver Tabela 3). A maioria dos sócios da ABRALIN é qualificada em nível de Pós-graduação, ou seja, 83,89% são Doutores (50,53%) ou Mestres (33,36%). Entendemos que a qualificação de alto nível dos associados da ABRALIN oferece condições para a realização de eventos organizados pela Associação que divulguem importantes resultados de pesquisa na área da Lingüística.

Os estados do Mato Grosso, Rio Grande do Norte, Acre, Rondônia, Piauí, Sergipe, Amazonas, Amapá, Maranhão e Roraima apresentam um baixo número de associados. Nenhum associado da ABRALIN reside no estado do Tocantins. Há, portanto, necessidade de captar mais associados em alguns estados do Brasil, na expectativa de ampliar o fórum de debate na área de *Lingüística* que é fomentado pela ABRALIN (ver Tabela 6).

Quanto às áreas de atuação dos associados da ABRALIN, há diferenças de índices em subáreas distintas que expressam o agrupamento contemporâneo e a emergência de algumas áreas.

5 Apêndice

Lista de especialidades

- | | |
|---|--|
| 1. Alfabetização e Letramento | 14. Ensino - aprendizagem de LI-BRAS |
| 2. Análise da Conversação | 15. Ensino a distância |
| 3. Análise do Discurso | 16. Estudos de Cognição e Linguagem |
| 4. Aquisição da Linguagem Escrita | 17. Estudos de Políticas e Planejamentos Lingüísticos |
| 5. Aquisição da Linguagem Oral | 18. Etnolingüística |
| 6. Aquisição da língua materna normal e deficitária | 19. Filologia |
| 7. Aquisição de Segunda Língua | 20. Filosofia da Linguagem |
| 8. Bilingüismo e Multilingüismo | 21. Fonética Acústica |
| 9. Crioulística | 22. Fonologia Articulatória |
| 10. Ensino e aprendizagem de L1 | 23. Fonologia Teórica |
| 11. Ensino e aprendizagem de L2: contextos indígenas e multilíngües | 24. Gramática Formal |
| 12. Ensino e aprendizagem de LE | 25. Gramática Funcional |
| 13. Ensino - aprendizagem de Português LE | 26. História da Lingüística |
| | 27. História das Idéias Lingüísticas e dos Estudos sobre Linguagem |

28. Identidade e Linguagem
29. Interacionismo
30. Interpretação
31. Legendagem
32. Língua Portuguesa
33. Línguas Africanas
34. Línguas Clássicas
35. Línguas de Sinais
36. Línguas Indígenas
37. Lingüística Antropológica
38. Lingüística Cognitiva
39. Lingüística Computacional
40. Lingüística de Corpus
41. Lingüística Românica
42. Lingüística Textual
43. Morfossintaxe
44. Morfologia Derivacional
45. Neurolingüística
46. Patologias da linguagem
47. Políticas do ensino de línguas
48. Processamento fonológico
49. Processamento sintático
50. Processamento semântico
51. Processamento do discurso
52. Prosódia
53. Representação e acesso lexical
54. Semiótica
55. Semântica Cognitiva
56. Semântica da Enunciação
57. Semântica Formal
58. Semântica Lexical
59. Teorias da Tradução
60. Teoria Lexical
61. Teoria da Linguagem Verbal
62. Terminografia
63. Tradução científica e literária
64. Tradução, Contexto e Cultura
65. Tradução Automática e Tecnologias de Tradução
66. Tradução Especializada
67. Variação e Mudança Sonora
68. Variacionismo

Histórico da tabela de áreas do conhecimento na área de linguística e necessidades de atualização na tabela vigente

Rosali Fernandez de Souza

Instituto Brasileiro de Informação em Ciência e Tecnologia

rosali@openink.com.br

Abstract

This article discusses the evolution of TACs (Tabelas de Áreas do Conhecimento - Areas of Knowledge Tables) as proposed in 1976, 1982 and 1984. Besides evaluating the chronological development of these tables the author also addresses how knowledge is viewed and organized in the area of Linguistics in order to provide information to Brazilian funding agencies.

1 Considerações Iniciais

A representação de áreas do conhecimento em tabelas ou esquemas de classificação para qualquer propósito de recuperação de informação é tarefa complexa e difícil. Este processo tem como pontos referenciais a natureza da informação a ser retratada e os objetivos de uso da classificação. No âmbito das agências de fomento em Ciência, Tecnologia e Informação (C,T&I), a classificação dos saberes é construída segundo o contexto político-administrativo definido em função das missões institucionais das agências, procurando levar em consideração as necessidades de uso de tal instrumento pela comunidade científica.

No ano de 2005, foi criada a Comissão de Estudos CNPq-CAPES e FINEP¹ com o objetivo de propor uma nova Tabela de Áreas do Conhecimento, daqui por diante TAC. A TAC atual, datada da década de oitenta, necessitava de atualização decorrente do desenvolvimento da ciência, tecnologia e inovação no país, nos últimos vinte anos. A metodologia adotada pela Comissão mobilizou a participação de pesquisadores em caráter individual, assim como das associações e sociedades científicas, programas de pós-graduação, fundações de amparo à pesquisa e comitês assessores das agências de fomento.

A inclusão de tal tema na Conferência Internacional da ABRALIN revelou o interesse na discussão da Tabela de Áreas do Conhecimento e o reconhecimento da importância de tal instrumento na comunicação entre os pesquisadores e as agências de fomento.

¹CNPq: Conselho Nacional de Desenvolvimento Científico e Tecnológico; CAPES: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior e FINEP: Financiadora de Estudos e Projetos.

No âmbito da organização do conhecimento para propósitos de gestão e avaliação de C,T&I, duas questões-chave são importantes de serem evidenciadas: o conceito de informação a ser considerado e as premissas básicas de classificação em função de necessidades e interesses de representação e recuperação de informação.

Com referência ao conceito de informação, é importante destacar que a classificação para finalidades de gestão e avaliação representa a informação *sobre* C&T e não a informação *de* C&T. Em outras palavras, a informação para gestão e avaliação em C&T se caracteriza como meta-informação de C&T.

Quanto à classificação, processo fundamental da natureza humana, é importante evidenciar como premissas básicas que: sem classificação, não poderia haver nenhum pensamento, ação ou organização que conhecemos; que os mesmos objetos e idéias podem ser classificados de maneiras diferentes; e que toda classificação é construída em função dos objetivos definidos de uso. Portanto, a relatividade da adequação de tabelas e esquemas de classificação está necessariamente condicionada aos propósitos específicos de construção e uso do instrumento.

O presente trabalho apresenta reflexões em torno e além da revisão da Tabela de Áreas do Conhecimento vigente levando em consideração a atuação da autora como secretária da Comissão de Estudos CNPq-CAPES e FINEP formada em 2005 e de resultados de pesquisa na área de Organização e Representação do Conhecimento no contexto da Ciência da Informação. O objetivo geral deste artigo é revelar um pouco da complexidade de elaboração de tabelas de classificação de áreas do conhecimento no contexto da gestão e avaliação de C,T&I. Como objetivo específico, pretende-se mostrar a representação da 'Linguística' como área do conhecimento nas diferentes versões da TAC revelando mudanças ocorridas no tempo. Como objetivo complementar, pretende-se apresentar, a título de comparação, como a área da 'Linguística' está representada em outras tabelas de áreas do conhecimento no Brasil e no exterior, construídas para objetivos diferentes de recuperação de informação: a tabela da SBPC (Sociedade Brasileira para o Progresso da Ciência) para a classificação das comunicações de pesquisa apresentadas às reuniões anuais da Sociedade e a tabela desenvolvida pelo Australian Bureau of Statistics para a classificação de projetos de pesquisa e desenvolvimento com a finalidade de geração de estatísticas nacionais.

2 A "Linguística" na Tabela de Áreas do Conhecimento do CNPq

A Tabela de Áreas do Conhecimento (TAC) do CNPq é adotada também pela CAPES e pela FINEP e por outras instituições de C&T no país, na íntegra ou com adaptações. Como já mencionado, a TAC é um instrumento de classificação de elaboração complexa. Como linguagem de organização e representação do universo de C,T&I visa atender as necessidades específicas de recuperação de informação para propósitos organizacionais de natureza político-administrativa das agências de fomento, buscando consonância com interesses específicos de representação de domínios de pesquisa, ensino e extensão da comunidade científica.

Ao longo do tempo, a TAC teve três versões: 1976, 1982 e 1984 e passou por várias tentativas, formais e informais, de revisão durante as duas últimas décadas. As diferentes versões da TAC apresentam, em sua introdução, a finalidade da classificação que a versão em vigor define como: "A classificação das áreas do conhecimento tem finalidade eminentemente prática, objetivando proporcionar aos órgãos que atuam em C&T, uma maneira ágil e funcional de agregar suas informações".

A seguir, apresenta-se uma breve indicação dos aspectos da estrutura adotada pelas diferentes versões da TAC no tempo e a representação da área da Linguística em cada versão.

2.1 A 'Linguística' nas versões TAC 1976, TAC 1982 e TAC 1984

A TAC 1976 apresenta uma estrutura hierárquica em três níveis: área, sub-área e especialidade (não apresenta grandes áreas). As quarenta e duas áreas do conhecimento são apresentadas em ordem alfabética. Entre elas, a Linguística está apresentada com Letras, formando a área "Linguística e Letras". A "Linguística" também está especificada como sub-área, no conjunto formado por "Ensino de Línguas, Gramática e Letras". Como sub-área, a Linguística apresenta as seguintes especialidades: Teoria Linguística e Semiologia; Filologia; Distúrbios e Correção da Linguagem; Psicolinguística; Sociolinguística e Etnolinguística; Linguística Computacional; Neurolinguística; Línguas Indígenas Brasileiras; Dialetoлогия; Outras (especificar). A sub-área "Outras" foi o recurso usado nesta versão para completar, quando necessário, a listagem das especialidades. Interessante observar que, na versão TAC 1976, a Literatura não aparece como área; aparece como sub-área, com Artes Cênicas, Artes Plásticas, Cinema e Música, na área Artes.

Quanto à TAC 1982, a sua parte introdutória esclarece que esta versão procura sanar omissões e inconsistências detectadas na edição anterior, observando as seguintes diretrizes: 1. não incidir em duplicações, ou seja, assuntos interdisciplinares são codificados em somente uma área do conhecimento, indicando a localização adotada pela remissiva 'ver'; 2. utilizar apenas termos relacionados com 'áreas do conhecimento' - embora não defina o que considera 'área do conhecimento', indica somente que "não codifica profissões, setores, tipos de atividades em C&T, etc."; 3. não utilizar o termo "Outras" em qualquer dos níveis da classificação.

A versão TAC 1982 apresenta estrutura hierárquica em 4 níveis: Grande Área, Área, Sub-área e Especialidade e compreende as seguintes seis grandes áreas ordenadas em ordem alfabética: Ciências Agrárias; Ciências Biológicas; Ciências da Engenharia; Ciências Exatas e da Terra; Ciências Humanas, Sociais e Artes; Ciências da Saúde. Nesta versão da TAC, a "Linguística" aparece como área, juntamente com Letras na grande área "Ciências Humanas, Sociais e Artes". A área "Letras e Linguística" compreende as sub-áreas: Língua, Literatura e Linguística. A sub-área "Linguística" apresenta as seguintes especialidades: Teoria e Análise Linguísticas; Linguística Aplicada; Sociolinguística e Dialetoлогия; Psicolinguística; Linguística Histórica; Pragmática. Além dessas especialidades, aparece a "Filosofia da Linguagem" com a nota "vide Filosofia", indicando que deve ser considerada como especialidade da Filosofia.

TAC 1976	
Área	Linguística e Letras
Sub-áreas	Ensino de Línguas, Gramática, Letras, Linguística
Especialidades	Teoria Lingüística e Semiologia Filologia Distúrbios e Correção da Linguagem Psicolinguística Sociolinguística e Etnolinguística Linguística Computacional Neurolinguística Línguas Indígenas Brasileiras Dialetologia Outras (especificar)
TAC 1982	
Grande Área	Ciências Humanas, Sociais e Artes
Área	Letras e Linguística
Sub-áreas	Língua, Literatura, Linguística
Especialidades	Teoria e Análise Lingüísticas Linguística Aplicada Sociolinguística e Dialetologia Psicolinguística Linguística Histórica Pragmática Filosofia da Linguagem (ver Filosofia)
TAC 1984, em vigor	
Grande Área	Linguística, Letras e Artes
Área	Linguística
Sub-áreas	Teoria e Análise Lingüísticas Fisiologia da Linguagem Linguística Histórica Sociolinguística e Dialetologia Psicolinguística Linguística Aplicada

Tab. 1: TAC Área Linguística no Tempo

A TAC 1984, versão em vigor, mantém a estrutura hierárquica em quatro níveis da versão de 1982, mas amplia o número de grandes áreas de seis para oito. Com o passar do tempo, incluiu a grande área nove, denominada 'Outros'. As oito grandes áreas da versão em vigor estão assim ordenadas: Ciências Exatas e da Terra; Ciências Biológicas; Engenharias; Ciências da Saúde; Ciências Agrárias; Ciências Sociais Aplicadas; Ciências Humanas; Linguística, Letras e Artes.

Em comparação com a versão anterior da TAC, de 1982, nota-se o desmembramento da grande área "Ciências Humanas, Sociais e Artes" em três grandes áreas: Ciências Sociais Aplicadas; Ciências Humanas; e Linguística, Letras e Artes, o que possibilitou uma ampliação considerável da classificação nessas áreas.

Na TAC em vigor, a "Linguística" aparece com Letras e Artes, formando a grande área "Linguística, Letras e Artes". Quanto à nomeação de área, é importante observar que na versão TAC 1984, quando a Linguística passa a ser área, as especialidades mencionadas nas versões anteriores passam a compor a hierarquia da tabela como sub-áreas. Assim, na TAC 1984, as sub-áreas da Linguística são: Teoria e Análise Linguística; Fisiologia da Linguagem; Linguística Histórica; Sociolinguística e Dialetoлогия; Psicolinguística; Linguística Aplicada. Na versão em vigor a "Linguística" não apresenta especialidades para nenhuma das sub-áreas.

Em relação à TAC em vigor, cabe uma observação em relação a Grande Área 9 'Outros'. A abertura dessa grande área pode ser interpretada como artifício de atualização em relação a itens não representados na Tabela cuja origem remonta, como já mencionado, aos anos oitenta. Esta grande área enumera 23 itens, apresentados em ordem alfabética nomeando tanto áreas do conhecimento (por exemplo, Engenharia Têxtil, Engenharia de Agrimensura, Engenharia Mecatrônica) como grandes áreas (por exemplo, Ciências e Ciências Sociais) sub-áreas e especialidades (por exemplo, Química Industrial, Administração Hospitalar) carreiras (por exemplo, Carreira Militar e Carreira Religiosa) e profissões (por exemplo, Secretariado Executivo).

Ainda em relação à grande área 'Outros', cumpre ressaltar que a CAPES, que também faz uso da TAC em sua versão em vigor - com apenas a alteração de não inclusão de áreas que não contavam até 2005 com cursos de mestrado - na classificação dos cursos de pós-graduação *stricto sensu* reconhecidos, acrescenta a grande área 'Outros' na classificação, com duas subdivisões: Ensino e Multidisciplinar. Na área "Multidisciplinar", inclui o curso: "Cognição e Linguagem".

Sintetizando a trajetória da área "Linguística" da TAC no tempo (Tabela 1), observamos as seguintes alterações nos quatro níveis hierárquicos da tabela. Na TAC em vigor, aparece na nomeação da grande área: "Linguística, Letras e Artes". Enquanto áreas, aparecem, nas diferentes versões da TAC, as nomeações: "Linguística e Letras" (TAC 1976), "Letras e Linguística" (TAC 1982) e "Linguística" (TAC 1984). Como sub-área, a nomeação "Linguística" aparece nas versões da TAC 1976 e TAC 1982.

Quanto à nomeação das especialidades da "Linguística", a sistematização a seguir resume a situação até a tabela em vigor. Cumpre lembrar que, em razão das mudanças na estrutura das versões da TAC no tempo, as subdivisões da "Linguística" na versão TAC 1984 correspondem ao nível de sub-áreas e não de especialidades. No entanto, para uma visão de conjunto, estas indicações foram

Dialetologia	TAC 1976		
Distúrbios e Correção da Linguagem	TAC 1976		
Filologia	TAC 1976		
Filosofia da Linguagem (ver Filosofia)		TAC 1982	
Fisiologia da Linguagem			TAC 1984
Línguas Indígenas Brasileiras	TAC 1976		
Linguística Aplicada		TAC 1982	TAC 1984
Linguística Computacional	TAC 1976		
Linguística Histórica		TAC 1982	TAC 1984
Neurolinguística	TAC 1976		
Pragmática	TAC 1982		
Psicolinguística	TAC 1976	TAC 1982	TAC 1984
Sociolinguística e Dialetologia		TAC 1982	TAC 1984
Sociolinguística e Etnolinguística	TAC 1976		
Teoria e Análise Linguísticas		TAC 1982	TAC 1984
Teoria Linguística e Semiologia	TAC 1976		

Tab. 2: Especialidades e Sub-áreas da Linguística no Tempo.

aqui agrupadas para compor o conjunto das subdivisões da “Linguística” nas versões da TAC de 1976, 1982 e 1984.

As especificações acima mostram que, no aspecto evolutivo da representação da Linguística no tempo, apenas a “Psicolinguística” permaneceu como sub-área nas três versões da TAC. Podemos observar também que o número de subdivisões diminuiu no tempo: nove em 1976, sete em 1982 e seis em 1984.

Uma vez que as dificuldades vão se tornando maiores na especificação das sub-áreas e nas especialidades na TAC, como dado complementar, é interessante apresentar como se configuraram essas subdivisões em duas tentativas de propostas de revisão da TAC em vigor até 2004: numa revisão interna no CNPq em 1998 (em base de contribuições dos membros Comitês Assessores da área naquela data) e numa proposta de revisão da Subcomissão de Reclassificação das Áreas do Conhecimento da Comissão Interministerial (CAPES-CNPq) para o Desenvolvimento da Pós-graduação e da Ciência e Tecnologia de 2003. Da comparação entre essas duas propostas, não está prevista a área “Linguística”, mas a área “Estudos de Linguagem” com as seguintes subdivisões comuns apresentadas nas duas propostas: Teoria e Análise Linguística; Teoria e Análise do Discurso; Lexicologia e Lexicografia; Linguística Histórica; Sociolinguística e Dialetologia; Psicolinguística; Linguística Aplicada; Filosofia da Linguagem.

Na tentativa de revisão da TAC em 1998 aparecem também como sub-áreas: Línguas Clássicas e Modernas, Línguas Indígenas e Historiografia Linguística. Na nomeação das sub-áreas da “Linguística” aparece “Historiografia Linguística” na proposta de revisão TAC 1998, e “Linguística Histórica” aparece nas propostas de revisão da TAC de 1998 e 2003, o que alerta para questões de natureza conceitual e de padronização.

No aspecto interdisciplinar é significativo comentar que a sub-área da “Filosofia da Linguagem” na versão TAC 1982 era remetida para a Filosofia, mas passa a ser indicada como sub-área da “Linguística” em propostas de revisão

da TAC de 1998 e de 2003. A sub-área "Lexicologia e Lexicografia" não parece em nenhuma das três versões da TAC, mas aparece nas propostas de revisão de 1998 e de 2003 e é indicada por pesquisadores para a revisão de 2005.

A seguir, considerações sobre a TAC em vigor e propostas de revisão da área da Linguística apresentadas pela comunidade à Comissão de Estudos CNPq-CAPEF-FINEP em 2005.

2.2 A "Linguística" na proposta de revisão da TAC 2005

Declarações da comunidade científica extraídas de trabalho de revisão da TAC pelo CNPq de 1998 e da Comissão CNPq-CAPEF-FINEP de 2005 revelaram problemas da TAC em vigor e apresentaram sugestões para o trabalho de revisão.

A TAC em vigor foi considerada pela comunidade científica como: defasada em relação à realidade da pesquisa e conseqüentemente sua representação na Tabela; ultrapassada em termos de categorização histórica de C&T no país; inadequada, anacrônica e pobre em termos de representação das pesquisas geradas no país; desatualizada e inútil para algumas áreas; deficiente em termos de evolução natural de algumas áreas e da forma como a pesquisa se desenvolveu no país; limitada em número de áreas e sub-áreas para a devida classificação de projetos de pesquisa, apresentando lacunas, imprecisões descritivas e abrangências insuficientes; não apresentando modernização nos tópicos.

Como sugestões para o trabalho de revisão da TAC, a comunidade aponta que a classificação não deve ser muito fragmentária, nem demasiadamente abrangente; que devem ser adotados critérios mínimos para inclusão de áreas; que a proporção de sub-áreas deve ter como princípio a prática da investigação vigente, procurando deter as denominações já consagradas entre os pesquisadores, mesmo que não reflita estritamente os recortes lógicos organizadores das disciplinas científicas; e que a indicação de sub-áreas não deve ser exaustiva. A Comissão de 2005, tanto quanto possível, procurou levar em consideração essas sugestões.

Os estudos da Comissão ressaltaram a relevância da TAC como instrumento para o acompanhamento e a avaliação de políticas públicas voltadas para a ciência, a tecnologia, a cultura, a arte e a inovação. Os estudos da Comissão também ressaltaram a importância da TAC, na preservação de séries estatísticas, em facilitar a coleta de dados das agências, em compatibilizar as informações de diferentes entidades, e em atualizar a terminologia utilizada pelas agências e pela comunidade de pesquisadores e professores.

Como evidenciado na divulgação da proposta preliminar na nova Tabela, os estudos quanto à definição da estrutura da TAC levaram a Comissão a propor a nova TAC em três níveis hierárquicos de Grande Área, Área e Sub-áreas, deslocando as Especialidades para uma listagem alfabética em separado, visando dar maior flexibilidade às temáticas interdisciplinares. Assim, segundo sua natureza, uma mesma especialidade poderia ser especificada em diferentes áreas do conhecimento.

Em relação à definição das grandes áreas, após analisar quatro propostas diferentes de agregação das áreas em nível de superordenação, a Comissão apresentou em sua proposta preliminar a decisão de manter as oito grandes áreas da TAC em vigor, excluindo a grande área nove "Outros". Em relação às áreas, foram mantidas as já existentes e foi proposta a criação de áreas novas. Houve mudanças nas sub-áreas em praticamente todas as áreas da TAC em vigor.

Os estudos da Comissão compreenderam também esforços na definição conceitual de área, grande área e sub-área e especialidade do conhecimento para representação na TAC. Assim, foi considerado: como “Área de conhecimento”, o conjunto de conhecimentos inter-relacionados, coletivamente construído, reunido segundo a natureza do objeto de investigação, com finalidades de ensino, pesquisa e aplicações práticas; como “Grande área”, o aglomerado de diversas áreas do conhecimento em virtude da afinidade de seus objetos, métodos cognitivos e recursos instrumentais refletindo contextos sócio-políticos específicos; como “Sub-área”, a segmentação da área do conhecimento estabelecida em função de estudo e de procedimentos metodológicos reconhecidos e amplamente divulgados; e como “Especialidade”, a caracterização temática da atividade de pesquisa e ensino, podendo estar relacionada a diferentes grandes áreas, áreas e sub-áreas. A seguir, tem-se a apresentação de sugestões recebidas como propostas para a área da “Linguística” para a nova TAC nos diferentes níveis hierárquicos.

Para a nomeação da grande área foram recebidas propostas no sentido de manter “Linguística, Letras e Artes”, como também propostas de mudar, incluindo e excluindo a palavra Linguística, tais como: Linguística, Estudos Literários e Artísticos; Linguística, Estudos Literários e Artes; Linguagens e Artes. Foram também sugeridas nomeações de grande área sem incluir Artes: Linguística e Literatura; Letras e Linguística; Linguagens, Literatura e Interfaces; Ciências da Linguagem e Cognição.

Quanto às especificações de áreas, a Comissão recebeu diferentes sugestões das entidades de classe das áreas de “Linguística” e “Literatura”. Entre as sugestões recebidas, uma apresentou as seguintes oito áreas: Teoria e Descrição Linguística; Teoria e Análise da Linguagem; Literaturas de Língua Portuguesa; Literaturas de Língua Estrangeira; Literaturas e Interfaces: Linguística Aplicada; Linguagens e Interfaces; Fundamentos dos Estudos Literários. Posteriormente, outra sugestão foi apresentada, compreendendo duas áreas, “Linguística” e “Linguística Aplicada”, divididas respectivamente em sub-áreas, como segue: área “Linguística”, sub-áreas: Fonética e Fonologia; Morfologia e Sintaxe; Semântica e Pragmática; Sociolinguística e Linguística Histórica; Psicolinguística e Aquisição da Linguagem; Teoria e Análise do Texto e do Discurso. Área “Linguística Aplicada”, sub-áreas: Ensino de Língua; Práticas Sociais da Linguagem; Tradução; Terminologia.

Interessante observar que no “V Congresso Internacional da ABRALIN”, conforme apresentado no Caderno de Programação, foram definidas as seguintes quatorze categorias para o enquadramento ou classificação dos trabalhos apresentados: Fonética e Fonologia; Morfologia e Sintaxe; Semântica e Pragmática; Sociolinguística e Dialetoлогия; Linguística Histórica; Psicolinguística e Aquisição da Linguagem; Análise do Texto e do Discurso; Alfabetização e Letramento, Línguas Indígenas; Ensino de Língua(s); Práticas Sociais da Linguagem; Tradução; Terminologia; Linguística de Corpus e Computacional.

A comparação entre as sub-áreas propostas para a revisão da TAC em 2005 e as categorias propostas para classificação dos trabalhos do V Congresso da ABRALIN em 2007 revela grande similaridade: das quatorze categorias, sete foram nomeadas exatamente da mesma maneira das duas classificações. Duas sub-áreas apresentaram diferenças de agregação na nomeação de temáticas compostas: sub-área “Sociolinguística e Linguística Histórica” (proposta revisão TAC 2005) e categorias “Sociolinguística e Dialetoлогия” e “Linguística Histórica” (V Congresso ABRALIN 2007). As outras diferenças entre as duas classificações

foram: a exclusão de “Teoria” na categoria “Análise do Texto e do Discurso”; a diferença de grafia entre “Ensino de Língua e Ensino de Língua(s)”; e a inclusão da categoria “Linguística de Corpus e Computacional”. No caso de “Línguas Indígenas Brasileiras”, é relevante indicar que esta categoria constou apenas como sub-área da Linguística na versão da TAC de 1976. Na TAC em vigor, “Línguas Indígenas” aparece como sub-área da área Letras.

As diferenças encontradas da especificação de sub-áreas nas várias versões da TAC no tempo e em propostas de revisão indicam que este nível de especificidade exige reflexão não só do panorama de desenvolvimento da área em si num universo determinado de representação, como das circunstâncias que envolvem os objetivos de construção e uso da classificação.

3 A ‘Linguística’ em outras classificações de áreas do conhecimento

Esta parte do trabalho apresenta a “Linguística” como área do conhecimento em outras tabelas de classificação, visando ilustrar a premissa de que os mesmos saberes podem ser classificados de maneiras diferentes, dependendo das finalidades de construção e do contexto de uso em que estão inseridos. O primeiro exemplo é o da Tabela da SBPC que representa as áreas do conhecimento no contexto da comunicação da pesquisa em desenvolvimento em ciência e tecnologia no país. Como as reuniões da Sociedade são anuais, foi possível seguir a trajetória da área da “Linguística” nos momentos marcantes das mudanças da classificação no tempo. O segundo exemplo mostra a representação da Linguística no contexto cultural de outro país - no caso, a Austrália, numa classificação construída para elaboração de estatísticas nacionais, em duas versões no tempo.

3.1 A área “Linguística” na tabela da SBPC

Do início dos anos cinquenta até meados dos anos setenta, a “Linguística” não era representada como denominação de Seção na Tabela adotada pela SBPC para a classificação das comunicações científicas apresentadas às reuniões anuais da Sociedade. Até o início dos anos setenta, como a classificação não apresentava, entre suas vinte e seis Seções, uma seção específica para a área da “Linguística, Letras e Artes”, as comunicações dessas áreas eram classificadas na Seção Z “Gerais”. Nos anos que representaram marcos de mudança na classificação da SBPC, a “Linguística” apareceu desta forma: em 1976, como “Linguística e Literatura” na grande área Ciências do Homem; no ano de 1984, como “Linguística e Semiótica” - “Letras e Literatura” formavam outra seção da classificação. No ano de 2001, “Letras, Linguística e Literatura” se apresentaram como áreas em separado na grande área Ciências Humanas. De 2003 a 2006, quando a tabela da SBPC adota a TAC do CNPq com poucas alterações, a “Linguística” apareceu na nomeação da grande área “Artes, Letras e Linguística” (em relação à TAC, alteração apenas na seqüência de ordenação em função da adoção da ordem alfabética - na TAC, a nomeação desta grande área é “Linguística, Letras e Artes”). Nessa grande área, a “Linguística” está incluída, como seção ou sub-área, com “Artes, Letras, Literatura e Semiótica”.

1972	
Seção Z	Gerais
1976	
Seção 5	Ciências do Homem
Seção 5.4	Lingüística e Literatura
1984	
Seção B	Ciências do Homem
Seção B.9	Lingüística e Semiótica
Seção B.10	Letras e Literatura
2001	
Seção A	Ciências Humanas
Seção A.15	Letras
Seção A.16	Lingüística
Seção A.17	Literatura
2003 - 2006	
Seção H	Artes, Letras e Lingüística
Seção H.1.0	Artes
Seção H.2.0	Letras
Seção H.3.0	Literatura
Seção H.4.0	Lingüística
Seção H.5.0	Semiótica

Tab. 3: SBPC e Área “Lingüística” no tempo

A trajetória da Lingüística na Tabela da SBPC (Tabela 3) revelou que sua representação nessa classificação partiu da denominação “Gerais” nos anos cinquenta, seguiu como área em grande área nas décadas de setenta, oitenta e noventa, e chega no início do século XXI como denominação de grande área, ocupando espaço próprio como área de conhecimento. Esta trajetória de alguma forma demonstra uma evolução significativa da “Lingüística”, enquanto posicionamento numa tabela de classificação de áreas do conhecimento que representa o desenvolvimento de ciência e tecnologia no país.

3.2 A área “Lingüística” na classificação da ASRC

A Australian Standard Research Classification (ASRC), apresenta um conjunto de classificações para projetos de pesquisa e desenvolvimento com o objetivo de geração de estatísticas, é aqui apresentada como uma visão a mais da “Lingüística” em tabelas de áreas do conhecimento. A seguir, apresentam-se comentários sobre a representação da área “Lingüística” na classificação de Campos de Pesquisa (FOR - Fields Of Research) na edição de 1993 que passa a se denominar classificação de Campos de Pesquisa, Cursos e Disciplinas (RFCD - Research Fields, Courses and Disciplines) na edição de 1998.

Na divisão das Ciências Sociais e Humanidades, da Classificação de Campos de Pesquisa de 1993, o grupo das Humanidades inclui a área “Linguagem e Literatura” com as subdivisões: Linguagem, Lingüística, Literatura e Estudos Literários. Como sub-áreas da “Lingüística”, encontramos “Lexicografia; Lingüística Aplicada; Semântica” e mais duas subdivisões assim especificadas: Lingüística: Estudos de Línguas Aborígenes Australianas e Lingüística: outros estudos australianos.

Na versão da ASRC de 1998, a classificação de Campos de Pesquisa, Cursos e Disciplinas apresenta um re-arranjo de áreas e sub-áreas em relação à versão de 1993. Na versão de 1998, a "Linguística" forma com a "Psicologia" e a "Ciência Cognitiva" as três disciplinas da divisão "Ciências Comportamentais e Cognitivas". As sub-áreas da "Linguística" são especificadas desta forma: Linguística Aplicada e Linguística Educacional; Linguística Computacional; Discurso e Pragmática, Fônica Laboratorial e Ciência do Discurso; Linguagem na Cultura e na Sociedade (Sociolinguística); Linguagem no Tempo e no Espaço (incluindo Linguística Histórica, Dialetologia); Estruturas Linguísticas (incluindo Gramática, Fonologia, Léxico, Semântica); Lexicografia.

Na parte da definição das classes da ASRC de 1998, na área de "Linguagem e Cultura", em que estão especificadas como sub-áreas "Estudos de Língua, Estudos de Literatura e Estudos Culturais", há a indicação de que a "Linguística" está incluída como área com a "Psicologia" e com a "Ciência Cognitiva" na grande área das "Ciências Comportamentais e Cognitivas". No entanto, "Processos Linguísticos" (incluindo Produção e Análise do Discurso) aparece como uma das sub-áreas entre as demais que compõem a área Ciência Cognitiva, que são: Filosofia da Cognição; Percepção por Computador, Memória e Atenção; Padrões Neurocognitivos e Redes Neurais; Representação do Conhecimento e Aprendizagem por Máquina; Planejamento e Solução de Problema.

Como podemos observar, a "Linguística" aparece como área em agregados conceituais diferentes na classificação da ASRC. Na edição de 1993, a "Linguística" aparece como área de "Linguagem e Literatura" no conjunto das "Humanidades". Na edição de 1998, a "Linguística" aparece como área de "Linguagem e Cultura" no conjunto das "Ciências Comportamentais e Cognitivas". Este fato merece destaque, uma vez que revela diferentes pontos de vista de inserção da "Linguística" em nível hierárquico mais alto de agregação numa mesma classificação de áreas do conhecimento no tempo.

4 Considerações Finais

A Tabela de Áreas do Conhecimento (TAC) se caracteriza como uma linguagem de organização do universo de Ciência e Tecnologia, construída para atender as necessidades de ordem prática das agências de fomento na administração de sua política científica. Nesta difícil tarefa, as agências buscam compatibilizar as necessidades da comunidade científica em representar singularidades e especificidades temáticas de desenvolvimento dos diversos campos do saber.

O fato de que as mesmas áreas de conhecimento podem ser ordenadas de maneiras diferentes em função das características históricas, culturais, educacionais, econômicas e sociais do desenvolvimento das atividades de pesquisa, ensino e extensão, indica a complexidade da construção ou atualização da TAC.

O esforço de instituições nacionais de vários países e de organismos internacionais devotados ao desenvolvimento de C&T tem sido buscar a maneira menos arbitrária possível de representar as áreas do conhecimento em tabelas e esquemas de classificação.

No caso brasileiro a expectativa em relação à TAC é de que tal classificação, tanto quanto possível, reflita referências habituais dos pesquisadores a tipos específicos de problemas, seja uma Tabela consensualmente aceita pela comunidade como rigorosa, abrangente e suficiente, isto é, sem lacunas; seja pensada

em termos estratégicos de futuro: ou apresentando compatibilidade com classificações internacionais, resguardadas as características nacionais.

Uma questão que cada vez mais tem sido discutida diz respeito aos aspectos multi-interdisciplinares do desenvolvimento de C&T. Esse aspecto acarreta questões não só de natureza conceitual, como de ordenação e representação das áreas, sub-áreas, mas principalmente das especialidades nos esquemas e tabelas de classificação do conhecimento. Por isso, as áreas interdisciplinares merecem atenção especial.

Duas outras características importantes a serem consideradas quanto a instrumentos de organização e representação das áreas do conhecimento dizem respeito à consistência das partes e ao equilíbrio do todo, buscando, tanto quanto possível, uma harmonia na classificação.

Para finalizar, as palavras de Edgar Morin, em sua obra *A Religião dos Saberes: o desafio do Século XXI* (2001), ilustram a singularidade, a complexidade e a necessidade da organização do conhecimento:

somos filhos do cosmos / trazemos em nós o mundo físico / trazemos em nós o mundo biológico / mas com e em nossa singularidade própria / em outras palavras: para enfrentarmos o desafio da complexidade, / precisamos de princípios organizadores do conhecimento

Morin evidencia o importante papel da organização do conhecimento em diferentes contextos de informação como uma arte, especificando atividades de natureza científica da classificação, tais como religar e diferenciar, contextualizar e globalizar. Morin ressalta ainda o fato de que informação e conhecimento são conceitos sempre relacionados a contexto e conjunto. Essas considerações evidenciam que a organização do conhecimento é um processo de abstrair, sistematizar e sintetizar.

Para concluir, podemos dizer que a construção de uma tabela de áreas do conhecimento para fins de administração e avaliação em ciência e tecnologia é uma atividade caracterizada como TÉCNICA, ARTE e CIÊNCIA. Portanto, uma atividade de natureza complexa e difícil. Porém fascinante!

Referências

- [1] BAILEY, K. D. *Typologies and Taxonomies: an introduction to classification techniques*. SAGE Publications, London, 1994.
- [2] BOWKER, G. C., AND STAR, S. L. *Sorting things out: classification and its consequences*. The MIT Press, Massachusetts, MA, 2000.
- [3] CARAÇA, J. *Do saber ao fazer: porque organizar a ciência*. Gradiva, Lisboa, 2003.
- [4] CASTLES, I. Australian standard research classification: a set of classifications for r&d projects fields of research. Tech. rep., Australian Bureau of Statistics, Canberra, 1993.
- [5] CNPQ. Tabelas de Áreas do conhecimento, 1976, 1982, 1984.

- [6] DARNTON, R. *O grande massacre de gatos e outros episódios da história cultural francesa*. Graal, Rio de Janeiro, 1986, ch. Os filósofos podam a árvore do conhecimento: a estratégia epistemológica da Encyclopédie, pp. 247-276.
- [7] DE SOUZA, R. F. Áreas do conhecimento. datagramazero. *Revista de Ciência da Informação* 5, 2 (Abril 2004), 1-28. Disponível em http://www.dgzero.org/abr04/Art_02.htm. Acesso em fevereiro de 2007.
- [8] DE SOUZA, R. F. Organização e representação de áreas do conhecimento em ciência e tecnologia: princípios de agregação em grandes áreas segundo diferentes contextos de produção e uso de informação. *Enc. Bibli. Bibliotconomia e Ciência da Informação* (2006), 27-21. Disponível em: http://www.Encontros-bibli.ufsc.br/bibesp/esp_03/3_GT2_souza.pdf. Acesso em fevereiro 2007.
- [9] MORIN, E. *A religião dos saberes: o desafio do século XXI*. Bertrand Brasil, Rio de Janeiro, 2001. ch. Jornadas temáticas idealizadas e dirigidas por Edgar Morin.
- [10] MORIN, E. *A cabeça bem feita: repensar a reforma, reformar o pensamento*. Bertrand Brasil, Rio de Janeiro, 2003.
- [11] SBPC. Reuniões anuais. tabelas de classificação das comunicações de 1956 a 2006.
- [12] SKINNER, T. J. Australian standard research classification: a set of classifications for r&d projects - research fields, courses and disciplines. Tech. rep., Australian Bureau of Statistics, Canberra, 1998.

Por uma descrição gramatical mais concreta: as funções sintáticas

Mário A. Perini

Universidade Federal de Minas Gerais, PUC-Minas

marioperini@hotmail.com

Abstract

Culicover and Jackendoff (2005, p. 5), giving expression to an idea which has been circulating among linguists for over a decade, proposed the Simple Syntax Hypothesis (SSH): The most explanatory syntactic theory is one that imputes the minimum structure necessary to mediate between phonology and meaning. In this paper I apply the SSH to some abstract relations common in grammar, namely, so-called syntactic functions, in particular the function of subject. I show that the subject can be reduced to concrete relations which are reducible to phonetic facts that must appear in any analysis of Portuguese.

1 Introdução: faltam dados

Para começar, gostaria de situar este trabalho dentro de uma concepção particular das tarefas atuais da pesquisa lingüística, concepção esta que norteia meu trabalho do momento (e que não é somente minha, mas que também não é de aceitação universal). Resumo essa posição em cinco pontos, a saber:

- (a) A preocupação com a formalização e a elaboração de teorias altamente especificadas deve ceder algum lugar a uma necessidade mais urgente, a de definir e esclarecer questões fundamentais de análise. Dentre essas questões, posso mencionar:
 - os princípios de taxonomia das formas lingüísticas;
 - a definição operacional dos papéis temáticos;
 - a questão da diferenciação entre adjuntos e complementos;
 - a explicitação dos fatores que governam a ordem das palavras;
 - a relação da ordem das palavras com as chamadas funções sintáticas.
- (b) Os estudos lingüísticos sofrem de falta de evidência para fundamentar as teorias. Ou seja, faltam dados cuidadosamente colhidos, sistematizados e descritos, cobrindo áreas significativamente grandes de línguas naturais particulares.
- (c) Para os objetivos acima discriminados, grande parte da estrutura da linguagem pode ser descrita em termos de estrutura formal aparente, representações semânticas acessíveis à introspecção e relações entre esses dois níveis: é preciso minimizar a postulação de níveis intermediários de análise, tais como os utilizados pela teoria gerativa clássica.

- (d) A lingüística se encontra atualmente em um estágio que pode ser chamado de “história natural” – ou seja, estamos mais ou menos na situação da física antes do século XVII, quando ainda não havia um paradigma que norteasse a interpretação dos fatos observados. Assim como os físicos daquela época, os lingüistas de hoje têm como tarefa principal levantar dados, sistematizá-los e encaixá-los, quando possível, em teorias parciais – em uma palavra, elaborar descrições (nossa tarefa é mais próxima da tarefa de Tycho Brahe e de Lineu do que da de Einstein e Darwin). A lingüística atual é dominada pela ilusão de que é possível queimar etapas e desenvolver teorias válidas sem um longo e laborioso trabalho prévio de levantamento e sistematização de dados. Oponho-me a essa atitude e não sou, nesse particular, nem pioneiro, nem original: muitos lingüistas estão e estiveram conscientes do problema, mas a meu ver não tiveram a repercussão merecida.
- (e) É urgente elaborar estudos amplos de grandes fatias da estrutura das línguas. Esses estudos devem se concentrar em línguas particulares, sem ceder à tentação de aplicar a uma língua a análise de outra. A comparação entre as línguas, assim como a procura de universais são empreendimentos válidos, mas só podem ser realizados em um segundo momento. Não devemos comparar pequenas áreas da estrutura em diversas línguas, mas fazer confrontos abrangentes de grandes porções das estruturas. E isso, a meu ver, não pode ser ainda realizado por falta de descrições confiáveis de um número suficiente de línguas.

Em virtude dessas convicções, não estou buscando um modelo de análise lingüística, muito menos uma teoria lingüística, mas apenas um modelo de descrição das valências verbais do português brasileiro – um modelo parcial, em harmonia com a idéia de que a lingüística não está preparada para um paradigma global. O modelo descritivo deve servir para orientar a coleta, sistematização, apresentação e análise preliminar de dados em quantidade apreciável.¹

Esse modelo descritivo não deve ser entendido como uma teoria da linguagem, nem mesmo da gramática portuguesa. Cheguei a ele através de um esforço no sentido de elaborar um catálogo das estruturas do português. O catálogo será um instrumento de verificação da adequação empírica de análises e teorias – em última análise, um repositório sistematizado de dados apresentados de forma esquemática. A discussão das funções sintáticas é parte da procura de uma linguagem na qual exprimir o catálogo. E a parte do catálogo a que pretendo chegar em minha pesquisa é a lista das diáteses verbais do português.

O modelo sacrifica, conscientemente, a elegância teórica em favor da adequação descritiva. No decorrer do meu trabalho de pesquisa, fui às vezes obrigado a descrever como um conjunto heterogêneo de fenômenos alguma coisa que é tradicionalmente considerada um fenômeno único – por exemplo, a ordem dos termos dentro do sintagma nominal. Talvez se possa dizer que com isso a análise ficou menos elegante. Mas eu gostaria de formular certas perguntas quanto à própria noção de “elegância” ou “simplicidade” de uma análise: será que não se trata, às vezes, de uma consequência da nossa ignorância? Como sabemos, afinal, que um fenômeno é único e não um conjunto de fatores? Aqui temos pouco

¹Gross (1975) e Levin (1993) também propõem modelos parciais, aos quais faço restrições por razões explicitadas no meu livro *Estudos de gramática descritiva* (Perini, inédito), que deve sair em breve.

mais que a simples tradição a nos orientar: e acho perigoso seguir a tradição e chegar a conclusões globais sobre hipóteses que versam sobre um objeto que só conhecemos de maneira tão fragmentária, como é o caso das línguas naturais. O critério de simplicidade é obviamente válido, mas não se aplica a fragmentos, se aplica a visões e interpretações muito mais abrangentes do que as que estão disponíveis no atual estágio do conhecimento em lingüística.

2 Hipótese da Sintaxe Simples

Culicover e Jackendoff (2005), dando expressão a uma idéia que circula há mais de uma década entre os lingüistas, formularam a Hipótese da Sintaxe Simples (*Simpler Syntax Hypothesis*)²: “A teoria sintática mais explicativa é a que atribui o mínimo de estrutura necessário para exprimir a mediação entre fonologia e significado.” (Culicover; Jackendoff, 2005, p. 5).

A motivação para essa hipótese vem da seguinte observação: o lingüista, em seu trabalho descritivo, lida com duas categorias de fatos, os fonéticos e os semânticos. Para dar um exemplo simples, a forma fonética que grafamos como *cheguei* tem duas sílabas, começa com uma fricativa palatal etc.; e se refere ao final de um deslocamento (real ou metafórico), que afeta o falante em um momento do passado. Isso são fatos, e nenhuma análise da língua pode passar sem eles: uma análise do português que deixe de exprimi-los é incorreta ou pelo menos incompleta.

Por outro lado, há outras afirmações que geralmente se fazem sobre a palavra *cheguei*: trata-se de um verbo, é da segunda conjugação, é regular, não tem objeto direto etc. Só que essas afirmações não se referem a fatos, mas antes a hipóteses: pode-se conceber uma análise do português que classifique *cheguei* junto com as palavras que chamamos de “adjetivos” – isso já foi feito, dentro da semântica gerativa dos anos 60:³ pode-se imaginar uma outra organização dos verbos que não a em conjugações, e assim por diante.

O que mostrei acima é uma distinção entre, de um lado, coisas que o lingüista precisa inevitavelmente incluir em sua análise do português e, de outro lado, coisas que dependem da teoria que ele adota, das conveniências descritivas que ele reconhece. É para dirigir o lingüista em suas escolhas (de hipóteses, não de fatos) que se elaboram as teorias. A hipótese da sintaxe simples faz isso minimizando o conteúdo hipotético das análises – não negando a necessidade de hipóteses, mas, digamos, colocando-as em seu devido lugar, que é o menor possível. Com isso, ela pretende evitar duplicações: descrições sintáticas ou morfológicas que simplesmente replicam informações disponíveis na fonologia e na semântica.

Não é essa a posição geralmente aceita em lingüística. Andrews (1985), por exemplo, faz uma afirmação no sentido oposto:

²Essa hipótese já tinha sido antecipada por Jackendoff (1990), e foi proposta independentemente por Langacker (1987, p. 26), Perini et al. (1996) e provavelmente outros. Aliás, talvez devêssemos ser mais arrojados e falar de “fonética”, não de “fonologia”, como fazem Culicover e Jackendoff.

³Ver, por exemplo, Givón (1970).

As funções semióticas [isto é, semânticas + pragmáticas / MAP] são sinalizadas por 'traços explícitos de codificação' tais como a ordem das palavras, as marcas de caso e a referência cruzada (concordância). Mas é difícil dar uma imagem coerente de como isso se dá em termos de uma conexão direta entre os traços de codificação e as funções semióticas que eles expressam. Antes, parece melhor postular um nível intermediário de estrutura gramatical: os traços de codificação indicam a estrutura gramatical da sentença, e a estrutura gramatical determina as funções semióticas. (Andrews, 1985, p. 63).

É justamente isso que questiono, e que me parece que Culicover e Jackendoff também questionam. A dificuldade mencionada por Andrews pode ou não ser real, e pode ou não ser limitada a certos pontos da análise. Como não se tentou seriamente representar a relação entre as funções semióticas de maneira direta em termos de traços explícitos de codificação, a afirmação de que isso é particularmente difícil é gratuita. Na verdade, minha pesquisa em gramática portuguesa até o momento sugere justamente o oposto: tanto no que se refere à descrição da estrutura interna do sintagma nominal, relatada em Perini et al. (1996), quanto no que se refere às funções sintáticas na oração, o papel do nível intermediário (composto de relações abstratas) parece mais restrito do que sugere a tradição.

Aqui pretendo examinar algumas conseqüências da hipótese da sintaxe simples para a definição das funções sintáticas. Tentarei mostrar que as funções sintáticas são em grande parte dispensáveis, e que o sistema de codificação explícita das funções semânticas é mais simples do que usualmente se acredita.

3 Para que funções sintáticas?

Não é certamente a primeira vez que alguém exprime insatisfação com a maneira pela qual as funções sintáticas são definidas ou, pior, são aceitas sem discussão. Acho que a passagem seguinte revela uma preocupação semelhante:

[...] tem havido uma conscientização cada vez maior [...] de que o reconhecimento das relações gramaticais depende inicialmente de certas regularidades observáveis. Segundo esse ponto de vista, os objetos e outras relações gramaticais não deveriam ser considerados dados observacionais primários, presentes de maneira inevitável e invariável em todas as línguas humanas. Em vez, deve-se esperar que se manifestem, na melhor das hipóteses, secundariamente (ou que nem se manifestem, segundo o caso) através de regularidades que envolvem os dados observacionais primários. (Plank, 1984, p. 1-2).

Não obstante essas observações de 23 anos atrás, não tenho notado um desenvolvimento no sentido sugerido por Plank, pelo menos não na proporção que seria de esperar. Mesmo nos artigos reunidos no volume organizado pelo próprio Plank, a função de "objeto" é tratada de maneira informal e às vezes nebulosa, e nunca diretamente questionada. O problema apontado por ele, portanto, continua de pé.

Em uma língua como o português, a motivação para se postular as funções sintáticas surge da necessidade de descrever os seguintes fenômenos:⁴

- (a) a atribuição de papéis temáticos aos diversos SNs da oração;
- (b) a concordância verbal;
- (c) a ordem dos termos na oração; e
- (d) a distribuição de itens como eu, em oposição a me.⁵

Se for possível dar conta desses fenômenos sem necessidade de postular relações abstratas entre o verbo e cada um dos sintagmas da oração ("sujeito", "objeto direto", "predicativo" etc.), essa solução deve ser preferida, em virtude da hipótese da sintaxe simples. Ou seja, a função sintática, em si, não tem utilidade nenhuma: falamos delas (se necessário) para podermos descrever os fenômenos acima que, eles sim, são necessários em si mesmos. A interpretação de *o cachorro em*

- (1) O cachorro me lambeu.

como Agente, a correlação desse sintagma com a terminação do verbo (- eu, não -i ou -emos), sua posição antes do verbo e a interpretação da forma *me* como Paciente são fatos da língua que precisam figurar na descrição. Se pudermos descrevê-los sem lançar mão de noções abstratas, tanto melhor. Estou explorando um meio de fazer exatamente isso. No momento, tendo a achar que a função de **sujeito** é a única que merece uma definição especial, que como veremos também não é propriamente abstrata: quanto às demais funções, tudo indica que podem ser totalmente reduzidas a fatores, tais como a ordem dos sintagmas na oração e a semântica das preposições.

Com o risco de ser redundante, vou dizer a mesma coisa de outra maneira: as funções sintáticas, em si, não estão presentes nem no sinal sonoro nem na representação semântica. Elas representam, na melhor das hipóteses, um elemento conector do signo; na pior das hipóteses, uma complicação descritivamente inútil. É importante, portanto, investigar até que ponto as funções sintáticas são realmente necessárias; e eu suspeito que muitas delas são dispensáveis para efeitos de descrição gramatical.

4 Funções sintáticas e descrição

O objetivo da descrição lingüística é, em última análise, relacionar interpretações semânticas com traços concretos do sinal lingüístico⁶, da maneira mais imediata possível. Aqui estamos encarando um problema básico da análise lingüística: descrever o caminho que o receptor precisa seguir para processar os enunciados que recebe. Ele parte de um sinal material (acústico ou visual) e processa esse

⁴ "Dados observacionais primários", no dizer de Plank.

⁵ Os casos em português se limitam a poucos itens, todos classificados tradicionalmente como "pronomes".

⁶ Deixando de lado os traços do contexto extralingüístico, que não são relevantes para o problema em pauta.

senal com base em seu conhecimento da língua: e é necessário relacionar eventualmente esse processo com o resultado final, que é a interpretação semântica. Mas as funções sintáticas (tal como se entendem tradicionalmente) não fazem parte do sinal material, e portanto não podem ser consideradas um ponto de partida.⁷

É claro que uma análise gramatical não se identifica com uma descrição dos processos cognitivos postos em ação pelo usuário da língua; mas sou de opinião de que precisa ser compatível com esses processos, e isso da maneira menos complicada possível. Ou seja: embora ninguém se iluda quanto à tremenda complexidade do fenômeno, qualquer elemento descritivo que contribua para essa complexidade precisa ser justificado.

5 Para que sujeito?

Vamos considerar em primeiro lugar a função de **sujeito**. Esta se define de diversas maneiras, segundo a teoria: para a teoria léxico-funcionalista e para a gramática relacional, trata-se de uma marca primitiva associada a certos constituintes da oração; para a teoria gerativa, seria uma função derivada da configuração estrutural da sentença, ou seja, o argumento "externo" do predicador. Aqui vamos deixar de lado essas definições teóricas para perguntar quais são as características do sujeito representadas materialmente na cadeia sonora ou presentes na interpretação semântica da oração. A idéia é que, qualquer que seja a definição, o sujeito deverá corresponder ao sintagma que tiver essas características.

Antes de mais nada, para que é que o receptor precisa de um sujeito?⁸ A tarefa do receptor pode ser resumida em atribuir uma interpretação semântica a uma cadeia sonora (ou gráfica). Há, é claro, muitas complicações, mas vou me limitar aos aspectos tradicionalmente ligados às funções sintáticas, a saber, a atribuição dos papéis temáticos, a ordem dos termos, a concordância e a distribuição dos poucos itens que mostram alteração de caso morfológico, como *eu, você x me, te*. Vou começar com o problema da atribuição de papel temático.

Digamos que o receptor tem que processar a frase

(2) O cachorro fugiu da cobra.

Ele vai ter que atribuir o papel de Agente a *o cachorro* e de Fonte (a origem de um movimento)⁹ a *da cobra*. Para isso ele dispõe, entre outras coisas, da própria cadeia sonora, mais seu conhecimento prévio da gramática e do léxico da língua, além do conhecimento geral do mundo. Vou deixar de lado este último fator, no que pese sua importância, porque não se aplica a nossos exemplos. Vamos ficar, portanto, com a realização material da frase (o enunciado) mais o conhecimento lingüístico do usuário.

⁷As funções podem ser entendidas de maneira mais concreta: assim, a meu ver, a posição na cadeia sonora é uma função de pleno direito. Nesse caso, a função faz parte do sinal material.

⁸Vou me referir sempre ao ponto de vista do receptor. Naturalmente, este se relaciona com o do emissor, *mutatis mutandis*, mas para os nossos objetivos, não é necessário discutir as diferenças.

⁹Ou Causa, se alguém preferir. Não faz diferença para nossa discussão.

No caso de (2), tudo é bastante óbvio. Reconhecido o verbo *fugir*, sua valência nos fornece boa parte das informações necessárias: esse verbo se associa, em uma de suas diáteses, a um SN Agente e a um sintagma preposicionado com *de* no papel de Fonte.

É importante notar que essa atribuição depende crucialmente da valência do verbo, porque se tivéssemos a frase

(3) O cachorro apanhou da cobra.

apesar da identidade sintática, os papéis temáticos seriam diferentes: *o cachorro* é Paciente e *da cobra* é Agente.

Por enquanto, não precisamos de funções sintáticas, porque a categoria de cada complemento nos permite atribuir corretamente o papel temático: o único SN presente é Agente em (2), Paciente em (3) e o único sintagma com *de* é Fonte em (2) e Agente em (3).

Tomando o caso do sintagma preposicionado, observe-se que não há necessidade de postular uma "função" para que o processo interpretativo chegue a bom êxito. Dados o verbo e a preposição, o papel temático pode ser atribuído sem problemas. Essa é a situação sempre que o complemento é regido de preposição. Nos raros casos em que há mais de um sintagma preposicionado e a preposição é a mesma, ou surge ambigüidade (o que mostra que não há função sintática resolvendo o problema), ou, mais freqüentemente, a ambigüidade é evitada através do uso de informação pragmática:

(4) Eu pensei em você em Curitiba.

Na pior das hipóteses, podemos reduzir a "função" a um fator de ordem. Assim, a frase seguinte não é ambígua:

(5) Eu pensei na fazenda em Curitiba.

Aqui o Local é sempre o segundo sintagma, de modo que a frase só pode significar "eu pensei a respeito da fazenda quando estava em Curitiba". Novamente, não há necessidade de distinguir os dois sintagmas usando funções no sentido tradicional.¹⁰ Ou seja, para efeitos de atribuição de papéis temáticos, não é necessário distinguir os dois sintagmas preposicionados como, digamos, "objeto indireto" e "adjunto adverbial de lugar" ou coisa que o valha. A valência de pensar e a semântica da preposição *em* (com suas muitas possibilidades) são suficientes para que a atribuição se faça corretamente.

Em frases como (5), o receptor dispõe da cadeia sonora (que inclui dois sintagmas introduzidos por *em*) e de seu conhecimento da preposição e da valência do verbo, mas não dispõe das funções, pois estas não têm representação material no enunciado - e, sendo relações sintagmáticas, não podem ser armazenadas previamente. Conclui-se que não estão ali e, se forem necessárias, terão que ser introduzidas pelo próprio usuário da língua. Mas, como vimos, pelo menos nesses casos elas não são necessárias.

¹⁰O mesmo se aplica se um dos SPreps for topicalizado: *em Curitiba, eu pensei em São Paulo* só pode ser entendida com *em Curitiba* no papel de Local. Na verdade, acho que aqui há outros fatores em jogo, talvez ligados à valência verbal. Assim, *eu pensei em Curitiba em você* é de aceitabilidade marginal, mas se mudarmos o verbo a inversão não afeta a aceitabilidade: *ele foi arrastado por 200 metros pelo carro e ele foi arrastado pelo carro por 200 metros* são ambas plenamente aceitáveis.

6 Identificando o sujeito

Nossa tarefa, daqui por diante, é examinar caso por caso, para ver até onde vai a possibilidade de atribuir papéis temáticos sem utilizar funções sintáticas. É uma tarefa longa e que mal começou: mas posso adiantar que o único caso que examinei e que depende de algo mais do que da ordem ou da semântica da preposição é o sujeito. Em outras palavras, o sujeito é a única função sintática que depende de definição específica. Mesmo assim, a definição terá que ser revista, conforme veremos.

Vamos considerar o caso do sintagma nominal. Sabemos que pode haver mais de um por oração, e não há preposição para diferenciá-los quanto ao papel de cada um. Por exemplo,

- (6) O cachorro matou a cobra.

A valência de *matar* nos informa que uma das diáteses compreende um Agente e um Paciente, ambos representados por SN.¹¹ No caso de (6), podemos lançar mão da ordenação e incluir na valência que **o SN pré-verbal é Agente, e o outro SN é Paciente**. Isso funciona para (6) e também para

- (7) A cobra, o cachorro matou.

Até o momento não temos grandes problemas: só precisamos da ordem dos sintagmas, que está concretamente representada no sinal sonoro, para atribuir os papéis temáticos de maneira correta. Mas agora temos pares como os seguintes:

- (8) Eu matei a cobra.

- (9) Matei a cobra.

Como atribuir o papel de Agente em (9), onde não há SN pré-verbal? É claro que o sufixo de pessoa-número do verbo *matar* indica um argumento que recebe o papel de Agente. Esse mesmo argumento pode ser sinalizado redundantemente por um SN, como em (8). Se chamarmos *eu* de "sujeito" de (8), poderemos dizer que

o sujeito é o SN que recebe o mesmo papel temático indicado pelo sufixo de pessoa-número do verbo.¹²

Para ser mais preciso, devo dizer que o sujeito **elabora o referente do receptor do mesmo papel temático indicado pelo sufixo**, porque em muitos casos o sujeito acrescenta informação que não está presente no sufixo. É o caso de

- (10) O cachorro matou a cobra.

¹¹É preciso estipular a representação formal porque, como se viu em (3), o Agente pode ser representado por sintagma preposicionado.

¹²Essa definição vale para o dialeto estudado, que é o português falado do Brasil.

em que a informação de que se trata do cachorro não consta do sufixo.

Assim, na valência do verbo *matar* tem que ficar explicitado que o sufixo de pessoa-número indica o Agente. E sabemos que o papel temático indicado pelo sufixo pode ser atribuído a um SN (que em geral elabora a indicação referencial feita pelo sufixo).

É pois característica do verbo português que um dos papéis temáticos a ele associados é sinalizado pelo sufixo de pessoa-número. Esse papel temático privilegiado pode ser também elaborado através de um SN, que se coloca em determinada posição na oração. Nenhum outro papel temático é sinalizado redundantemente dessa maneira.¹³

A indicação feita pelo sufixo varia de verbo para verbo; assim, em

(11) Apanhei da minha mulher

é Paciente, e em

(12) Ouvi um estrondo.

é Experienciador. Isso quer dizer que na valência de cada verbo é preciso especificar o papel temático que se associa ao sufixo de pessoa-número. Por exemplo, uma das diáteses de *matar* é

(13) *matar* : SufixoPN (Agente), SN (Paciente)

e uma das diáteses de *apanhar* é

(14) *apanhar* : SufixoPN (Paciente), de+SN (Agente)

A indicação também depende da construção (isto é, da diátese verbal), porque com o verbo *engordar* o sujeito é Agente em

(15) Essa comida vai engordar o Ronaldo.

mas é Paciente em

(16) O Ronaldo engordou.

Uma regra geral nos diz que pode haver um SN com o mesmo papel temático do sufixo de pessoa-número. Como as referências sempre se superpõem, não surge a situação de haver o mesmo papel temático atribuído a duas entidades distintas.¹⁴

Mas agora surge um problema: qual dos sintagmas presentes na oração é o que vai receber o papel temático indicado pelo sufixo de pessoa-número? Se conseguirmos caracterizar esse sintagma em termos de sua posição, poderemos chegar a uma definição bastante concreta de sujeito, sem lançar mão de vínculos não representados foneticamente. Acredito que, para o português brasileiro, é

¹³E a redundância apontada muitas vezes não se verifica, seja porque não há SN que elabore o papel temático indicado pelo sufixo (*cheguei ontem*), seja porque não há sufixo de pessoa-número (o que ocorre em orações de gerúndio e de infinitivo impessoal).

¹⁴Essa possibilidade é excluída por um mecanismo especial, equivalente à metade do Critério Teta dos gerativistas (porque fica de pé a possibilidade de um só constituinte receber mais de um papel temático).

possível realizar essa caracterização, e que portanto, até onde podemos ver, não há necessidade (para essa língua) do sujeito enquanto função sintática abstrata.

É o que mostro em meu texto *Estudos de gramática descritiva* (Perini, inédito), e que aqui só posso esboçar. Por ora, bastará dizer que a identificação do sujeito é clara e unívoca em todas as frases: ou só existe um SN na oração, ou, quando há mais de um, o sujeito é claramente identificado a partir de sua posição (tipicamente, pré-verbal). Frases em que a diferença entre sujeito e objeto depende de informação pragmática não ocorrem no português falado atual, embora existam na língua escrita clássica.¹⁵

Um caso que merece discussão é o de

(17) Matei a cobra.

em que só há um SN, que no entanto não é o sujeito, porque não recebe o papel temático de Agente reservado ao sujeito de *matar*. Como evitar que *a cobra* seja rotulado de “sujeito” nesses casos?

Creio que o que bloqueia a possibilidade é que *a cobra* não pode elaborar a indicação do sufixo de pessoa-número por incompatibilidade de pessoa. O sufixo indica a primeira pessoa do singular (não o sujeito, mas, no caso, o Agente), ao passo que *a cobra* não pode incluir essa referência. Não posso desenvolver plenamente esse ponto dentro dos limites desta discussão. Vou apenas dizer que é possível dar conta de todos os casos de maneira concreta, jogando com a posição do sintagma, com suas possibilidades referenciais e alguns outros fatores que me parecem razoavelmente claros.

É importante observar que esse sistema nos fornece uma interpretação concreta do fenômeno da concordância verbal: ela se reduz agora a uma exigência de compatibilidade referencial entre o sujeito (identificado da maneira descrita) e o sufixo de pessoa-número. Quero dizer, a inaceitabilidade de frases como

(18) * O cachorro matei a cobra.

surge da indicação contraditória do Agente: “eu” segundo o sufixo, “o cachorro” (ou seja, “ele”) segundo o SN pré-verbal.

7 Definição de sujeito

O sujeito é, portanto,

um SN de pessoa e número compatíveis com a pessoa e número indicados pelo sufixo de pessoa-número do verbo.

Ele ocorre em determinadas posições na oração, e esse posicionamento faz parte do conhecimento gramatical dos usuários da língua.

O processo de atribuição do papel temático a um sintagma em uma oração (seja ele considerado “complemento” ou “adjunto”) se baseia apenas na ordem em que ele ocorre (relativamente ao verbo), mais a semântica da preposição que

¹⁵Um exemplo, tirado de Varnhagen (1854): *essa exposição ou relatório publicou Cândido Mendes de Almeida*.

o introduz (no caso de haver tal preposição). Isso vale para todos os constituintes,¹⁶ **menos** para o sujeito. No caso do sujeito, não basta procurar um SN em determinado lugar; é preciso ainda verificar sua compatibilidade em pessoa e número com o sufixo verbal.¹⁷

Mas mesmo assim estamos lidando com elementos concretamente presentes no enunciado: a ordem, a preposição (quando há) e o sufixo. Por outro lado, cada sintagma é categorizado no léxico ou então através de regras, de maneira que se possui a informação de que *o cachorro* é não só um SN, mas ainda de terceira pessoa, masculino e singular. Esse conhecimento é certamente de domínio dos falantes da língua.

8 Funções sintáticas: resumo

Escapamos, assim, da necessidade de definir “funções sintáticas” enquanto relações abstratas. Essas relações representariam uma espécie de “plano intermediário” entre a forma e o significado, plano esse que parece dispensável. Termos, como “objeto direto”, “sujeito” ou “adjunto adverbial de tempo”, são apenas maneiras alternativas de denotar especificações de categoria e posição na oração; não possuem significado gramatical próprio e só podem ocorrer na descrição a título de economia e facilidade de referência.

9 Conclusão

Vamos voltar agora aos fenômenos que as funções sintáticas pretendem ajudar a descrever. Repito aqui a lista:

- (a) a atribuição de papéis temáticos aos diversos SNs da oração;
- (b) a concordância verbal;
- (c) a ordem dos termos na oração; e
- (d) a distribuição de itens como eu, em oposição a me.

A atribuição dos papéis temáticos fica resolvida da maneira ilustrada acima. Ela é codificada na valência de cada verbo, composta cada uma de uma ou mais diáteses, como por exemplo para *matar*:

- (19) *matar* : SufixoPN (Agente), SN (Paciente)
Exemplo: o cachorro matou a cobra
- (20) *apanhar* : (diátese 1) SufixoPN (Paciente), de+SN (Agente)
Exemplo: o cachorro apanhou da cobra
- (21) *apanhar* : (diátese 2) SufixoPN (Agente), SN (Paciente)
Exemplo: apanhei algumas jabuticabas

¹⁶Para ser exato, estou desprezando diversos constituintes, que em geral não se analisam, como complementos ou adjuntos: a partícula negativa, o verbo auxiliar etc.

¹⁷Aqui há complicações, que são devidamente tratadas em Perini (inédito). Estas têm a ver com construções de infinitivo impessoal e gerúndio, em que não ocorre sufixo de pessoa-número.

Essa especificação de diáteses é feita para cada verbo da língua e é uma parte essencial do nosso conhecimento léxico-gramatical.

A **concordância verbal** é descrita em termos de congruência do argumento que ocupa o papel temático indicado pelo sufixo de pessoa-número. Uma frase é marcada como mal formada se essa congruência não se verificar, como em

(22) * Eu compramos um carro novo.

em que o sufixo indica “nós” (o falante mais alguma pessoa) e o SN na posição de sujeito indicam apenas “eu” (o falante).

A ordem dos termos na oração é parte da definição de cada diátese. Assim, sabemos que matar inclui um SN pré-verbal e um SN pós-verbal, como em

(23) O cachorro matou a cobra.

Já no caso de

(24) A cobra, o cachorro matou.

poderíamos, em princípio, acrescentar mais uma diátese ao verbo *matar*, com o Paciente no início da frase. Mas acontece que essa possibilidade não tem nada a ver com o verbo *matar*, sendo governada por fatores estruturais. Por isso, deve ser objeto de uma regra separada. Essa regra tem o efeito da transformação de Topicalização da gramática gerativa clássica, embora deva a meu ver ser formulada em termos de “via-rule”, ou seja, de implicação de existência: sempre que houver uma frase da forma $SN_1 V SN_2$, há também uma frase da forma $SN_2 SN_1 V$.¹⁸ Note-se que a ordem seqüencial é uma função sintática de pleno direito; e que, conseqüentemente, a função de *a cobra* em (23) é diferente da função do mesmo sintagma em (24).

Finalmente, temos a **distribuição de itens como eu em contraposição a me**. Ainda estou pensando em como descrevê-la convenientemente, e aqui não posso dar mais que algumas sugestões preliminares.

É bastante claro que a noção de “sujeito” vai ser relevante aqui. Vimos que o sufixo verbal é elaborado por um SN compatível com ele em pessoa e número. Essa condição não é suficiente, porque em certos casos é necessário levar em conta a ordem dos termos: isso acontece quando há mais de um SN na oração. Quando só há um SN, e este é compatível com o sufixo, esse SN é o sujeito.

Mas não quando se trata de um pronome oblíquo. Assim, na frase

(25) Eu pulei.

o SN *eu* é marcado como sujeito – e portanto Agente – da maneira habitual. Mas em

(26) Me lambi.

o SN *me* tem que ser o Paciente. Note-se que não é possível acrescentar um SN Paciente, e é possível acrescentar um Agente:

(27) * Me lambi o braço dela.

¹⁸Se há restrições (não me ocorre nenhuma), estas se formulam em termos estruturais e não léxicos.

(28) Eu me lambi.

Uma maneira de descrever esse comportamento seria dizer que a forma *me* é impedida de elaborar o papel temático indicado pelo sufixo de pessoa-número. mesmo quando é compatível com ele (isso equivale a dizer que essas formas não podem ser sujeito). As formas *me, nos, te, se* seriam então marcadas idiossincraticamente dessa maneira.

Essa é apenas uma análise possível: como disse, a questão ainda não está clara para mim. Mas gostaria de observar que é pouco provável que se possam analisar essas formas sem considerá-las idiossincráticas de algum ponto de vista.

Em nenhum desses casos é preciso lançar mão de funções sintáticas abstratas. No caso do sujeito, que é o único mais complexo, faz-se necessário um mecanismo de verificação da congruência em pessoa e número entre o sufixo verbal e um dos SNs. Nos outros casos, basta a categoria do sintagma (por exemplo, SN), a ordenação dos mesmos, a presença de uma preposição (por exemplo, *de* no complemento de *apanhar*) e a semântica da mesma (como no verbo *morar*, que tem um complemento de Lugar marcado por *em*).

Concluo que é possível descrever todos os fenômenos citados sem a utilização de funções sintáticas abstratas. Resta confirmar ou não essa hipótese levantando e analisando mais dados.

10 Especulando: outras línguas

Eu disse acima que não é hora de fazer comparações interlingüísticas. Mas, se me permitem sair do sério por alguns instantes, vou fazer alguns comentários comparativos que podem vir a ser interessantes eventualmente.

Vimos que as funções sintáticas são relações muito mais concretas do que muitas vezes se entende, redutíveis a relações de ordem ou à semântica das preposições. O sujeito é a grande exceção (e a única, até onde posso ver). Mesmo assim, se define em termos bastante imediatos, baseados na categorização de pessoa e número do SN e do sufixo verbal, além da posição na seqüência.

Suponho que, em muitas línguas sem concordância, o sujeito se reduz à situação das outras funções, sendo identificável apenas pela posição. O português (como muitas outras línguas) decidiu privilegiar um dos papéis temáticos da oração, codificando-o redundantemente. Isso vale para todas as variedades do português brasileiro. No mínimo, distingue-se a primeira pessoa do singular do restante, porque nenhuma variedade aceita **eu vai* ou **nós vou*. Outras línguas, como o náhuatl do México, privilegiam duas funções, representando-as tanto no verbo quanto, em forma elaborada, em um SN autônomo. Já o latim clássico privilegia o sujeito, como o português, mas o codifica em termos de desinências nominais, deixando uma grande liberdade para a ordenação dos sintagmas. Os demais termos são marcados não redundantemente, através das desinências e de preposições. Em latim como em náhuatl, e como em português, essas informações devem constar da valência do verbo.¹⁹

¹⁹Há mesmo casos de representação tripla de um dos papéis temáticos. Isso se verifica em alguns dialetos do norte da Itália; assim, em trentino se diz *la mama la magna el pan*, onde *la mama* é 'a mamãe', *la* é 'ela', e o verbo *magna* é marcado como terceira pessoa (em oposição a *magno*, 'como' etc.). O Agente de 'comer' é, portanto, sinalizado três vezes. Um fenômeno

E em todos esses casos as funções podem ser definidas em termos bastante concretos, e a distância entre a fonologia e o significado é bem pequena, não envolvendo relações abstratas do tipo "A é o sujeito de V".

11 Nota sobre as funções sintáticas na literatura

Não é aqui o lugar de fazer uma resenha detalhada da literatura a respeito das funções sintáticas. Vou me limitar a citar duas passagens que me parecem instrutivas.

Primeiro, vimos acima que as funções sintáticas tradicionais são na verdade noções mistas, definidas em termos parcialmente sintáticos, parcialmente semânticos. Na face sintática, temos a categoria dos sintagmas e a sua ordem na oração, mais a presença de preposição; e a face semântica é representada pela exigência de compatibilidade em pessoa e número entre elementos de mesmo papel temático para a identificação do sujeito. Isso parece se opor a uma posição encontrada na literatura, exemplificada pela passagem seguinte: "Como o nome indica, as relações gramaticais [isto é, as funções/MAP] são em princípio bem independentes dos papéis semânticos [isto é, papéis temáticos] veiculados por esses elementos." (Trask, 1992, p. 123).

Acho curiosa essa afirmação, em face da tendência muito geral de definir as funções sintáticas, explícita ou implicitamente, em termos de significado.

Quanto à análise da concordância verbal que esbocei acima, é necessário observar que uma análise semelhante se encontra em Pollard e Sag (1988). Mas na maior parte da literatura, a concordância se analisa como uma regra que adapta a forma do verbo aos traços do sujeito. Culicover e Jackendoff preferem não se comprometer:

Mencionamos de tempos em tempos fenômenos em outras línguas (e em inglês) sem oferecer uma análise séria. Como é que o presente quadro de referências trata fenômenos sintáticos bem conhecidos como a concordância intra-sentencial [...]? (Culicover; Jackendoff, 2005, p. 545)

12 Como funciona o sistema gramatical

Para terminar, direi uma palavra final sobre o funcionamento dos processos gramaticais. Quando consideramos um processo que consiste de diversas etapas – digamos, verificar a categoria a que pertence um sintagma, verificar se ele é introduzido por preposição, checar sua compatibilidade com o sufixo de pessoa-número etc. –, pode parecer que essas operações são ordenadas, mais ou menos como se ordenavam as regras nas análises gerativas até os anos 70 ou 80. Mas essa impressão é falsa e decorre apenas de necessidades textuais de exposição. Acredito que o sistema gramatical funciona de maneira radicalmente diferente.

Vamos lembrar que as regras gramaticais são parte do nosso conhecimento do mundo, ao lado de coisas como nosso conceito de "cachorro" e a disposição dos cômodos da nossa residência. Todos esses conceitos se organizam no que

semelhante ocorre em português com frequência, em frases topicalizadas do tipo *essa menina, ela ainda vai se machucar*.

alguns autores chamam de **esquemas**; e uma das características fundamentais dos esquemas é que eles são todos interligados, podendo ser acessados quando necessário.

Assim, para compreender uma afirmação como

(29) Não dormi direito porque teve uma festa no apartamento de baixo.

precisamos fazer diversas ligações, de diversas naturezas. Temos que usar o conhecimento de que festas podem ser barulhentas; que festas às vezes duram até tarde da noite; que geralmente as pessoas dormem de noite; que o barulho atrapalha o sono; que o apartamento de baixo é muito próximo do meu; que quanto mais próxima a fonte do barulho, maior o incômodo; que o som se propaga mais facilmente entre apartamentos do que entre casas.

Não faz muito sentido perguntar qual dessas informações precisa ser posta em jogo primeiro. Elas estão todas disponíveis simultaneamente, e nossa mente faz as conexões de acordo com as necessidades do momento; certamente, deve haver várias maneiras de abordar essa tarefa. Não sabemos grande coisa de como a nossa mente realiza essa tarefa, mas é bem claro que não é um processo linear, e que o resultado é uma rede complexa de conexões entre vários esquemas presentes em nossa memória, que colaboram para construir nosso entendimento dessa sentença.

Acredito que o processamento gramatical funciona basicamente da mesma forma. O receptor tem acesso simultâneo às regras, categorias e demais componentes de seu conhecimento da língua, e faz uso de tudo isso (ao lado de seu conhecimento do mundo) para construir a compreensão das frases. As regras, itens léxicos e demais recursos são também esquemas, e compreender uma frase é estabelecer as conexões certas entre os esquemas certos.

Para que isso funcione, é preciso que as estratégias estejam disponíveis mais ou menos em pé de igualdade. Não é plausível que o usuário tenha que tentar primeiro uma, e só recorrer à outra quando a primeira não funciona - ou seja, a noção habitual de ordenação não se aplica aqui.

Referências

- [1] ANDREWS, A. *Language typology and syntactic description*. Cambridge University Press, 1985, ch. The major functions of the noun phrase.
- [2] CULICOVER, P. W., AND JACKENDOFF, R. S. *Simpler syntax*. Oxford University Press, 2005.
- [3] DE VARNHAGEN, F. A. *História geral do Brasil*, 5 ed. Melhoramentos, São Paulo, 1956.
- [4] GIVÓN, T. Notes on the semantic structure of english adjectives. *Language* 46 (1970).
- [5] GROSS, M. *Méthodes en syntax*. Hermann, Paris, 1975.
- [6] JACKENDOFF, R. S. *Semantic structures*. MIT Press, Cambridge, MA, 1990.

- [7] LANGACKER, R. *Foundations of Cognitive Grammar, Vol. 1, Theoretical Prerequisites*. Stanford University Press, Stanford, 1987.
- [8] LEVIN, B. *English Verb Classes and Alternations: a Preliminary Investigation*. The University of Chicago Press, Chicago, 1993.
- [9] PERINI, M. A. Estudos de gramática descritiva: as valências verbais. Manuscrito. Inédito.
- [10] PERINI, M. A. O sintagma nominal: estrutura, significado e função. *Revista de Estudos da Linguagem* (1996). Edição especial.
- [11] PLANK, F. *Objects: towards a theory of grammatical relations*. Academic Press, London, 1984. ch. Ces obscurs objets du désir.
- [12] POLLARD, C., AND SAG, I. *Parasession on Agreement in Grammatical Theory*. Chicago Linguistic Society, 1988, ch. An information-based theory of agreement.
- [13] TRASK, R. L. *A dictionary of grammatical terms in linguistics*. Routledge, London, 1992.

The bipartite structure of verbs cross-linguistically

Heidi Harley
University of Arizona

hharley@email.arizona.edu

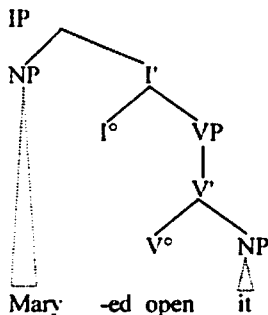
Resumo

Neste artigo demonstrarei a necessidade de uma estrutura complexa, bipartida para o sintagma verbal, que se aplica inclusive ao verbo. Tal estrutura configura-se através de um elemento V° e de outro 'V'. Revisarei algumas das motivações teóricas e empíricas para essa estrutura complexa. A seguir, apresentarei fenômenos que essa abordagem estendida cobre, fenômenos esses que até então não haviam sido sintaticamente tratados. Finalmente mostrarei como essa abordagem permite uma nova e, talvez, mais abrangente explicação para uma peculiaridade da alternância dativa do inglês; isto é, o fato de essa alternância não se aplicar à porção de base latina do inventário verbal inglês. A mesma restrição se aplica às construções de verbo + partícula do inglês, assim como a construções resultativas de maneira semelhante. Como a teoria trata esses três casos de forma semelhante, ela pode ser estendida para justificar a aplicabilidade da restrição aos verbos latinos para todos os casos em pauta.

1 Introduction

Not so long ago, in a universe very similar to this one except for the beliefs held by Chomskyan linguists about VPs, *Mary opened it* was assigned the following syntactic structure, within an X-bar approach to syntax (the internal structures of the NPs are not shown):

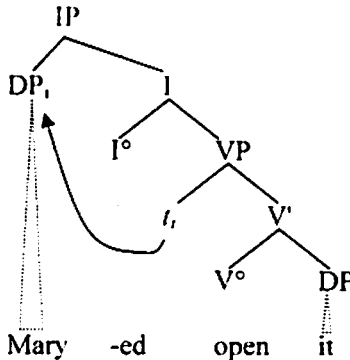
(1)



(The suffix *-ed* and *open* would get put together by some process ('Affix Hopping', e.g.) to produce the surface linear form of the sentence.)

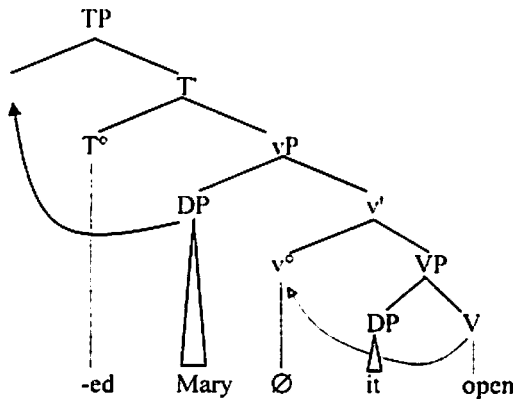
Then for a few years it looked like the structure illustrated below, where the salient change is that the subject argument is structurally related to the VP, reflecting the semantic (thematic) relationship between the V and its subject, and the nonthematic nature of the content of *I'*.

(2)



Nowadays (Hale: Keyser, 1993, as adopted in Chomsky, 1995), the structure (focusing only on the VP part) has minimally at least one more projection in it:

(3)



In this yet more abstract structure, the surface form is derived via two operations. Head-movement from $\sqrt{\quad}$ to v^0 simultaneously creates the causative verb

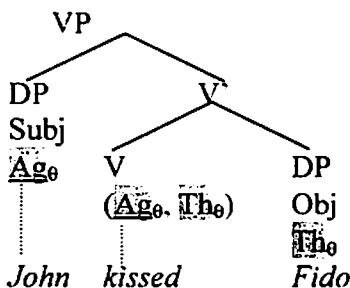
open, and moves *open* into the right position with respect to *it*; movement of *Mary* to spec-TP for the usual case-related or Extended Projection Principle-related reasons puts *Mary* in the sentence-initial position and allows *-ed* to be adjacent to the verb *open*, permitting *opened*.

The crucial aspect of this proposal is the claim that the portion of syntactic structure previously identified simply as the verb phrase exhibits a (minimally) bipartite structure. In fact, verbs themselves, according to this approach, are necessarily complex, being made up of a *v'* part and a 'V' part. In this paper, I will review some of the theoretical and empirical motivation for this more complex structure. Then I will present some of the phenomena which this account has been extended to cover, phenomena which previously had not been treated syntactically. Finally, I will show how the approach allows a new and perhaps more general account of a particular peculiarity of the English 'dative shift' alternation, namely its famous failure to extend to the Latinate portion of the English verbal inventory. The same constraint applies to the English verb-particle constructions and resultative constructions as well, and insofar as the theory treats all three of these constructions in a similar way, it may be extended to account for the applicability of the Latinate restriction to all of these cases.

2 Bipartite Verbs: Syntactic arguments

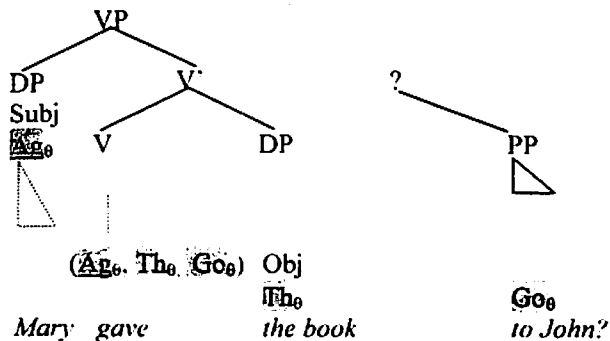
The first argument for a bipartite verb structure was presented in Larson's (1988) treatment of Barss and Lasnik's (1986) c-command facts in ditransitive constructions. Verbs with two internal arguments were problematic theoretically and empirically. The VP had gotten somewhat crowded when the external argument was moved down to its specifier position by the VP-internal subject hypothesis (Koopman; Sportiche, 1991); in ditransitive verbs, it was more crowded still:

- (4) VP-internal subject hypothesis:

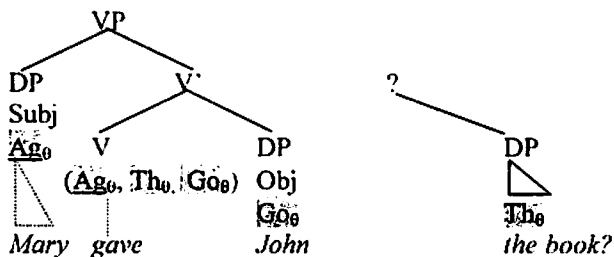


- (5) Ditransitive verbs?

- a Mary gave the book to John.
- b Mary gave John the book.
- c give (*Ag₀*, *Th₀*, *Go₀*)



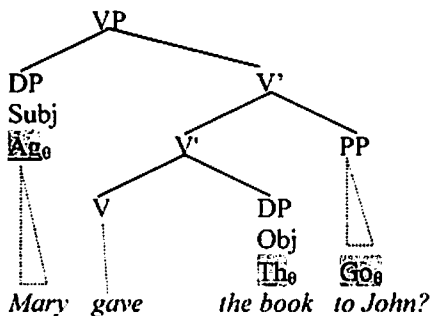
d



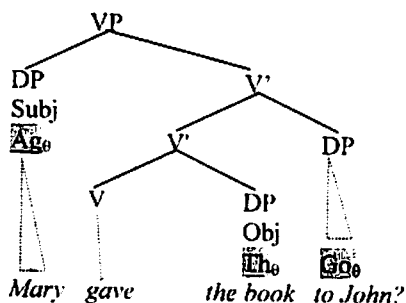
The X-bar schema did not allow for ternary branching, nor for more than one specifier and one complement. As discussed by Barss and Lasnik, it would be possible to right-adjoin the second DP to V', as follows:

(6) Adjunction to V':

a



b



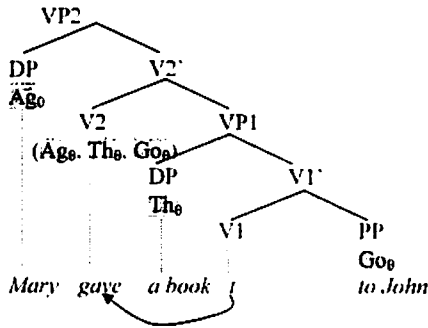
However, this structure resulted in the wrong c-command relations, as shown by Barsz and Lasnik (1986). In both *to*-dative and double object structures, the leftmost internal argument c-commands the rightmost internal argument, as shown by binding and negative polarity tests:

- (7) a John showed Bob₁ himself₁ (in the mirror)
 b *John showed himself₁ Bob₁ in the mirror
 c Mary gave no one anything
 d * Mary gave anyone nothing
 e The cruel boss denied [each worker₁][his₁ paycheck]
 f The cruel boss denied [it₂s owner] [every paycheck₂]

Furthermore, the bar-level adjunction position was supposed to be reserved for optional adjuncts, and to the extent that the argument/adjunct distinction was thought to have structural reflexes, the adjunction-to-V' analysis predicted that the rightmost internal argument should behave like an adjunct, not an argument (when in fact, if anything, it's the leftmost internal argument that has a few adjunct-like properties).

Larson's solution to this problem was to propose that the external argument was base-generated in the specifier of an upper VP "shell". The verb would project its two internal arguments in the specifier and complement position of the lower shell, and then head-move to the upper shell to establish its thematic connection to the external argument, as illustrated in (8):

(8) Larson's solution: Two VPs:



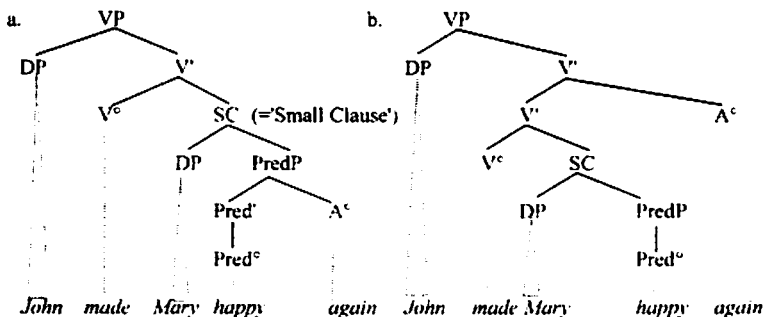
This structure made the right empirical predictions, as the leftmost argument c-commands the rightmost, and conformed to X-bar theory, as the X-bar theoretic template applies to each VP independently.

It does, however, raise several new questions. How can one verb project two VPs? What is the nature of the 'late' theta-role assignment to the external argument? What is the nature of the upper V^o head? Larson asserted that the upper V^o was purely structural in nature, containing no semantic information. The 'mediated' assignment of the external theta-role, however, seemed problematic within the extant θ -theory.

2.1 Semantic reasons to split VP

In a sentence like *John made Mary happy again*, the adverbial *again* can be interpreted in two ways, as modifying *happy* or as modifying *make*. This ambiguity receives a straightforward structural analysis, since *again* can have two loci of adjunction: one on the embedded (small clause) predicate *happy* and one on the matrix predicate *make*, corresponding to the two interpretations. On the former, Mary was happy before (independently of John), had become sad, and then she became happy again, thanks to John. On the latter, Mary had been made happy by John in the past, had become sad, and then been made happy by John again. The two structures are illustrated in (9) below (TP omitted to economize on space):

(9)

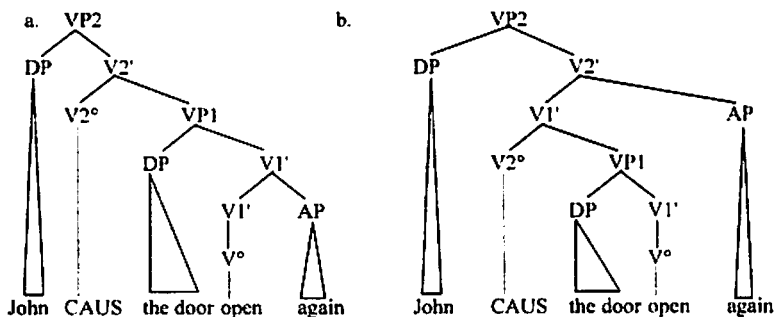


As shown by the original Generative Semantics literature from the late 60s and early 70s, however, similar scopal ambiguities are present with simple ditransitive and other change-of-state verbs, as illustrated in (10) below:

- (10) a John opened the door again
- i The door was open before, and now it's open again
 - ii John opened the door before, and he did it again
- b Mary gave the book to Sue again
- i Sue had the book before and now she has it again
 - ii Mary had given the book to Sue before and now she gave it to her again

Von Stechow (1995) argued strongly for a Generative Semantics-type analysis of variable scope for adverbials like *again* in these sentences, according to which the causative verb *open* is made up of a (null) predicate CAUSE syntactically taking a propositional complement headed by the (intransitive) predicate *open*. The scope of *again*, then, depends on whether it is adjoined to the embedded predicate or the matrix CAUSE predicate.

(11)



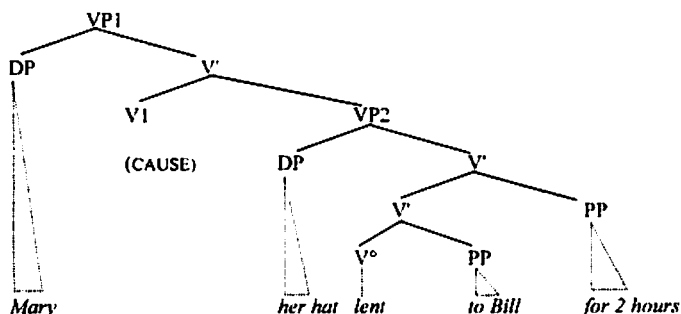
Beck and Johnson (2004) framed the same argument for the double object verbs, where again modifying the upper Larsonian VP-shell is interpreted as iterated causation of the event, and *again* attached to the lower VP shell is interpreted as the iterated result. With a ditransitive verb, the result denoted by the lower VP shell seems clearly to be stative location or possession. This can clearly be seen in another generative semantics argument from McCawley (1968, 1974) and Ross (1976), also resurrected by Beck and Johnson: the interpretation of temporal modifiers with ditransitive verbs:

(12) Temporal modifiers modifying the result of the event:

a Mary gave Bill the car until 3 o'clock (earlier this morning)

b Mary lent her hat to Bill for 2 hours

(13)



Here, it is not the action itself that lasts for two hours, but the state of the hat's being lent to Bill, or Bill's having the car. A similar effect can be seen with *open* and related change-of-state verbs:

(14) a John opened the window for five minutes

b Mary turned the tap to 'cold' for five minutes.

If the resultant state is represented in the structure independently of the initiating action, in a VP-shell structure like those above, it is easy to treat the modification of that resultant state by a temporal adverbial like those above; if it is not present, the syntax/semantics interface becomes remarkably complicated, as argued by von Stechow.

On this view of the contribution of the upper and lower V^o s – the 'little' v^o and the 'big' V^o – the upper v^o has its own semantic content, being related to event initiation and causation. The external argument, then, is related semantically to this upper v^o , and is in fact not 'selected' by the root V^o at all (though obviously the nature of the causation or event-initiation in which the external argument engages will be affected by the content of the V^o head, since different events require different types of initiation).

The intuition that the external argument is in some way semantically independent of the big V^o is supported by an independent observation originally made by Marantz (1984), and later treated within a compositional semantic framework by Kratzer. Kratzer (1996) points out that if external, agent arguments are in fact arguments of a separate v^o functional projection (with a semantics like that of a neo-Davidsonian predicate Agent(x, e)), then Marantz's (1984) generalization about the restrictions on idiomatic composition can be explained. Marantz noted that while verb-object idioms/special interpretations are ubiquitous cross-linguistically, verb-agent idioms (that exclude the object) are close to nonexistent.

(15)

kill a bug	=	cause the bug to croak
kill a conversation	=	cause the conversation to end
kill an evening	=	while away the time span of the evening
kill a bottle	=	empty the bottle
kill an audience	=	entertain the audience to an extreme degree

Kratzer notes that if the subject and the object both compose directly with the verb *kill*, there is no principled semantic reason why there shouldn't be as many subject-verb idioms as there are verb-object ones. For example, *A bug killed the boy* could have one special interpretation (a non-'kill' meaning), while *The food killed the boy* could have another. However, these kinds of idioms, with free object positions and bound (idiomatic) agentive subjects, do not seem to occur.

If, however, agents compose with a separate light verb and then have their interpretation composed with that of the lower predicate via a process Kratzer calls Event Identification (essentially, conjunction of the functions and simultaneous binding of the event variables), the semantic independence of Agent arguments is expected. Event Identification combines the denotation of v (type $\langle e, \langle s, t \rangle \rangle^1$) with the (argumentally saturated) denotation of the lower VP, (type $\langle s, t \rangle$), to produce a function of type $\langle e, \langle s, t \rangle \rangle$. This operation can be seen below (Kratzer's ex. 19):

(16)

(Voice)	(VP)	→	Voice'
f	g		h
$\langle e, \langle s, t \rangle \rangle$	$\langle s, t \rangle$		$\langle e, \langle s, t \rangle \rangle$
$\lambda x \lambda e, e, \text{Agent}(x)(e)$	$\lambda e, \text{wash}(\text{the clothes})(e)$		$\lambda x \lambda e, [\text{Agent}(x)(e) \& \text{wash}(\text{the clothes})(e)]$

(It's important to note that Kratzer's treatment of Marantz's generalization only works if the object of the verb is truly an argument of the verb, composing with it directly. A truly Neo-Davidsonian analysis where there is, e.g., a function Patient(x)(e) and a separate Manner function that introduces the contribution

¹e = individuals, s = events, t = truth values

of the verb root won't work, or rather, will make the wrong predictions about idiomatic interpretations of the $\sqrt{\cdot}$.)

For at least certain verb classes, then we have some evidence that the verb is made up of a matrix external-argument introducing projection, v^o involving causation or initiation, and a formally independent downstairs V^o projection, which a) seems to denote a result state (as shown by the interpretation of the examples in (12) and (14) and b) selects the internal arguments of the verb, and contributes the 'root' meaning of the predicate. Next we turn to crosslinguistic morphological evidence that verbs are indeed composed of more than one element: in many (perhaps most?) languages, change-of-state verbs are clearly morphologically complex.

3 Bipartite Verbs: Morphological evidence

Hale and Keyser (1993, 2002 and other works) also argue for syntactically complex (agentive) verbs, composed of an external-argument-introducing v^o and an independent lower V^o which contains the verb root itself. Their initial reason for thinking so was the fact that in many languages, verbs are visibly complex, being made up of a root element (itself, in fact, *not* verbal) plus some type of verbalizer. Their examples of morphologically complex unergative verbs from Jemez are given below:

(17)

a.	sæ-a	b.	zæi-a	c.	se-ʔa
	word-do		song-do		word-do
	"work"		"sing"		"speak"

In fact, this kind of bipartite morphological structure in 'simple' verbs, especially in change-of-state verbs, is crosslinguistically pervasive. Below I give just a few groups of examples from some familiar and less-familiar languages that I have worked with; examples from unrelated languages can be easily multiplied.

3.1 Japanese

Japanese has a large class (containing hundreds of exemplars) of derived causative (transitive) / inchoative (intransitive) verbs which share a common bound root, are generally semantically related in their entailment consequences in the usual way in such pairs, and exhibit a complex morphological structure in either or both the causative or inchoative forms. Below I give a table illustrating the various types of morphological alternations that these verbs exhibit, as identified and classified by Jacobsen (1981, 1992):

(18)

Class/# ²	√	Intr	Tr	Rough √ gloss
I: e/θ 30 pairs	hag hirak	hag-e-ru hirak-e-ru	hag-θ-u hirak-θ-u	'peel off' 'open' ³
II: θ/e 44 pairs	ak hikkom	ak-θ-u hikkom-θ-u	ak-e-ru hikkom-e-ru	'open' 'draw back'
III: ar/c 71 pairs	ag aratam	ag-ar-u aratam-ar-u	ag-e-ru aratam-e-ru	'rise' 'improve'
IV: ar/θ 8 pairs	hasam husag	hasam-ar-u husag-ar-u	hasam-θ-u husag-θ-u	'catch between' 'obstruct (clog, jam?)'
V: r/s 27 pairs	ama hita	ama-r-u hita-r-u	ama-s-u hita-s-u	'remain' 'soak'
VI: re/s 18 pairs	arawa hana	arawa-re-ru hana-re-ru	arawa-s-u hana-s-u	'show (up)' 'separate from'
VII: ri/s 2 pairs	ka ta	ka-ri-ru ta-ri-ru	ka-s-u ta-s-u	'borrow/(lend)' 'suffice/(supplement)'
VIII: θ/as 38 pairs	hekom her	hekom-θ-u her-θ-u	hekom-as-u her-as-u	'dent' 'decrease'
IX: e/as 45 pairs	bak bar	bak-e-ru bar-e-ru	bak-as-u bar-as-u	'turn into/bewitch' 'come/bring to light'
X: i/as 8 pairs	ak dek	ak-i-ru dek-i-ru	ak-as-u dek-as-u	'tire' 'come/bring into existence'
XI: i/os 6 pairs	horob ok	horob-i-ru ok-i-ru	horob-os-u ok-os-u	'(fall to) ruin' 'get up'
XII: θ/se 6 pairs	abi ki	abi-θ-ru ki-θ-ru	abi-se-ru ki-se-ru	'pour over (self/other)' 'put on (self/other)'
XIII: e/akas 4 pairs	obi hagur	obi-e-ru hagur-e-ru	obi-(y)akas-u hagur-akas-u	'take fright/frighten' 'stray/evade'

²The number of pairs does not include other pairs derived from a root already on the list even when these are not transparently semantically related; the number of items on each list, then, is actually somewhat larger.

³Mamoru Saito and Yosuke Sato (p.c.) inform me that the forms listed as meaning 'open' here, hirakeru~hiraku, are not used (the pair from class II, aku~akeru, is the appropriate one). Some other particular items in Jacobsen's lists also seem to not currently be in use, for example bakasu, dekasu, and nukumeru.

XIV: or/e	kom	kom-or-u	kom-e-ru	'be fully present/fill'
2 pairs	nukum	nukum-or-u	nukum-e-ru	'warm'
XV: are/e	sut	sut-are-ru	sut-e-ru	'fall into disuse/discard'
3 pairs	wak	wak-are-ru	wak-e-ru	'divide'
XVI: Misc	nigiwa	nigiwa- h -u	nigiwa-s-u	'(make) prosper'
25 pairs	nob	nob-i-ru	nob-e-ru	'extend'

This is clearly a salient and pervasive pattern within Japanese.

3.2 Hiaki (Yaqui) derived verbs

A similar contrast can be seen in the Uto-Aztecan language Hiaki, spoken in northern Mexico and the southern United States. In Hiaki, there are many causative/inchoative pairs of verbs which again share a common bound root and semantic content and are related to each other via distinct suffixation patterns. Consider the following pairs, taken from Jelinek (1997), Jelinek and Escalante (2001):

bwasa	"cook"	bwase	"cook, ripen"
chakukta	"bend"	chakukte	"bend"
chakta	"drip"	chakte	"leak"
chihakta	"smash"	chihakte	"shatter"
hamta	"break"	hamte	"break"
heokta	"melt"	heokte	"melt"
chu'akta	"stick on"	chu'akte	"adhere"
chukta	"cut loose"	chukte	"come loose"
chupa	"finish"	chupe	"come to end"
ko'okta	"pull apart"	ko'okte	"come undone"
kowiikta	"make crooked"	kowiikte	"get crooked"
kitokta	"deform"	kitokte	"shrivel"
kotta	"break"	kotte	"break"
kutta	"tighten"	kutte	"get tight"
kuuta	"stir"	kuute	"mix"
luuta	"use up"	luute	"run out"
mana	"place before"	mane	"be before"
mohta	"grind"	mohte	"break up"
mohakta	"take apart"	mohakte	"crumble"
nasonta	"damage"	nasonte	"get damaged"
patta	"shut"	patte	"shut"
pesta	"burst"	peste	"burst"
pitta	"press"	pitte	"settle down"
pohta	"boil"	pohte	"boil"
potta	"stretch out"	potte	"rise, expand"
rauta	"rinse"	raute	"rinse"
resta	"spread out"	reste	"spread out"
revekta	"break apart"	revekte	"come apart"
riuta	"split"	riute	"split"

ropta	"sink"	ropte	"sink"
ro'akta	"roll over"	ro'akte	"roll along"
sihho'ota	"sprinkle"	sihho'ote	"drizzle"
sipa	"cool"	sipe	"cool"
siuta	"tear"	siute	"tear"
teita	"trip"	teite	"trip"
tohta	"discolor"	tohte	"fade"
topakta	"turn over"	topakte	"flip over"
totta	"bend"	totte	"collapse"
tuhta	"press"	tuhte	"settle"
tuucha	"put out (fire)"	tuuke	"go out"
veeta	"burn"	veete	"burn"
vi'ita	"twist"	vi'ite	"twist"
vohta	"pour out"	vohte	"drop out"
vutta	"undo"	vutte	"come undone"
weeyya	"carry"	weeye	"move"
wiokta	"untangle"	wiokte	"untangle"
wiuta	"spend"	wiute	"run out"
wohokta	"dig up; puncture"	wohokte	"get a hole in"
woita	"untie"	woite	"come untied"
wo'ota	"spill"	wo'ote	"spill"
yohta	"drop"	yohte	"drop"
yooka	"paint"	yooke	"change color"

Again, seemingly a salient and pervasive pattern in the language (this list contains only about half the pairs given by Jelinek; further, these do not include verbs that contain a lexicalized version of the productive causative suffix *-tua* or the (separate) creation suffix *-te*). The roots themselves are not independently used, as in Japanese; they must always come bound with one of their suffixes. Nonetheless it seems clear that they exist as morphological units, and the verbalizing suffixes are independent recurring entities exhibiting specific morphosyntactic behavior.

3.3 Persian

Persian (Farsi) exemplifies a final type of 'bipartite' verb construction, one in which the 'verb' is usually so bipartite it is not in fact a verb at all, but a complex predicate construction consisting of a light verbal head and a complement phrase of any other syntactic category, including nouns, adjectives, prepositions and prepositional phrases. The internal arguments introduced by these complement phrases behave like the regular internal arguments of the few 'heavy' verbs of the language; that is, they do not show any apparent evidence of being further embedded than regular internal arguments (Megerdooimian, in press).

The verbal lexicon of Persian has undergone 'separation' in its documented history, from mostly 'heavy' verbs with fewer complex predicates to the current situation in which nearly all verbal meanings are commonly expressed via complex predicates, with the simple 'heavy' verbs restricted to written or elevated language. Some examples illustrating the historical derivation from heavy verbs to complex predicate construction are given below. All examples and discussion below taken from Folli, Harley and Karimi (2005).

(19)

Simple	Complex		
lasidan	las zadan	(flirtation doing)	'to flirt'
raghsidan	raghs kardan	(dance doing)	'to dance'
agahanidan	agah kardan	(informed making)	'to inform'
aghazidan	aghaz kardan	(start doing)	'to start'

A more comprehensive sampling of Persian complex predicate types is provided below:

(20)

a N+LV

kotak zadan/xordan	(beating hitting/colliding)	'to beat, to get beaten'
xar kardan/shodan	(donkey doing/becoming)	'to fool, become fooled'
dust dāshtan	(friend having)	'to love'

b A+LV

sabok kardan/shodan	(light making/becoming)	'to degrade' (tr & intr)
pahn kardan/shodan	(wide making/becoming)	'to spread, to widen' (tr & intr)
deráz keshidan	(long pulling)	'to lie down, to take a nap'

c Particle+LV

birun kardan	(out doing)	'to dismiss, to fire (someone)'
bâlâ ávardan	(up bringing)	'to vomit'
bâlâ keshidan	(up pulling)	'to steal'

d PP+V

be yád dāshtan	(to memory having)	'to have in memory'
bejá ávardan	(to place bringing)	'to recognize'
be bād dādan	(to wind giving)	'to waste'

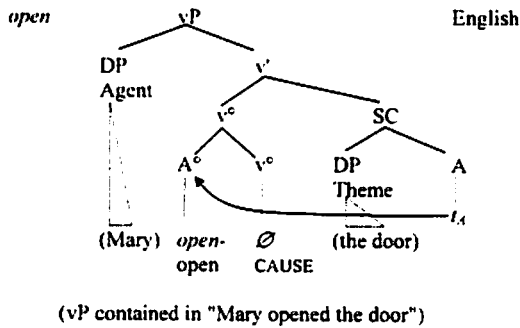
Clearly the bipartite nature of the Persian verb is strongly evident in the language, and supported by the syntactic independence of the two subparts of the construction, despite their semantic unity.

Similar examples of (derivational) morphological complexity in broad segments of the verb inventory of unrelated languages could be multiplied almost indefinitely. Indeed, we will next see examples of this type from English. However, in the complex English cases, an interesting difference can be seen in the Persian, Japanese and Hiaki cases, related to the locus of the contentful verb root in the structure.

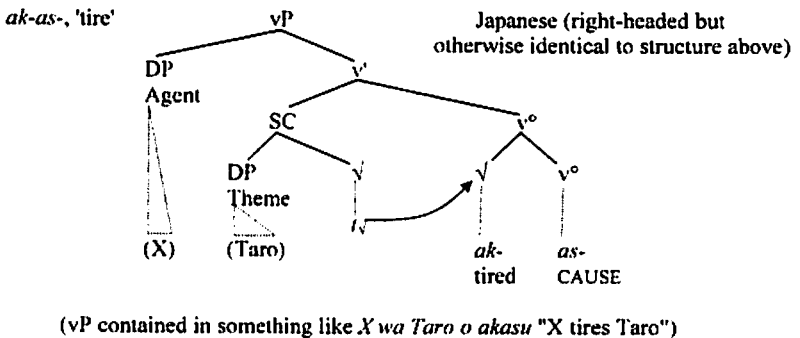
3.4 English 'High' roots

In the complex verbs described above, it seems clear that the verb root originates lower in the verbal structure, identifying the result of the change-of-state indicated by the complex form: this then is composed with a verbalizing morpheme higher in the structure with an abstract meaning like 'cause' or 'become/happen' via head-to-head movement. Trees illustrating this structure for causative open in English, causative *ak-as-u*, 'tire' in Japanese, causative *hanta*, 'break' in Hiki and causative *pahn kardan*, 'widen' in Persian, are provided in 3.4-3.4 below. Two important changes have been made to the initial Chomskyan structure illustrated in 1 above. The lower, contentful 'root' part of the projection has been relabeled with its syntactic category (where this is clear from elsewhere in the language), and with the label ' $\sqrt{\quad}$ ' where the item is a bound, acategorical root which cannot be used freely. In addition, the projection of this root, labeled 'VP' in the previous structure, has been relabeled 'SC' for 'Small Clause', reflecting its predicational and propositional nature - see, e.g. Hoekstra (1988) for discussion. The upper v^0 selects for a state-denoting complement, and the semantic combination of the causative v^0 and SC creates a change-of-state meaning for the whole (compare. *Mary made [John happy]*).

(21)

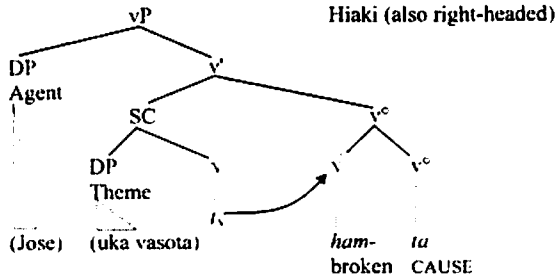


(22)



(23)

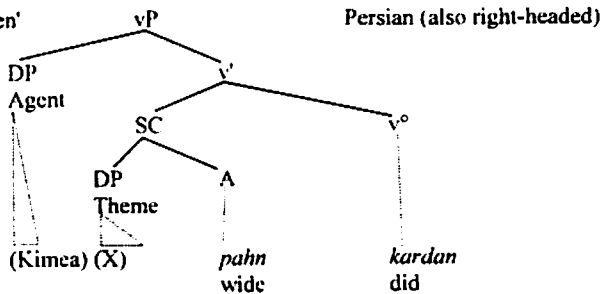
ham-ta, 'break'



(vP contained in something like *Hose uka vasota hamtak*, "Jose broke the glass.")

(24)

pahn kardan, 'widen'



(vP in something like *Kimeca X-ro pahn kard*, "Kimeca widened X")

In short, change-of-state verbs in all these languages can profitably be analyzed with the same complex vP syntactic structure, where the semantic structure of the complex involves a causation portion expressed by the v' morpheme, and a resultant state portion expressed by morpheme denoting the predicate of the SC. The predicate may or may not incorporate into the v° to form a 'heavy' (contentful) verb.

In English, there is no morphological evidence for complex structure in verbs like *open*, since the causative and inchoative variants share identical morphology. There is, however, a large class of morphosyntactic bipartite verbs in English in which it seems most of the same processes are at work, but in which the 'verb root' part of the structure—the contentful part—seems to be located in v° , and the resultative predicate part of the structure lower down is more functor-like, less contentful. These of course are the verb-particle constructions. Just a few examples are given in (25); the pattern is very common in the language:

- (25)
- | | | |
|---|-----------------------|-----------------------|
| a | throw the garbage out | throw out the garbage |
| b | pick the paper up | pick up the paper |
| c | push the needle in | push in the needle |

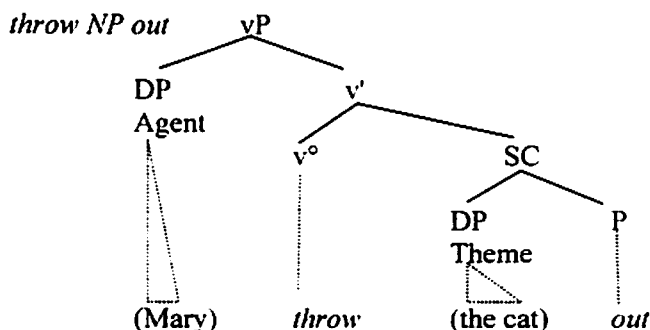
- | | |
|---------------------------|-------------------------|
| d pass the dish around | pass around the dish |
| e fight the attackers off | fight off the attackers |
| f paste the stamp on | paste on the stamp |
| ... | |

A common approach to the structure of these English constructions within this general framework holds that the particle identifies the resultant state, while the verbal portion identifies the causative portion of the meaning. When one considers the results of, e.g., the *again* modification test with compositional examples of verb-particle constructions like the above, it is clear that on the 'result modification' reading, the crucial thing is that the resultant state be iterated, not the action denoted by the verb root. Consequently, e.g., the result-modification reading of (26a) below requires only that the cat has been outside before, not that he has been *thrown* outside before. A similar observation applies to the endstate modification reading of the temporal adverbial in (26b) - it's the state of the cat being outside that lasted for two hours, not the throwing portion of the event:

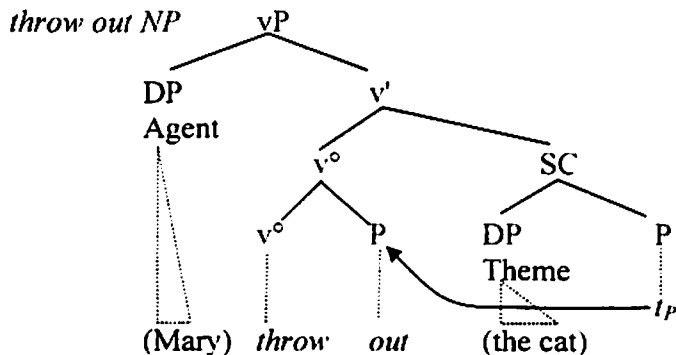
- (26) a (After letting him in only an hour before,) Mary threw the cat out again.
- b Mary threw the cat out for two hours (then she let him in again).

This suggests that the structure associated with these verbs involves having the verb root, identifying the causation portion of the structure, in v^0 , and allowing the particle to sit lower in the structure, as the predicate of the small clause, illustrated in (27). (Particle shift arises when the particle head-moves to adjoin to the verb in v^0 , illustrated in (28)):

(27)



(28)



Of course, verb-particle constructions are famous for having idiosyncratic, 'lexical' or idiomatic interpretations, where the independent semantic content of the verb or particle is not obviously present, as in the following cases:

- (29) a see NP through 'to persevere with NP'
b chew NP out 'to scold NP'
c piss NP off 'to anger NP'
d fill NP in 'to brief NP'
e work NP over 'to heat NP'
f while NP away 'to pass NP(=time)'
...

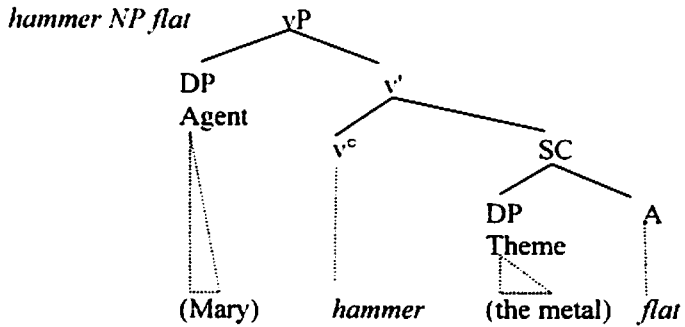
Clearly we have an overtly bipartite verb structure in these cases; the pervasive verb-particle construction of English is thus another excellent example of the way in which verbs can have internally complex morphological and syntactic structure. However, these bipartite verbs differ in an important respect from the Persian, Japanese and Hiaki cases discussed above, as well as from cases like simple *open*. In English verb-particle constructions, the 'contentful' part of the complex verb—the 'root'—is associated with the v° head, rather than with the lower-down predicate of the SC .

4 Manner Incorporation: Productive insertion of roots into v°

Similar observations apply to the less syntactically and semantically idiosyncratic English resultative construction. The complex predicate wipe NP *clean* clearly involves a causative portion and a result state portion (consider the result-modification interpretation of *Mary wiped the table clean again*). As with the verb-particle constructions, a contentful verbal element is present as the

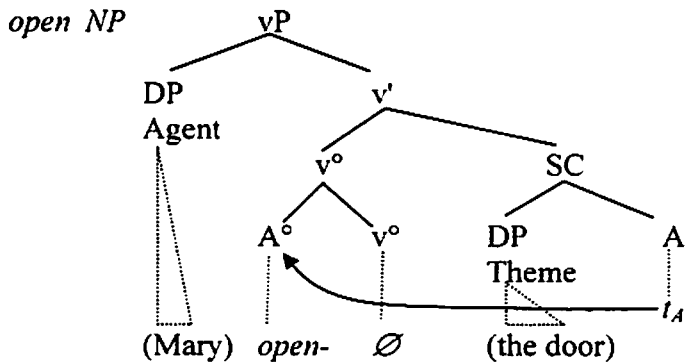
v° head, not the SC predicate, which is independently realized by a separate (contentful) resultative adjective. The structure for *Mary hammered the metal flat* is illustrated below:

(30)

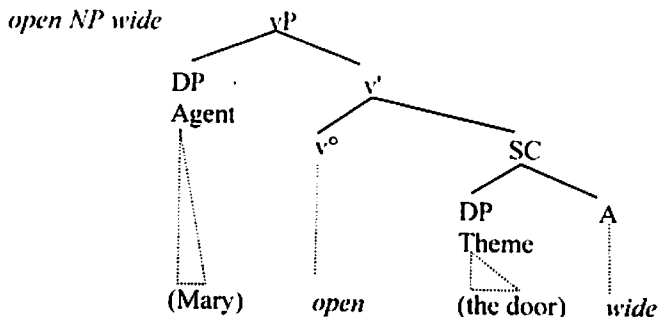


Indeed, it seems that in English, the same verbal root can alternately appear as the predicate element of a small clause, or in a v° head, with an independent element present realizing the SC predicate. Compare *Mary opened the door* (where *open* is the SC predicate describing the endstate of the *door*) and *Mary opened the door wide* (where *wide* is the SC predicate, and *open* is present in the v° head):

(31)

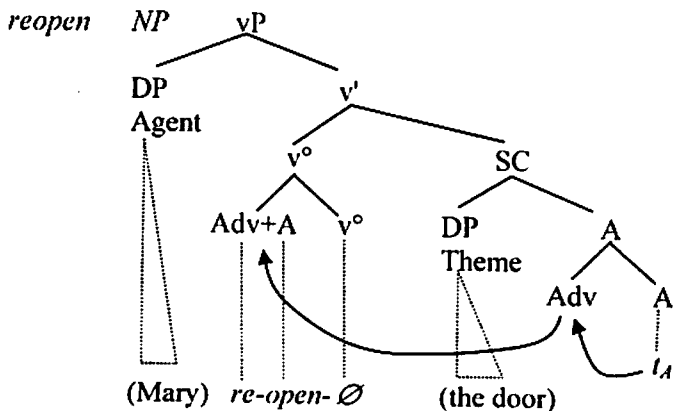


(32)



The distinct structural sources for *open* in the two structures are revealed by their different behaviors with *re-* affixation, discussed by Keyser and Roeper (1992). English iterative *re-* has essentially the meaning of result-modifying *again*. It applies to change-of-state predicates to indicate that the resultant state, previously holding at some point and then (up to reference time) not holding, has been brought about again. In short, *re-* modifies the endstate-denoting part of the bipartite verb structure, as illustrated in 4, and is then incorporated with the verb root into the v° position via head-to-head movement:

(33)



With the resultative or verb-particle construction version of *open*, however, *re-* affixation is impossible, as shown by the examples in 4. If *re-* must attach to a SC predicate, which then incorporates into a v° , then with the resultative, *re-* affixation to *open* will be impossible because *open* is never the SC predicate. Consequently, the *re-* can never be attached to *open* when a resultative predicate or verb particle co-occurs with it—in such cases, *open* starts off in the v°

position, too high in the structure to be modified by *re-*.⁴ See Harley (2004) for a somewhat fuller discussion of these cases in the context of head-movement.

- (34) a *Mary reopened the door wide.
 b *Mary reopened the door up.

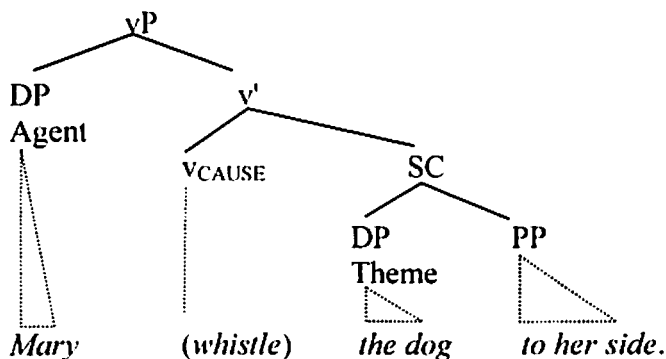
The adjectival verb root *open*, then, can be associated either with the endstate-denoting small clause predicate, or with the activity-denoting *v*^o head, in English. It seems that in general, English *v*^o is able to host either a null causative or a full verb root.

This flexibility of English roots allows a nice account of causative manner-of-motion constructions in the language. Verb roots which do not entail any motion whatsoever in their 'basic' use may occur in causative motion-denoting verb phrases, often with NP direct objects which the verbs themselves would never select, as discussed by Levin and Rappaport-Hovav (1999), among many others. Consider the examples below:

- (35)
- | | | |
|---|---|-------------------------|
| a | Mary whistled the dog to her side. | *Mary whistled the dog. |
| b | Mary shoved her way to the front of the line. | *Mary shoved her way. |
| c | Mary talked herself onto the stage. | *Mary talked herself. |
| d | Mary pitched her team into the finals. | *Mary pitched her team. |

In such cases, the assumption that *v*^o may contain roots indicating manner-of-causation allows us to treat these as garden-variety causative constructions, where the underlying structure of something like (35a) involves a caused change of location in which the causation is accomplished in a *whistling* manner:

(36)



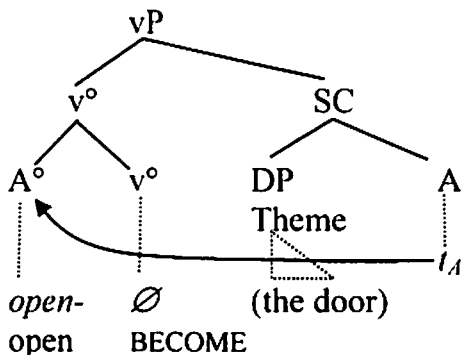
⁴Why can't *re-* attach to the resultative adjective or to the particle, remaining in situ, to produce *Mary opened the door rewide* or *Mary opened the door reup*? For the moment, I conclude that *re-* is subject to a constraint that requires that it ultimately incorporate into a *v*^o head. If a more principled explanation exists, it will have to wait for future research.

The internal arguments of the construction are introduced by the small clause predicate *to her side*. This explains why these sentences are ill-formed without the PP (as shown in (4) above) and why *whistle* thus may co-occur with these ‘unselected’ objects as a purely manner-denoting root—no selectional conflict occurs because *whistle* here is behaving like a causative verb requiring a small clause complement, not like its unmarked behavior as a root verb where it can optionally select only for a tune-denoting DP complement (*Mary whistled “Three Blind Mice”*).

This general approach can also be applied to non-causative manner-of-motion constructions. In our lists of causative/inchoative alternating verbs above, a change in the v^o portion of the bipartite morphemic structure is associated with the absence of an external argument; this can be interpreted as reflecting the change from the external-argument-selecting v^o meaning CAUSE to the inchoative v^o meaning BECOME, illustrated for English *open* (intr), Japanese *ak-i-ru*, ‘tire (intr)’, Hiaki *hamte*, ‘break (intr)’, and Persian *pahn shodan*, ‘widen (intr)’:

(37)

open (intr)

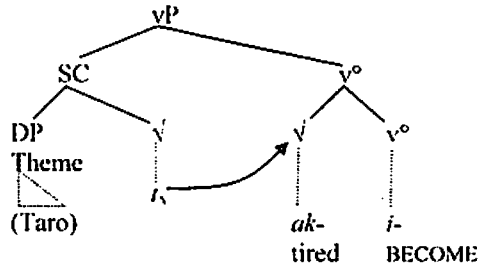


(vP contained in “The door opened”; the DP the door moves to subject position in Spec-TP to check nominative case and produce the final word order)

(38)

ak-i-, 'tire' (intr)

Japanese

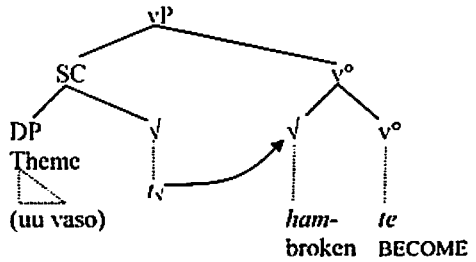


(vP contained in something like *Taroo wa akiru* "Taro gets tired")

(39)

ham-te, 'break (intr)'

Hiaki

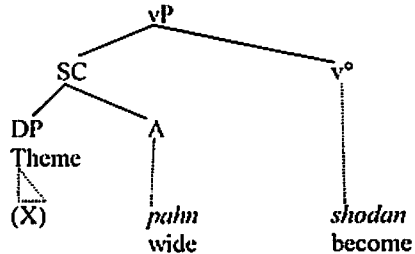


(vP contained in something like *Uu vaso hamtak*, "The glass broke")

(40)

pahn shodan, 'widen'

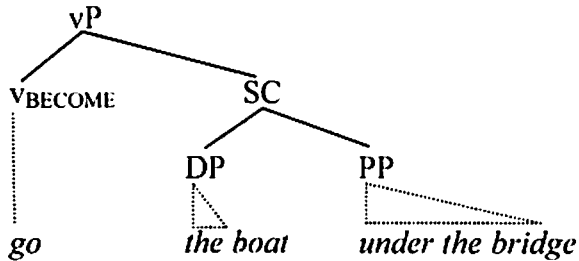
Persian



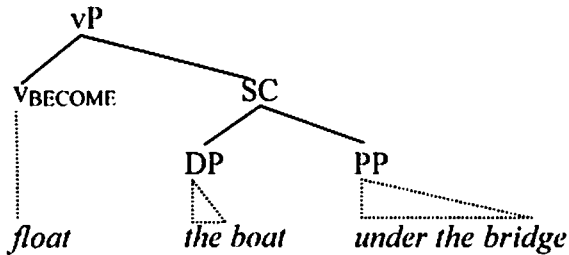
(vP in something like *X pahn shod*, "X widened")

By analogy, we can consider the inchoative version of a manner-of-motion construction to involve a BECOME change-of-state v'' predicate and small clause predicating the goal-of-motion of the Theme argument. With no manner element specified, this v'' shows up as *go* or *get*. However, other verbs—even non-motion-entailing ones may appear in the v'' head position to specify the manner in which the change-of-location occurred:

- (41) a. The boat went under the bridge. (# The boat went in place)



- b. The boat floated under the bridge. (cf. The boat floated in place)



- c. The ball bounced across the road. (cf. The ball bounced in place)
- d. The bullet whistled through the window.
- e. The water swirled down the drain.
- f. ... etc

In the motion constructions discussed here, then, a verb root is used in v'' to identify the manner in which a change (either caused or simply inchoative) occurs. Cross-linguistically, however, it is seemingly common for the verb root to realize the resultant state portion of a change-of-state predicate, and to have the v'' head represented by a more functional derivational morpheme or a light verb (and, indeed, such structures are also available in English as well).

Given this characterization of manner-of-motion constructions, we can recast Talmy's (1985) generalization about 'verb-framed' vs. 'satellite-framed' languages in syntactic terms. Recall Talmy's observation: Romance languages (and others) do not allow 'manner' elements to appear as verb roots in motion constructions. So while (42a) is a perfectly good description of a motion event in English, its word-by-word translation in Spanish (42b) cannot denote a motion event. Rather, the PP is interpreted as a location adverbial, and the verb receives its usual, non-motion entailing reading. In order to describe both the manner and the motion in a single sentence of Spanish, a motion verb is required as the main verb, with a manner-denoting participial adjunct, as in (42c):

(42)

- a. The bottle floated into the cave.
- b. La botella flotó a la cueva.
 The bottle floated to the cave.
 "The bottle floated in the cave" (no motion involved)
 "# The bottle floated into the cave" (motion interpretation impossible)
- c. La botella entró a la cueva, flotando
 The bottle entered to the cave, floating
 "The bottle entered the cave, floating."

The difference between English and Spanish with respect to manner-of-motion constructions, then, can be characterized as follows: English allows verbs roots to be inserted into v^o and behave as causative predicates – 'high' in the bipartite structure – as well as in the canonical position in the lower part of the bipartite structure. Spanish does not allow verb roots to appear in the v^o position.

This predicts, as well, that languages like Spanish will not allow resultative constructions nor verb-particle constructions, both of which have also been analyzed above as involving a verb root appearing in the v^o position. Indeed, we expect the availability of manner-of-motion constructions, verb-particle constructions, and resultative constructions to correlate with each other, cross-linguistically: if a language allows one, it should allow them all, and if it forbids one, all should be impossible. Indeed, at least as far as the Germanic and the Romance languages go, this prediction appears to be borne out. See Harley (1999, 2005), Mateu (2001), Folli and Harley (2006) and McIntyre (2003) for alternative formulations of this approach and additional discussion.

5 The two English lexicons

Spanish and English differ, then, as to whether they allow an independent verb to realize the v^o element in causative and inchoative constructions. What has also long been recognized, but not received a convincing analysis, is that a certain subpart of the English verbal lexicon behaves like Spanish, in failing to allow those particular verb roots optional placement in v^o as well as in the small clause. In fact, complex Latinate verbs in English behave for the most part as if they are still in a Romance language, failing to occur in verb-particle constructions or in resultatives. Consider the following contrasts:

- (43) Verb-particle constructions fine with Anglo-Saxon-type verbs but not Latinate ones

write it up	*compose it up /*arrange it up
eat it up	*consume it up
finish it up	*complete it up
throw it out	*discard it out
lie down	*recline it down
hand it out	*distribute it out
show it off	*exhibit it off /*reveal it off
fire it up	*ignite it up
slice it off	*incise it off
tidy it up	*arrange it up
hide it away	*conceal it away
cut it apart	*dissect it apart
figure it out	*calculate it out
move it over	*displace it over
go away	*depart away
clear it up	*clarify it up
write it up	*compose it up
cast it off	*release it off
dig it up	*excavate it up
swell up	*expand up
trade it in	*exchange it in

- (44) Resultatives fine with Anglo-Saxon verbs, not many Latinate ones

cut it apart	*divide it apart
fill it full	*inflate it full
walk yourself tired	*perambulate yourself tired
work yourself ragged	*decide yourself ragged
squeeze it empty	*compress it empty
stab it dead	*impale it dead
train yourself fit	*condition yourself fit
freeze solid	*congeal solid
dance yourself pink	*exert yourself pink
eat yourself sick	*devour yourself sick
drink yourself unconscious	*imbibe yourself unconscious
scrape it raw	*abrade it raw
break it short	*divide it short
grow big	*expand big
burn black	*combust black

The resultative construction has additional constraints on it, so it is not completely productive in any case, but the resultatives that *are* permitted obey the constraint against Latinate-type verbs.⁵

⁵In the interests of full disclosure, there are certain exceptions to the overall pattern. For example, to my ear and those of a few other English speakers, *divide up*, *collect up*, and *calculate up* are quite natural. The generalization does not hold for motion-verb constructions

The other case in which the literature has repeatedly noted a significant difference between complex Latinate verbs and Anglo-Saxon ones is with respect to 'dative shift', aka the double object construction. Ditransitive verbs which occur with an accusative theme and to-DP Goal can generally also occur in the ditransitive construction, providing certain semantic conditions are met and, importantly, providing that the verb is Anglo-Saxon. Complex Latinate verbs refuse to occur in the double object construction, as shown in the following examples from Pesetsky (1995):

- (45)
- a Susie gave Oxfam some canned food.
 - a' Susie gave some canned food to Oxfam.
 - b *Susie donated Oxfam some canned food.
 - b' Susie donated some canned food to Oxfam.
 - c Bill sent Sue his regards.
 - c' Bill sent his regards to Sue.
 - d *Bill conveyed Sue his regards.
 - d' Bill conveyed his regards to Sue.
 - e Mary showed the committee her findings.
 - e' Mary showed her findings to the committee.
 - f *Mary displayed the committee her findings.
 - f' Mary displayed her findings to the committee.
 - g Tom told Ben the story.
 - g' Tom told the story to Ben.
 - h *Tom recounted Ben the story.

PP-resultatives, or way-constructions. For example, in X's way-manner-of-motion constructions the only restriction is that the verb has an atelic interpretation available, or can be coerced to one (the same restriction holds for non-Latinate verbs):

- i ...complained his way out the door.
- ...exclaimed his way through the museum
- ...composed his way into the hall of fame
- ...narrated his way to an Oscar
- ii *consumed his way out the door.
- iii *discarded his way to Buddhism
- iiii *completed his way through the degree
- iv *reimbursed his way to a healthy bank balance.

The explanation presented in the text, which involves a simple morphological constraint on v-insertion, will need to be revised to account for the above.

h' Tom recounted the story to Ben.

This restriction shows up as early as it can be reliably investigated. Gropen et al. (1989) demonstrated that this morphophonological restriction on dative shift shows up in both adults and children (4-5 yrs old) using made-up verbs in an experimental context. The kids were taught novel ditransitive verbs that had either 'Latinate' or 'Anglo-Saxon' shapes, where the verb was introduced to them only in a *to*-dative frame. Then the verbs were elicited from the kids in context, manipulated to favor a double-object production (i.e. with a familiar, topical Goal argument). The kid's utterances were recorded, and the number of extensions to a novel double-object frame were tabulated for each verb. Anglo-Saxon verbs were extended significantly more frequently to the double-object frame than the Latinate verbs were, even in children. The children already were able to innovate novel double object constructions, and able to avoid doing so with Latinate verbs.

6 The nature of the restriction

Of course, "Latinate" is not the correct generalization, if it is understood to refer to etymological origin. Many of the verbs in the verb-particle or resultative examples above are in fact historically borrowed from a Romance language, e.g. *finish* (in *finish up*), *train* (in *train yourself fit*) and even *dance* (in *dance yourself pink*). One much commented-upon crucial factor is their morphophonological structure, in particular, their prosody. Many of the constrained verbs exhibit a weak-strong stress pattern.

It is clear that something about the complex prosodic structure of Latinate verbs is involved in this phenomenon. If the initial weak syllable of a Latinate verb may be clipped, its occurrence with a particle becomes fine: **confess up* is bad, but '*fess up*' is perfectly idiomatic. Similarly, although the canonical judgment is that *donate* may not occur in a double object frame (i.e. that **John donated the library the books* is illformed), many of my students accept it in this frame-with the stress pattern [\gg downejt], rather than [dow \gg nejt] (both patterns are listed in the Oxford American Dictionary).

Most researchers are agreed that some morphophonological restriction on dative shift exists, but most are mute about why such a restriction should hold for *this particular syntactic operation*. For instance, Grimshaw (1986) proposes a prosodic account of the dative shift restriction, according to which words which undergo dative shift must consist of a single prosodic foot. But no explanation is offered as to *why* such a constraint should hold for the double object construction, particularly. Further, both the Latinate and prosodic-foot accounts are leaky, as argued by Boguraev (1989).

Pesetsky (1995) suggests a purely morphotactic solution, arguing that the double object structure involves a special affix with special morphophonology. He claims that the double object structure involves affixing a null preposition which he calls G to the main verb, while the *to*-structure involves no such affixation. (Hence, *John gave Mary the book* is really *John gave-G Mary the book*, while *John gave the book to Mary* just is as it seems, morphologically.) Morphophonological restrictions on affixation are not uncommon (Pesetsky cites the famous syllable-counting and stress-sensitive nature of the comparative *-er*

affix in English), so he proposes that *G* is subject to affix-specific morphophonological restriction on the roots to which *G* can attach, namely roots with typically Anglo-Saxon prosodic structure, not Latinate.

One could wonder whether a null morpheme should be affected by the morphophonology of the root to which it affixes. Furthermore, for Pesetsky's account to be completely general, the special *G* morpheme would have to be involved not only in the double-object construction, but also in the verb-particle construction.

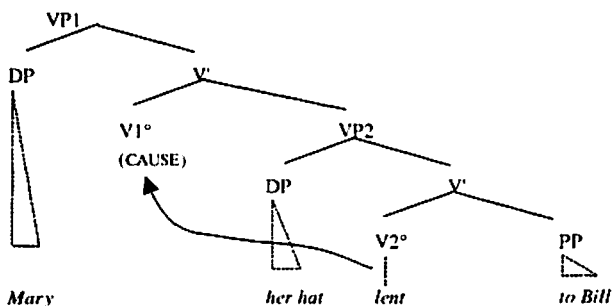
I propose here that this is a morphosyntactic effect, rather than a morphophonological one. In preliminary support of this idea, consider the case of *clarify*, which may not occur with particles (**clarify up the situation*) or in dative shifted frames (*clarify the situation to/for Bill* but not **clarify Bill the situation*). This verb doesn't fit the prosodic WS pattern of other restricted verbs, but it is morphologically complex. My claim is that if the relevant Latinate verbs are actually treated as synchronically morphologically complex, in the same way that verb-particle constructions clearly are, then their failure to occur with particles and in the double-object construction is simply due to the fact that their morphologically bipartite structure prevents them from being inserted straight into v^o as a causative root, which is what would be necessary for them to occur with particles and in the double-object construction.

In order to understand this proposal, it will be briefly necessary to reconsider the structure of ditransitive verbs, and adopt a specific hypothesis about the nature of dative shift.

7 The structure of ditransitive verbs

Let us revisit the structure for a to-dative, based on Larson's split-VP proposal, that we considered above:

(46)



Here, in the canonical frame, the verb originates in the lower part of the VP structure, as the head of a (complex) SC predicate, *(be) lent to Bill*. It head-moves to the null causative v^o in the upper part of the structure to generate.

Larson derives the double-object frame from the one above by a 'passivization' operation applied to VP2. However, his account runs into a number

of fairly subtle semantic problems, notably the famous animacy constraint on the double object frame. In the *to*-dative frame, the Goal argument can be animate or inanimate. However, in the double object frame, the Goal argument must usually be animate, and most of the inanimate Goals that are possible in the *to*-dative frame are peculiar, as illustrated in (47)-(48) - Oehrle's (1976) generalization; see also Green(1974):

- (47) a The editor sent the article to Sue.
 b The editor sent the article to Philadelphia.
 c The editor sent Sue the article.
 d ??The editor sent Philadelphia the article.
- (48) a Susan sent Harry to Max/down the hall/to his room/away.
 b Susan sent Max/*the hall/*his room/*away Harry.
 c Susan kicked the ball to Max/down the hall/out the window/upward.
 d Susan kicked Max/*the hall/*upward/*the window the ball.

Further, there is a subtle implication of possession present in the double object frame that is absent in the *to*-dative construction. In (49a), there is an implication that the students actually learned some French, while that is absent in (49b). In (50a), there is a presupposition that *the baby* actually already exists - this sentence would be slightly odd if uttered by a woman only intending to become pregnant, e.g. - while (50b) does not have this flavor.

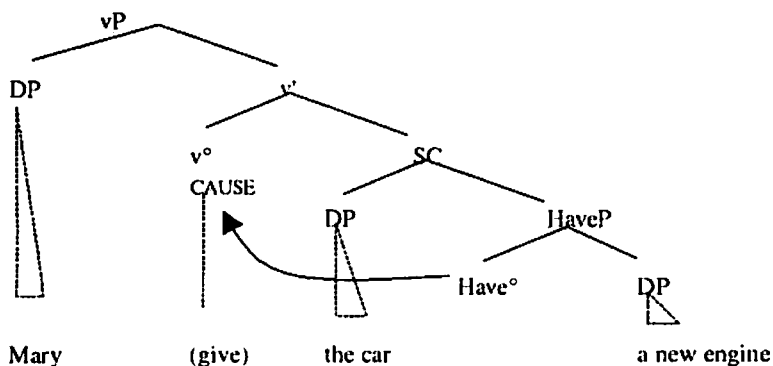
- (49) a John taught the students French
 b John taught French to the students
- (50) a I knitted this sweater for our baby.
 b I knitted our baby this sweater.

The notion that 'possession' is somehow involved in the double object construction is supported by the fact that the only inanimate Goals permitted in this construction are those which inalienably possess the Theme argument. This constraint is identical to the inalienable possession constraint on inanimate subjects of *have* in straightforward possession expressions of English (see Belvin, 1996): An inanimate subject can 'have' its inalienable parts - things to which it is meronymically related - but not anything else. Animate subjects, on the other hand, can 'have' both inalienably and alienably; this same constraint seems to be in effect in the double object construction. In (51a), an inanimate possessor is grammatical as the subject of *have* with an inalienably possessed object, but in (51b), the same possessor is ungrammatical with an alienable object. (In (51b), we see that *have* can be used as long as a locative coindexed PP is added to the expression, turning it into a location, rather than a possessor). In (52), we see that the same constraint is true of the double object construction: if the Goal object is inanimate, the Theme must be inalienably possessed by it (i.e. in a part-whole relation to it); non-inalienable Theme objects trigger the animacy constraint:

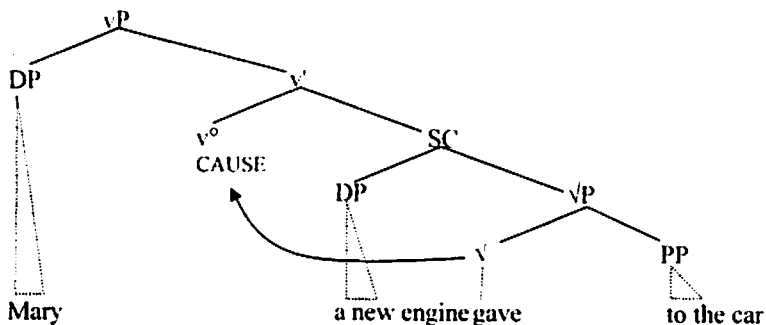
- (51) a The car has an engine.
 b # The car has a sweater.
 c The car has a sweater in it.
- (52) a Mary gave the car a new engine. (used, e.g., if she's a mechanic)
 b # Mary gave the car a sweater

These facts, plus a few other more involved arguments, have suggested to many researchers that the double object construction is actually a causative of a possession relationship (see, among others, Richards, 2001; Harley, 2002). That is, the structure for a double object construction, rather than being derived from the *to*-dative construction in (46), is really based on the syntactic structure in (53) below. In this structure, the predicate heading the SC is an abstract 'HAVE' relation. Following Kayne (1993) and Freeze (1991), I assume that it is a prepositional relation that is a subconstituent of the verb *have* in English (Kayne and Freeze argue that 'have' is the result of incorporating a possession P into the copula BE). In the double object construction, the verb gets its content in the same way as in the verb-particle construction and resultative construction above: a verb is inserted into v° as an expression of the manner of causation of the transfer of possession.

(53)



That is, in the double object construction, the verb is a manner element, exactly as in the verb-particle and resultative and motion constructions discussed above. This contrasts with the *to*-dative construction, in which the verb is itself the contentful source of the predicative material in the small clause, as (re)-illustrated below:



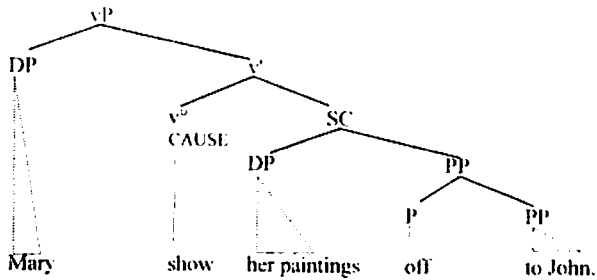
Now we can see what the double object construction has in common with the resultative and the verb-particle construction: in the double-object construction, but not the to-dative, the verb root is in the 'high' part of the bipartite structure, identifying the manner of causation.

8 The failure to alternate

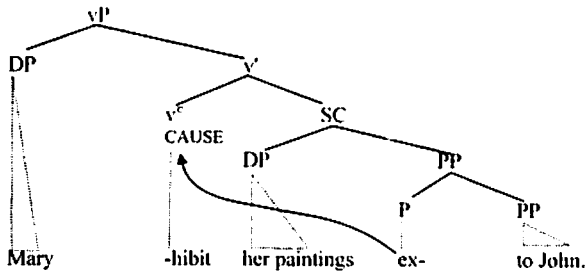
The generalization we have arrived at so far, then, is that the Latinate verbs may not be inserted as manner elements into v° , while Anglo-Saxon verbs can be so inserted. But we thus far have not understood why this would be.

I argue that it is because the Latinate verbs are bimorphemic. Each of the two morphemes in a verb like *exhibit* is associated with a particular part of the bipartite verb structure. In particular, these verbs, historically composed of a prepositional particle prefixed to an independent verb root, are in fact *synchronically* so composed, in the minds of English speakers. That is, Latinate verbs are fundamentally verb-particle constructions, where a) the particle mandatorily incorporates into the v° and b) the whole is interpreted idiomatically, as so many verb-particle constructions are. The verb roots and particles are *cran-morphs*, but morphemes nonetheless, each associated with a terminal node of the syntactic structure. Marantz (2001) comes to this conclusion based on independent evidence. The structure of *Mary exhibited her paintings (to John)*, then, will be exactly the same as the structure of *Mary showed her paintings off (to John)*, with the additional requirement that the particle *ex-* incorporate into the bound verb stem in v° , *-hibit*. The structures are illustrated below:

(55)



(56)



In other words, in the Latinate verbs, the head of the lower part of the bipartite structure is built right into the verb, in the form of a separate morpheme (in this case, *ex-*). I assume that this morpheme, like an English particle, is specified for its category, so that when it appears, it entails the appearance of its particular P projection. It cannot be used within a 'manner' modifier of a causative *v^o*, since its appearance forces the projection of the resultant state SC headed by the PP headed by the particle. Hence such Latinate verbs cannot appear in double object constructions, or in verb-particle or resultative constructions.

Verb-particle constructions with resultative meanings should thus also fail to occur in double object constructions, for the same reason, and the literature generally concurs that this is the case, although scattered counter-examples turn up now and then. Certainly there is a contrast between (57a) and (57b) below, and between (57c) and (57d); I assume that the relative degradation of the double object construction, when the particle appears, submits to the same explanation.

- (57) a Mary showed John her paintings.
 b *Mary showed off John her paintings. (cf. Mary showed off her paintings)
 c Mary passed the kids the potatoes.

d *Mary passed around the kids the potatoes. (cf. Mary passed around the potatoes)

Finally, the account for the failure of *clarify* in verb-particle constructions will follow the same lines, with the twist that in the case of this change-of-state verb, the root is the small clause predicate (*clar-*), while the verbalizing suffix *-ify* is associated with the *v*^o head. The prediction is the same, however: the complex structure, with subparts associated both with the upper and lower portions of the bipartite verb phrase, means that the whole verb may not be inserted as a simple causation modifier; the bipartite structure will be parsed and the subparts assigned to the upper and lower parts of the VP, preventing the appearance of a particle in the lower part of the VP (in the case of a verb-particle construction) or the appearance of null *Have* there (in the case of the double-object construction).

9 Synchronic segmentation?

But is it reasonable to suppose that English speakers aggressively segment words into morphemes whose semantics are not at all compositional or introspectively apparent? Hammond (1999) argues for extremely aggressive morphological segmentation, on the basis of phonological distribution data. Further, he suggests (in Hammond, 2000) that such segmentation is the only explanation for the word-internal phonotactic violations you find in Latinate verbs like *adhere*: if there's a morpheme boundary between *ad* and *here*, then the word is well-behaved according to English phonotactics, but if there isn't, an otherwise very clear generalization about word structure fails. Furthermore, a line of masked-priming work begun with Taft and Forster (1975) shows evidence for morphological priming ('affix stripping') across Latinate words with shared roots, such as *exhibit* and *inhibit*; this effect can be most readily understood if speakers do in fact decompose these forms into their subconstituents, despite their semantic opacity.

Considering things from the learner's point of view, then, there are several cues in the speech stream that these verbs are made up of more than one part:

(58) Synchronic motivations for decomposition

- a phonotactic clue (like *adhere*)
- b phonological clue: (word-final stress)
- c familiar with the morpheme from elsewhere (like *deceive* / *receive* / *conceive*, *clar-ify*...)

To confirm the plausibility of this account (as well as to test the empirical generalization, I consulted the lists of verbs in Levin (1999) that are listed as occurring in both double-object and the *to*-dative construction, as well as those which only occur in the *to*-dative (nonalternators). If these synchronic clues are adequate to sort verbs into monomorphemic and bimorphemic classes, then the alternators should exhibit none or at most one of the traits listed in (58) above, while the non-alternators should exhibit two or more. In (59) I provide Levin's complete list of alternating verbs as she gives them, sorted by

semantic class. (These are the ones which should be purely 'Anglo-Saxon' in character.) Potential counterexamples are in boldface, and addressed in the discussion below. There are a few odd ones that according to my judgments really shouldn't alternate; these are followed by question marks.)

(59) Alternating verbs, sorted by semantic class (from Levin, 1999)

sending : forward, hand, mail, post, send, ship, slip, smuggle, sneak

giving : feed, give, lease, lend, loan, pass, pay, peddle, **refund**, render, rent, **repay**, sell, serve, trade

throwing : bush, bat, bunt, **catapult**, chuck, flick, fling, flip, hit, hurl, kick, lob, pass, pitch, punt, shoot, shove, slam, slop, sling, throw, tip, toss

telling : ask, cite, pose, preach, quote, read, relay, show, teach, tell, write

instrument : cable, email, fax, modem, phone, radio, relay, **se-maphore(??)** **satellite (??)**, sign, signal, **telephone**, **te-lecast**, **telegraph**, telex, wire, **wireless(??)**

All the non-boldfaced verbs are clearly Anglo-Saxon in terms of stress pattern, and most are clearly monomorphemic. All the boldfaced cases except the re- ones are multimorphemic, hence their boldfacing, but are a) zero-derived from nouns or adjectives and b) have initial stress. These should pose no problem for the learner.

It seems likely that the 're' cases are actually not a problem, either, despite having a weak-strong stress pattern. These re- morphemes, unlike the re- in receive or recite, are attached to verb roots which are actually themselves free verbs in English, *fund* and *pay*-in short, these words are compositional. English speakers know that this re- morpheme means "again" or "reverse", and composes compositionally with the verb root. Consequently, they don't assign it to the P category, but rather to an adverbial position which is free to occur wherever these roots may occur.

Next I provide Levin's complete list of ditransitive non-alternators. These are the ones that should be multimorphemic, exhibiting one or more of the determining characteristics of Latinate particle verbs listed above in (58). Again, there are a few potential problem cases, highlighted in bold.

(60) Non-alternating "Latinate" verbs.

address, administer, **broadcast**, convey, contribute, delegate, deliver, denounce, **demonstrate**, describe, donate, elucidate, exhibit, express, explain, **forfeit**, **illustrate**, introduce, **narrate**, portray, **proffer**, recite, recommend, refer, reimburse, remit, restore, return, **sacrifice**, submit, surrender, transfer, transport.

Among the group, the boldfaced verbs have initial stress, so they would perhaps be expected to alternate. However, some seem clearly to be identifiably multimorphemic on the basis of their recurring subparts alone (e.g. the *-ate*

verbs, the compound *broad-cast*, the bipartite *for-feit*, with pieces potentially recognizable from *forgive*, *forbid*, *surfeit*, *counterfeit*). The two remaining counterexamples are more problematic, although since these lists are based on corpus data, non-occurrence of the double object frame does not necessarily indicate ungrammaticality. For what it's worth, in my judgment *proffer* is equally acceptable or odd in both the *to*-dative and double object frames - it more normally occur without any Goal argument at all, to my ear. *Sacrifice* is more puzzling, though its denominal character may be important.

Levin isolates one group of alternating verbs which poses a real problem for the whole account, which she characterizes semantically as 'verbs of future having'; the complete list is in (61) below:

(61) **A Class of Problematic Alternating Verbs**

advance, allocate, **allot**, **assign**, award, **bequeath**, cede, **concede**, **extend**, grant, guarantee, issue, leave, offer, owe, promise, refuse, vote, will, yield.

In this list, the non-boldfaced verbs are not a problem because they're either not Latinate, or from nouns, or have initial stress. However, non-underlined, boldfaced ones are the real problems; as far as I can tell they are genuine counterexamples. It's quite puzzling how they all cluster in this semantic class. It would be very easy to tell a semantic story about why these *shouldn't* alternate, that is, why they should all be double-object-only class, and then they'd easily be accounted for as necessarily containing forms of Have. But in fact, they do all alternate, and they're probably multimorphemic. What's needed is a story of why these double-object ones should all be happy as *to*- forms but not vice versa. For the moment, these will have to be left as an issue for future research.

10 Conclusions

In short, a cross-linguistically valid bipartite verb structure is supported by evidence from several independent converging sources. It can shed light on several puzzling morphological patterns, including the restriction on re-affixation with respect to resultative and particle constructions, and the constraint against Latinate verbs occurring in double object, resultative, and particle constructions. Many questions remain. For example, the prohibition does not extend to Latinate verbs occurring as manner elements in unselected-object motion causatives, although typologically the failure of certain languages to exhibit that property as well as all three of the others (double objects, verb-particle constructions and resultatives) suggests that they should all be subject to the same restriction both within and across languages. Another issue is the fact that the account proposed here concerning Latinate verbs in double object constructions breaks down for several examples in a particular semantic subclass; the behavior of the Latinate verbs in this verb class really has yet to be grappled with.

References

- [1] BARSS, A., AND LASNIK, H. A note on anaphora and double objects. *Linguistic Inquiry* 17 (1986), 347-354.

- [2] BECK, S., AND JOHNSON, K. Double objects again. *Linguistic Inquiry* 35, 1 (2004).
- [3] BELVIN, R. S. Inside events: the non-possessive meanings of possession predicates and the semantic conceptualization of events. Master's thesis, University of Southern California, Los Angeles, 1996.
- [4] CHOMSKY, N. *The Minimalist Program*. MIT Press, Cambridge, MA, 1995.
- [5] FOLLI, R., AND HARLEY, H. On the licensing of causatives of directed motion: Waltzing matilda all over. *Studia Linguistica* 60, 2 (2006), 1-35.
- [6] FOLLI, R., HARLEY, H., AND SIMIN KARIMI. Determinants of event structure in persian complex predicates. *Lingua* 115, 10 (2005), 1365-1401.
- [7] FREEZE, R. Existentials and other locatives. *Language* 68, 3 (1992), 553-595.
- [8] GREEN, G. M. *Semantics and Syntactic Regularity*. Indiana University Press, Bloomington, 1974.
- [9] GRIMSHAW, J. Prosody and dative shift? Master's thesis, Rutgers University, 1986.
- [10] GROPEN, ET AL. The learnability and acquisition of the dative alternation in english. *Language* 65, 2 (1989), 203-257.
- [11] HALE, K., AND KEYSER, S. J. *The view from Building 20: essays in honor of Sylvain Bromberger*. MIT Press, Cambridge, 1993. ch. On argument structure and the lexical expression of syntactic relations.
- [12] HALE, K., AND KEYSER, S. J. *Prolegomena to a theory of argument structure*. MIT Press, 2002.
- [13] HAMMOND, M. *The phonology of English*. Oxford University Press, Oxford, 1999.
- [14] HARLEY, H. *Subjects, Events and Licensing*. PhD thesis, Massachusetts Institute of Technology, Cambridge, 1995.
- [15] HARLEY, H. *The Linguistic Variation Yearbook*, vol. 2. John Benjamins, Amsterdam, 2004, ch. Possession and the double object construction, pp. 29-68.
- [16] HOEKSTRA, T. Small clause results. *Lingua* 74 (1988), 101-39.
- [17] JACOBSEN, W. M. *Transitivity in the Japanese Verbal System*. PhD thesis, University of Chicago, 1981.
- [18] JACOBSEN, W. M. The transitive structure of events in japanese. *Studies in Japanese Linguistics* 1 (1992).
- [19] JELINEK, E. *The projection of arguments: lexical and compositional factors*. Palo Alto: CSLI, Stanford University, 1997, ch. Voice and transitivity as functional projections in Yaqui.

- [20] JELINEK, E., AND ESCALANTE, F. *Uto-Aztecan: structural, temporal and geographic perspectives*. Hermosillo: Universidad Autónoma de Sonora, 2001, ch. Unergative and Unaccusative Verbs in Yaqui.
- [21] KAYNE, R. Towards a modular theory of auxiliary selection. *Studia Linguistica* 47 (1993).
- [22] KEYSER, S. J., AND ROEPER, T. The abstract clitic hypothesis. *Linguistic Inquiry* 23, 1 (1992), 89-125.
- [23] KOOPMAN, H., AND SPORTICHE, D. The position of subjects. *Lingua* 85 (1991), 211-258.
- [24] KRATZER, A. *hrase Structure and the Lexicon*. P. Dordrecht: Kluwer, 1996, ch. Severing the external argument from its verb, pp. 109-137.
- [25] LARSON, R. K. On the double object construction. *Linguistic Inquiry* 19, 3 (1988), 335-391.
- [26] LEVIN, B., AND RAPPAPORT-HOVAV, M. Two structures for compositionally derived events. In *SALT* (Ithaca, NY, 1999), Cornell Linguistics Circle Publications, Cornell University, pp. 199-223. Santa Cruz.
- [27] MARANTZ, A. *On the Nature of Grammatical Relations*. MIT Press, Cambridge, MA, 1984.
- [28] MARANTZ, A. No escape from syntax: Don't try a morphological analysis in the privacy of your own lexicon. In *Proceedings of the 21st Annual Penn Linguistics Colloquium* (1997), D. A. and S. L. et. al., Eds., vol. 4.2, University of Pennsylvania Working Papers in Linguistics, pp. 201-225.
- [29] MATEU, J. Syntactically-based lexical decomposition: the case of climb revisited. In *Meeting of the Berkeley Linguistics Society* (2001), vol. 26. Paper presentation.
- [30] MCCAWLEY, J. *Regional Meeting of the Chicago Linguistics Society*. University of Chicago, 1968, ch. Lexical insertion in a transformational grammar without Deep Structure.
- [31] MCCAWLEY, J. D. On identifying the remains of deceased clauses. *Language Research* 9 (1974), 73-85.
- [32] MCINTYRE, A. *The German dative, HAVE, and the syntactic decomposition of verbs*. University of Leipzig, 2003. Available: <http://www.uni-leipzig.de/angling/mcintyre>.
- [33] MEGERDOOMIAN, K. *The status of the nominal in Persian complex predicates*. In press.
- [34] OEHRLE, R. T. *The grammatical status of the English dative alternation*. PhD thesis, Massachusetts Institute of Technology, Cambridge, 1976.
- [35] PESETSKY, D. *Zero Syntax: Experiencers and Cascades*. The MIT Press, Cambridge, MA, 1995.

- [36] RICHARDS, N. An idiomatic argument for lexical decomposition. *Linguistic Inquiry* 32, 1 (2001), 183-192.
- [37] ROSS, J. R. *Linguistic and literary studies in honor of Archibald A. Hill*, vol. 1. The Hague: Mouton, 1976, ch. 'To have 'have' and to not have 'have'', pp. 263-270.
- [38] SNYDER, W., AND STROMSWOLD, K. The structure and acquisition of english dative constructions. *Linguistic Inquiry* 28:2 (1997), 281-317.
- [39] TAFT, M., AND K. FORSTER, K. Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior* 14 (1975), 638-647.
- [40] TALMY, L. *Language Typology and Syntactic Description: Grammatical Categories and the Lexicon*, vol. 3. Cambridge: Cambridge University Press, 1985, ch. Lexicalization patterns: semantic structure in lexical forms, pp. 57-150.
- [41] VON STECHOW, A. *Lexical Knowledge in the Organization of Language*. Amsterdam: John Benjamins, 1995, ch. Lexical decomposition in syntax, pp. 81-118.

Prosodic complexity and processing complexity: evidence from language impairment

John Harris, Nichola Gallon and Heather van der Lely
University College London

johnh@linguistics.ucl.ac.uk

Resumo

Este artigo discute se uma concepção de complexidade prosódica faz previsões válidas sobre o grau de complexidade ou dificuldade em processamento fonológico on-line. Através de métodos experimentais o autor considera a influência da complexidade métrica e interna à sílaba no processamento fonológico. Os resultados oferecem evidência de que a complexidade métrica e interna à sílaba são relevantes no processamento fonológico independente do número de sílabas ou de segmentos envolvidos.

1 Introduction

Given two related speech sounds, it can usually be established on the basis of cross-linguistic comparison that one is more favoured or less marked than the other. For example, a front nonround vowel can be judged less marked than a front round counterpart on the grounds that all languages have vowels of the first type whereas only some languages also have vowels of the second type.

Two widespread assumptions are often made regarding markedness and its relation to phonological representation and on-line phonological processing. First, more marked sounds are understood as being phonologically more complex than their unmarked counterparts, in the sense that they bear more feature specifications. Second, the phonological complexity of a sound stands in direct relation to the processing effort required to produce and perceive it in speech.

The link between complexity and markedness cannot be directly extended to prosodic structure. This is because we often find apparently complex syllabic or metrical structures that are actually less marked than apparently simpler counterparts. For example, although a binary branching foot is on the face of it representationally more complex than a degenerate (non-branching) foot, it is nevertheless less marked.

* Authors' UCL affiliations: Harris (Department of Phonetics and Linguistics); Gallon and Lely (UCL Centre for Developmental Language Disorders and Cognitive Neuroscience). The study featured in §5 of the paper is based on a much fuller presentation in Gallon, Harris and Lely (2007). Versions of the present paper were presented at Miyagi Gakuin University (Sendai), UC Berkeley Linguistics Colloquium and V Congresso Internacional da Associação Brasileira de Linguística (Belo Horizonte). Thanks to participants who provided valuable comments at each of these meetings.

The link between complexity and markedness in prosody needs to be understood as being mediated by constraints on the canonical syllabic and metrical shapes of different types of morpheme. One of these constraints requires lexical heads to be prosodically heavy, that is to branch, with the result that words are minimally bimoraic or bisyllabic. What then contributes to the prosodic complexity of a word is not branchingness *per se* but rather any structure that deviates from this minimal binary shape.

Does this conception of prosodic complexity make valid predictions about degrees of complexity or difficulty in on-line phonological processing? This paper presents evidence bearing on this question. The evidence derives from an English non-word repetition experiment designed to illuminate the relative influence of syllabic and metrical complexity on phonological processing. One group of subjects had previously been identified as presenting with a grammatical type of specific language impairment. The reason for including these subjects in the study was to investigate whether a deficit in processing structural complexity in morphology and syntax might extend to prosodic phonology. As will be seen below, the results support the conclusion that metrical and syllable-internal complexity can affect repetition accuracy independently of the number of segments or syllables in a non-word.

§2 discusses the relation between prosodic complexity and constraints on canonical morpheme shape. §3 briefly reviews what is known about the phonological aspects of specific language impairment. §4 describes a non-word repetition test in which stimuli are systematically varied in ways that allow us to investigate the impact of syllabic and metrical complexity on phonological processing. The test provides the stimuli for the study that is presented in §5. §6 summarises the main conclusions.

2 Prosodic complexity and canonical morpheme shape

In segmental phonology, there is a long tradition of linking markedness with representational complexity (Chomsky; Halle, 1968; Trubetzkoy, 1939 *et passim*). The relative markedness of a given segment is established initially on the basis of its distribution within and across different languages. An unmarked segment is one that enjoys a wider distribution than a marked counterpart. Forging a link between markedness and representational complexity has usually been achieved by assuming that a marked segment bears more feature specifications than an unmarked counterpart (Archangeli; Pulleyblank, 1994; Chomsky; Halle, 1968 *et passim*).

The marked nature of more complex representational entities is often assumed to correlate with some notion of functional complexity or difficulty. For example, compared to an unmarked counterpart, a marked segment is acquired later and places a heavier burden on articulatory effort, auditory perception or phonological processing.

A moment's reflection shows that the link between complexity and markedness does not obviously carry over into the prosodic hierarchy, especially when we compare syllabic and metrical structure. For example, while a binary branching syllable onset is both more marked and more representationally complex

than a non-branching onset, a binary branching foot is less marked than a degenerate foot. To put it concretely: while the two-consonant cluster in *pra* is more marked than the single consonant in *pa*, the bisyllabic foot constituted by *páta* is less marked than the monosyllabic foot *pa*.

The apparent mismatch between complexity and markedness in prosodic structure can plausibly be attributed to a strong tendency for prosodic structure to mirror morphological structure in certain ways (Downing, 2006; Inkelas; Zoll, 2005). The structural homology between the two domains can be understood as resulting from the action of two grammatical constraints, set out in (1) (formulated in terms based on Downing (2006), Dresher and Hulst (1998), Russell (1997)).

- (1) a MORPHEME-SYLLABLE: morphemes are coextensive with syllables.
- b HEADS BRANCH: lexical heads (roots/stems) must prosodically branch.

There is a clear cross-linguistic preference for lexical-category morphemes to be prosodically heavier than functional-category morphemes. This is expressed in the constraint HEADS BRANCH and is evidently related to the fact that lexical heads generally carry greater semantic weight. Moreover, being open-class items, they need to be phonologically more complex in order to maintain lexical distinctiveness.

Acting alone or in unison, the constraints in (1) derive the canonical prosodic shapes of different types of morphology. An affix is canonically monosyllabic (by MORPHEME-SYLLABLE) and monomoraic (i.e. non-branching). A root is also canonically monosyllabic (by MORPHEME-SYLLABLE) but bimoraic (branching, by Heads Branch). A stem (root plus affix) is minimally bisyllabic (by both MORPHEME-SYLLABLE and HEADS BRANCH).

We now have a motivation for saying that branchingness in and of itself is not enough to render a prosodic structure marked. It only counts as marked if it violates constraints on canonical morpheme shape. Take for example stems, which for our present purposes we may consider coterminous with words. Prosodic branching does not contribute to complexity when it provides a word with its minimal shape. In languages such as English, this minimal structure coincides with the stress foot. What does contribute to complexity is any structure that exceeds this minimum. For example, the word *banana* in English counts as prosodically complex under this definition. The complexity does not arise from the branching structure of the foot *nána*, since that satisfies the minimal word requirement. Rather it arises from the presence of the unstressed syllable *ba*, since the structure attaching this to the word results in the minimal shape being exceeded.

A question arises as to whether there is a correlation between prosodic complexity as defined in these terms and complexity or difficulty in phonological processing. This paper reports the results of a non-word repetition experiment that was specifically designed to answer this question. One group of subjects chosen for the study had previously been identified as presenting with specific language impairment (SLI). One of the initial motivations for including this

group came from earlier studies which indicate that, compared to the developmental norm, individuals with SLI often show a deficit in processing phonologically complex forms (see the references below). However, complexity in these studies is largely calculated on the basis of the number of segments or syllables in a word, with no account taken of the kind of prosodic factors being discussed here.

In the present study, non-word stimuli were systematically varied along three prosodic parameters, three syllabic and two metrical, each of which expresses a binary opposition between an unmarked and a marked structure. The parameters are: branching versus non-branching onset; open versus closed rime; word-final vowel versus consonant; presence versus absence of an unstressed syllable left-adjoined to the word; presence versus absence of an unstressed syllable right-adjoined to the word.

3 Phonological aspects of Specific Language Impairment

SLI is a disorder of language acquisition in an otherwise typically developing child (Leonard, 1998). Children with SLI are reported to have difficulty in producing and repeating polysyllabic words compared with language-matched controls (Bishop et al., 1996; Botting; Conti-Ramsden, 2001; Gathercole; Baddeley, 1990). It has been claimed that this difficulty can be attributed to a short-term memory deficit: either the capacity of the phonological store is unusually small (for example, limited to one or two syllables), or the contents of the store decay unusually rapidly (Gathercole; Baddeley, 1990). Based on this hypothesis, a speech assessment tool, the Children's Test of Non-word Repetition (CNRep, Gathercole; Baddeley, 1996), has been designed specifically to detect a deficit in phonological short-term memory. It is a production test in which subjects repeat non-word stimuli that are varied along a single dimension of complexity defined in terms of number of phonemes or syllables.

One drawback of CNRep is that it takes no account of the potential impact of prosodic factors on repetition accuracy (see the comments in Dollaghan et al. (1995), Lely and Howard (1993), Sahlen et al. (1999)). Moreover, it has been shown that not all children with SLI have short-term memory deficits (Lely; Howard, 1993). This strongly suggests that other factors are in play when difficulties emerge in non-word repetition.

Consider the processing tasks involved in repeating a non-word stimulus (Lely; Howard, 1993; Snowling et al., 1991): a phonological representation of the non-word input must be constructed; the representation must be maintained in phonological memory; the non-word must be reconstructed in speech output. In interpreting poor performance results in non-word repetition tasks, we thus need to pay very close attention to the nature of the phonological representations that are being manipulated by the subject.

The group of speakers with SLI included in the study reported on below present with a specific form of the disorder known as grammatical SLI (G-SLI), characterised as a relatively discrete, persisting core deficit in syntax and morphology (Lely *et al.*, 1998). None of the subjects suffered from any articulation problems, and their non-verbal IQ was within normal limits (Bishop

PARAMETER	SETTING	DESCRIPTION	REAL WORD	NON-WORD
Onset	Unmarked	No onset cluster	city	pifi
	Marked	Onset cluster	pretty	pifi
Rime	Unmarked	Open syllable	city	pifi
	Marked	Closed syllable	filter	pilfi
Word-end	Unmarked	V-final	city	pifi
	Marked	C-final	sit	pif
Left-adjunction	Unmarked	Initial footed syllable	city	pifi
	Marked	Initial unfooted syllable	banana	spifi
Right adjunction	Unmarked	Final footed syllable	city	pifi
	Marked	Final unfooted syllable	Canada	pifit.

Tab. 1: The five prosodic parameters in the Test of Phonological Structure.

et al., 2000). The main purpose of the study was to investigate whether a G-SLI impairment in processing morphosyntactic complexity extends to prosodic complexity. The data for the study are drawn from a non-word repetition task known as the Test of Phonological Structure, in which stimuli are systematically varied along the five prosodic parameters outlined above.

4 Test of Phonological Structure

The major typological characteristics of syllabic and metrical structure in a given language can be described in terms of a relatively small number of binary parameters. For each parameter, one of the settings is marked relative to the other, as determined by the asymmetric distribution of the relevant structures across the world's languages. Each of the preferences revealed in these asymmetries can be expressed in terms of a unidirectional implication. For example, single-consonant onsets can be identified as unmarked on the basis of the observation that they occur in all languages, while only some languages also have consonant-cluster onsets.

Five prosodic parameters of this type feature in the Test of Phonological Structure (ToPhS, Lely; Harris, 1999). The test assesses phonological abilities by requiring subjects to repeat non-word stimuli that are systematically varied along the three syllabic and two metrical parameters set out in Table 1. These prosodic permutations were applied to a set of four segmentally distinct base forms, each consisting of a CVCV structure with stress on the first syllable: **dépo**; **pifi**; **kéto**; **fipl**. This structure is maximally unmarked in terms of all five of the chosen parameters. The total number of -words in the ToPhS stimulus set is 128 (five binary parameters times four base forms).

The onset parameter controls whether an onset contains one consonant (unmarked) or two, e.g. **pifi** versus **prifi**. The rime parameter controls whether the stimulus has an open (unmarked) or a closed syllable, e.g. **pifi** versus **pilfi**. The word-end parameter controls whether the stimulus ends in a vowel (unmarked) or a consonant, e.g. **pifi** versus **pif**. The rime and word-end parameters are typically conflated in traditional accounts of English syllable structure. However, there are good cross-linguistic reasons for keeping them separate; see Harris (1994). Each of the stimuli contains one metrical foot, consisting either of a bimoraic monosyllable (e.g. **pif**) or a bisyllabic trochee (left-stressed, e.g. **pifi**).

Non-word Real word	dé̌pə city	dré̌pə pretty	dré̌mpə plenty	bə̌dré̌p suppress	bə̌dré̌mp deflect
Onset	0	1	1	1	1
Rime	0	0	1	0	1
Word end	0	0	0	1	1
Left adjunction	0	0	0	1	0
Right adjunction	0	0	0	0	0
Total marked	0	1	2	3	4

Tab. 2: The relative markedness of bisyllables in the Test of Phonological Structure.

The two adjunction parameters control whether or not an unfooted (and therefore unstressed) syllable is attached to the beginning or end of this foot, e.g. *pifi* versus *spifi* (left adjunction) and *pifitə* (right adjunction). The appearance of an unfooted syllable increases the metrical complexity of a word and is therefore the marked option.

An important design property of ToPhS is that it allows the prosodic complexity of words to be manipulated independently of the number of segments or syllables they contain. For example, *dé̌pə* is segmentally longer than *dɛp* but is prosodically less complex (per the word-end parameter). We are thus in a position to pinpoint the potential effect of prosodic complexity on test performance, independent of string complexity. This point is more fully exemplified in Table 2. This shows the varying complexity of bisyllabic non-words, which can contain anything from zero to four marked structures.

5 ToPhS non-word repetition study

5.1 Hypotheses

The ToPhS study discussed in this section draws on the much more detailed exposition presented by Gallon, Harris and Lely (2007). The study set out to test the following hypotheses. (i) Non-words with marked, complex prosodic structures are more difficult to repeat accurately than those with unmarked, simplex structures. (ii) The greater the number of complex prosodic structures a non-word contains, the greater will be the difficulty in repeating it accurately. (iii) Prosodic complexity influences repetition accuracy independently of string complexity.

5.2 Subjects and stimuli

The study compares three groups of children, matched according to language ability rather than chronological age. One group was composed of subjects presenting with G-SLI. The other two were composed of subjects with typically developing language (LA1, LA2). The aim of matching the G-SLI group to younger, typically developing children was to control for the effects of vocabulary and morphosyntactic abilities.

The G-SLI group was made up of 13 subjects (11 male, two female), with a mean chronological age of 15;8. Every individual in this group showed significant impairment on one or more standardised language tests. The non-verbal abilities of each individual was measured at or above average on standard IQ tests (British Ability Scales and Raven's Progressive Matrices). None showed deficits in articulation (such as dyspraxia) or psycho-social skills (such as pragmatic or attention deficits).

The language-ability control groups comprised 24 children drawn from a London primary school. These were evenly split into two age groups equally balanced by sex. All showed average verbal and non-verbal abilities as assessed by standard tests. A younger group (LA1), with an average chronological age of 5;6, was matched to the G-SLI group on the basis of morpho-syntactic ability. An older group (LA2), with an average chronological age of 7;4, was matched to the G-SLI group on the basis of vocabulary.¹

The stimuli for the study consisted of 96 audio-recorded ToPhS non-words produced by a female native English speaker. These were presented to subjects via headphones, once each, in an uninterrupted sequence and in random order. Each stimulus was separated from the next by a three-second silent interval, during which the subject was required to repeat the non-word. Subjects' responses were audio-recorded.

5.3 Errors and scoring

Two independent coders transcribed subjects' responses, reaching 94% word-by-word agreement. The initial analysis on which the discussion below is based consisted of a simple binary classification of each response as either correct (score of 1) or incorrect (0). At this stage, no account was taken of the number and nature of errors in each response (for a more detailed analysis of specific segmental and prosodic errors occurring in the same set of responses, see the case studies in Marshall *et al.* (2002)). However, to make the discussion below more concrete, it will be useful to look briefly at illustrations of some of the more common types of error. The examples given in (2) come from two subjects in the G-SLI group.

(2) Subject DS

	STIMULUS	RESPONSE		STIMULUS	RESPONSE
(a)	dɪfrɪpələ	dɪfɪpələ	(b)	dɪfrɪpl	frɪpl
	bədɛmpə	fədɛpə		kɛstələ	kɛstə
	dɪfɪmpl	dɪfrɪpl		dɪfrɪp	dɪfrɪp
				bədɛmpəri	pədɛmbri

¹The standardised tests used to match the G-SLI group to the control groups were: the Grammatical Closure subtest, Illinois Test of Psycholinguistic Abilities (a test of expressive morphology); the Test of Reception of Grammar (sentence understanding); the British Picture Vocabulary Scale (receptive vocabulary). For more details of the matching procedure, see Gallon, Harris and Lely (2007).

(3) Subject TF

	STIMULUS	RESPONSE		STIMULUS	RESPONSE
(a)	badrépa	drépa	(b)	badrépa	dadéfrí
	dífípa	fípa		badrémpa	badréfrí
	démpa	démpa			
(c)	klétala	čččala			
	spífíta	bátífata			
	dífímpala	díbrí ala			

Each of these examples contains some combination of errors involving syllabic structure, metrical structure and segmental content. Syllabic errors include the simplification of a complex onset or rime (see for example (2a) and (2b)). Metrical errors include weak syllable deletion (as in (2b) and (3a)) or refooting (evident in the introduction of a secondary stress in *dídríp* > *díyfríp* in (2b)). (3b) shows a combination of syllabic and metrical errors, while (3c) shows a combination of syllabic and segmental errors.

5.4 Results

5.4.1 Overall scores

Averaged scores for each of the three subject groups are presented in Figure 1. This shows the relation between the number of correct repetitions and the number of marked/complex prosodic structures per non-word. There are significant main effects for group, prosodic complexity and an interaction between the two (all $p < 0.001$, 3x5 ANOVA). That is, prosodic complexity, as measured by the number of marked structures per non-word, has a significant influence on the performance of all groups.

As suggested in Figure 1, the most dramatic effect of complexity on performance is seen in the G-SLI group, who found the task significantly harder than the two control groups. For the G-SLI group, there is a clear decrease in the number of correct responses as the number of marked structures increases.

Planned comparisons between groups across the different degrees of prosodic complexity revealed no significant differences between LA1 and LA2. In contrast, the G-SLI group performed significantly worse than both control groups at every level of complexity. (Results derived from the Games-Howell test showed the significance of the differences between the G-SLI group and the other two groups to be: $p=0.13$ (G-SLI versus LA1 on words containing one complex structure), $p=0.001$ (G-SLI versus each of LA1 and LA2 on words with two complex structures), $p=0.000$ (G-SLI versus each of LA1 and LA2 on words with three complex structures), and $p=0.000$ (G-SLI versus each of LA1 and LA2 on words with four complex structures).

5.4.2 Syllabic versus metrical complexity

Let us now separate syllabic complexity and metrical complexity and consider their relative influence on repetition accuracy. Figure 2 shows how the two types of complexity impact on the performance of the G-SLI and LA1 groups. An

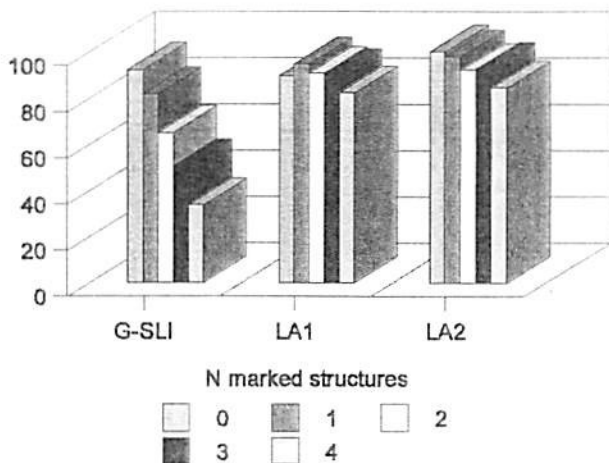


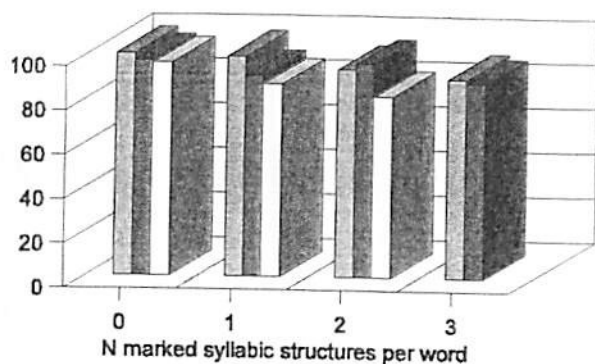
Fig. 1: Mean % correct responses per number of marked prosodic structures per ToPhS non-word.

analysis by group, syllabic complexity and metrical complexity (3x4x3 ANOVA) shows significant main effects for all three factors (all $p < 0.001$). There are also significant two-way interactions for group and syllabic complexity ($p < 0.01$), group and metrical complexity ($p < 0.001$), and syllabic and metrical complexity ($p < 0.001$).

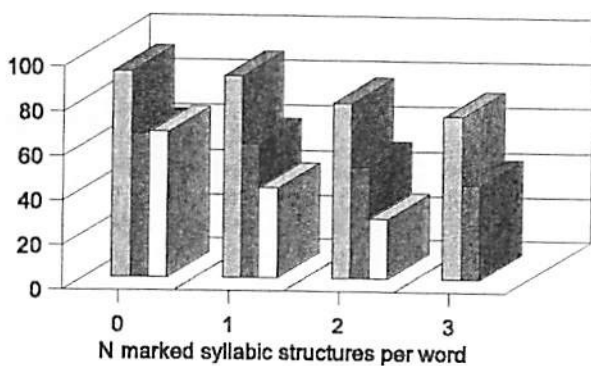
In order to investigate the interaction between syllabic and metrical complexity in more detail, a series of analyses (one-way ANOVAs) was performed to investigate what happens when (i) metrical complexity is increased while syllabic complexity is held constant and (ii) syllabic complexity is increased while metrical complexity is held constant.

First, what effect does increasing metrical complexity have at each level of syllabic complexity? When words contain no complex syllable structures, the metrical effect is significant only for the G-SLI group ($p < 0.05$). With more complex syllabic structures, there are significant metrical effects for all three groups (G-SLI $p < 0.001$, LA1 $p < 0.05$, LA2 $p < 0.05$ for one complex syllabic structure; G-SLI $p < 0.001$, LA1 $p < 0.01$, LA2 $p < 0.05$ for two complex structures).

Now, what is the effect of increasing syllabic complexity while keeping metrical complexity constant? In the case of the G-SLI group, there are significant effects on performance at all levels of metrical complexity. For example, for words containing two marked metrical structures, there are significant differences between those containing zero versus one complex syllabic structure ($p < 0.05$) and zero versus two ($p < 0.001$) (though not for one versus two). In the case of the LA1 group: the only significant effect of syllabic complexity was in words marked for two metrical structures ($p < 0.01$). As for LA2, the only significant effect of syllabic complexity was in words marked for zero or one metrical structure ($p < 0.05$) (though inexplicably not in those marked for two).



LA2



G-SLI

Fig. 2: Mean % correct responses by number of marked metrical and syllabic structures per ToPhS non-word: G-SLI versus LA2.

What these two series of comparisons tell us is that increasing metrical complexity has a greater negative effect on overall performance than increasing syllabic complexity, particularly for the G-SLI group.

5.4.3 String complexity versus prosodic complexity

Let us now turn to the question of whether prosodic complexity affects repetition accuracy independently of string complexity. String complexity, recall, is based on a linear count of the number of segments or syllables in a word. Figure 3 allows us to compare the proportions of correct repetitions across words with different numbers of syllables. There are significant differences based on group and number of syllables ($p < 0.01$, 3x3 ANOVA). The performance of the G-SLI group deteriorates markedly as the number of syllables per word increases. On the face of it, this is consistent with a string-based definition of complexity.

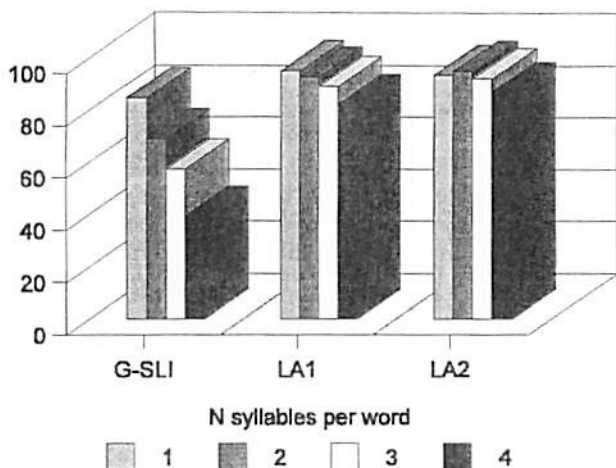


Fig. 3: Mean % correct responses by number of syllables per ToPhS non-word.

However, syllable-internal complexity can also be shown to affect repetition accuracy. This is illustrated in Figure 4, where we can see what happens when syllable-internal complexity is increased while the number of syllables is held constant, in this case in monosyllabic words. There are significant main effects for group, syllabic structure and an interaction between the two (all $p < 0.01$, 3x3 ANOVA).

A series of analyses (one-way ANOVAs) was performed to investigate the impact of increasing syllable-internal complexity in monosyllables within the different subject groups. The effect is significant for the G-SLI group but not for the two LA groups. In the case of the G-SLI group, there are significant differences between monosyllables marked for one versus two syllabic structures and one versus three (both $p < 0.001$). Thus, for this group, syllable-internal complexity can be seen to affect repetition accuracy independently of the number of syllables in a word.

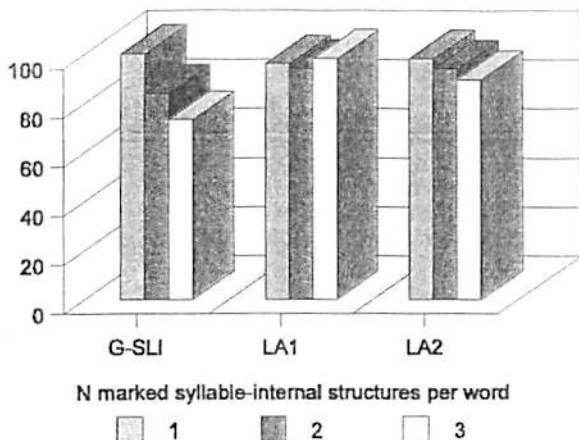


Fig. 4: Mean % correct responses for monosyllabic ToPhS non-words.

Now let us consider whether metrical complexity affects repetition accuracy independently of string complexity. In bisyllabic words containing one stress, i.e. one foot, the stress can fall either on the first syllable (the downhill pattern in for example *dépa*) or on the second (the uphill pattern in for example *bédép*). The downhill pattern is unmarked: both syllables are contained within the foot, i.e. $(\sigma\sigma)$. The uphill pattern is marked: only the second syllable falls within the foot, while the first syllable is adjoined to the left, i.e. $\sigma(\sigma)$. Since the two patterns are identical in terms of the number of syllables they contain, a string-based conception of complexity predicts that one should cause no more difficulty than the other in a repetition task. In contrast, under a prosodic account of complexity, partially footed $\sigma(\sigma)$ is metrically more complex than fully footed $(\sigma\sigma)$ and should thus be more difficult to repeat accurately.

In Figure 5, we can compare how accurately subjects repeated fully versus partially footed disyllables (the *dépa* versus *bédép* contrast). Here there are significant main effects for group, metrical structure and an interaction between the two (all $p < 0.001$, 3×2 ANOVA). Planned comparisons (using the Games-Howell test) revealed that the G-SLI group performed significantly worse than the LA2 group on fully footed words ($p < 0.05$) and significantly worse than both control groups on partially footed words (both $p < 0.001$). There was no significant difference between the LA1 and LA2 groups under both metrical conditions. As suggested in Figure 5, the G-SLI group, unlike both control groups, repeated partially footed disyllables significantly less accurately than those with full footing ($p < 0.05$, t-test). It is clear then that, at least in disyllables, metrical complexity impacts on the G-SLI group's performance independently of string complexity.

5.4.4 Error types

Most of the subjects in the G-SLI group found the ToPhS task much harder than the control subjects. The average proportion of correct responses produced

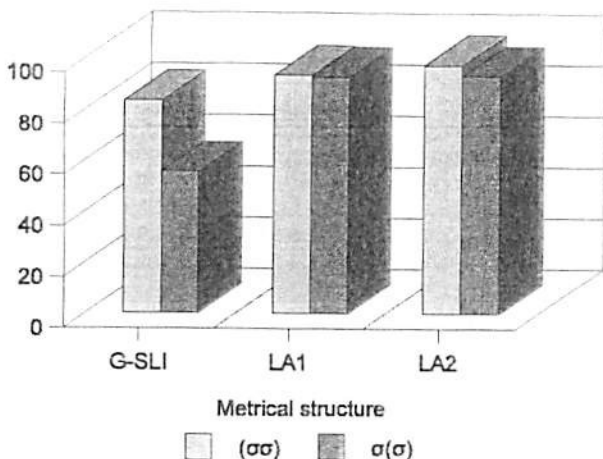


Fig. 5: Mean % correct responses by metrical structure in bisyllabic ToPhS non-words.

by the two control groups was very high – 90% for LA1 and 93% for LA2. This is consistent with the view that prosodic structure is largely in place by three years of age (Demuth, 1996). Only three individuals from the G-SLI group fell within or close to this range. Three subjects from the G-SLI group found the task especially difficult, producing scores of 31%, 36% and 41%.

The majority of repetition errors made by all groups took the form of cluster reduction in syllable onsets (e.g. *bødréþa* > *bøðéþa*), a process that is amply attested in younger, normally developing children. In the case of the G-SLI group, the most errors occurred in the prosodically most marked words. Table 3 shows the seven words that caused the most difficulty for this group. Each of these words contains at least three marked prosodic structures. All of them contain a complex onset. And all of them are also metrically marked, containing at least one weak syllable adjoined to the word edge. Of the 36 ToPhS words containing both a complex onset and two complex metrical structures, the G-SLI group got only half right (52%), compared to LA1 (87%) and LA2 (90%).

What is striking about the nature of the errors made by the G-SLI group in metrically complex words is that it is not the metrical structure itself that is primarily affected. Rather, the errors mainly affect syllable structure and to a lesser extent segmental content.

5.5 Discussion

The main findings of the ToPhS study presented in this section may be summarised as follows. Increasing prosodic complexity has an adverse effect on the accuracy with which subjects are able to repeat ToPhS non-words. The prosodic complexity of a word influences repetition accuracy independently of string complexity. In and of itself, marked syllable structure has little impact

Non-word	Mean % correct	N syllables	N marked prosodic structures per non-word					
			Syllabic			Metrical		Total
			Onset	Rime	Word-end	L-adj	R-adj	
<i>difripələ</i>	15	4	1	0	0	1	1	3
<i>fəklətələ</i>	15	4	1	0	0	1	1	3
<i>fəklətələ</i>	15	4	1	1	0	1	1	4
<i>difrimpl</i>	23	3	1	1	0	1	0	3
<i>difrimp</i>	23	2	1	1	1	1	0	4
<i>fəkl'ist</i>	23	2	1	1	1	1	0	4
<i>difrimpələ</i>	23	4	1	1	0	1	1	4

Tab. 3: ToPhS non-words with the most errors made by the G-SLI group.

on repetition accuracy. The source of most repetition difficulties lies in marked metrical structure, represented in this study by weak-syllable adjunction. Metrical structure itself, however, is faithfully repeated most of the time, reflected in the relatively low incidence of weak-syllable omission. Yet it is precisely in metrically marked words that the most syllabic and segmental errors occur. That is, metrical complexity gives rise to processing difficulties at lower levels of the prosodic hierarchy.

6 Conclusion

It seems intuitively reasonable to expect that the more complex a prosodic structure is the greater will be the burden it places on phonological processing. At first sight, this expectation is confounded by the observation that branching structure is marked in some areas of the prosodic hierarchy but not others. However, to understand the relation between markedness and structural complexity in prosodic phonology, we need to take account of how much of the structure is motivated by constraints on canonical morpheme shape. Once prosodic complexity is understood in these terms, we are better placed to understand its potential impact on phonological processing.

References

- [1] ARCHANGELI, D., AND PULLEYBLANK, D. *Grounded phonology*. MIT Press, Cambridge, MA, 1994.
- [2] BORTOLINI, U., AND LEONARD, L. B. Phonology and children with specific language impairment: status of structural constraints in two languages. *Journal of Communication Disorders* 33 (2000), 131-150.
- [3] BOTTING, N., AND CONTI-RAMSDEN, G. Non-word repetition and language development in children with specific language impairment (g-sli). *International Journal of Language & Communication Disorders* 36 (2001), 421-432.
- [4] CHOMSKY, N., AND HALLE, M. *he Sound Pattern of English*. Harper and Row, London, 1968.

- [5] DEMUTH, K. Alignment, stress and aprsing in early phonological words. In *Proceedings of the International Conference on Phonological Acquisition* (Sommerville, MA, 1996), B. BERNHARDT, J. GILBERT, and D. INGRAM, Eds., Cascadilla Press, pp. 113-124.
- [6] DOWNING, L. *Canonical forms in prosodic morphology*. Oxford University Press, Oxford, 2006.
- [7] DRESHER, E., AND VAN DER HULST, H. G. Head-dependent asymmetries in prosodic phonology: visibility and complexity. *Phonology* 15 (1998), 317-52.
- [8] ET AL., B. S. Non-word repetition in children with language impairment – pitfalls and possibilities. *International Journal of Language and Communication Disorders* 34 (1999), 337-352.
- [9] ET AL., C. A. D. Lexical influences on non-word repetition. *Applied Psycholinguistics* 16 (1995), 211-222.
- [10] ET AL., C. M. Investigating the impact of prosodic complexity on the speech of children with specific language impairment. *UCL Working Papers in Linguistics* 14 (2002), 43-66.
- [11] ET AL., D. V. M. B. Non-word repetition as a behavioural marker for inherited language impairment: evidence from a twin study. *Journal of Child Psychology and Psychiatry and Allied Disciplines* 37 (1996), 391-403.
- [12] ET AL., D. V. M. B. Grammatical g-sli: a distinct subtype of developmental language impairment? *Applied Psycholinguistics* 21 (2000), 159-181.
- [13] ET AL., M. S. Words, non-words and phonological processes: some comments on gathercole, willis, emg-slie and baddeley. *Applied Psycholinguistics* 12 (1991), 369-373.
- [14] GALLON, N., HARRIS, J., AND VAN DER LEIJ, H. J. K. Non-word repetition: an investigation of phonological complexity in children with grammatical sli. *Clinical Linguistics and Phonetics* 21 (2007), 435-455.
- [15] GATHERCOLE, S. E., AND BADDELEY, A. D. Phonological memory deficits in language disordered children: is there a causal connection? *Journal of Memory and Language* 29 (1990), 336-360.
- [16] GATHERCOLE, S. E., AND BADDELEY, A. D. Children's test of non-word repetition. Tech. rep., The Psychological Corporation, Harcourt Brace, 1996.
- [17] HARRIS, J. *English sound structure*. Blackwell, Oxford, 1994.
- [18] INKELAS, S., AND ZOLL, C. *Reduplication: doubling in morphology*. Cambridge University Press, Cambridge, 2005.
- [19] LEONARD, L. *Children with specific language impairment*. MIT Press, Cambridge, MA, 1998.

UFPA - Faculdade de Letras

- [20] RUSSELL, K. *Optimality Theory: an overview*. Oxford: Blackwell, 1997, ch. Optimality Theory and morphology, pp. 102-133.
- [21] TRUBETZKOY, N. Grundzüge der phonologie. *Travaux du cercle linguistique de Prague* 7 (1939).
- [22] VAN DER. H. K. J., LELY, ROSEN, S., AND MCCLELLAND, A. Evidence for a grammar-specific deficit in children. *Current Biology* 8 (1998), 1253-1258.
- [23] VAN DER LELY, H. K. J., AND HARRIS, J. Test of phonological structure. Tech. rep., Centre for Developmental Language Disorders and Cognitive Neuroscience, UCL, London, 1999.
- [24] VAN DER LELY, H. K. J., AND HOWARD, D. Children with specific language impairment: linguistic impairment or short-term memory deficit? *Journal of Speech and Hearing Research* 36 (1993), 1193-1207.

The role of production in the development of phonological representations

Marilyn May Vihman & Tamar Keren-Portnoy
University of York

mv509@york.ac.uk

Resumo

Este trabalho considera relações entre a produção e percepção durante a aquisição da linguagem. As autoras discutem a relevância do léxico do adulto em ciclos de aprendizagem baseados em tipos (types). As autoras discutem o modelo “Whole-word Phonology” avaliando dados do inglês e do português brasileiro. Ao discutirem dados sobre a percepção relativos a reconhecimento e aprendizagem da linguagem apresentam resultados sobre a produção e memória fonológica.

1 Introduction: The need to account for production as well as perception

The findings of experimental studies of infant speech perception are sometimes taken to be sufficient in themselves to account for phonological development (Peperkamp, 2003). Yet evidence bearing on perceptual processing and representation alone cannot suffice to account for language learning. A long tradition of both diary and planned observational studies have established the fact of wide individual differences in the rate and pathway of emergence of word production and phonological knowledge across children developing normally and learning a range of different languages (see Vihman, 1996); experimental group studies of word recognition and learning shed little light on this critical aspect of phonological development. It is evident that both lexical and phonological learning depend on the development of representations that integrate perception and production;¹ this remains a central issue which has so far attracted insufficient attention.

Pierrehumbert (2003) proposed that the phonological system is “initiated bottom-up from surface statistics over the speech stream, but refined using type statistics over the lexicon” (p. 118). She does not elaborate on the source of the lexical knowledge that supports the second cycle of statistical learning, however. We argue below that the missing link is production experience, which brings the specific adult lexicon to which the child is exposed into focus and into partial or incipient mastery, leading, as Pierrehumbert says, to a new cycle of statistical learning based on types, not tokens. We provide examples from the

¹Although we leave aside the question of syntactic learning here, the constraints on infant vocal production are relevant within this domain as well: e.g., Behrens and Cut (2005).

learning of geminate consonants to support that contention and then explore the notion of 'whole word phonology' with reference to the origins of the idea in the 1970s and with examples from children learning both English and Brazilian Portuguese. We then consider conflicting data from the current experimental perception literature on infant word recognition and word learning and conclude by presenting some findings from a recent study of the role of production in the construction of phonological memory.

2 Interaction of perceptual processing and production: The case of geminates

The earliest evidence of the interaction of perceptual processing and vocal production comes from studies designed to look for ambient language influence on production. Despite a long period of inconclusive studies based on adult judgements of the vocal production of infants exposed to different languages, beginning with Atkinson, MacWhinney and Stoel (1968) and continuing into the present century with Engstrand, Williams and Lacerda (2003), several studies have been able to show through acoustic or transcription-based analyses of naturalistic recordings that by 10 months infants who have barely begun to produce words (if at all) are biasing their vowel space (Boysson-Bardies; Hallé; Sagart; Durand, 1989), prosodic contours (DePaolis; Vihman; Kunnari, in press; Whalen; Levitt; Wang, 1991) and consonant production (Boysson-Bardies; Vihman, 1991) in the direction of the patterns most frequently heard in the ambient language.

Since production is only a secondary reflection of the child's perception of adult speech in these cases, the evidence of ambient language influence is less direct than in perception studies demonstrating, for example, child sensitivity to increasingly small prosodic units from 4 to 11 months and to increasingly segmental aspects of the native language over roughly the same time period (Jusczyk, 1997). Implicit pattern learning – also commonly referred to as 'statistical' or 'distributional' learning – must underlie these effects of perception of the ambient language on infant vocal production. That is, to account for the reported effects we must assume that infants are biased to selectively reproduce with greatest frequency the vocalizations that they implicitly experience as better matches to what they are hearing with greatest frequency in input speech. If this interpretation of the mechanism that mediates input language pattern effects on infant vocal output is correct, the effect should be seen only in patterns that have a solid grounding in the infants' production repertoire, i.e., that are also produced with sufficient frequency to be subject to such a 'pruning' effect of the perception/production match. This supposition is borne out by the evidence. For example, phonological categories such as coda consonants, which are not commonly produced by most infants in babbling, come into alignment with the distributional frequencies of the adult language at a later stage than do phonetic categories that come under infant control earlier (Vihman; Boysson-Bardies, 1994).

The differential production of long consonants by children exposed to Finnish and Welsh provides an instructive test of the adequacy of implicit, 'statistical' or frequency-based learning to account for the emergence of phonological

knowledge. Finnish phonology includes a categorical long: short contrast that affects most consonants (as well as all vowels), whereas in Welsh consonants are lengthened medially after a stressed syllable² falling under nuclear accent, a non-contrastive effect that is nevertheless a pervasive gradient phonetic aspect of Welsh speech. Children exposed to Welsh thus tend to hear more long consonants than do children exposed to Finnish, where they occur in only about a third of the content words mothers use with their children and account for only 10% of all consonants used in running speech, based on infant directed speech (this difference in usage results in a median length in Welsh mothers' child-directed disyllables of 118ms vs. 75ms for Finnish, a highly significant difference (Vihman, 2001; Vihman; Kumari, 2006). As predicted, in babble and first words Welsh infants produce longer consonants on average than do Finnish infants (170ms vs. 116 ms, resp.) – presumably due to the overall greater impact of long consonants experienced in Welsh input speech, particularly in prosodically salient word forms. By the time that the children reach the end of the single word period the situation has reversed. Direct measurement reveals that Finnish medial consonants attain a mean of 223ms, while Welsh children show little change from the earlier stage (Vihman, 2001; Vihman; Kumari, 2006).

Long medial consonants are within the infants' production repertoire, given the relatively slow articulatory production that characterizes early speech. This is evident from the fact that children exposed, on the one hand, to English and French, in which there is no systematic occurrence of long consonants, and, on the other hand, to Finnish show similar ranges of medial consonant duration at the earliest stages of word production, with some children in each group producing far longer consonants than are typical of adult English or French (Vihman; Velleman, 2000). The Welsh children followed in the later study summarized above showed the effect of adult long consonant production already at this early stage, reflecting *both* the ease of production and the relatively high input frequency of this phonetic feature in Welsh. As word learning progresses in each group, infants exposed to English and French cut back (or 'prune') the length of their medial consonants in accordance with adult norms while Finnish children increase theirs, but the Welsh children show no further lengthening, since in their case lexical learning does not increase the salience of the (purely phonetic) long medial consonants in the input.

Given that there is no need for motoric advances to increase long consonant production over the course of early word learning, since Welsh children are producing a good many long medial consonants already from the onset of identifiable word use, how do Finnish children gain the additional phonological knowledge that results in the increase in mean length of medial consonants observed in their production at the later lexical stage? In fact, by the end of the single word period Finnish children are 'overshooting' the target in producing even more long consonants, proportionately, than are found in input speech (a mean of 47% of their words are transcribed as including a long consonant as compared with 38% of words with geminates in the mothers' child-directed speech).

²The stressed vowel itself is short in Welsh. For a more complete description and illustration of the adult model as well as the effect on early word production, see Vihman, Nakai and DePaolis (2006).

It is here that production must be understood to be providing an impetus to change not found in perceptual learning alone. Note that the difference between the early and the later stages analyzed was defined by a production milestone, not by age – as is appropriate, given the wide differences in linguistic knowledge of infants of the same age, as demonstrated by the variability in performance in age-based perception studies (Werker; Fennell; Corcoran; Stager, 2002). At the early stage the children had produced at least four distinct word types spontaneously within a half-hour recording session, in play interaction with a parent and in the presence of a familiar observer (the ‘4 word point’, or 4wp); at the later stage 25 or more different word types were produced under the same conditions (25wp); these milestones correspond to parental vocabulary counts of roughly twice as many words (Vihman; Miller, 1988). The median age for both the Welsh and the Finnish children was 13 months at the 4wp and 17 months at the 25wp.

The change in the extent of production of long consonants, then, is plausibly traceable to some aspect of this increase in production knowledge of the target lexicon. Specifically, the Finnish children’s increase in long consonant production with lexical learning can be understood to reflect not only the interest of common words with medial geminates but also their notoric accessibility. Once one or two of the common words with geminates (e.g., *anna* ‘give’, *kukka* ‘flower’, *loppu* ‘finished, all done’, *pallo* ‘ball’, *tyttö* ‘girl’) have been produced, production practice can be taken to lend further salience to words with long medial consonants in Finnish despite their relatively low overall frequency of occurrence in the input speech stream. That salience is thus due to a kind of ‘top-down’ (cognitive, lexical) processing in contrast to the ‘bottom-up’ or purely ‘signal-based’ salience of long medial consonants in a language like Welsh, where they occur more frequently due to their status as phonetic markers of accent.

Pierrehumbert (2003) has argued that “the application of type statistics to a system initiated with surface statistics [i.e., with statistics based on tokens, not types] does not cause a fundamental reorganization of the system” (p. 115). In her terms, the Finnish children’s shift to overproduction of long consonants reflects the addition of “internal feedback from type statistics over the lexicon” (p. 115) – but the real impact of lexical learning, we argue, is mediated through production experience, an element that is mentioned but minimized in her account.

3 The ‘whole word phonology’ arguments from production

The view that phonological development begins with ‘whole word representations’ was first put forward in the 1970’s (e.g., Ferguson; Farwell, 1975; Macken, 1979; Menn, 1971, 1983; Waterson, 1971). Waterson based her hypothesis on the linguistic theory of J. R. Firth:

It is the whole-unit or holistic approach of prosodic phonology which brought to light the patterned relationships between a child’s hardly recognizable words and the adult models and thus made it possible to explain the somewhat bizarre forms of early words, as well as why

sounds that children are capable of producing are not used in all the contexts in which they occur in adult forms. (Waterson, 1971, p. 2)

One of many sets of examples of the 'bizarre forms of words' provided by Waterson from her son P's data is the 'nasal structure', a group of words that share reduplicated forms with palatal nasal onset in the child's production, although not in the adult targets:

Child Form	Adult Target
[ɲaɲa]	another
[ɲe ɲe:], [ɲe:ɲ]	finger
[ɲaɲə]	Randall
[ɲe ɲe:]	window

Tab. 1: The 'nasal structure' (Adapted from Waterson, 1971)

Here the child has settled on a motoric pattern that results in the production of nasal patterns and is deploying it, with freely varying vowels, to reproduce the remembered pattern of various adult word forms whose only common feature is a word-medial nasal, in coda position in the syllable in three of four instances. The notion of 'whole word representation' seems to fit the case very well.

The three primary arguments used to support the claim that the earliest phonological structures are whole-word based are summarised in Vilhman and Croft (in press, p. 8f.): we review and illustrate them here.

- (1) *Variability of segment production*: A child may produce the same sounds differently in different words, and some words may be more variable than others. This suggests that the child has knowledge of particular words but has not yet developed abstract *categories of sounds* for production (Ferguson; Farwell 1975).

Ferguson and Farwell (1975) famously presented 12 widely varying pronunciations of the word *pen* produced in the course of a single session at about 15 months by K, one of the two American children they had observed, with alternate production of labial or alveolar stop or nasal or neither, with a range of oral or nasal low to mid vowels:

[mā^a (imit.), [ã] (imit.) de^{du}, lɪn, ^mbō, p^hɪm, t^hn (3X), ba^h, d^hauⁿ, buā]

A similar example of a 'hard word', attempted six times by an English child, Jude (also aged 15 months), in his 25wp session is *circle*, variously produced, in a full or partial whisper, as:

[tsɹu, tsɹ^hɔ (x2), t^h t^h, tɹɹɹɹɹɹ (imit.), k^hɹɹɹɹɹɹ (imit.)]

Here we see evidence of child attention to the sibilant and its co-occurrence with a stop and a lateral, although the identity of the stop appears to be uncertain as does the sequencing of the various segments, despite the presence of an adult model in two cases (note that two of K's productions of *pen* were also in direct imitation).

UFPA - Universidade de Leobers

The child's 'underlying representation' cannot easily be inferred from his production efforts, but it could be described as dynamic or fleeting rather than set or stable, with an apparent influence on the momentary remembered form of the word from co-existing production patterns in the child's repertoire - patterns which must be accessed for vocal expression. In Jude's case the repertoire of onset consonants used in more than one other word in this recording session includes both alveolar and velar voiceless stops, while [l] occurs once as an onset (*flower* [la]) and once as a coda (in the word combination *mummy ball* [ʌmʌmi :b ə]).³ it occurs medially (as a voiceless lateral) only in this word.

It is evidently not the sounds themselves that Jude cannot accurately reproduce, nor is there reason to believe that he cannot *perceive* the adult segments, since he produces them all in one attempt at the word or another. Instead, his difficulty appears to involve the planning and production of the word pattern as a whole, in sequence, with its rapidly changing series of consonantal gestures. Note that Jude's remaining words tend to be one or two syllables in length. Only two other word forms of the 43 he produced in this session include a change in consonantal place or manner (and not both) within the word. *Caterpillar* and *chocolate* are attempted just once each, as [p^hɪt^hɪp^hʌp^hk^hæ...] and [ʃyt^hə], respectively.⁴ Target sibilants are attempted also in the words *fish* [s:] and *trousers*, imitated as [tətʃ:]. We see that in his spontaneous attempts at *circle* (those unsupported by immediate or echoic memory of the adult form) Jude is here again deploying the syllable [tʃ]. This provides a clue that motoric practice may play a role in supporting 'representation' for word forms, or what we might better call emergent lexical memory.

- (2) *Relationship of child word to adult target:* The relation of early child words to their adult models is often found to be difficult to account for on a segment-by-segment basis. Instead, the child seems to be targeting a whole gestalt (Waterson, 1971).

P's 'nasal structure' words, K's attempts at pen and Jude's attempts at circle all provide excellent examples of this principle.

- (3) *Relationship between child words:* The interrelation between the child's own words may be more evident than the relation to the adult models (Macken, 1979).

In order to illustrate this principle a larger number of contemporaneous child forms is needed - ideally, the full set of words used in a single recording session. The early words of one child acquiring Brazilian Portuguese may be taken as an example.⁵ These word forms, like those provided from Jude's data, derive from the end of the single word period (the 25wp), which typically coincides with the very first word combinations or precedes these by a few weeks at most. Note again that, for the study of production, particularly on a cross-linguistic basis, chronological age is not a good basis for comparison. The lexical milestone

³This is the only use of a coda in the session.

⁴Notice that both of these words, like circle, have laterals in the adult target, although the child's production gives no evidence of his having registered or represented the lateral.

⁵We are grateful to Daniela Oliveira, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, for sharing these data with us.

Word pattern	Selected words			Adapted words		
	adult word	gloss	child word form	adult word	gloss	child word form
(C)V	é	'it is'	ɛ	(em) bora [bora]	'gone, away'	bo
	né (im.)	'yes; okay'	nɛ	pão (im.)	'bread'	pã
	tá	'okay'	ta	pato (im.)	'duck'	pa
				pé	'foot'	pa
				praia	'beach'	pa
				tartaruga (im.)	'turtle'	ta
				tchau	'bye'	ta
(C)VCV	Cacá	[name]	ka'ka	Fernanda (im.)	[name]	ve'v
	ovo (im.)	'egg'	'ofu	Ízabel (im.)	[name]	pe'pe
	vovó [vovo]	'grand-father'	vo'vo	papel [pa'pɛ]	'paper'	pe'pe
	vovó vova	'grand-mother'	vo'va	Pedro [pe'dru]	[name]	du'du
	Xixi [ʃi ʃi]	[name]	zi'zi	Viviane	[name]	vi'vi
(C)VCVV	mamãe	'mama'	mã'mã	chapéu (im.)	'hat'	pe'peũ
	neném [ne'nɛ]	'baby'	[ne'nɛ]	e não	'is not'	ɛnãũ
	papai	'papa'	pa'pai			
(C)VV	boi	'bull'	boi	água	'water'	ãũ
(C)VC	abre (im.) [abri]	'open'	ab	Lucas (im.) [lukas]	[name]	us
	desce (im.) [des]	'go down'	dɛS	não	'no'	nãũ
	dois	'two'	dos	uva (im.)	'grape'	uf
	peixe [peʃi]	'fish'	pes			
CVIV				tira [tʃira]	'take off'	'dili
				parabéns	'birthday congratulations' (im.)	pala

Tabella 2: Primary production patterns at the 25wp, Lucas, aged 24 months, acquiring Brazilian Portuguese

Journal of Child Language
1998, 25, 107-128

a cumulative diary record of about 50 to 75 words or the identification of 25 or more different spontaneous word types produced in a half-hour session provides a more reliable developmental index: At this point children can be seen to 'select' words, or to produce relatively accurate forms for adult words that fall within their production repertoire as regards both phonotactic shape and segmental inventory, and also to attempt, as do both Jude and Lucas, words which exceed their production abilities but which can be made to conform to an existing word production template or routine.

We divide the words into those that are roughly 'accurate' or 'selected', based on a match between the adult target and one of the child's existing word production routines, and those that are less accurate as a result of being 'adapted' from less similar adult targets to fit those same routines. Table 2 shows one token of each of the non-onomatopoeic word forms used in this session (selecting the spontaneous or most frequently produced tokens where variant forms were transcribed for the same word); a total of 40 words were used in the session, of which 26 word types were produced spontaneously, or in the absence of any immediate adult model. Words produced only as imitations are indicated in the table; adult forms are given in IPA only where the form is not easily determined from the spelling, which is generally transparent.

Lucas' word forms provide a good illustration of both the second and the third arguments for whole word phonology. The child's forms fit a small number of patterns, with 'practice' in adult word production afforded by the 'selected' words, which are readily pronounceable with the child's existing word production routines. The words 'adapted' to fit the child's patterns exemplify the 'somewhat bizarre forms' to which Waterson alluded. The extreme reduction of *tartaruga* to [ta], for example, or the use of [pa] for a range of different adult word forms, or the reduplicated form [ve've] for *Fernanda* all indicate that the child is drawing as much on his own remembered word production patterns as on the adult target. That it is the whole word that inspires the choice of child pattern for production is indicated particularly clearly in the case of [uf] for (initial stressed) *Lucas*, for example: Most of the disyllabic forms the child uses are iambic; the proper names with a trochaic pattern are all converted into iambs in the child form, and in this case only the vowel is retained from the stressed syllable of the adult form. Note, finally, that there are only two disyllabic forms with a change in onset consonant across the two syllables (although there are several monosyllables with codas - all fricatives, as is commonly the case, with the exception of [nam], an anomaly for Brazilian Portuguese which has no coda /m/). The two exceptions are both 'adapted' (and one is imitated) - and both have [l] as the medial consonant, suggesting an advance in production that can be expected to lead to greater use of [l], and of the new pattern CVIV, in subsequent sessions. The [l] is used for target tap [r] in both cases.

4 How detailed are early perceptual representations?

The first experimental studies to explore the level of 'phonetic detail' in infants' lexical representations were carried out in the mid-nineties, using mispronounced consonants as a test. Jusczyk and Aslin (1995) first trained children aged 7.5

months on one of two pairs of basic words (*cap, dog; bike, feet*) and found that the infants maintained a longer head turn in response to a short narrative passage with multiple repetitions of the trained pair than to a matched passage with the untrained pair. They determined in a follow-up experiment that if the first consonant was changed, the infants no longer responded differentially to the two passages, which they were therefore taken to be unable to segment, based on failure to recognize the mispronounced trained word forms. This was interpreted as reflecting 'rather detailed' (newly formed) representations of the sound patterns of the target words - although in fact only initial consonant changes were tested.

A study by Hallé and Boysson-Bardies (1994, 1996) used the head-turn paradigm to explore this issue further with French children, but the stimuli were lists of untrained words, many of which could be expected to be familiar from the home as they are among the first words produced by French children (*bonjour* 'hello', *canard* 'duck', *chapeau* 'hat'). Younger infants did not appear to recognize enough words to show a greater attention to the familiar words than to matching rare words (*busard, cobaye* 'guinea pig', *bigot*), but 11-month-olds showed significantly longer headturns in response to the familiar words. In the follow-up study using mispronounced words, a change in voicing or in manner to the initial consonant did not block the response, but omission of the initial consonant did. Change to the medial consonant produced mixed results. The authors proposed two hypotheses to explain these results, which conflicted with those of Jusczyk and Aslin: (1) the greater age of the children might mean that they were now listening in a 'lexical mode' rather than simply processing the speech signal 'bottom-up'; this could result in their using a 'global' or relatively unspecified lexical representation rather than the 'phonetically analytic mode' used by the younger children; (2) the children were tested on French words, with an iambic (weak-strong) accentual pattern, whereas the earlier study with American children tested monosyllabic English words; this could result in differences in infant response depending on the salience of the initial consonant under the differing accentual conditions - either in the experiment itself or in the infants' differing general experience with the two ambient languages. A later study used the same experimental design to test the second hypothesis with 11-month-old children acquiring British English in North Wales, using disyllabic words as stimuli. The results complemented and extended those of Hallé and Boysson-Bardies: Changes to initial but not to medial consonants in trochaic (strong-weak) words blocked word recognition for the English children (Vilman; Nakai; DePaolis; Hallé, 2004), supporting the second hypothesis.⁶ Interestingly, a shift in stress pattern on familiar words also failed to block word recognition, suggesting greater infant reliance on segmental aspects of their representations than on prosody at that age.

The extent of phonetic detail in infant lexical representations has also been tested with other experimental paradigms. Werker and her colleagues (Fennell; Werker, 2003; Werker; Stager, 2000) initially used the nonwords /b/ - /d/ in a habituation paradigm in which the children are trained to associate one of these 'nonwords' with a moving image of a novel object and are then shown the same

⁶Vilman et al. (2004) also tested 9-month-olds, who failed to show the familiarity effect, as did 10-month-olds in a subsequent study (Vilman; Thierry; Lum; Keren-Portnoy; Martin, in press).

UPPER RECORDS de Letras
M 1 1 1 6 1 2 2 2

object with the minimally different 'name' (the 'switch' trial). At 17 and 20 months infants respond with recovery from habituation on the switch trial, but at 14 months they do not - although they do respond to non-minimally contrasting forms (/hf/ vs. /nim/). Furthermore, children with a larger vocabulary, by parental report, already showed dishabituation to the minimally contrasting nonword at 14 months (Werker; Fennell; Corcoran; Stager, 2002) and the younger children as a group discriminated the minimally contrasting forms when tested with a checkerboard, which is unlikely to provoke an attempt at word learning. Werker and her colleagues interpreted the infants' failure to dishabituate to the minimally contrasting form in line with the first hypothesis of Hallé and Boysson-Bardies, namely, that the word learning task demands more resources, leaving little space for attention to the subtle difference between the forms. A later study tested 14-month-olds with the familiar words *ball* and *doll* in the same paradigm; in this case the infants did dishabituate to the minimally contrasting form even at this early age (Fennell; Werker, 2003).

These studies agree in showing, as had earlier work by Barton (1980), that children respond differently to small differences in known than in unknown words - or, as Fennell and Werker (2003) put it, access to phonetic detail is more difficult in word learning than in word recognition. Fennell and Werker emphasize the importance of the varied experience with the words that infants are likely to have gained in their daily lives, hearing them produced by different speakers in different contexts, in contrast with the numerous repetitions of the word forms (up to 120) as produced by a single unfamiliar speaker in the experimental setting. Finally, they argue that words are "fully specified from the beginning and [...] the [infants'] difficulty [or failure to dishabituate] lies in access rather than in the representation" (p. 258).

On the other hand, Swingley and Aslin have used eye-tracking experiments with children aged 14-23 months (Swingley, 2005; Swingley; Aslin, 2000, 2002), presenting words likely to be known to the children (*apple, baby, ball, car, dog, kitty*) with and without changes to the vowel ([ɒp], [kr]) or to the initial consonant ([veɪbi, gəl, tɔg, pti]) along with two images, only one of which matches the spoken word. (In the study with 14-15-month-olds both 'close' and 'distant' changes were made to the words; there was no effect of mispronunciation type.) Swingley and Aslin found more rapid recognition of the correctly pronounced words (although all the words were recognized), with no variation across age or reported vocabulary size. Furthermore, the extent to which the words did or did not have close 'lexical neighbors', or similar words likely to be known to the children, did not affect the results. Swingley and Aslin interpreted their results as supporting continuity between early phonetic learning in the first year and the use of that knowledge in word learning in the second year.

Finally, Swingley (2005) tested 11-month-old children acquiring Dutch on their preference, in a head-turn procedure, for CVC words with correct or incorrect onset or offset consonants. Onset consonants were recognized only when correctly pronounced, replicating the findings of Vihman et al. (2004) for English infants. However, the results for offset consonants were more complex. While words with correctly pronounced offset consonants were preferred over similar nonwords, they were not preferred when directly contrasted with the same words with mispronounced offsets. This suggests, again in agreement with Vihman et al., that offsets are less well represented than onsets, at least in stressed

syllables, which accounts for the lack of an effect when mispronounced familiar words are contrasted with similar nonwords in the same experiment. The unexpected failure of the same familiar words to be 'preferred' over (listened to with greater attention than) the mispronounced forms may be due to the weaker representations of the offset consonant being just "sufficient to obscure a preference for words over these mispronunciations" under this condition (p. 440). That is, as first suggested by Hallé and Boysson-Bardies (1996), hearing correctly produced known words may prime the words in the child's memory, so that the mispronunciation of [vif] for [vis] 'fish' may be overlooked. As Swingle points out, however, the evidence that, in at least some conditions, "infants respond differently to correct and mispronounced forms is consistent with the possibility that by 11 months infants have begun to build a phonological system that will help distinguish lexical minimal pairs [...]" (p. 441) i.e., "well before they say words reliably" (p. 442).

How, then, can these two sets of data and interpretations – one set from early word production, the other from experimental studies of children in roughly the same age range – be reconciled? 'Phonological representation' in infancy is clearly complex and as yet poorly understood. We have seen different inferences depending on accentual pattern and task demands – word learning, word recognition and word production. The differences in demands are important: In the case of word learning, as Werker and her colleagues have argued, significant attentional resources must be allocated to the problem of retaining the arbitrary sound-meaning link, making the task difficult, we would argue, for children who lack a stock of existing well-practiced production patterns or routines to support memory of the new word. In the case of word recognition, on the other hand, both the word form and the contextual situation or image may be expected to prime memory for the word and its associations. Word production itself is more or less demanding depending on the context. In the case of imitation, particularly of a known word form, priming is again available. In the case of spontaneous word production, context-based word use is clearly less challenging than context-flexible use, in which a child may request an object not currently in view, for example, or refer to an event that occurred earlier in the day (Bates et al., 1979; Vihman; McCune, 1994). This latter use, which is generally accepted to reflect the onset of referential or symbolic word use, is the most demanding of all. To reconcile the contradiction between perception experiments and the implications of production, then, we would argue that word production requires memory and planning as well as perception and articulatory capacity, leading to the use of holistic, temporarily activated representations for production which show regression in accuracy as the child comes to make use of more word types and which are clearly not characterized by 'fine phonetic detail'

5 A study of the effect of production practice on phonological memory

We have been arguing that production practice supports child memory for word forms (see also McCune; Vihman, 2001; Storkel, 2001). In a recent study we tested the hypothesis directly with two-year-old children, using a specially designed Nonword repetition test (NRT), a variant on the kind of test usually

UFPA - Faculdade de Letras
21112100

used to assess individual differences in phonological memory, on the assumption that this capacity reflects innate 'talent' or 'skill', which cannot be modified. Previous research with quite young children (aged 2-3 years) tested the effect of individual differences in phonological memory on productive language (Adams; Gathercole, 1995). We ask the reverse question: What is the effect of individual differences in type or extent of production practice on phonological memory? In order to answer the question we carried out a study of 17 children, 12 of whom had been followed longitudinally since age 11 months. A sample of each child's naturalistic speech was obtained at the 25wp or, if the child achieved that lexical milestone by 15-20 months, in a later session, before age two. All of the children were given our NRT at 26 months.

The nonword test that we devised consisted of two parts. The first part included 18 familiar real words and 20 'standard' nonwords which were rearrangements, following English phonotactics, of the phonemic sequences of the real words (Roy; Chiat, 2004). The second part was individually designed for each of the children and consisted of nine 'in'-words and nine 'out'-words, based on common vs. infrequently used sounds for each of the children, as transcribed from their most recent recording (Keren-Portnoy; Vihman; Armstrong, in prep.). All 'out' sounds were 'in' for at least one child. Of all of the tailor-made words (e.g., [tepi]), 53% were 'out' for some children but 'in' for others, ensuring that all of the words were potentially within the production repertoire of children of this age. Figure 1 (taken from Keren-Portnoy et al., in prep.) shows the results for the 15 individual children whom we were able to test (two children were excluded, one due to an overly rich consonantal repertoire, which made it difficult to design 'out' words, the other due to an overly impoverished one, which made it difficult to design 'in' words).

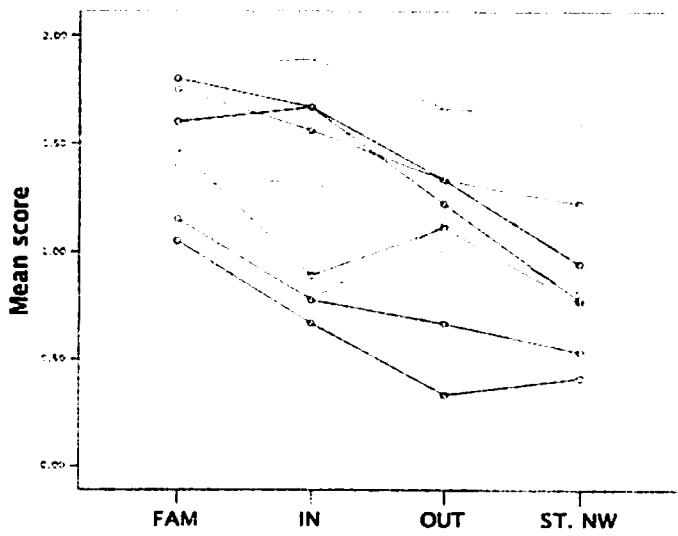


Fig. 1: Mean score on Familiar, In, Out and Standard Nonwords for 15 English children.

As can be seen in Figure 1, virtually all of the children achieved their best scores on Familiar words, followed by In, Out and Standard nonwords, in that order ($p < 0.01$). Performance was significantly better on In than on Out words, as predicted ($p = 0.03$, one-tailed). Thus, as predicted, production practice had a positive effect on the children's ability to reproduce the nonwords. However, the fact that they achieved even greater success with familiar words, some of which included 'out' sounds for a given child, demands an explanation.

We reasoned that if words are first learned as wholes, then by 26 months familiar words would constitute whole familiar chunks, while 'In nonwords', albeit containing familiar units (syllables, segments), could not be familiar as wholes. To test this interpretation we compared numbers of errors in In and Out nonwords in comparison to Familiar words. As shown in Figure 2, the error score for Familiar words was considerably lower.

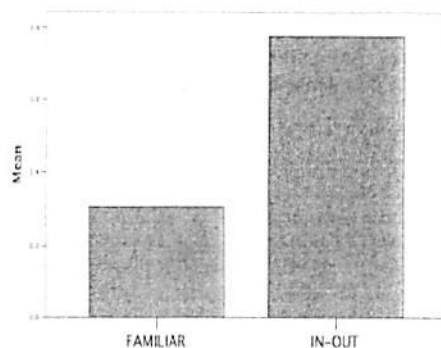


Fig. 2: Mean errors in Familiar vs. In-Out nonwords.

These findings lead us to conclude that patterns with well practiced sounds are more easily recognized, represented and accessed, while less practiced sounds exert more load on memory, resulting in less precise memory for the whole sequence.

6 Conclusion

Our findings on the relationship of production to phonological memory validate the idea that early word representations for production are holistic and strengthen with use. Statistical learning provides a basis for language acquisition - but production supports memory, which in turn supports the development of phonological knowledge and further lexical learning. Lexical representations begin to form early (as shown by the recognition of untrained word forms by 11-month-olds), but they are at first accessible only through priming by sound, image or situational context. The distinction sometimes made between 'representation' and 'performance' (e.g., Sinha, 2007) may not be well justified: For children just learning to talk, at least, the 'performance' or production of speech is implicated in the on-line creation of a representation (cf. also McCune, in press). Months of use may be needed to create stable representations that are

accessible at will, transforming what is at first new and fragile into the robust and automatically accessible word forms that we experience as adults.

References

- [1] ADAMS, A.-M., AND GATHERCOLE, S. E. Phonological working memory and speech production in preschool children. *Journal of Speech and Hearing Research* 38 (1995), 403-414.
- [2] ATKINSON, K., MACWHINNEY, B., AND C.STOEL. An experiment on the recognition of babbling. *Language Behavior Research Laboratory Working Paper 14* (1968).
- [3] BARTON, D. *Child Phonology: Perception*, vol. 2. Academic Press, New York, 1980, ch. Phonemic perception in children.
- [4] BATES, E., ET AL. *The emergence of symbols: cognition and communication in infancy*. Wiley, New York, 1979.
- [5] BEHRENS, H.; GUT, U. The relationship between prosodic and syntactic organization in early multiword speech. *Journal of Child Language* 32 (2005), 1-34.
- [6] DE, B. B.-B., AND VIHMAN, M. M. Adaptation to language: evidence from babbling and first words in four languages. *Language* 67 (1991), 297-319.
- [7] DE BOYSSON-BARDIES, B., HALLÉ, P., SAGAKI, L., AND DURAND, C. A crosslinguistic investigation of vowel formants in babbling. *Journal of Child Language* 16 (1989), 1-17.
- [8] DEPAOLIS, R. A., VIHMAN, M. M., AND KUNNARI, S. Prosody in production at the onset of word use: a cross-linguistic study. *Journal of Phonetics*. In press.
- [9] ENGSTRAND, O., WILLIAMS, K., AND LACERDA, F. Does babbling sound native? listener responses to vocalizations produced by swedish and american 12- and 18-month-olds. *Phonetica* 60 (2003), 17-44.
- [10] FENNELL, C. T.; WERKER, J. F. Early word learners' ability to access phonetic detail in well-known words. *Language and Speech* 46 (2003), 245-264.
- [11] FERGUSON, C. A., AND FARWELL, C. B. Words and sounds in early language acquisition. *Language* 51 (1975), 419-439.
- [12] HALLÉ, P.; BOYSSON-BARDIES, B. D. Emergence of an early lexicon: infants' recognition of words. *Infant Behavior and Development* 17 (1994), 119-129.
- [13] HALLÉ, P.; BOYSSON-BARDIES, B. D. The format of representation of recognized words in infants' early receptive lexicon. *Infant Behavior and Development* 19 (1996), 435-451.

- [14] JUSCZYK, P. W. *The Discovery of Spoken Language*. MIT Press, Cambridge, MA, 1997.
- [15] JUSCZYK, P. W., AND ASLIN, R. N. Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology* 29 (1995), 1-23.
- [16] JUSCZYK, P. W., LUCE, P. A., AND CHARLES-LUCE, J. Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language* 33 (1994), 630-645.
- [17] MACKEN, M. A. Developmental reorganization of phonology: a hierarchy of basic units of acquisition. *Lingua* 49 (1979), 11-49.
- [18] MCCUNE, L. *How Children Begin to Learn Words: dynamic systems in development and action*. Oxford University Press, Oxford.
- [19] MCCUNE, L., AND VIHMAN, M. M. Early phonetic and lexical development. *Journal of Speech, Language and Hearing Research* 44 (2001), 670-684.
- [20] MENN, L. Phonotactic rules in beginning speech. *Lingua* 26 (1971), 225-251.
- [21] MENN, L. *Language Production*, vol. 2. Academic Press, London, 1983, ch. Development of articulatory, phonetic, and phonological capabilities.
- [22] PEPPERKAMP, S. Phonological acquisition: recent attainments and new challenges. *Language and Speech* 46 (2003), 87-113.
- [23] PIERREHUMBERT, J. Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech* 46 (2003), 115-154.
- [24] ROY, P., AND CHIAT, S. A prosodically controlled word and nonword repetition task for 2- to 4-year olds: evidence from typically developing children. *Journal of Speech, Language, and Hearing Research* 47 (2004), 223-234.
- [25] SINHA, C. Language as a biocultural niche and social institution. *ABRILIN Conference Plenary* (2007). This volume.
- [26] STORKEL, H. L. Learning new words: phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research* 44 (2001), 1321-1337.
- [27] SWINGLEY, D. 11-month-olds' knowledge of how familiar words sound. *Developmental Science* 8 (2005), 432-443.
- [28] SWINGLEY, D., AND ASLIN, R. N. Spoken word recognition and lexical representation in very young children. *Cognition* 76 (2000), 147-166.
- [29] SWINGLEY, D., AND ASLIN, R. N. Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science* 13 (2002), 480-484.
- [30] VIHMAN, M. M. *Phonological Development: the origins of language in the child*. Blackwell, Oxford, 1996.

[U]NIVERSITÄT WÜRZBURG
FACULTÄT FÜR
HUMANWISSENSCHAFTEN
LINGUISTIK

[31] VIHMAN, M. M. A cross-linguistic study of the acquisition of geminate consonants. *End of award report. Economic and Social Research Council Award R000223134* (2001).

[32] VIHMAN, M. M., AND CROFT, W. Phonological development: toward a 'radical' templatic phonology. *Linguistics*. In press.

[33] VIHMAN, M. M., AND DE BOYSSON-BARDIES, B. The nature and origins of ambient language influence on infant vocal production and early words. *Phonetica* 51 (1994), 159-169.

[34] VIHMAN, M. M., AND KUNNARI, S. The sources of phonological knowledge: a cross-linguistic perspective. *Recherches Linguistiques de Vincennes* 35 (2006), 133-164.

[35] VIHMAN, M. M., LUM, G. T. J., KEREN-PORTNOY, T., AND MARTIN, P. Onset of word form recognition in english, welsh and english-welsh bilingual infants. *Applied Psycholinguistics*. In press.

[36] VIHMAN, M. M., AND MILLER, R. *The Emergent Lexicon: the child's development of a linguistic vocabulary*. Academic Press, New York, 1988, ch. Words and babble at the threshold of lexical acquisition.

[37] VIHMAN, M. M., NAKAI, S., DEPAOLIS, R. A., AND HALLÉ, P. The role of accentual pattern in early lexical representation. *Journal of Memory and Language* 50 (2004), 336-353.

[38] VIHMAN, M. M., AND VELLEMAN, S. L. *Phonological knowledge: its nature and status*. Oxford University Press, Oxford, 2000, ch. Phonetics and the origins of phonology, pp. 305-339.

[39] WATERSON, N. Child phonology: a prosodic view. *Journal of Linguistics* 7 (1971), 179-211.

[40] WERKER, J. F., FENNELL, C. T., CORCORAN, K. M., AND STAGER, C. L. Infants' ability to learn phonetically similar words: effects of age and vocabulary size. *Infancy* 3 (2002), 1-30.

[41] WERKER, J. F., AND STAGER, C. L. *Papers in Laboratory Phonology: acquisition and the lexicon*, vol. 5. Cambridge University Press, Cambridge, 2000, ch. Developmental changes in infant speech perception and earl word learning: is there a link?, pp. 181-193.

[42] WHALEN, D. H., LEVITT, A. G., AND WANG, Q. Intonational differences between the reduplicative babbling of french- and english-learning infants. *Journal of Child Language* 18 (1991), 501-516.

How might a rapid serial visual presentation of text affect the prosody projected implicitly during silent reading?

Eva M. Fernández

Queens College and Graduate Center, City University of New York

eva.fernandez@qc.cuny.edu

Resumo

Neste artigo revisarei algumas investigações que testam como as prosódias explícita e implícita influenciam a resolução de ambigüidades. Descreverei os resultados de cinco experimentos que examinam se a "prosódia visual" afeta a prosódia projetada durante a leitura silenciosa. O Experimento 1, que apresenta estímulo acústico, demonstra que a prosódia explícita afeta a resolução de ambigüidade durante a escuta. O Experimento 2, com estímulo escrito, replica efeitos que são apresentados em outras publicações como estando ligados à prosódia implícita. O Experimento 3 explora um paradigma RSVP com velocidade fixa para desorganizar a projeção da prosódia implícita. Finalmente, os Experimentos 4 e 5 usam uma velocidade RSVP variável para recuperar em parte a projeção da prosódia implícita, eliciando efeitos que parecem se assemelhar àqueles observados na leitura silenciosa e na escuta.

1 From a broad context to a theoretically interesting question

Rapid serial visual presentation (RSVP), self-paced reading, and other related text-based experimental paradigms have of late come to be regarded as second-class methods for research on the processes that underlie the comprehension of sentences. First-class techniques include those that record eye movements (Van Gompel, 2007) or examine neural activity during reading or listening (Brown; Hagoort, 2000). The superiority of these methods is attributable to the fact that they provide fine-grained data about the moment-by-moment processes at play during sentence comprehension, in tasks that can avoid the need for presenting linguistic stimuli in writing, and might not even require conscious awareness of the stimuli. We now therefore have ways to study in great detail the very early stages of sentence processing, while at the same time eliminating confounding variables (such as individual differences in literacy or metalinguistic awareness) and expanding the range of testable populations (to include preliterate children, or even preverbal infants; Sekerina; Fernández; Clahsen, in preparation). Why would any self-respecting psycholinguist insist on using second-class techniques?

UNIVERSIDADE DE LISBOA
BIBLIOTECA

A lack of ecological validity is another criticism of the family of techniques that rely on presenting sentences word-by-word or chunk-by-chunk on a computer screen. Reading in such a mode, it is claimed, is not much like reading in real life. In a self-paced reading experiment, participants read sentences that appear in small (or big) chunks on a computer screen. With RSVP displays, words appear one by one on the screen in a rapid sequence. Most of the reading we perform on a daily basis hardly resembles this type of unnatural exercise. Normally, words for one or multiple sentences are arranged on a page in single lines or paragraphs, where they are available for later inspection, and where they may be insignificant enough (or parafoveally discernible enough) so as to be occasionally skipped (Drieghe; Brysbaert; Desmet; De Baecke, 2004). However, ubiquitous small-screen technologies call this generalization into question (Mitchell, 2004). Word-by-word and chunk-by-chunk displays are now part of the reading experience for regular users of cellular telephones, portable digital assistants, and digital music players. Such gadgets have sophisticated text-display functions, with screens that are often no bigger than a few square centimeters. Small-screen devices display not just short titles for musical selections or telephone numbers and addresses for contacts, but also driving directions, email messages, newspaper articles, perhaps even complete novels. The size of the screen, as well as the size of the font, determines how much text can be displayed on each screen, but the user interface differs from system to system. In some, the reader paces herself manually from chunk to chunk of text, by pressing a button or sliding a roller. In others, chunks of text are displayed serially at a pre-set rate that the user may be able to speed up or slow down. Technologies like this are prompting a renewed interest in experimental work that examines what reading is like under such circumstances (Castelhano; Muter, 2001; Goldstein; Öqvist; Bayat-M; Ljungstrand; Björk, 2001). The experimental data described below are well framed within this context of small screen technologies, since the data could potentially inform on their design, to improve reader experience. The questions explored below, however, are not formulated about reading experience.

The experimental series focuses on whether the prosodic representation that people build mentally while reading a sentence (implicit prosody) can be manipulated by variably timed visual displays. I will argue that “visual prosody”—that is, structured timing patterns for serial visual presentations of text—can influence the resolution of syntactic ambiguities, because the timing of the visual signal affects the prosody that is hypothetically projected mentally during silent reading. This type of work is of both practical and theoretical interest. On the one hand, we have a new technique to study prosody, one that avoids the multiple problems associated with using acoustic stimuli, such as ensuring that test utterances are natural and uniformly instantiate the desired prosodic manipulations. On the other hand, we have new data to add to the set of investigations concerned with implicit prosody in particular, to refine the theories that describe the syntax-phonology interface in sentence processing.

In what follows, I will first review some relevant investigations probing how explicit and implicit prosody influence ambiguity resolution. I will then describe the outcome of five experiments examining whether “visual prosody” affects the prosody projected during silent reading. Experiment 1, which features acoustic stimuli, demonstrates that explicit prosody affects ambiguity resolution during listening. Experiment 2, with written stimuli, replicates effects that have been

argued elsewhere to be linked to implicit prosody. Experiment 3 exploits a fixed-rate RSVP paradigm to disrupt the projection of implicit prosody. Finally, Experiments 4 and 5 use variable-rate RSVP to recover in part the projection of implicit prosody, eliciting effects that begin to resemble those observed with silent reading and with listening.

2 Does prosody affect the interpretation of sentences, while listening?

Explicit prosody—that is, the prosody present in a spoken utterance—can affect ambiguity resolution during listening, because phrasings of one type or another, as well as intonation contours of one type or another, can bias listeners to understand strings as having one syntactic structure or another (among others, Lehiste, 1973; Price; Ostendorf; Shattuck-Hufnagel; Fong, 1991; Carlson; Clifton; Frazier, 2001). The supra-segmental properties of speech collectively represent the explicit prosody of an utterance. This level of representation is signaled in an utterance by phrasing (durational) and intonation (pitch modulation) cues, to indicate the boundaries and internal properties of prosodic constituents (intonational phrases, pitch accents, prosodic words, syllables, and so forth). But while these units play a role in the phonological representation of an utterance, how, if at all, do they affect the syntactic representation that is built while listening? This question has proved to be surprisingly difficult to answer in experimental psycholinguistics, but the growing consensus (e.g., Cutler; Dahan; Van Donselaar, 1997) is this: prosody heard while listening to ambiguous strings can bias listeners toward one or another interpretation. And this makes intuitive sense: ambiguous strings like *ladies and gentlemen from Oklahoma*, frequently encountered in readings of all kinds, are almost never uttered such that the intended syntactic representation remains obscure. When one listens, interlocutors let one know what they mean: their prosody betrays them. In the example just cited, a pause after *ladies* will lead to one interpretation, a pause after *gentlemen* to another.

One problem that makes the study of prosody difficult is the non-categorical nature of well-formedness judgments for prosodic structures. In syntax, the grammaticality of a string of words is almost always clear (e.g., **lady and gentlemen from Oklahoma*). In prosody, though, the location of a pause or the shape of an intonational contour might go against the grain of the prosodic phonology without being outright ungrammatical (e.g., if I were to produce a phrasal break between each of the syllables of *Oklahoma*). It is therefore desirable to think of specific prosodic contours as felicitous or infelicitous, to borrow terminology from the study of pragmatics and information structure. This indeterminate nature of prosody has consequences for the syntax-prosody interface. For instance, a particular prosodic contour for a syntactically ambiguous string will boost one interpretation without ruling out the other, and a particular prosodic contour might fail to disambiguate altogether. Kang and Speer (2004) demonstrate that for sentences like *Aaron followed a poor guy drinking his soda*, a phrasal break after the verb *followed* strongly biases interpretation of the subject of the ambiguous clause, *drinking his soda*, as referring to a *poor guy*. However, a phrasal break before the ambiguous clause, *drinking his soda*, does

not have the mirror effect, but rather results in participants responding that the person drinking the soda could be either Aaron or the poor guy. So, certain prosodic attempts to deal with syntactic ambiguity can be more effective than others.

Let us turn to two experiments demonstrating that explicit prosody influences ambiguity resolution. Both are particularly relevant because they deal with the same construction used in the experiments reported below, the relative clause attachment ambiguity. Schafer et al. (1996) studied the preferred interpretation of sentences like (1):

- (1) The sun sparkled on the propeller of the plane that the mechanic was so carefully examining.

The attachment of the relative clause, *that the mechanic was so carefully examining*, is globally ambiguous: the mechanic could be examining the *propeller* (N1) or the *plane* (N2). Schafer and colleagues asked participants to listen to utterances of sentences like these, in which either N1 (*propeller*) or N2 (*plane*) received a pitch accent. Interpretation was measured by asking a comprehension question like *What was the mechanic examining?* Participant responses indicated that pitch accents attract modification: when *propeller* was accented, N1 interpretations were reliably more frequent (65%) than when *plane* was accented (53%).

Schafer and colleagues thus demonstrate that focus, realized prosodically by means of pitch accents, can affect the interpretation of the ambiguity. Continuing with the same ambiguous construction, let us examine a different aspect of prosody: phrasing. Igoa and Teira (2003) report a series of experiments carried out using the translation-equivalent in Spanish of the construction in (1), but with morphology that disambiguated attachment to N1 (2a), or to N2 (2b). (The literature variously describes N1 attachment also as high or non-local attachment, and N2 attachment as low or local attachment.)

- (2) a Ellos llamaron al tío de los niños que estuvo tan antipático.
They called the_{SG} uncle_{SG} of the_{PL} children_{PL} that was_{SG} so unpleasant_{SG}
- b Ellos llamaron al tío de los niños que estuvieron tan antipáticos.
They called the uncle_{SG} of the_{PL} children_{PL} that were_{PL} so unpleasant_{PL}

Igoa and Teira's audio materials included a prosodic phrasing manipulation meant to bias the interpretation of the ambiguity. A phrasal break ("|") before the preposition (*...the uncle || of the children that...*) supports the low attachment interpretation, while a boundary before the relative clause (*...the uncle of the children || that...*) supports the high attachment interpretation. (The prosodic boundaries were realized in the materials by phrase-final lengthening and pausing as well as a rising boundary tone, typical for Spanish sentence-medial phrasal breaks.) The experimental design crossed the two variables of syntactic disambiguation and prosodic bias, such that the prosody was either congruent or incongruent with the syntax. Igoa and Teira asked listeners to perform a well-formedness judgment task. Mean errors and mean reaction times revealed the same type of data pattern: incongruent prosody was costly (it

UFPA - Faculdade de Letras
Biblioteca

resulted in more errors and longer reaction times), but much more so when materials forcing N1-attachment syntactically (2a) contained prosody supporting a low attachment interpretation (a prosodic boundary before the preposition). In contrast, congruent prosody had a facilitating effect, particularly when materials forcing N2-attachment syntactically (2b) contained prosody supporting a low attachment interpretation. These results demonstrate that two different prosodic phrasings for this construction bias (or at least facilitate) interpretation, but the two contours have asymmetric force: the prosody that signals low (N2) attachment seems to be a more reliable indicator than the prosody that signals high (N1) attachment.

3 Does prosody affect the interpretation of sentences, while reading silently?

We have just seen that the interpretation of the relative clause attachment ambiguity can be affected by explicit prosodic manipulations, like pitch accenting (Schafer et al., 1996) and phrasing (Igoa; Teira, 2003). We now turn to whether the structural choices made in the face of syntactic ambiguity can be affected by a prosody that is hypothesized to be projected mentally. During silent reading, explicit prosody is absent, and few graphemic cues provide information about the prosodic structure that might accompany the sentence (on the role of punctuation in silent reading, see Hirota, Frazier and Rayner (2006)). In spite of the absence of explicit cues to guide readers, a great deal of evidence has confirmed that phonological representations are built during completely silent reading.

Slowiczek and Clifton (1980) demonstrated that silent reading comprehension is impaired when the subvocalization is blocked by a verbal task. In these classic experiments, subvocalization was blocked by having participants count, or repeat a word or words (e.g., *coca-cola cola*). One way to understand this effect invokes the model of working memory of Baddeley and Hitch (1974; Baddeley, 2002), in which a phonological loop stores phonological information, presumably based on a phonological representation projected mentally, generated using default phonological settings for the language. The information stored in the phonological loop decays very rapidly if it is not rehearsed. During silent reading, rehearsal is achieved by subvocalization. But if the system is busy performing a verbal task, it has a harder time building and retaining a phonological representation of the sentence it is reading.

So, subvocalization permits projecting a prosodic structure, and that prosodic structure feeds into sentence processing, facilitating the extraction of meaning from sentences. But can subvocalization and the implicit prosody it generates actually bias the way an ambiguity is processed? Two proposals about this were offered at approximately the same time: Bader (1998) examined how focus in German affected the reanalysis of certain local syntactic ambiguities, and Fodor (1998) examined the resolution of ambiguous relative clauses in multiple languages. We will focus on Fodor's proposal, the *Implicit Prosody Hypothesis* (IPH), because it deals precisely with the ambiguity under consideration.

Fodor (1998, 2002) hypothesized that the prosody projected during silent reading could affect ambiguity resolution, offering this idea as an explanation

for two robust effects observed in a number of studies on relative clause attachment. One effect is that relative clauses are more likely to be attached high when they are long (e.g., Fernández, 2003). The other is that relative clauses are more likely to be attached high by readers of some languages (Spanish among them), compared to readers of other languages (English among them; e.g., Cuetos; Mitchell, 1988). The evidence on whether the IPH can account for cross-linguistic differences is still not very abundant, so we will not consider this problem further (though some suggestive data are reported by Fernández and Bradley (2004); Fernández, Bradley, Igoa and Teira (2003); see also Fernández (2005)). We focus instead on the matter of length effects and the compelling explanation that links them to implicit prosody (Fodor, 1998, 2002).

Adding length to a sentence increases the probability of sentence-internal prosodic discontinuities (phrasal breaks), and for English speakers, producing sentences with the relative clause attachment construction, added length anywhere promotes phrasal breaks before the relative clause rather than elsewhere in the sentence (Bradley; Fernández; Taylor, 2003). A phrasal break before the relative clause (i.e., following N2) is the type of phrasing that promotes high-attachment interpretations. With a short relative clause, in contrast, no sentence-internal break is needed, so the (computationally simpler) attachment to the local N2 is preferred.

4 Examining implicit prosody using RSVP

One way to demonstrate that implicit prosody, and not some other variable, is the force behind the length effects just described is to show that length effects disappear when implicit prosody is disrupted during silent reading. Slowiaczek and Clifton (1980) disrupted implicit prosody by asking readers to engage in a verbal production task while reading silently. A different approach, exploited here, induces disruption visually: reading naturally (and, by extension, building a mental phonological representation of the material being read) will be difficult if the visual presentation of the words is serial, isochronous, and unbearably slow.

RSVP (rapid serial visual presentation) is an experimental paradigm introduced by Forster (1970; see also Potter, 1984) to study how far the sentence processor could be taxed before its routines broke down. Later work focused on some of the advantages of RSVP presentations, compared to static displays, including the fact that the fixed display of RSVP reduces eye strain (Potter, 1984). (On the limitations of the RSVP paradigm, in comparison to other text-display based techniques, see Just, Carpenter and Woolley (1982)). As mentioned earlier, small screen technologies used to display text have prompted recent research examining reader experience when using RSVP presentations of text with small displays. Castelhana and Muter (2001), for example, report that RSVP presentations in which punctuation is accompanied by pauses improve user preference ratings.

Potter (1984) notes that RSVP readers can handle speeds of 12 words per second without a severe compromise in comprehension. The RSVP presentation used here is nowhere near as rapid as that. The speed of the display in Experiment 3—clocking in at 2 words per second—is dramatically slower (and perhaps would warrant our calling the method of display *Slow-SVP*, rather than

Rapid-SVP). But this is not a design flaw, since the objective of using a serial visual presentation of text is to determine what happens to interpretation in the presence of an isochronous and very slow incoming signal, one which I claim disrupts implicit prosody.

In line with conditions under which RSVP reading has been found to lead to improved user preference (Castelhano; Muter, 2001), Experiments 4 and 5 introduce a faster rate of display (225 ms, base rate) that varies based on word length (20 ms per character is added to the base rate). Crucially, the objective is to recover effects linked to implicit prosody by presenting text at rates that resemble the experience during reading and listening, both of which involve looking at or listening to the stimulus in a serial and variably timed way. The display timing in Experiments 4 and 5 is of approximately 3 words per second (an average of 311 ms per word, for target materials). This rate is not much slower than the timing of eye fixations during reading (which hover around 225 ms, modulated by word length (Rayner, 1998)). It is also close to average speaking rates, which tend to be around 3 words per second (e.g., Yuan; Liberman; Cieri, 2006), though speaking rate is highly variable (along parameters that include age, sex, topic of the interaction, nativeness).

5 From listening to reading, to reading without and with “visual prosody”

We now turn to the experimental report, which omits details regarding the design and analyses due to space considerations. The primary design components series of the experiments are summarized in Table 1. We will first examine what they have in common (namely, participants and materials) and then describe their differences (procedures) as well as the outcomes.

5.1 Participants

Participants – a different sample for each of the five experiments – were native speakers of North American English. They were recruited from the undergraduate student body at Queens College (Flushing, New York). Data were rejected and replaced for any participant who was not a native speaker or who claimed to speak a language other than English fluently, as revealed by answers on a language history questionnaire. Such rejections made up a negligible proportion of the total sample (less than 3%).

Experiments 2 and 3, with a design calling for two experimental lists, required less participants overall ($N=18$) than the other experiments ($N=60$), each with six lists. Notably, performance on fillers is comparable between the experiments with identical designs. The lower error rates in Experiments 2 and 3 might have to do with the fact that the lexical material remained on the screen longer (whole sentences with a time out of 9000 ms for Experiment 2; 500 ms per word for Experiment 3).

5.2 Materials

The sentences employed as targets and fillers were very close to identical throughout. Given their smaller design, Experiments 2 and 3 employed fewer

UFPA - Faculdade de Letras
 Biblioteca

	Stimulus presentation (procedure); design variables	N Targets	N Fillers	N Participants	Errors on Fillers
Experiment 1	Audio (listening); length and prosody	36 × 6	72	60	11%
Experiment 2	Whole-sentence (reading); length	32 × 2	64	18	5%
Experiment 3	Fixed-rate RSVP (reading); length	32 × 2	64	18	6%
Experiment 4	Variable rate RSVP with visual "accenting" (reading); length and prosody	36 × 6	72	60	11%
Experiment 5	Variable rate RSVP with visual "pauses" (reading); length and prosody	36 × 6	72	60	12%

Tab. 1: Primary design components and filler accuracy data for Experiments 1-5.

items (a subset of the total 108). Importantly, the ratio of targets to fillers was identical across experiments, one to two. Practice materials (N=5) were the same for all five experiments, as were list protector items (N=5 at the beginning of the experiment, N=5 at the end).

The target sentences contained the relative clause attachment construction placed in object position, with a relative clause that could permissibly modify either the higher noun (N1) or the lower noun (N2), on morphosyntactic, semantic, and pragmatic grounds. Targets, as the example below, had relative clauses that were short (1 prosodic word, (3a)) or long (3 prosodic words, (3b)), the length variation systematically produced by adding lexical material to the short version, as in the example in (3):

- (3) a Amelia knows the clerk of the accountant who was married.
- b Amelia knows the clerk of the accountant who was married in the large cathedral.
- c Who was married?
 the clerk, the accountant

For targets short or long, interpretation was measured by asking questions probing the resolution of the ambiguity. The example in (3c) is illustrative of how questions were always formulated, citing only the relative clause verb, regardless of relative clause length.

Filler items had associated questions with only one correct answer, appearing on the right or left, counterbalanced; in the sample filler in (4), the correct answer is placed on the right:

- (4) a Miranda had an ivory box where she kept her great-grandmother's jewelry.

- b. What used to belong to Miranda's great-grandmother?
the ivory box, the jewelry

The order of the answers coincided with the order (left-to-right) that the related lexical material appeared in the stimulus sentence. Thus, for target materials, N1 was always the choice on the left, and N2 the choice on the right. This systematic relationship between stimulus sentences and the form of the questions was meant to prevent participants from becoming distracted by the order of the elements in the answer.

Sentence comprehension questions for all items, in all experiments, were presented visually, on two lines centered on the screen. Participants responded by pressing one of two buttons on a keyboard (the right and left shift keys), each corresponding to one of the two answers.

In Experiments 1, 4 and 5, there were sentence internal phrasing manipulations in the target materials. These manipulations, described in greater detail below, involved one variably located phrasal boundary (Experiment 1), a variably located word that remained longer on display (Experiment 4), or a blank screen variably located between words (Experiment 5). To match the proportion of these manipulations in the targets, two-thirds of the fillers contained comparable sentence-internal phrasing manipulations. For example, in Experiment 1 the stimulus sentence for the filler item in (5) was uttered with a phrasal break after *saw*, in Experiment 4 it was presented with *saw* remaining on the screen three times longer than other words, and in Experiment 5 it was presented with a 675 ms blank screen after *saw*.

- (5) a. Julie called her husband as soon as she saw that her son had
escaped again.
- b. Who had escaped?
Julie's husband, Julie's son

Notice that in all three cases, there are events affecting a location where a phrasal break is permissible and perhaps even desirable, before the sentential complement, *that her son had escaped again*. The visual manipulations in Experiments 4 and 5 were devised to indicate visually what might be signaled acoustically by either a pitch accent or word-final lengthening (the increased duration of *saw* in Experiment 4) or by a pause (the blank screen after *saw* in Experiment 5).

We now turn to the experimental findings, which will be presented along with remarks on the experiment-specific properties of the procedures. The experiments together aim to address three questions: First, Experiment 1, a listening comprehension study, examines whether overt prosody influences syntactic ambiguity resolution. Second, Experiment 2, a whole-sentence reading comprehension study, is designed to address whether implicit prosody during silent reading influences syntactic ambiguity resolution. Third, Experiments 3, 4, and 5 use variations on RSVP presentations to examine whether "visual prosody" manipulations affect ambiguity resolution, in ways that emulate the effects caused by explicit prosody.

5.3 Experiment 1: Listening Comprehension

For this first experiment, recordings of the materials were made by a female native speaker of standard North American English drawn from the same population as the listeners. The speaker, a professional musician with some background knowledge of phonetics and phonology, trained herself to produce the sentences at a natural and uniform speech rate, and with uniform intonation, using the declination pattern normal for English declaratives. Sentence-internal phrasal breaks were produced with a falling boundary tone and a brief pause.

Each item in the target materials was recorded with three distinct prosodic forms, in each of its long and short versions: utterances contained no phrasal breaks at all ("No-Breaks"), a phrasal break after N1 ("N1-Break"), or a phrasal break after N2 ("N2-Break"). A detailed acoustic analysis of the materials is underway, to confirm what informal ear inspection of the utterances suggests, that the phrasing and intonation characteristics for each of the three prosodies were realized uniformly across length manipulations. Preliminary acoustic data on one-third of the target items are reassuring that intonation and durational patterns were similar across the two different relative clause lengths. In N=11 targets analyzed so far, the average pitch for N1 and N2 is roughly equal for all three prosodic forms of the sentences; in N1-Break materials (produced with a mean 180 ms pause after N1), N1 is on average 121 ms longer in duration than N1 in either No-Breaks or N2-Break materials (for the three prosodies, durational differences between short and long relative clause versions, $p > .50$); in N2-Break materials (produced with a mean 247 ms pause after N2), N2 is on average 49 ms longer in duration than in either No-Breaks or N2-Break materials (again, durational differences between short and long versions, $p > .50$).

The materials were presented over headphones. Upon participants' request, the comprehension question appeared visually, and was answered by pressing one of two buttons, corresponding to the two interpretations of the relative clause.

This experiment is designed to confirm, for English, that the two different types of prosodies, N1-Break and N2-Break, have different types of effects on interpretation: a break after N1 should decrease N1 interpretations, a break after N2 should increase N1 interpretations. The No-Breaks prosody should establish the baseline interpretation preferences, in the absence of prosodic bias. There should be no effect of relative clause length in this experiment, if it is exclusively caused by implicit prosody.

Figure 1 displays the data for Experiment 1, reported as % N1 choices in responses to the comprehension question. The distribution of N1 choices revealed a modest (5%) but reliable difference between No-Breaks and N2-Break materials ($F(1,54)=4.62, p < .05$; $F(1,30)=3.99, p = .055$), which did not interact with length ($F(1, F(2) ; 1)$). Just as expected, the length of the relative clause made no difference to interpretation with No-Breaks and N2-Break materials, but the presence of a phrasal break after N2 increased the likelihood of N1 (high attachment) interpretations.

UNIVERSITY OF TORONTO LIBRARY

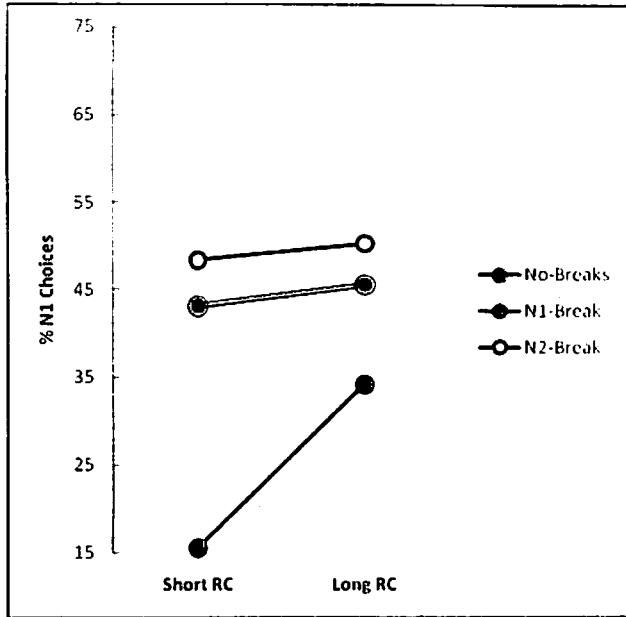


Fig. 1: % N1 choices for No-Break, N1-Break, and N2-Break materials as a function of relative clause length, Experiment 1.

Figure 1 also depicts a sizable difference (19.4%) between No-Breaks and N1-Break materials ($F_1(1,54)=86.39, p < .001$; $F_2(1,30)=44.18, p < .001$). (Recall that N1-Break prosody is the type that biases toward low attachment interpretations.) The placement of a phrasal break after N1 substantially lowered N1 attachment interpretations. That N1-Break prosody is stronger, somehow, than N2-Break prosody is in line with the effects reported by Igoa and Teira (2003), discussed earlier.

The length effect with N1-Break materials (a difference of 18.6%; $F_1(1,54) = 47.01, p < .001$; $F_2(1,30)=38.24, p < .001$) remains a puzzle for future work to tackle. This effect could in principle be linked to artifacts in the audio stimuli, a possibility that on-going analyses of the materials should help rule out. This effect is more likely sourced in as yet not well understood variables that affect how phonological representations decay in the presence of lengthy and therefore potentially semantically and pragmatically complex lexical material.

5.4 Experiment 2: Silent whole-sentence reading

In Experiment 2, rather than listening, participants were asked to read. All sentences in the experiment appeared on a single line, displayed in the center of a computer screen, and followed (upon participants' request) by their respective comprehension question. In this experiment the only variable was length of the relative clause, included to identify the default length effect for these materials, when they are read silently.

As expected, in Experiment 2 the length effect had a size of 9.1%: materials with short relative clauses were interpreted as high-attached 54.5%, compared to 63.6% for those with long relative clauses ($F_1(1,16)=7.37, p < .02$; $F_2(1,30)=7.65, p < .01$). Arguably, when the relative clause was long, the implicit prosody projected during silent reading was more likely to contain a sentence internal phrasal break. Such breaks are more likely sited before the relative clause, the type of prosody that modestly encourages N1 interpretations, as in Experiment 1.

Let us compare the overall rate of attachment in Experiment 2, 59.1%, to the overall rate in Experiment 1's No-Breaks condition, 44.3%. Experiment 1's result reflects intrinsic biases for these materials (e.g., based on their lexical content), in the absence of biasing prosody. In contrast, Experiment 2's overall higher rate of attachment is likely the result of readers more frequently projecting sentence-internal phrasal breaks: in principle, they had as much time as necessary to read and re-read (phrase and re-phrase) the sentence. This is speculative, but supported by the evidence we now turn to, which reveals what happens when we disrupt (by visual means) the likelihood of naturally projected prosody during silent reading.

5.5 Experiment 3: Fixed-rate RSVP

The sentences used in Experiment 2 were presented in RSVP format, at 500 ms per word, without intervals between words. The expectation was that the isochrony of the presentation would prevent readers from building a useful prosody or from building a prosody at all, and would therefore reduce or eliminate the relative clause length effect. This is precisely the outcome: short and long relative clauses had very low and very similar rates of N1 attachment, 43.7% and 45.1% respectively ($F_1, F_2 < 1$).

Figure 3 re-plots the data for No-Breaks materials in Experiment 1, and plots the data for whole-sentence reading in Experiment 2 and fixed-rate RSVP reading in Experiment 3. Whole sentence reading (black squares) stands out above the others as the reading condition with the highest rate of attachment (59.1%) and with the strongest length effect (9.1%). It is also remarkable that listening (No-Breaks materials) and fixed-rate RSVP produce such similar rates of attachment and fail to obtain a length effect. I have so far argued that this is all linked to prosody. But perhaps the culprit is working memory. Listening and fixed-rate RSVP have the following in common: neither permits looking back, and N2 is always closer to the end so it is by default more memorable.

Experiments 4 and 5 address this problem empirically, using a combination of variable timing manipulations. We will be able to trace to either prosody or working memory (or both) the effects that these manipulations have on the interpretation of the ambiguity.

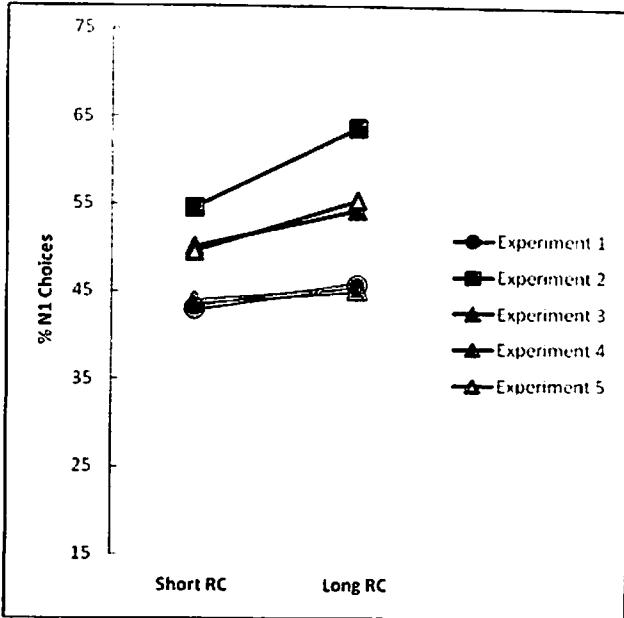


Fig. 2: %N1 choices for listening (No-Breaks materials, Experiment 1), whole-sentence reading (Experiment 2), fixed-rate RSVP (Experiment 3) and variable-rate RSVP (No-Breaks materials, Experiments 4 and 5), as a function of relative clause length.

5.6 Experiments 4 and 5: Variable-rate RSVP

Experiments 4 and 5 presented words at a rate of 225 ms per word plus 20 ms per character. As pointed out earlier, this speed of display resembles moderate speech rates and falls within the range of eye fixation durations during normal reading, thus facilitating a comparison to data from Experiments 1 (listening) and 2 (whole sentence reading). Informal piloting of this and faster as well as slower presentation rates, using readers from the same pool of participants, indicated that this was a comfortable speed: not too fast, not too slow.

No-Breaks data combined for Experiments 4 and 5 are plotted above in Figure 2, which shows that the variable timing of the display has two evident effects. Firstly, overall attachment is higher (7.8%) than it was in two directly comparable conditions: No-Breaks listening in Experiment 1 and fixed-rate RSVP in Experiment 3. Nevertheless, the 7.8% attachment rate of Experiments 4 and 5 is still lower than it was with whole-sentence reading in Experiment 2. Secondly, the slope of the solid triangles indicates a moderate restoration (5%) of the length effect ($F(1,108)=5.34, p < .025$; $F(1,30)=3.39, p < .10$). Together, these two outcomes suggest that the variable-rate RSVP in Experiments 4 and 5 slightly restored readers' ability to project implicit prosody, which in turn modulated their attachment preferences.

Experiment 4 included a "visual prosody" manipulation by which N1 or N2 were "accented", that is, their duration was lengthened to 3 times their

ordinary rate, simulating the rhythm that results from producing N1 or N2 with a pitch accent denoting focus (Schafer et al., 1996). The data for N1- and N2-Longer materials in Experiment 4 are plotted in the left panel of Figure 3. In Experiment 4, visual "accenting" had no effect (no difference between No-Breaks, N1-Longer or N2-Longer materials) and did not interact with relative clause length ($F_1, F_2 < 1$). But clearly, both N1- and N2-Longer materials in Experiment 4 produced a robust effect (8.2% in size) of relative clause length ($F_1(1,54)=11.11, p < .002; F_2(1,30)=13.87, p < .001$).

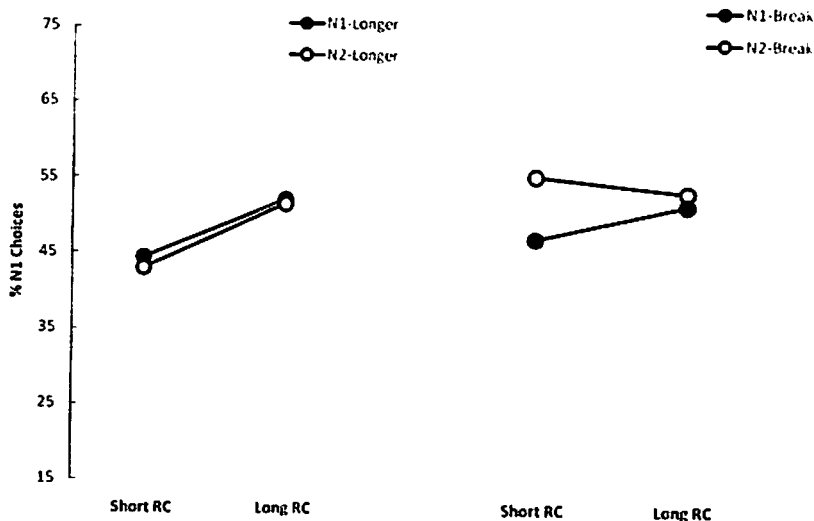


Fig. 3: % N1 choices for "visual prosody" manipulations in Experiments 4 and 5, as a function of relative clause length. The left panel displays data for N1- and N2-Longer materials in Experiment 4; the right panel displays data for N1- and N2-Break materials in Experiment 5.

A strict working memory account of Experiment 4's data is not feasible: leaving N1 or N2 on the screen longer should have made them more memorable, and therefore produced differences in rate of attachment, but clearly this did not happen. Notice also that we cannot interpret the "visual prosody" manipulation as one that emulates the effects of pitch accenting, since lengthened nouns did not attract modification as they did in the listening of Schafer et al. (1996).

The design of Experiment 5 was identical to that of Experiment 4, and included a "visual prosody" manipulation, only instead of visual "accenting", two thirds of the materials contained visual "pauses". In such sentences, a 675 ms blank screen (three times the base rate of 225 ms) was placed either after N1 or after N2. These visual "pauses" were meant to perform the same function as the phrasal breaks in the explicit prosody of Experiment 1: N1-Break pauses should decrease N1 attachment rates, while N2-Break pauses should increase them, if prosody is what is affected by this visual manipulation. If, in contrast, what is at stake is working memory disruptions, both N1- and N2-Break pauses should decrease N1 attachment rates, since both increase the distance between N1 and the end of the relative clause by the same amount of time.

The data for the two "pausing" conditions in Experiment 5, plotted in the right panel of Figure 3, are strikingly similar to those for the corresponding conditions in Experiment 1. One important difference has to do with the size of the effect, notably smaller with "visual prosody". "Pauses" after N2 decrease N1 interpretations, while "pauses" after N1 increase them, when the relative clause is short ($F_1(1,54)=4.66$, $p < .05$; $F_2(1,30)=3.30$, $p=.079$). Thus, visual manipulations emulate the effects observed with explicit prosody.

Further work is required to explain what underlies the fact that the data patterns supporting some of the arguments made here are cleanest with short relative clauses. Perhaps these effects are so rapid, within the working-memory buffer, that with longer sentences, "visual prosody" decays too quickly. No doubt insights about this are to be found by studying the interface between the phonological loop and the visual processor in working memory.

6 Summary, conclusions, and limitations

I have discussed a number of effects related to prosody and the resolution of syntactic ambiguity, placed in the larger context of research on explicit and implicit prosody. The experimental series presented in the previous section is anchored by the findings of Experiments 1 and 2. Experiment 1 shows that the location of a phrasal break, when the prosody is explicit, can influence the interpretation of syntactic ambiguities. With the relative clause attachment construction, a break after N1 strongly promotes low attachment interpretations, while a break after N2 modestly encourages high attachment interpretations. In whole sentence reading, as in Experiment 2, the availability of implicit prosody raises overall attachment rates and produces a length effect, materials with long relative clauses prompting more high attachment interpretation than materials with short relative clauses. But Experiments 1 and 2 merely replicate effects that are already well known.

The novel data come from the three RSVP experiments. Experiment 3, using a fixed-rate RSVP presentation, made the length effect vanish and attachment rates plummet. The variable-rate RSVP of Experiments 4 and 5 recovered the length effect and boosted N1 attachment interpretations, though not with the strength observed with whole sentence reading. Visual "accents" (Experiment 4) had no effect, but visual "pauses" (Experiment 5) influenced interpretation in ways that emulate effects observed with explicit prosody in Experiment 1, albeit effects were smaller in size and restricted to materials with short relative clauses.

Overall, the investigation offers data that support a central role for implicit prosody in the resolution of the relative clause attachment ambiguity during silent reading. It also offers a new technique for studying prosody in sentence processing, one that is particularly attractive because it does not require producing explicit prosody. The technique is therefore especially valuable for studying phrasing patterns that are particularly difficult to elicit, even from trained speakers (e.g., prosodic contours in code-switched utterances).

I must stress how small the "visual prosody" effects actually are. This should not come as a surprise, but it does highlight the fact that explicit prosody is a solid cue exploited by the sentence processing machinery. The same machinery also reacts to the timing of visually presented text in ways that I have argued

are linked to implicit prosody. In addition to its theoretical import, this investigation offers insights that could be exploited not only in the laboratory, but also in the real world, where more and more small screens are being used to display text.

7 Author Note

This investigation grew out of a problem identified by Kerri Smith in her Master's thesis (2004). I am indebted to Kerri for the idea and for assistance in the early stages of the project. I also thank the research assistants who were instrumental with data collection: Shukhan Ng, Yukiko Koizumi, and Nur ("Fitri") Gebers. Fitri deserves additional thanks for producing the audio stimuli. Finally, my deepest gratitude goes to the organizers of ABRALIN 2007, for offering me the opportunity to present this work at the conference. This project was funded in part by PSC-CUNY 36 Grant No. 67230-00 36.

References

- [1] BADDELEY, A. D. Is working memory still working? *European Psychologist* 11, 56 (2002), 851-64. Reprinted from *American Psychologist*.
- [2] BADDELEY, A. D., AND HITCH, G. J. *Recent advances in learning and motivation*. Academic Press, New York, 1974, ch. Working memory, pp. 47-90.
- [3] BADER, M. *Reanalysis in sentence processing*. Kluwer Academic Publishers, Dordrecht, NL, 1998, ch. Prosodic influences on reading syntactically ambiguous sentences, pp. 1-46.
- [4] BRADLEY, D., FERNÁNDEZ, E. M., AND TAYLOR, D. Prosodic weight versus information load in the relative clause attachment ambiguity. *Annual CUNY Conference on Human Sentence Processing, 16th, 2003* (2003).
- [5] BROWN, C. M., AND HAGOORT, P. *The neurocognition of language*. Oxford University Press, Oxford, UK, 2000.
- [6] CARLSON, K., CLIFTON, C., AND FRAZIER, L. Prosodic boundaries in adjunct attachment. *Journal of Memory and Language* 45 (2001), 58-81.
- [7] CASTELHANO, M. S., AND MUTER, P. Optimizing the reading of electronic text using rapid serial visual presentation. *Behavior and Information Technology* 20, 4 (2001), 237-247.
- [8] CUETOS, F., AND MITCHELL, D. C. Cross-linguistic differences in parsing: restrictions on the use of the late closure strategy in spanish. *Cognition* 30 (1988), 73-105.
- [9] CUTLER, A., DAHAN, D., AND DONSELAAR, W. A. V. Prosody in the comprehension of spoken language: a literature review. *Language and Speech* 2, 40 (1997), 141-202.

- [10] DRIEGHE, D., BRYSHAERT, M., DESMET, T., AND BAECKE, C. D. Word skipping in reading: on the interplay of linguistic and visual factors. *European Journal of Cognitive Psychology*, 16 (2004), 79-103.
- [11] FERNÁNDEZ, E. M. *Bilingual sentence processing: relative clause attachment in English and Spanish*. John Benjamins Publishers, Amsterdam, 2003.
- [12] FERNÁNDEZ, E. M. The prosody produced by spanish-english bilinguals: a preliminary investigation and implications for sentence processing. *Revista da ABRALIN* 1, 4 (2005), 109-141.
- [13] FERNÁNDEZ, E. M., AND BRADLEY, D. Exploring the prosody of the re-attachment construction in english and spanish. Poster presentation, 2004.
- [14] FERNÁNDEZ, E. M., BRADLEY, D., IGOA, J. M., AND TEIRA, C. Prosodic phrasing in the re-attachment ambiguity: effects of language, re- length, and position. *9th AMLaP - Architectures and Mechanisms for Language Processing* (2003).
- [15] FODOR, J. D. Learning to parse? *Journal of Psycholinguistic Research*, 27 (1998), 285-317.
- [16] FODOR, J. D. Psycholinguistics cannot escape prosody. *Proceedings of the 1st International Conference on Speech Prosody* (2002), 83-88.
- [17] GOLDSTEIN, M., ÖQVIST, G., BAYAT-M, M., LJUNGSTRAND, P., AND BJÖRK, S. Enhancing the reading experience: using adaptive and sonified rsvp for reading on small displays. *Proceedings of IHM-HCI* (2001).
- [18] GOMPEL, R. P. G. V. *Eye movements: a window on mind and brain*. Elsevier, Oxford, UK, 2007.
- [19] HIROTANI, M., FRAZIER, L., AND RAYNER, K. Punctuation and intonation effects on clause and sentence wrap-up: evidence from eye movements. *Journal of Memory and Language*, 54 (2006), 425-443.
- [20] IGOA, J. M., AND TEIRA, C. The contribution of prosody to relative clause attachment in spoken sentences in spanish. Poster presentation, 2003.
- [21] JUST, M. A., CARPENTER, P. A., AND WOOLLEY, J. D. Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, 111 (1982), 228-238.
- [22] KANG, S., AND SPEER, S. R. Prosodic disambiguation of participle constructions in english. *Proceedings of the 2nd International Conference on Speech Prosody* (2004).
- [23] LEHISTE, I. Prosodic disambiguation of syntactic ambiguity. *Glossa*, 7 (1973), 107-121.
- [24] MITCHELL, D. C. *The on-line study of sentence comprehension: eyetracking, ERP, and beyond*. Psychology Press, Brighton, UK, 2004, ch. On-line methods in language processing: introduction and historical review, pp. 15-32.

UFPA - Faculdade de Letras
B 13112122

[25] PRICE, P. J., OSTENDORF, M., SHATTUCK-HUFNAGEL, S., AND FONG, C. The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America* 6, 90 (1991), 2956-2970.

[26] RAYNER, K. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin* 3, 124 (1998), 372-422.

[27] SCHAFER, A. J., ET AL. Focus in relative clause construal. *Language and Cognitive Processes*, 11 (1996), 135-163.

[28] SEKERINA, I., FERNÁNDEZ, E. M., AND CLAHSSEN, H. Language processing in children. In preparation.

[29] SLOWIACZEK, M. L., AND CLIFTON, C. Subvocalization and reading for meaning. *Journal of Verbal Learning and Verbal Behavior*, 19 (1980), 573-582.

[30] SMITH, K. M. *The influences of reading fluency and prosody on length effects of attachment preferences of ambiguous relative clause constructions*. PhD thesis, CUNY Graduate Center, New York, 2004. Unpublished.

[31] YUAN, J., LIBERMAN, M., AND CIERI, C. Towards an integrated understanding of speaking rate in conversation. *9th International Conference on Spoken Language Processing (Interspeech 2006 - ICSLP)* (2006). Available at: http://papers.ldc.upenn.edu/Interspeech2006/Interspeech.2006_Speech_Rate_Paper.pdf.

Reflexos segmentais da organização rítmica do português do Brasil

Maria Bernadete Marques Abaurre
Universidade de Campinas, CNPq

babaurre@matriz.com.br

Abstract

This purpose of this paper is to analyze the rhythmic organization of Brazilian Portuguese (BP) in a text read by seventeen informants from the state of São Paulo. The distribution of secondary stresses and the application of vowel sandhi processes across the boundaries of prosodic words were considered. The analysis of the data showed a high frequency of occurrence of binary trochees, which points to a correlation between the occurrence of these phenomena and the implementation of a binary rhythmic pattern in BP.

1 Introdução

Fazemos, neste trabalho, algumas considerações preliminares sobre a possível relação entre a organização rítmica do português do Brasil (doravante PB) conforme refletida em dados de um *corpus* específico, a distribuição dos acentos secundários e a aplicação de processos de sândi vocálico de degeminação, elisão e ditongação (Cf. Bisol, 1990, 1992, 1994, 1995, 1996, 1999, 2000, 2002, 2003) no interior de um domínio prosódico específico, a palavra prosódica: ω (Nespor; Vogel, 1968).¹

2 O contexto da investigação

O *corpus* utilizado para detecção e análise de dados faz parte do banco de dados do Projeto Temático Interdisciplinar FAPESP/PRONEX 03/09930-9, *Comportamento estocástico, fenômenos críticos e identificação de padrões rítmicos nas línguas naturais*.²

Esse projeto temático – que reúne probabilistas, matemáticos, físicos e linguistas – tem por objetivos investigar, com base na Teoria dos Processos Estocásticos:

1. A questão da existência ou não de padrões rítmicos nas línguas naturais;
2. A existência de uma tipologia discreta, caracterizada por pontos críticos bem definidos, em oposição a um contínuo rítmico;

¹Alguns autores preferem referir-se a este domínio prosódico como “palavra fonológica”.

²Coordenado, no Ime/USP, por Antonio Galves (<http://www.ime.usp.br/~tycho/>).

3. A existência de marcas do ritmo no sinal acústico e em textos escritos.

Teoria dos Processos Estocásticos estuda a evolução (temporal ou espacial) de sistemas com comportamento aleatório.³ Suas técnicas permitem extrair o comportamento "coletivo" de sistemas constituídos de um grande número de componentes. Pelo fato de pressupor a interação de uma série de variáveis, os padrões rítmicos das línguas naturais podem ser modelados a partir dos pressupostos dessa teoria. Não é esse o exercício que nos propomos fazer aqui, no entanto. O objetivo deste texto é contribuir para a identificação dos fatos de natureza lingüística que precisam necessariamente ser levados em conta por qualquer investigação que objetive modelar padrões rítmicos.

3 Processos segmentais e ritmo no português do Brasil: hipóteses iniciais

Neste texto, vamos discutir a possível correlação entre a ocorrência de acentos secundários e de processos de sândi vocálico e a implementação de uma determinada organização rítmica em Português do Brasil (PB). Focaliza-se, portanto, o primeiro aspecto da questão 3 do projeto temático (mencionada na seção 2, acima), a saber: a investigação de marcas do ritmo nos enunciados orais.

A implementação dos processos de sândi será investigada, em um primeiro momento, apenas no interior de um dos domínios da hierarquia prosódica, a palavra prosódica.

Para fins do trabalho, foram consideradas palavras prosódicas todas as palavras lexicais portadoras de acento próprio, precedidas ou não de palavras funcionais acentualmente inertes: (*me inclua, o antigo, se expressabilizando*).⁴ O trabalho está baseado nas seguintes hipóteses:

1. A implementação de uma determinada organização rítmica decorre da interação entre os acentos lexicais e da localização de acentos secundários em palavras prosódicas com duas ou mais sílabas pré-tônicas;
2. A aplicação de processos segmentais de apagamento/redução e de sândi pode otimizar a implementação da organização rítmica preferencial da língua;
3. O PB tende a construir pés binários.

Exemplos como os seguintes, extraídos do banco de dados do projeto temático, serviram de base para a formulação das hipóteses aqui investigadas:

³Cf. Berger, 1992; Feller, 1957; Karlin; Taylor, 1972; Lipschutz, 1994; Papoulis, 1991; Parzen, 1972; Resnic, 1999; Todorovic, 1992.

⁴O domínio da palavra prosódica ou palavra fonológica tem sido muito discutido nos trabalhos de fonologia prosódica, considerando-se as hierarquias prosódicas propostas. Para autores como Nespor e Vogel (1986), o que aqui se toma como palavra prosódica corresponde ao domínio do grupo clítico. Para fins do presente trabalho, tomaremos como definição desse domínio a que foi apresentada no corpo do texto. Para uma discussão específica sobre a palavra prosódica em português europeu, Cf. Vigário (2003). Ver também, para o português brasileiro, a discussão feita por Schwindt (2000) e Bisol (2005).

(1) (a) mo **der** **n(i)** za **ção** (apagamento/redução de vogal)

[dɛrn]

(b) car **tei** ra **de** **ha** bi li ta **ção** (ditongação)

[dʒa]

(c) de que **ti** nha ha vi do um **rou** bo (degeminação)

[ɲa] [du]

Os dados em (1) referem-se a trechos de uma narrativa espontânea produzida por um falante do Estado do Rio de Janeiro (Niterói). Na primeira linha, em (a), (b) e (c), as sílabas marcadas em **negrito** foram produzidas com acento secundário. As sílabas em **negrito** e **sublinhadas** são as portadoras de acento lexical. Na segunda linha, o ponto indica a ocorrência de acento secundário, e o asterisco indica a ocorrência do acento lexical. Na terceira linha, entre colchetes, indica-se a pronúncia após a aplicação do processo de apagamento/redução de vogal (a), ou após a aplicação de um processo de sândi vocálico (ditongação, em (b) e degeminação, em (c)).

Em (a), observa-se que o apagamento da vogal [i] leva à eliminação de uma sílaba, o que permite a realização de um pé troqueu (binário), [der n za], em vez do dátilo inicial, [der ni za], com três sílabas; em (b), a ditongação que se aplica às sílabas [dʒi] e [a], criando a sílaba inicial [dʒa] do pé trocaico [dʒa bi], impede a realização do dátilo [tei ra dʒi]; em (c), a degeminação (crase) que toma como contexto as sílabas [ɲa] e [a] força a realização do troqueu [tʃi ɲa] e evita o dátilo [tʃi ɲa a]. Nos dois primeiros exemplos, o processo se dá no interior de uma palavra prosódica: (modernização)_ω, (de habilitação)_ω; no terceiro exemplo, ocorre na fronteira de duas palavras prosódicas, portanto no interior de uma frase fonológica: ((tinha)_ω (havido)_ω) φ.

4 Sobre acento secundário e ritmo em PB

Hulst (1997) defende que a estrutura rítmica é dada pela atribuição de acentos secundários, possivelmente marcados pós-lexicalmente em termos de pés, ou seja, o autor defende que a atribuição de acentos secundários está associada à organização rítmica. Estas afirmações são confirmadas para o português brasileiro por Collischonn (1993, 1994), Abaurre e Galves (1998), Sandalo, Abaurre, Mandel e Galves (2006), dentre outros autores.

Segundo Abaurre e Galves (1998), Sandalo, Abaurre, Mandel e Galves (2006) e outros autores, o PB privilegia uma construção de unidades rítmicas binárias na implementação de acentos secundários.

A maneira como essa proeminência secundária é implementada em PB, favorecendo a construção de pés binários, determina a organização rítmica da língua, diferenciando-a da organização rítmica do português europeu (D'Andrade; Laks, 1999; Castelo, 2004).

O acento secundário tem sido estudado, em PB, no quadro da Fonologia Métrica (Collischoun, 1993) ou da Teoria da Otimalidade (Saudalo; Abaurre; Mandel; Galves, 2006).

Nas seções seguintes, verificaremos, a partir da identificação e análise de dados do *corpus* utilizado neste trabalho, de que maneira a implementação de proeminências secundárias e de processos de sândi vocálico, que levam à ressilabificação, pode estar relacionada à otimização de um padrão binário de organização rítmica em PB.

5 Corpus analisado: justificativa e metodologia

Optamos por investigar as questões indicadas na seção anterior em dados relativos à leitura de um texto escrito feita por vários informantes, pelos motivos seguintes: 1) nos trabalhos relativos à atribuição de acentos secundários feitos no quadro das fonologias atuais (em abordagens derivacionais como a da Fonologia Métrica, ou representacionais, como naquelas baseadas na Teoria da Otimalidade), costuma-se definir a janela (no caso, a palavra prosódica) no interior da qual se atribuem/realizam as proeminências secundárias, bem como o algoritmo/restrições que dão conta de sua implementação, mas não se analisa sistematicamente, em textos orais, a maneira como “interagem” as pautas acentuais das palavras prosódicas concatenadas sequencialmente no interior de domínios prosódicos hierarquicamente superiores; 2) por conseqüência, não se chega a tomar como relevante a questão de saber como as opções de implementação de acentos secundários e de reduções segmentais/sândi podem eventualmente otimizar um determinado padrão de ritmo; 3) também não se costumam analisar as proeminências secundárias do ponto de vista da maneira como interagem com as ocorrências de ênfase, pausas e determinados contornos entoacionais. Por acharmos que a discussão sobre a organização rítmica dos textos orais, em qualquer língua, não pode deixar de levar em conta tais fatores, optamos por realizar uma investigação preliminar em um texto lido por vários informantes.

A opção por analisar inicialmente um texto lido – e não textos orais espontaneamente produzidos por vários sujeitos – deveu-se ao fato de a leitura de um texto previamente selecionado possibilitar um maior controle das diferenças entre as escolhas, feitas pelos informantes, por implementar ou não, nos diferentes contextos, os acentos secundários e os processos de sândi vocálico. A mesma investigação deverá ser posteriormente estendida para textos orais espontâneos.

Justifica-se, ainda, a escolha por trabalhar apenas a partir de dados de informantes provenientes do estado de São Paulo, pelo fato de ser necessário manter também um certo controle sobre as diferenças que caracterizam, também no plano prosódico, as diversas variedades regionais do português do Brasil. Embora, por hipótese, a tendência geral que caracteriza a organização rítmica do PB com relação ao PE deva se manifestar no conjunto das variedades regionais faladas no país, é importante, do ponto de vista metodológico, que se proceda a partir de análises voltadas, em um primeiro momento, para a caracterização da organização rítmica em variedades distintas de PB, para depois comparar essas

variedades. Em uma etapa subsequente do trabalho, a mesma análise poderá também ser realizada para as variedades de PE.⁵

O *corpus* da pesquisa mais ampla, em desenvolvimento no interior do Projeto Temático, é constituído pela leitura, por 17 (dezesete) falantes nativos de PB (estado de São Paulo) e por 5 (cinco) falantes nativos de PE (Lisboa), de uma crônica publicada em uma revista semanal do Brasil. Essa crônica foi escolhida pela alta frequência de palavras com mais de duas sílabas pré-tônicas, o que resulta na necessária diversidade de contextos para verificação da implementação das proeminências secundárias. O texto sofreu pequenas adaptações (em termos lexicais e sintáticos), feitas por falantes nativos de PE, para a leitura pelos informantes portugueses. Neste trabalho, foram analisados apenas os dados referentes à leitura dos informantes de PB. Siga-se o texto da crônica:

Complicabilizando

Não, por favor, nem tente me disponibilizar alguma coisa, que eu não quero. Não aceito nada que pessoas, empresas ou organizações me disponibilizem. É uma questão de princípios. Se você me oferecer, me der, me vender, me emprestar, talvez eu venha a topar. Até mesmo se você tornar disponível, quem sabe, eu aceite. Mas, se você insistir em disponibilizar, nada feito. Caso você esteja contando comigo para operacionalizar algo, vou dizendo desde já: pode tirar seu cavalinho da chuva. Eu não operacionalizo nada para ninguém. Tampouco compactuo com quem operacionalize. Se você quiser, eu monto, eu realizo, eu aplico, eu ponho em operação. Se você pedir com jeitinho, eu até implemento. Mas, operacionalizar, jamais. O quê? Você quer que eu agilize isso para você? Lamento, mas eu não sei agilizar nada. Nunca agilizei. Está lá no meu currículo: faço tudo, menos agilizar. Precisando, eu apresso, eu priorizo, eu ponho na frente, eu dou um gás. Mas agilizar - desculpe, não posso, acho que matei essa aula.

Outro dia mesmo queriam reinicializar meu computador. Só por cima do meu cadáver virtual! Prefiro comprar um computador novo a reinicializar o antigo. Até porque eu desconfio que o problema não seja assim tão grave. Em vez de reinicializar, talvez seja o caso de simplesmente reiniciar, e pronto.

Por falar nisso, é bom que você saiba que eu parei de utilizar. Assim, sem mais nem menos. Eu sei, é uma atitude um tanto quanto radical da minha parte, mas eu não utilizo mais nada. Tenho consciência de que a cada dia que passa mais e mais pessoas estão utilizando, mas eu parei. Não utilizo mais. Agora eu só uso. E recomendo. Se você soubesse como é muito mais elegante, também deixaria de utilizar e passaria a usar.

Sim, estou me associando à campanha nacional contra os verbos que acabam em "ilizar". Se nada for feito, daqui a pouco eles serão

⁵Esse é um projeto de investigação de longo prazo que pressupõe, evidentemente, a participação de vários pesquisadores de instituições brasileiras e portuguesas. O que aqui se pretende, na verdade, é levantar um conjunto de hipóteses que se podem revelar produtivas em uma investigação dessa natureza, bem como sugerir princípios metodológicos que poderão nortear as pesquisas futuras sobre esse tema.

mais numerosos do que os terminados simplesmente em “ar”. Todos os dias os maus tradutores de livros de marketing e administração disponibilizam mais e mais termos infelizes, que imediatamente são operacionalizados pela mídia, reinicializando palavras que já existiam e eram perfeitamente claras e eufônicas.

A doença está tão disseminada que muitos verbos honestos, com currículo de ótimos serviços prestados, estão a ponto de cair em desgraça entre pessoas de ouvidos sensíveis. Depois que você fica alérgico a disponibilizar, como você vai admitir, digamos, “viabilizar”? É triste demorar tanto tempo para a gente se dar conta de que “desincompatibilizar” sempre foi um palavão.

Precisamos reparabilizar nessas palavras que o pessoal inventabiliza só para complicabilizar. Caso contrário, daqui a pouco nossos filhos vão pensabilizar que o certo é ficar se expressabilizando dessa maneira. Já posso até ouvir as reclamações: “Você não vai me impedibilizar de falabilizar do jeito que eu bem quilibiliser”. Problema seu. Me inclua fora dessa. (Freire, 2003).

Com relação à metodologia utilizada no trabalho, ressaltamos os seguintes procedimentos:

1. Para possibilitar a descrição da maneira como foram implementadas as proeminências secundárias no texto de PB, procedeu-se:
 - à identificação auditiva da ocorrência de acentos secundários em palavras prosódicas com 2 ou mais sílabas pré-tônicas;⁶
 - à tabulação dos dados e ao levantamento das semelhanças e diferenças na atribuição dos acentos secundários por parte dos vários informantes;
2. Para possibilitar o levantamento das ocorrências de sândi vocálico (elisão, degeminação, ditongação) no texto de PB, procedeu-se:
 - à identificação auditiva das ocorrências de sândi vocálico;
 - à tabulação das ocorrências de sândi e ao levantamento das semelhanças e diferenças na implementação desses processos segmentais que levam à ressilabificação, por parte dos vários informantes.

⁶ A gravação das leituras e a transcrição inicial dos dados do corpus aqui analisado foram realizadas como parte do trabalho de Iniciação Científica de Natália Cristina Grosso (IC-FAPESP, 2006/2007), Os acentos secundários no português do Brasil (2007). As transcrições foram posteriormente conferidas por outros dois juízes.

UFPA
FACULDADE DE LETRAS
20110122

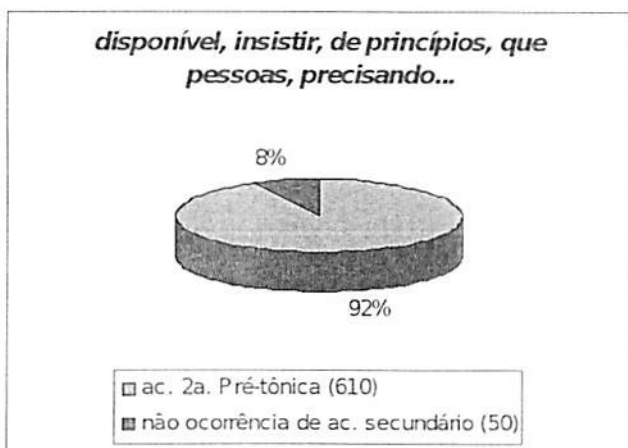
A partir das transcrições e tabulações, fez-se a análise dos dados e, com base nos percentuais obtidos para cada contexto investigado, procedeu-se então à discussão e elaboração de hipóteses acerca da relação entre a implementação dos acentos secundários e dos processos de sândi e a manifestação de um determinado padrão responsável pela configuração rítmica do texto analisado.

6 Resultados: posição dos acentos secundários

Para entender os gráficos que serão apresentados ao longo desta seção, é importante observar que os números foram atribuídos às sílabas a partir da primeira sílaba pré-tônica que antecede imediatamente a sílaba portadora do acento lexical. Assim, a 1ª, 2ª, 3ª, 4ª, 5ª, 6ª e 7ª sílabas são contadas da direita para a esquerda. Ver a numeração das sílabas no exemplo *desincompatibilizar* (sílabas em negrito: portadoras de acento secundário; sílabas em negrito, sublinhadas: portadoras do acento de palavra):

- (2) De (7) **sin** (6) com (5) pa (4) ti (3) bi (2) li (1) zar
 De (7) **sin** (6) com (5) pa (4) ti (3) bi (2) li (1) zar

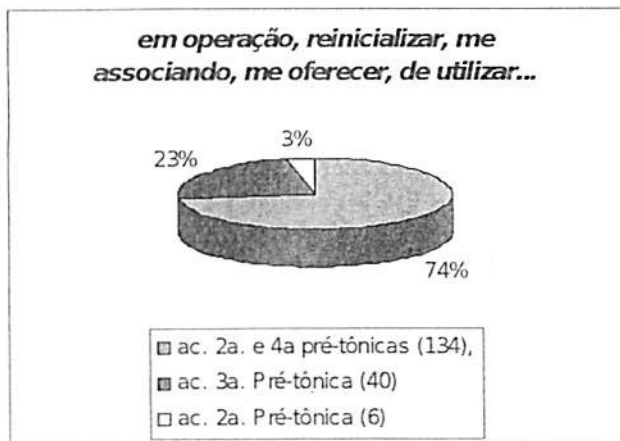
6.1 Palavras prosódicas com 2 sílabas pré-tônicas (660 ocorrências)



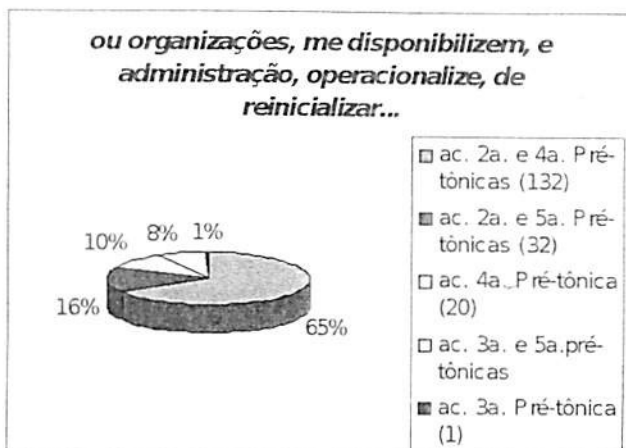
6.2 Palavras prosódicas com 3 sílabas pré-tônicas (314 ocorrências)



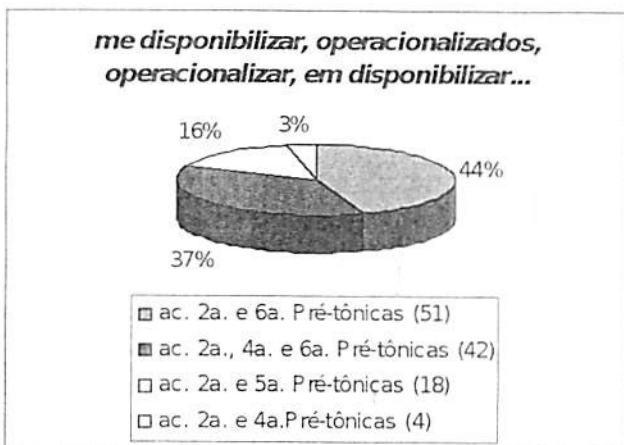
6.3 Palavras prosódicas com 4 sílabas pré-tônicas (180 ocorrências)



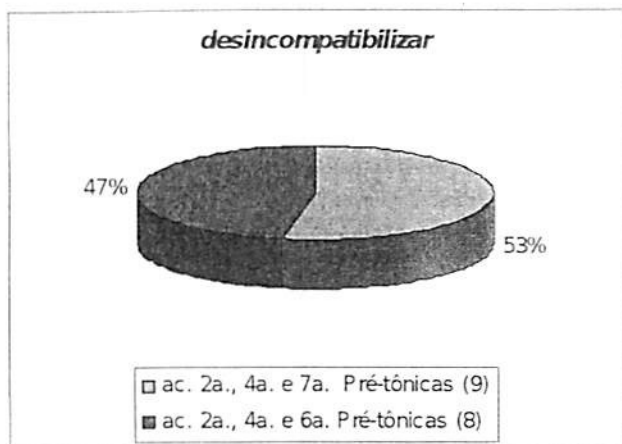
6.4 Palavras prosódicas com 5 sílabas pré-tônicas (200 ocorrências)



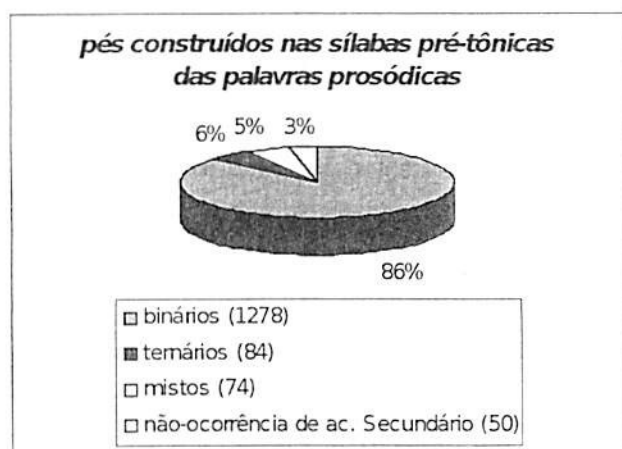
6.5 Palavras prosódicas com 6 sílabas pré-tônicas (115 ocorrências)



6.6 Palavras prosódicas com 7 sílabas pré-tônicas (17 ocorrências)



6.7 Tipos de pés e acentos secundários (1486 ocorrências)



A observação relevante a fazer sobre os gráficos apresentados nas seções 6.1 a 6.6, acima, cujos resultados percentuais estão resumidos em 6.7, é que é significativamente maior a opção pela implementação de pés de natureza binária (86%), ainda que, por circunstâncias que devem ainda ser identificadas, ocorram também pés que se manifestam como ternários (dátilos), no nível da implementação fonética.⁷

⁷No caso das palavras prosódicas com seis sílabas pré-tônicas, 44% das opções foi por colocar o acento secundário na 2ª. e na 6ª. sílabas, mas não na 4ª., o que resultou em um pé fonético inicial com quatro sílabas, mantendo-se porém a opção pelo número par de sílabas no interior do pé. Optamos por considerar esses pés como sendo também de orientação

7 Resultados: implementação dos processos de sândi vocálico

Os gráficos aqui apresentados referem-se à implementação, no *corpus* analisado, dos processos de: degeminação ou crase (por meio do qual duas vogais idênticas em juntura de palavra reduzem-se a uma, e.g., *nunca agilizei* → *nun[ka]gilizei*), elisão (por meio do qual, em uma seqüência de duas vogais diferentes em juntura, realiza-se apenas a segunda, e.g., *contra os* → *com[trôs]*) e ditongação (por meio do qual uma seqüência de vogais em dois núcleos silábicos em juntura de palavra realiza-se como ditongo, em um único núcleo silábico, e.g., *me associando* → [mja]ssociando).⁸

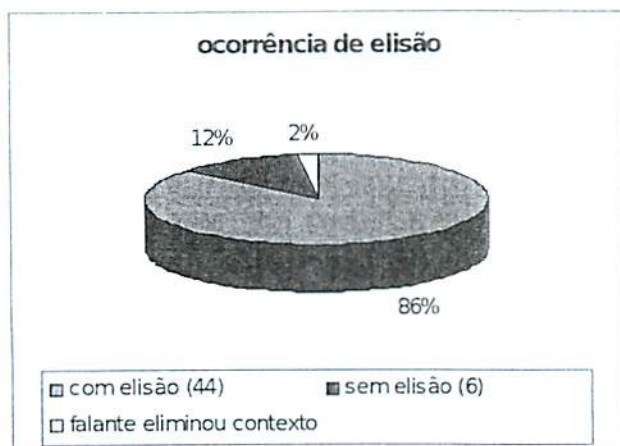
7.1 Degeminação (136 contextos)



binária, baseando-nos na hipótese (a ser verificada) de que os falantes podem não ter realizado foneticamente a proeminência na 4^a. sílaba da seqüência pré-tônica, embora reconheçam essa sílaba como possível lugar de pouso da proeminência secundária, nesse tipo de seqüência.

⁸A propósito desses processos de sândi vocálico, Cf. Bisol (1990, 1992, 1994, 1995, 1996, 1999, 2000, 2002, 2003).

7.2 Elisão (51 contextos)



9

7.3 Ditongação (238 contextos)



7.4 Os processos de sândi e a organização rítmica: análise de algumas ocorrências registradas no corpus

Da análise dos gráficos apresentados de 7.1 a 7.3, acima, conclui-se que a tendência geral manifestada pelos informantes foi a de implementar os três processos de sândi vocálico nos contextos relevantes (68% de aplicação de degeminação, 86% de elisão e 93% de ditongação). O interessante, naturalmente,

⁹Neste e no gráfico seguinte, a menção a casos em que o falante eliminou o contexto refere-se a modificações introduzidas na sequência original de palavras, durante a leitura.

com relação à aplicação desses processos no *corpus* analisado é saber se é possível afirmar que existe alguma correlação entre a aplicação dos processos e a otimização de uma organização rítmica binária (resultante da implementação de pés binários no nível da implementação fonética).

Nas seções seguintes, vamos analisar, sob essa perspectiva, algumas seqüências retiradas do *corpus*, de forma a exemplificar, em contextos em que os processos podem se aplicar, os resultados obtidos do ponto de vista da organização rítmica dos trechos considerados.

7.4.1 Observações sobre ocorrências de degeminação

(3) ... o certo é ficar (se expressabilizando) ω dessa maneira.

Com relação ao exemplo (3), registramos as seguintes ocorrências:

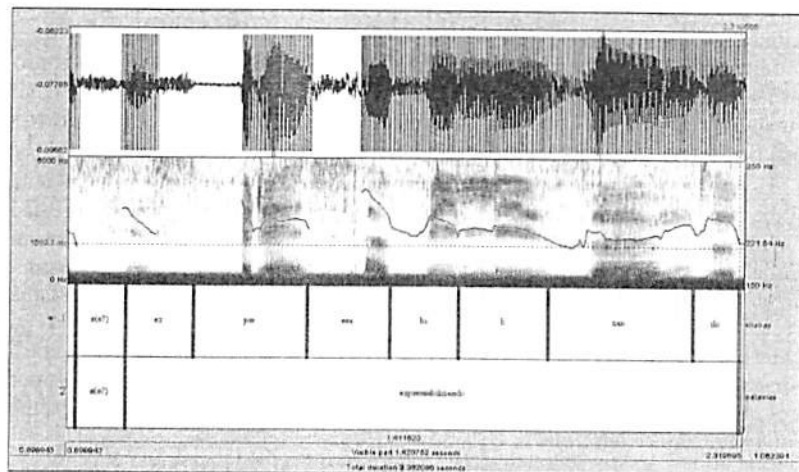
(4) [sis] pre ssa bi li zan do (9 informantes)

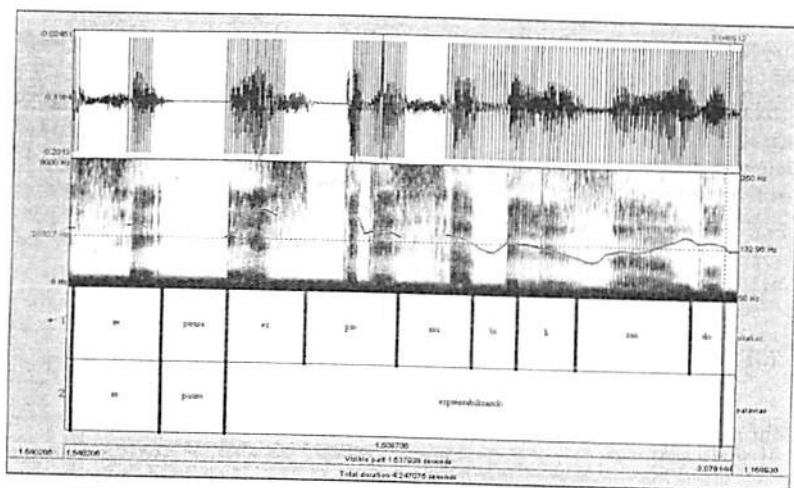
(5) [sɪ] [ɪs] pre ssa bi li zan do (7 informantes)

(6) [se] [es] pre ssa bi li zan do (1 informante – com ênfase)

Nos espectrogramas abaixo, registram-se, em (7), uma das nove pronúncias com a degeminação e, em (8), a pronúncia do informante que realizou a seqüência com ênfase:

(7)





Do ponto de vista do resultado rítmico, observa-se que os nove informantes que optaram pela aplicação da degeminação na seqüência [si is], que foi realizada como [sis], colocaram acento secundário nas sílabas 'pre' e 'bi', o que resultou em dois troqueus binários antes do pé iniciado pela sílaba 'zan', portadora do acento lexical. No caso dos sete informantes que optaram por não aplicar o processo, a escolha foi por colocar uma proeminência secundária na sílaba inicial da palavra prosódica, 'se', o que também resulta em uma seqüência de troqueus binários no interior da palavra prosódica. Se considerarmos, no entanto, que a palavra prosódica anterior, (*ficar*) *o*, termina com sílaba lexicalmente acentuada, vemos que as duas escolhas não são equivalentes, do ponto de vista da otimização do padrão binário no texto como um todo. De fato, o que se obtém foneticamente, no caso da não implementação da degeminação, nos casos em (5), é uma colisão acentual envolvendo as sílabas 'car' e 'se', a primeira portadora de acento lexical e a segunda, de acento secundário.

É necessário, para melhor entender opções como essas, verificar as estratégias utilizadas pelos falantes para possibilitar a realização com o choque de acentos. Podem ter sido feitas pequenas pausas e/ou modificações locais no contorno entoacional. Não nos ocupamos ainda, em nossa pesquisa, da análise sistemática dessas estratégias, no caso dos choques de acento, mas esse é um aspecto que terá de ser considerado em futuros trabalhos.

Houve apenas um caso em que a opção foi por colocar uma proeminência na sílaba 'ex', o que resultou em um dátilo, ou seja, um pé com três sílabas e acento inicial. Nesse caso, o informante teve a intenção clara de enfatizar a primeira sílaba da palavra 'expressabilizando', talvez como estratégia para conseguir realizar foneticamente uma palavra com a qual não está muito familiarizado. É interessante observar que, após a primeira sílaba da palavra prosódica, 'se', o falante realizou uma pausa relativamente longa, conforme se pode observar claramente no espectrograma em (8), acima. Realizou também a sílaba com a proeminência secundária, 'ex', com um tom H(igh), de contorno perfeitamente visível no espectrograma.

A existência de casos como (6), relativamente freqüentes, permite levantar a hipótese de que a subversão da tendência rítmica preferencial da língua (uma organização de base binária) pode contribuir para a obtenção de um efeito de ênfase.

Há um outro dado do *corpus* que também permite fazer considerações interessantes sobre a possível correlação entre a ocorrência de troqueus ou dátilos e a marcação de ênfase, em PB. Observe-se o trecho abaixo:

- (9) ...para a gente se dar conta de que "desincompatibilizar" sempre foi um palavrão.

Com relação ao exemplo (9), trecho em que ocorre a palavra fonológica com maior número de sílabas pré-tônicas do *corpus* ('desincompatibilizar', com 7 pré-tônicas), registraram-se as seguintes ocorrências:

- (10) (de sin com pa ti bi li zar) ω (9 falantes - 53%)

- (11) (de sin com pa ti bi li zar) ω (8 falantes - 47%)

A primeira observação a fazer é que, no texto, a palavra vem grafada entre aspas, o que reforça o fato de que está sendo mencionada, ressaltando-se o seu estatuto de palavrão (o autor parece explorar, aqui, os dois sentidos possíveis de 'palavrão'). O texto já cria, portanto, um contexto para a leitura enfática da palavra. Considerando-se as realizações dos falantes, pode-se talvez supor que os nove que produziram o dátilo inicial optaram por uma leitura enfática, conforme sugerido pelas aspas, colocando um acento secundário já na primeira sílaba, do que resultou o dátilo inicial. Seguem-se, depois, dois troqueus binários. No caso dos outros oito falantes, a opção foi pela realização que otimiza o ritmo binário no interior da palavra prosódica, o que implica realizar a segunda sílaba da seqüência com acento secundário e uma série de três troqueus.

O exemplo deixa evidente, portanto, que é necessário também analisar a maneira como os falantes recorrem, nos textos orais, a estratégias de ênfase / focalização, e como distribuem pausas e determinados contornos de entoação que certamente interagem com a implementação do padrão rítmico preferencial da língua.

7.4.2 Observações sobre ocorrências de elisão

- (12) ...que eu parei (de utilizar)ω

Com relação ao exemplo (9), registramos as seguintes ocorrências:

- (13) [dʒv] ti li zar (15 falantes)

- (14) [dʒi] u ti li zar (1 falante)

- (15) [dʒi] u ti li zar (1 falante)

Uma sequência de letras

A opção de quinze dos dezessete informantes foi pela aplicação da elisão, o que resulta em uma otimização evidente da realização de pés fonéticos trocaicos, considerando-se que a palavra prosódica anterior termina com a sílaba lexicalmente acentuada 'rei'. Tem-se, portanto, nesse trecho da seqüência, (...rei) ω ([dʒu] ti li zar), com dois troqueus antes da sílaba final acentuada 'zar'. Dois informantes, apenas, optaram por não realizar a elisão e manter a vogal [u] como núcleo de uma sílaba. Um dos dois falantes colocou acento secundário na sílaba 'u', em (14); o outro atribuiu proeminência secundária à primeira sílaba da palavra prosódica, 'de' ([dʒi]), em (15). No exemplo em (14), pode-se observar que a opção por acentuar o 'u' permite a realização de um troqueu fonético na associação das sílabas final e inicial das duas palavras fonológicas, mas o pé fonético seguinte será inevitavelmente um dátilo (já que a sílaba 'li' não poderia receber acento, pois isso produziria um choque acentual com a sílaba lexicalmente acentuada, 'zar': (...rei) ω ([dʒi] u ti li zar) ω. Quanto ao falante que optou por um acento secundário na sílaba inicial da palavra prosódica, 'de' ([dʒi]), em (15), vê-se que conseguiu uma boa solução rítmica no interior da palavra prosódica (dois troqueus), mas acabou provocando um choque entre o acento lexical da palavra prosódica anterior, 'rei', e a sílaba inicial da palavra prosódica seguinte, 'de' ([dʒi]). Conclui-se, portanto, que a melhor solução rítmica para esse trecho foi aquela adotada pelos quinze falantes que optaram pela realização com elisão e acento secundário na sílaba 'ti', como no dado (13).

A análise desse exemplo deixa bastante evidente a necessidade de avaliar a implementação rítmica de um texto levando também em conta as reestruturações feitas pelos falantes através de junturas de palavras fonológicas, bem como a interação entre suas opções de colocação de acentos secundários e de aplicação de processos segmentais (como o sândi vocálico), e a utilização de estratégias de reparação de possíveis choques acentuais resultantes dessas opções.

7.4.3 Observações sobre ocorrências de ditongação

(16) Sim, estou (me associando) ω ...

Com relação ao exemplo (16), registramos as seguintes ocorrências:

(17) [mja] sso ci an do (15 falantes)

(18) [mu] a sso ci an do (1 falante)

(19) [m] a sso ci an do (1 falante)

Este caso é semelhante ao caso da aplicação da elisão, analisado na seção 7.4.2, acima. Os quinze falantes que optaram por aplicar o processo de ditongação, no exemplo (17), conseguiram o melhor resultado do ponto de vista da estruturação rítmica desse trecho, considerando-se que a palavra prosódica anterior termina com a sílaba acentuada 'tou'. Assim, a aplicação da ditongação reduz as sílabas 'e' e 'a' a uma só sílaba, [mja], que será a segunda sílaba do pé trocaico (...tu)ω ([mja] sso ci an do)ω formado pela associação da última sílaba da primeira palavra prosódica à primeira sílaba da segunda, na implementação fonética. Na seqüência, realizaram-se mais dois pés trocaicos, um iniciado pelo acento secundário na sílaba 'sso' e o outro iniciado pela sílaba portadora do acento lexical da palavra. A opção (18) levou à realização de um dátilo na juntura das

duas palavras prosódicas, ao passo que a opção (19) manteve a estruturação rítmica binária no interior da palavra prosódica, mas criou um choque acentual entre a sílaba final da palavra prosódica anterior e a sílaba inicial de ([m] a sso ci an do)6. Mais uma vez, a grande maioria dos informantes optou pela melhor solução rítmica, (17), de orientação binária, considerando-se a interação entre a pauta acentual das duas palavras prosódicas, no nível da implementação fonética dos pés.

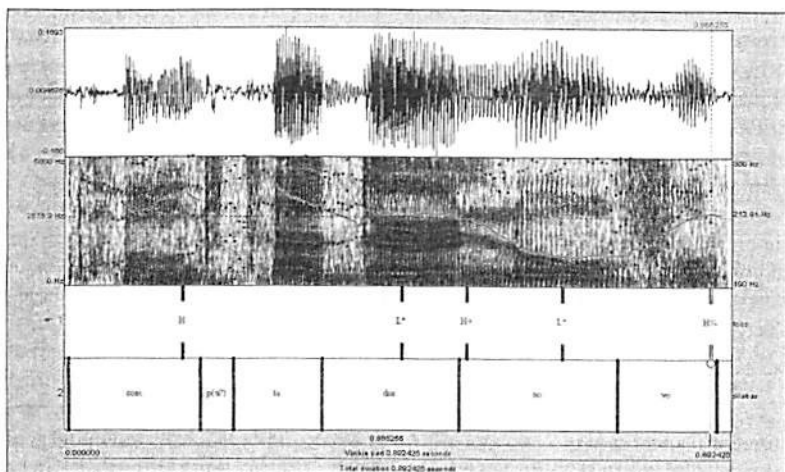
8 Uma observação sobre outros casos de redução vocálica e sua relação com o acento secundário e a entoação

Em PB, a cadeia segmental sofre também outros reajustes, como o apagamento/redução de vogais pré-tônicas em sílabas internas de um radical. Tais reajustes podem também otimizar uma construção rítmica binária, evitando dátilos e levando à implementação de troqueus. Este tipo de reestruturação rítmica é encontrado no texto do PB aqui analisado. Cf. '*computador*', em:

(20) Prefiro comprar (um **comp(u)**tador)6 novo...

Reduções vocálicas como essa, freqüentes no PB nos estilos mais coloquiais de fala, são muitas vezes determinadas pelo próprio contexto segmental. A ocorrência da vogal alta posterior arredondada [u] no núcleo da sílaba que tem como onset a consoante bilabial [p] é, na maioria das vezes, realizada apenas com um arredondamento dos lábios durante a produção da consoante. Tais reduções levam a uma reestruturação da seqüência silábica, o que pode, por sua vez, interferir na estrutura interna dos pés. No exemplo em questão, todos os dezessete informantes reduziram drasticamente a vogal, o que resultou em um pé binário.¹⁰

¹⁰Nem sempre, no entanto, tais processos levam a uma otimização do ritmo. Por serem determinadas mais pelo contexto segmental do que pelo próprio ritmo, tais reduções podem ocorrer independentemente de levarem ou não a uma otimização do padrão rítmico preferencial da língua. Esse é um assunto para futuras investigações, e dele não nos ocuparemos aqui.



No espectrograma acima, referente à realização de um dos informantes, verifica-se que na palavra 'computador' a sílaba 'pu' (que recebe acento secundário pelo algoritmo, antes da implementação fonética) sofre redução drástica de seu núcleo vocálico, ao ser produzida. Verifica-se também que a sílaba 'com' porta um tom H(igh) (Cf. Tenani, 2002).

Estes dois fatos podem, por hipótese, ser pistas da presença de acento secundário na sílaba 'com', nesse dado. Observe-se que a ocorrência de proeminência secundária em 'com' otimiza a construção de pés binários nessa seqüência, por evitar a ocorrência de um dátilo fonético: (...prar) ω (*um com...*) ω. Tal dátilo resultaria da associação da sílaba acentuada final de *comprar*, da palavra prosódica anterior, às duas sílabas iniciais da palavra prosódica seguinte, para fins de implementação fonética da seqüência rítmica:

(22) (Pref^o)ω (comprar)ω (um comp(u)tador)ω (novo)ω

9 Conclusão

Os dados aqui analisados permitem sustentar a hipótese de que o PB privilegia a construção de pés binários, que estão na base da organização rítmica da língua. Pode-se concluir, também, desses dados, que os processos de sândi vocálico e outros processos de redução vocálica tendem a otimizar a organização binária do ritmo. A confirmação dessas hipóteses depende, evidentemente, da ampliação da investigação para *corpora* de fala espontânea representativa de vários graus de formalidade (foneticamente implementados por meio de diferentes taxas de elocução) e de outras variedades regionais de PB.

Por fim, cabe mencionar que há vários fenômenos, além dos acentos secundários, que interagem na construção do ritmo e que estão a merecer estudo

sistemático: pausas e ênfase (por vezes marcada por tom II) podem determinar outras proeminências e levar à construção de pés ternários (dátilos). Uma discussão mais aprofundada dessas questões será matéria de futuros trabalhos.

Referências

- [1] ABAURRE, M. B., AND GALVES, C. Rhythmic differences between european and brazilian portuguese: an optimalist and minimalist approach. *DELTA* 14 (1998), 377-403.
- [2] BERGER, M. A. *An Introduction to Probability and Stochastic Processes*. Springer-Verlag, New York, 1992.
- [3] BISOL, L. Degeminação e elisão. In *Anais do Congresso Internacional da Associação de Linguística e Filologia da América Latina* (Campinas: IEL/Unicamp, 1990), no. 9.
- [4] BISOL, L. Sândi vocálico externo: degeminação e elisão. *Cadernos de Estudos Linguísticos* 23 (1992), 83-101.
- [5] BISOL, L. *Gramática do Português Falado: convergências*, vol. 5. Ed. da Unicamp, Campinas, 1994, ch. Sândi externo: o processo e a variação, pp. 55-96.
- [6] BISOL, L. O sândi e a ressilabação. In *Congresso Internacional da ABRALIN* (Salvador, 1996), pp. 370-376. Boletim da ABRALIN.
- [7] BISOL, L. *Estudos Fonéticos e Fonológicos*. UEL, Londrina, 1999, ch. O sândi e a sintaxe, pp. 1-10.
- [8] BISOL, L. A elisão, uma regra variável. *Letras de Hoje* 35, 1 (2000), 319-330.
- [9] BISOL, L. *A degeminação e a elisão no VARSUL*. EDIPUCRS, Porto Alegre, 2002.
- [10] BISOL, L. Sandhi in brazilian portuguese. *Probus* 15, 2 (2003), 177-200.
- [11] BISOL, L. O clítico e seu hospedeiro. *Letras de Hoje* 40, 3 (2005), 163-184.
- [12] CASTELO, A. Proeminência secundária no português europeu. Master's thesis, Faculdade de Letras, Universidade de Lisboa, 2004.
- [13] COLLISCHONN, G. Um estudo do acento secundário em português. Master's thesis, Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 1993.
- [14] COLLISCHONN, G. Acento secundário em português. *Letras de Hoje - Fonologia: análises não-lineares* 29, 4 (1994), 43-53.
- [15] D'ANDRADE, E., AND LAKS, S. B. Na crista da onda: o acento de palavra em português. In *Encontro da Associação Portuguesa de Linguística* (1991), no. 7, pp. 15-26.

- URBES, ESTADOS DA LINGUA PORTUGUESA
- [16] FELLER, W. *An Introduction to Probability Theory and its Applications*, vol. 1. John Wiley & Sons, New York, 1950.
- [17] FREIRE, R. Xongas. *Época*. 275 (25 ago. 2003).
- [18] GROSSO, N. C. Os acentos secundários no português do brasil: identificação e análise das ocorrências em um corpus de leitura. Tech. rep., FAPESP, 2007. Relatório final do Projeto de Iniciação Científica.
- [19] HULST, H. V. D. Primary accent is non-metrical. *Rivista di Linguistica* 8, 1 (1997).
- [20] KARLIN, S., AND TAYLOR, H. M. *A First Course in Stochastic Processes*. Academic Press, New York, 1972.
- [21] KOVACS, Z. L. *Teoria da probabilidade e processos estocásticos: com aplicações em engenharia de sistemas e processamento de sinais*. Acadêmica, São Paulo, 1996.
- [22] LIPSCHUTZ, S. *Probabilidade*. Makron Books, São Paulo, 1994.
- [23] NESPOR, M., AND VOGEL, I. *Prosodic Phonology*. Foris, Dordrecht, 1986.
- [24] PAPOULIS, A. *Probability, random variables, and stochastic processes*, 3 ed. McGraw-Hill, New York, 1991.
- [25] PARZEN, E. *Stochastic Processes*. Holden-Day, San Francisco, 1962.
- [26] RESNIC, S. I. *Adventures in Stochastic Processes*. Birkhauser, Basel, 1999.
- [27] SANDALO, F., ABAURRE, M. B., MANDEL, A., AND GALVES, C. Secondary stress in two varieties of portuguese and the sotaq optimality based computer program. *Dordrecht 18* (2006), 97-125.
- [28] SCHWINDT, L. C. *O prefixo no português brasileiro: análise morfofonológica*. PhD thesis, Faculdade de Letras, Pontifícia Universidade Católica do Rio Grande do Sul. 2000.
- [29] TENANI, L. *Domínios prosódicos no Português do Brasil: implicações para a prosódia e para a aplicação de processos fonológicos*. PhD thesis, Instituto de Estudos da Linguagem. Universidade Estadual de Campinas, Campinas, 2002.
- [30] TODOROVIC, P. *An Introduction to Stochastic Processes and their Applications*. Springer-Verlag, New York, 1992.
- [31] VIGARIO, M. *The prosodic word in European Portuguese*. Mouton de Gruyter, Berlin/New York, 2003.

L'enonciation Aphorisante

Dominique Maingueneau

Université Paris XII

maingueneau.dominique@numeris.fr

Resumo

Na presente conferência, abordaremos os chamados enunciados “destacados”, isto é, enunciados curtos, na forma oral ou escrita, representados, em geral, por uma única frase, que são retirados de seu contexto original. Ora, o “destacamento” não acontece, necessariamente, sem a intervenção do autor do texto original, fenômeno que aparece, com relativa frequência, graças à ajuda de procedimentos ligados a diversos domínios (sintáticos, prosódicos, semânticos, tipográficos etc.) Esses procedimentos são convocados pelo autor, que parece assinalar, assim, que certos fragmentos de seu texto são suscetíveis à operação de “destacamento”. Propusemos, num primeiro momento, que esse fenômeno fosse chamado de “sobre-asserção”. No entanto, dada a insuficiência dessa noção para operar a descrição do funcionamento efetivo dos enunciados “destacados”, introduziremos um novo conceito: a “aforização”, que designará o regime enunciativo específico dos chamados “enunciados destacados”. Deslocaremos, pois, nosso interesse para o estudo das propriedades linguísticas de uma enunciação aforizante. Não se trata aqui de questionar se o enunciado “original” foi alterado pela operação de “destacamento” ou se o “destacador” está (ou não) respeitando as intenções do autor; trata-se, sim, de considerar tudo aquilo que implica a descontextualização de tais enunciados. Após ter mencionado os traços gerais dessa problemática, buscaremos dar uma atenção particular às diversas dimensões da noção de aforização, uma vez que ela envolve problemas de várias ordens (modal, textual, hermenêutica, genérica), que, a nosso ver, não foram ainda suficientemente estudados.

Je vais évoquer une question, l'enonciation « aphorisante », qui se trouve au croisement de la linguistique de l'enonciation et de l'analyse du discours, en adoptant une perspective qu'on pourrait dire « anthropolinguistique ». Cette communication prolonge un des chapitres (« **Citação e destacabilidade** ») du livre « **Cenas de enunciação** », récemment publié¹, où j'ai introduit le terme « aphorisation », mais sans véritablement le développer.

Cette question de l'enonciation « aphorisante » s'inscrit dans une problématique plus vaste, celle des « énoncés détachés ». L'existence de tels énoncés ne fait aucun doute. Il en circule en effet dans la société un nombre infini. On songe tout de suite aux proverbes ou aux slogans, c'est-à-dire à des énoncés autonomes, décontextualisés par nature; mais la majorité ont été extraits de textes. On voit donc apparaître une ambiguïté du terme « détachement » : 1) détachement de tout contexte originel (c'est le cas des proverbes et de toutes les formules sentencieuses, qu'elles soient ou non attribuées à un auteur identifié),

¹ Curitiba, Criar, 2006.

2) détachement d'un texte particulier. Pour ce dernier type de détachement, on entre dans une logique plus traditionnelle de citation, mais il ne s'agit pas de n'importe quel type de citation : seulement des citations qui constituent de brefs énoncés - le plus souvent des phrases - qui reçoivent divers noms dans l'usage : « formules », « pensées », « maximes »...

1 Détachabilité et surassertion

On ne peut cependant se contenter de constater que certains énoncés ont été détachés d'un texte. Une telle opération ne s'exerce pas sur n'importe quel matériau verbal : bien souvent il s'agit d'énoncés qui dans le texte où ils figureraient étaient déjà présentés comme détachables. Considérons par exemple ce paragraphe qui conclut un article figurant dans les pages économiques du quotidien français «le Figaro» :

France Télécom devient une entreprise privée « [...] Et comme deux autres formes symboles, Renault et Air France, deux groupes publics sauvés grâce aux deniers publics privatisés pour leur permettre de devenir de vrais champions mondiaux, France Télécom illustre à son tour la difficile métamorphose de « France Entreprise ». Car en ce début du *XXI^e* siècle, il est impossible de faire de la bonne industrie si on n'est pas capable d'être aussi un bon actionnaire. » (« Le Figaro économie », 2 septembre 2004, p. 11 (c'est nous qui le soulignons)).

Le lecteur perçoit aisément que la dernière phrase est présentée comme détachable ; elle contraste avec ce qui précède par ses propriétés énonciatives, aussi bien par son lien avec la thématique centrale de l'article, donnée dans le titre. Il s'agit d'un énoncé conclusif par la place privilégiée qu'il occupe, un énoncé généralisant, et qui pose une norme avec un ethos légèrement solennel. C'est le type d'énoncé dont on imagine qu'il pourrait facilement être repris en titre.

Autre exemple, cet entretien qu'un mathématicien, Gilles Dowek, a donné à un journaliste du quotidien gratuit français « 20 minutes ». Le titre de l'article est « L'âge d'or des mathématiques, c'est aujourd'hui » ; cet énoncé résulte d'un détachement, on le retrouve dans le texte de l'entretien, qui le présente comme détachable : il est en effet placé à la fin d'une réponse, introduit par le connecteur de reformulation « autrement dit » et constitue un paradoxe (« age d'or » = « aujourd'hui »)

« (...) on pense trop souvent qu'elles (= les mathématiques) appartiennent au passé, alors que la moitié des mathématiciens qui ont sévi au cours de l'Histoire sont...vivants et en exercice. *Autrement dit, l'âge d'or des mathématiques, c'est aujourd'hui.* » (« 20 minutes », 18 octobre 2004, p.39 (c'est nous le qui soulignons)).

Pour ce formatage d'un fragment comme détachable, candidat à une reprise citationnelle, on ne peut pas parler de « citation », ni même d'embryon de citation : c'est seulement une mise en relief qui est opérée par rapport au reste des énoncés. Pour désigner ce type de phénomène nous avons introduit la notion de « surassertion ». De manière informelle, on peut dire qu'une séquence surassertée dans un texte

- Est relativement brève.
- Possède une structure prégnante sur le plan du signifiant (symétrie, syllepse...) et/ou du signifié (métaphore, paradoxe...);
- se trouve en position saillante, le plus souvent en début ou en fin de texte ou de partie de texte, de manière à lui donner le statut d'un condensé sémantique du texte dont elle fait partie :
- sa thématique doit être en relation avec l'enjeu du texte concerné et il s'agit d'une prise de position sur un point controversé ;
- elle implique une sorte d' « amplification » de la figure de l'énonciateur.

La surassertion peut se marquer en outre :

- au niveau de l'embranchement énonciatif : en conférant à l'énoncé surasserté une valeur généralisante ou générique ;
- dans le métadiscours : en indiquant que tel énoncé joue un rôle clé (ainsi par une reprise catégorisante : « cette vérité essentielle... »), ou en l'introduisant par des connecteurs comme « en fin de compte », « autrement dit », « bref », etc.

2 Détachements forts et faibles

La notion de détachement d'un fragment de texte recouvre en fait deux phénomènes différents: on peut parler d'un détachement *fort* lorsque l'auditeur ou le lecteur n'ont plus aucun moyen de revenir au texte-source, et d'un détachement *faible* lorsque l'énoncé détaché est extrait d'un texte qui se trouve contigu.

Le détachement « faible » est très fréquent dans la presse écrite : il correspond en général à un titre, un sous-titre, une légende de photo. C'est pourquoi, si le lecteur voit un article ayant un titre entre guillemets, il pensera immédiatement que ce titre est une citation extraite du corps de l'article. Par exemple :

- « Au Quai d'Orsay : 'Les déclarations prêtées au ministre ne sont pas crédibles' » (« **Le Monde** », 24-1-2004, p.8)
- Jean-Louis Borloo, ministre de la ville, sur les zones urbaines sensibles :

« Les cités doivent devenir des quartiers ordinaires » (« **Libération** », 10-11-2003, p.14).

Mais la presse écrite contemporaine fait aussi usage du détachement « fort » : dans ce cas, on a affaire à des citations qui sont totalement dissociées de leur texte-source. Ce caractère de détachement fort est souligné par des étiquettes très diverses ; par exemple « la citation du jour » :

La citation du jour: « Il y a une panne européenne, il y a une crise, mais ce n'est pas l'explosion. » Le commissaire européen Michel Barnier, hier. (« Métro », 15-1-2003)

Ou encore « la phrase qui tue »:

C'est dit ! « Tous ceux qui vivent en France doivent se soumettre aux règles et coutumes de la société française. » - Le Conseil représentatif des institutions juives de France a salué, hier, le discours du chef de l'Etat. (« 20 minutes », 18-12-2003)

Le détachement fort est particulièrement visible quand il s'agit de pages entières de citations, qui forment une sorte de patchwork. On peut ainsi évoquer dans « **Veja** » la double page de la rubrique « **Veja essa** ». Celle du 3 septembre 2003, par exemple, contient dix-huit énoncés détachés venant des locuteurs les plus divers. Celle ci, par exemple :

« **Eu me acho linda.** » (Preta Gil, cantora, a filha robusta do ministro da Cultura, Gilberto Gil, que poson nua para o encarte do seu CD) (3 septembre 2003, p.34-35)

Ici le commentaire placé après l'énoncé détaché contextualise celui-ci très partiellement; en effet, **Veja** explicite le lien entre la phrase et la nudité, mais pas du tout la situation de communication dans laquelle cette phrase a pu être dite.

Si l'on observe de nombreux exemples de détachement faible, on note que la plupart du temps l'énoncé subit une altération quand il est détaché. Voici un exemple d'altération particulièrement simple, puisqu'il y a seulement suppression d'une partie de l'énoncé surasserté :

[Titre] **Hussein Chalayan:** « **Je suis très sexuel** »
(...) Les gens pensent que, parce que vous intellectualisez votre travail, vous ne pouvez pas être quelqu'un de très physique. Les deux ne sont pas antagonistes ! Je suis, et j'ai toujours été, quelqu'un de très sexuel.
(« **Jalouse** », n° 58, mars 2003, p.159)

Mais on peut aller plus loin dans l'altération. C'est le cas dans cet entretien qui a pour titre :

Alexandra Kazan : « **Pour durer dans ce métier, il faut être costaud** »

Le texte placé en dessous est assez différent de l'énoncé détaché :

« Les gens ne se rendent pas compte, ils ont l'impression que lorsqu'on est connu, on est arrivé. Mais c'est difficile de durer. Il faut être très costaud psychologiquement. Parfois, je le suis, parfois non. » (p.91)

On le voit, l'énoncé détaché ne correspond à aucune surassertion. L'énoncé original est un mouvement argumentatif complexe, réparti sur quatre phrases, avec des modulations du locuteur ; avec le détachement, il se trouve transformé en une phrase unique généralisante, une sorte de sentence.

De telles altérations ne sont en rien réservées au détachement faible ; il en va de même pour le détachement fort, mais il est beaucoup plus difficile de s'en apercevoir.

3 L'aphorisation

Ce type de divergence entre l'énoncé de départ et l'énoncé détaché souligne, à son niveau, que le détachement et la surassertion sont deux phénomènes très différents. La surassertion, qui pose un fragment comme détachable, est un phénomène qui relève du fonctionnement textuel, alors que l'énoncé détaché implique un régime énonciatif spécifique, non textuel, que nous désignons par le terme « aphorisation ». Autrement dit, on ne confondra donc pas une logique de *surassertion*, qui fait ressortir une séquence sur un fond textuel, et une logique d'*aphorisation*² – ou plus exactement d'énonciation aphorisante –, qui confère un statut pragmatique spécifique à un énoncé dépourvu de contexte.

L'aphorisation n'est pas là un phénomène qui serait totalement étranger au système linguistique. Il existe en effet dans de nombreuses langues des marquages qui ne sont pas sans rapport avec les caractéristiques de l'énonciation aphorisante. A ce propos on peut évoquer la problématique de la phrase nominale en indo-européen, telle que l'a analysée E. Benveniste dans un article célèbre, repris dans ses *Problèmes de linguistique générale*³.

La phrase nominale est un phénomène qui est attesté dans de nombreuses langues du monde. Benveniste soutient qu'une phrase nominale latine telle que « homo homini lupus » (= l'homme est un loup pour l'homme) n'est pas une variante de la phrase avec verbe (« homo homini lupus est ») mais que ces deux phrases correspondent à « deux modes d'énonciation distincts » (1966 : 166). Pour lui, la phrase à prédicat nominal, sans verbe ni copule, dans des langues qui disposent par ailleurs d'une structure de phrase à verbe être, constitue « une assertion nominale, complète en soi, pose l'énoncé hors de toute localisation temporelle ou modale et hors de la subjectivité du locuteur » (1966 : 159-160).

Il étudie plus particulièrement la phrase nominale en grec ancien :

« étant apte à des assertions absolues, la phrase nominale a valeur d'argument, de preuve, de référence. On l'introduit dans le discours pour agir et convaincre, non pour informer. C'est hors du temps, des personnes et de la circonstance, une vérité proférée comme telle. C'est pourquoi la phrase nominale convient si bien à ces énonciations où elle tend d'ailleurs à se confiner, sentences ou proverbes, après avoir connu plus de souplesse. » (1966 : 165).

²Le choix de ce terme n'est pas totalement satisfaisant ; en grec *aphorizo* signifie avant tout une opération de détermination, et *aphorisma* une définition. Nous préférons nous appuyer sur l'usage contemporain qui voit dans l'aphorisme, de manière plus large, « une phrase d'allure sentencieuse, qui résume en quelques mots une vérité fondamentale ». (*Grand Larousse de la langue française*).

³« La phrase nominale », *Problèmes de linguistique générale*, Paris, Gallimard, 1966, pp.151-167.

Dès lors, la phrase nominale et la phrase avec le verbe être au présent « n'assertent pas de la même manière et n'appartiennent pas au même registre. La première est du discours ; la seconde de la narration. L'une pose un absolu ; l'autre décrit une situation. Ces deux traits sont solidaires. »

Benveniste valide cette thèse par une étude de corpus. Pour ce faire, il compare un texte poétique, les « Pythiques » de Pindare, et la prose narrative d'Hérodote. L'étude montre que les poèmes contiennent un grand nombre de phrases nominales, et la prose narrative très peu. Les phrases nominales de Pindare ont toujours un caractère sentencieux et sont toujours du discours direct. La phrase nominale sert à convaincre, non à informer, elle permet d'appuyer une argumentation : « C'est, hors du temps, des personnes et de la circonstance, une vérité proférée comme telle. »

On notera à quel point la démarche suivie par Benveniste ressemble à celle qu'il adopte dans sa célèbre étude sur les deux plans d'énonciation, « histoire » et « discours ». Mais ici il n'oppose pas de régimes énonciatifs en se fondant sur leur relation avec leur situation d'énonciation, c'est-à-dire sur les éléments déictiques. En traitant de la phrase nominale, il associe étroitement les dimensions référentielles, modales et textuelles : la phrase nominale grecque est à la fois un énoncé non embrayé et un énoncé qui fait autorité, dont la responsabilité est attribuée à une instance qui ne coïncide pas avec le producteur empirique de l'énoncé. C'est aussi un énoncé (proverbe, adage, sentence...) qui ne constitue pas à proprement parler un texte.

4 L'énonciation aphorissante

Dans cette étude, Benveniste ne s'intéresse pas directement à la textualité : il étudie des énoncés sentencieux, qui, même s'ils apparaissent dans des textes, sont par nature hors de la textualité : « Homo homini lupus » énonce une vérité absolue. Mais quand on aborde la problématique plus large des énoncés détachés, on est bien obligé de mettre en avant cette question de la textualité pour définir deux régimes d'énonciation, de façon à distinguer *énonciation aphorissante* et *énonciation textualisante*. La seconde inscrit chaque énoncé dans l'horizon global d'un texte relevant d'un genre de discours, que celui-ci soit monologal ou dialogal. En revanche, la première n'entre pas dans la logique du genre de discours.

Énonciation

Aphorissante

textualisante

sentencieuse

personnelle

L'énonciation sentencieuse (proverbes, adages, dictons, slogans...) concerne des énoncés autonomes par nature et à valeur généralisante ; elle est foncièrement polyphonique puisque son locuteur attribue la responsabilité de son dire à une

instance anonyme, un « hyperénonciateur ». Quant à l'énonciation aphorissante « personnelle », elle est attribuée à des individus, en général indexés par des noms propres, et n'est pas nécessairement généralisante.

Entre un énoncé aphorisé et un texte, il n'y a pas une différence de taille mais d'ordre. L'énoncé aphorisé ne se laisse pas enfermer dans la géométrie usuelle, qui fait de la phrase un constituant du texte, lui-même rapporté à un genre de discours, seule réalité pour les locuteurs.

L'énonciation aphorissante institue une scène de parole où il n'y a pas d'interaction entre deux protagonistes placés sur un même plan : l'instance responsable de l'énonciation aphorissante se tient dans un ailleurs. Dans le cas de l'aphorisation de type sentencieux c'est particulièrement évident ; mais il en va de même pour l'aphorisation personnelle, qui convoque des paroles dites sur une autre scène par des locuteurs autorisés, qui ne sont pas n'importe qui. Le travail de détachement des énoncés en suppose en effet un autre, celui par lequel un nombre restreint d'individus ont été détachés de la communauté, placés dans la sphère des geus qui « comptent », qui ont de l'importance.

Bien entendu, il n'existe aucun énoncé qui soit hors d'un texte, et donc d'un genre. Cette parole tenue sur une autre scène intervient dans des textes, lesquels relèvent d'un genre de discours déterminé. L'aphorisation « *Eu me acho linda* », par exemple, se trouve dans une rubrique du magazine « *Veja* ». Dire que l'énonciation aphorissante n'entre pas dans la logique du genre de discours, ce n'est donc pas dire qu'elle soit proférée hors de tout genre. Cela veut seulement dire qu'il faut prendre acte de la *prétention illocutoire* attachée à l'aphorisation, celle d'être une parole absolue, sans contexte, convoquée dans un texte qui, lui, n'est pas absolu. L'essentiel est ainsi la tension, qui s'établit entre une aphorisation et le texte qui l'accueille. On pourrait tenir un raisonnement semblable à propos des textes (par exemple scientifiques ou romanesques) qui effacent toute relation à leur situation d'énonciation, ces textes où il semble que « *personne ne parle* », pour reprendre une formule de Benveniste. En réalité, un texte scientifique ou un roman impersonnel ont bien un auteur, mais le linguiste est obligé de prendre acte leur prétention à ne pas marquer leur relation avec leur situation d'énonciation.

L'aphorisation relève en principe de la citation. Ce n'est pas une première parole, mais la répétition d'une parole. C'est évident pour les énoncés sentencieux, qui sont voués à la répétition. Mais cela vaut aussi pour les aphorisations personnelles, qui affichent clairement leur statut de citation. On peut néanmoins se demander s'il existe des exceptions, c'est-à-dire des aphorisations en quelque sorte primitives, énoncées comme telles par un locuteur. De tels énoncés existent. On peut penser à ceux que produisait la Pythie antique, ou plus largement n'importe quel devin ou gourou qui, interrogé, ne répond pas par une argumentation ou par un récit mais en proférant un énoncé bref, énigmatique, absolu. On ne peut pas parler ici de citation, mais on observera que l'on a affaire à des locuteurs très particuliers, qui se présentent comme parlant au nom d'un Esprit qui les inspire ; leur statut fait qu'ils n'appartiennent pas pleinement au monde de ceux à qui ils s'adressent. Ainsi, même s'il se trouve que l'immense majorité des aphorisations personnelles résultent d'un travail de détachement textuel, on ne peut pas exclure que dans certaines situations de communication très particulières on ait affaire à des aphorisations qu'on pourrait dire « *originelles* », données d'emblée comme telles. En fait, tout se passe comme si l'aphorisation « *originelle* » poussait à l'extrême la logique de la citation : la personnalité

charismatique produit d'emblée une aphorisation vouée au commentaire, dans une sorte d'autocitation originelle.

Avec l'aphorisation, on touche à l'archaïque. A travers elle, c'est la Vérité qui se manifeste par la bouche d'un Sujet en majesté. Par là il s'agit de revenir en deçà, ou au-delà, de la diversité infinie des genres de discours et de la spatialité des textes. D'une part, l'aphorisation implique un énonciateur qui se pose en Sujet de plein droit, d'autre part le Sujet se manifeste comme tel par sa capacité à aphoriser. Il s'agit fondamentalement de faire coïncider *sujet d'énonciation* et Sujet au sens *juridique et moral* : quelqu'un se pose en responsable, affirme des valeurs et des principes à la face du monde, s'adresse à toute une communauté, par-delà les allocutaires empiriques qui sont ses destinataires.

On comprend ce lien entre Sujet et aphorisation. Idéalement, le Sujet, le sub-jectum, c'est-à-dire ce qui est placé au dessous, ce qui ne varie pas, échappe à la relativité des contextes ; il peut répondre de ce qu'il dit à travers la pluralité des situations de communication. De là son lien étroit avec la juridicité : quand on veut condamner quelqu'un pour ses propos, en général ce n'est pas un texte - toujours relatif à un contexte - mais une aphorisation ou un ensemble d'aphorisations que l'on condamne. On peut en voir une illustration dans cette scène archétypale qu'est le procès de Jésus :

Quelques-uns se levèrent pour porter contre lui ce faux témoignage :
« Nous l'avons entendu qui disait : « Je détruirai ce Temple fait de main d'homme, et en trois jours j'en rebâtirai un autre qui ne sera pas fait de main d'homme. » (Marc, 14, 57-59).

Ceux qui veulent le faire condamner exhibent une aphorisation sacrilège, dont l'absence de contexte n'a rien de contingent : c'est la mauvaise nature du Sujet qu'il faut dénoncer, et non un énoncé inscrit dans un genre de discours.

La décontextualisation des aphorisations va de pair avec une opacification du sens, qui exige un travail interprétatif. Quand « *Veja* » cite « *Eu me acho linda* », le lecteur est contraint d'attribuer à cet énoncé apparemment trivial un sens au-delà de son sens immédiat. On se retrouve ici dans une situation herméneutique typique. Certes, il ne s'agit pas d'une herméneutique aussi riche et institutionnalisée que celle qui accompagne les textes philosophiques, religieux ou littéraires, mais c'est bien une « attitude herméneutique » qui amène les lecteurs ou les auditeurs à mobiliser un certain nombre de stratégies interprétatives : partant du postulat que l'aphorisation est pertinente, le lecteur doit construire des interprétations permettant de la justifier. Peu importe quelle interprétation il va construire, l'essentiel est qu'il postule un au-delà du sens immédiat et agisse en conséquence.

5 Conclusion : le visage du Sujet aphorisant

Cette problématique de l'énonciation aphorisante a un certain nombre de conséquences. Il en ressort que des notions comme « parole », « énoncé », « phrase », etc. ne sont sans doute pas si claires qu'on pourrait le penser. Parler est en effet une activité qui peut se déployer selon deux régimes d'énonciation bien distincts : l'un, textualisant, assume le caractère foncièrement situationnel de l'énonciation et la diversité des modes de subjectivation énonciative attachés aux genres de discours ; l'autre, aphorisant, pose des *propositions*, et non des

textes, référées à un Sujet de plein droit, qui n'est pas asservi au jeu des interactions discursives. Le point de vue des spécialistes du discours – pour qui, dans la lignée de Bakhtine, il n'est de parole qu'enfermée dans l'horizon d'un genre de discours – doit donc être relativisé.

Dans cette perspective, la présence très fréquente de photos du visage des locuteurs à côté des aphorisations n'apparaît pas comme un accident mais comme la manifestation de quelque chose de constitutif : la photo du visage authentifie l'aphorisation du locuteur comme étant sa parole, celle qui en fait un Sujet pleinement responsable. Un manifeste à visée éthique publié dans le magazine de gauche « **Le Nouvel Observateur** » (23-29 octobre 2003, p.27) par trois personnalités du parti socialiste (Jean-Luc Mélenchon, Vincent Peillon, Mamel Valls) est ainsi précédé par deux aphorisations extraites, par détachement faible, du texte contigu :

« On ne tric pas les citoyens en fonction de leur origine ou de leur religion »

« **Monsieur Ramadan
ne peut pas être des nôtres** »

L'énoncé de ces principes moraux s'accompagne tout naturellement des photos du visage de chacun des trois locuteurs, placées entre les énoncés détachés en titre et le texte proprement dit. Le visage est la partie noble de l'individu, celle qui exprime le mieux sa conscience, alors que le reste du corps fait passer au premier plan toute la diversité sociale, ethnique, géographique...

L'aphoriseur est un locuteur qui prend de la hauteur, qui montre l'ethos d'un homme autorisé, au contact d'une Source transcendante, au-delà des interactions immédiates. L'effacement du cotexte et du contexte qu'implique l'énonciation aphorisante renforce l'engagement illocutoire. Quand Marx dit « La religion est l'opium du peuple », ou Preta Gil « Eu me acho linda », ils sont censés énoncer leur vérité, soustraite à la négociation l'expression d'une totalité vécue : une doctrine, une conception de l'existence.

Lingüística comparativa e pré-história dos povos indígenas sul-americanos: a hipótese Tupí-Karíb e as relações genéticas entre Tupí, Karíb e Macro-Jê

Aryon Dall'Igna Rodrigues
Universidade de Brasília LALI
Laboratório de Línguas Indígenas

aryon@unb.br

Abstract

The author argues for genetic relations amongst Tupi, Karib e Macro-Jê languages by using the Comparative Method. His proposal was formerly presented in 1978 at the American Anthropological Society annual meeting and it was later published in Klein and Stark 1985. Recently, there has been some controversy regarding this proposal. The author then considers these recent proposals and argues that further research is still needed for a complete understanding of the relationship amongst the languages under discussion.¹

O método comparativo da lingüística histórica, desenvolvido pelos pesquisadores de línguas indo-europeias no século XIX e início do século XX (v. Meillet, 1925) e desde então aperfeiçoado e aplicado a outras famílias lingüísticas (semítica, bântu, ugro-finica, uralo-altaica, dravídica, austronésia, algonquina, etc.), é o instrumento metodológico mais importante para demonstrar que um conjunto de línguas tem um ancestral comum no passado pré-histórico (para uma exposição mais recente, v. Hock, 1986). Baseado fundamentalmente na regularidade das alterações fonológicas que afetam as línguas no curso do tempo e que, quando divergentes, refletem a redução ou perda de comunicação entre segmentos de uma comunidade humana que se afastaram no espaço, o método comparativo depende crucialmente do exame de dados lexicais comparáveis das línguas investigadas. Um dos fatores que têm dificultado o desenvolvimento de pesquisas de lingüística histórica na América do Sul é justamente a documentação lexical, muito limitada para a maioria das línguas. Essa deficiência na documentação, devida em boa parte a que ainda são muito poucos os pesquisadores de línguas indígenas e muito limitadas as possibilidades de trabalho extensivo, é agravada pelo fato de que muitas línguas têm desaparecido antes que tenha podido ser feita qualquer documentação lexical mais ampla.

¹O Professor Sérgio Meira foi convidado por um associado da ABRALIN para participar do V Congresso Internacional da Associação a partir de demanda aberta pela Comissão Organizadora do Congresso. Ao cientificar-se da proposta de conferência apresentada pelo Prof. Sérgio Meira a Comissão Organizadora achou por bem convidar o Prof. Aryon D. Rodrigues para participar junto com o prof. Meira de um debate sobre as línguas Tupi-Karib. A Comissão Organizadora convidou adicionalmente a Profa. Yonne Leite para coordenar a discussão dos Professores Meira e Rodrigues.

Mas, mediante utilização criteriosa dos dados disponíveis, é possível reconhecer afinidades entre línguas aparentadas. Mesmo neste continente tem sido possível não só identificar afinidades, mas definir mais claramente as relações de parentesco genético entre certas línguas, reconhecê-las como integrantes de uma mesma família e iniciar a reconstrução de propriedades da língua ancestral de que elas descendem. Esta é, entre outras, a situação do tronco lingüístico Tupí e das famílias Karíb e Jê. Há cerca de trinta anos, observando línguas desses três grupos genéticos, percebi que, embora claramente distintos, apresentavam correspondências fonológicas recorrentes em seus inventários lexicais e também em alguns elementos gramaticais. Isto era mais imediatamente observável entre Tupí e Karíb e a comparação lexical das línguas desses dois conjuntos foi o foco principal do trabalho que então empreendi. O léxico de boa parte das línguas do tronco Tupí já me era bastante conhecido e tratei de obter dados publicados e inéditos de membros da família Karíb. Dentre as centenas de palavras comparadas a partir de sua identidade ou semelhança semântica, foi possível identificar mais de 120 casos de correspondências associadas por regularidade fonológica, que é um dos principais critérios para atribuir-lhes origem comum. A maioria dessas correspondências era de palavras com significados mais universais, normalmente presentes em todas as línguas e que, por isso, dificilmente são tomados por empréstimo de outras línguas. A conclusão só podia ser a de que havia muito alta probabilidade de que tais palavras constituíssem herança de um passado comum às línguas de ambos os grupos. Cumpre ressaltar que esse conjunto de correspondências lexicais (que chamei de Lista A) abrangia todo o tronco Tupí, que é um grupo de dez famílias lingüísticas (Tupí-Guaraní, Awetí, Mawé, Jurúna, Mundurukú, Arikém, Tuparí, Mondé, Ramaráma e Puroborá), todas situadas ao sul do Rio Amazonas (com exceção de três línguas da Tupí-Guaraní, que se deslocaram para o norte em conseqüência da penetração portuguesa na Amazônia) e toda a família Karíb, cujas línguas estão em sua maioria ao norte daquele rio, mas com algumas ao sul dele, na bacia do Rio Xingu.

O estudo comparativo então feito revelou também um conjunto menor de palavras comuns só à família Tupí-Guaraní (e não às outras nove famílias do tronco Tupí) e às línguas norte-amazônicas da família Karíb. Ao contrário da Lista A, esse outro conjunto (Lista B) era formado por nomes de animais, de plantas e de alguns objetos culturais e não de conceitos universais, o que indica que devem ter passado por empréstimo de uma(s) língua(s) à(s) outra(s) numa situação de contacto entre os respectivos povos. Foi notado, além disso, que parte dessas palavras são vocábulos longos, de três, quatro e cinco sílabas (p. ex., *marukajá*, *urukure'á*), o que as caracteriza como estranhas no léxico da família Tupí-Guaraní, no qual predominam radicais de uma e duas sílabas. Isso indica que a vertente Tupí-Guaraní deve ter sido a receptora dos empréstimos e não sua fornecedora. Acresce que essas palavras da Lista B são comuns a todas as línguas da família Tupí-Guaraní, as quais já desde a época do descobrimento do Brasil estão distribuídas entre o Rio Amazonas ao norte e o Rio da Prata ao sul e desde o Rio Madeira a oeste até Rio Gurupi a leste, assim como ao longo da costa atlântica. Para explicar a presença das palavras em questão em línguas assim espalhadas, temos de admitir que o contacto que as introduziu deve ter sido com a língua que deu origem à família Tupí-Guaraní, isto é, com o Proto-Tupí-Guaraní, para o qual é estimada uma antiguidade de bem mais

de 1.000 anos, portanto em época pré-colombiana. E como a família Tupí-Guaraní é essencialmente sul amazônica – como, aliás, todo o tronco Tupí –, o ponto de contacto deve ter sido ao sul do Rio Amazonas, o que implica a migração da maioria das línguas Karíb do sul para o norte, tal como já havia sido postulado pelo etnólogo Karl von den Steinen. Como se vê, foram levantadas duas hipóteses diferentes, ambas de alcance histórico: uma, baseada na Lista A, sobre a origem comum do tronco Tupí e da família Karíb e a outra, baseada na Lista B, de um contacto entre o Proto-Tupí-Guaraní e línguas Karíb que hoje estão ao norte do Rio Amazonas.

Naquele estudo incluí também três outras listas com correspondências devidas certamente a contato recente entre línguas das famílias Tupí-Guaraní e Karíb mais ou menos vizinhas geograficamente e que são casos mais seguros de empréstimos: Lista C.1 – empréstimos da Língua Geral em línguas Karíb norte-amazônicas, Lista C.2 – empréstimos do Wayampí ou da Língua Geral no Wayána, Lista C.3 – empréstimos do Wayána no Wayampí.

Além disso, depois de mencionar que a possibilidade de parentesco genético entre Tupí e Karíb tinha sido conjecturada já em 1909 pelo pesquisador holandês C. H. de Goeje, observei que nenhuma das classificações gerais das línguas sul-americanas havia considerado essa possibilidade. Daí o cuidado que tive em distinguir situações de possíveis empréstimos e situações que apontam com mais segurança para uma provável origem comum. Assinalei também a existência de características estruturais idênticas nas fonologias e nas gramáticas das línguas Tupí e Karíb.

Por fim, acrescentei que várias palavras ou morfemas incluídos na Lista A parecem pertencer a um conjunto mais amplo de relações que incluí as línguas do tronco Macro-Jê, tendo apresentado alguns exemplos com dados de duas famílias deste tronco, a Jê e a Boróro. A esse propósito, comentei a inadequação dos agrupamentos propostos por Joseph H. Greenberg e salientei que, nas situações em que podemos utilizar dados comparativos fonologicamente controlados, surge como possível um grupo genético que incluí Tupí, Karíb e Macro-Jê. Esse estudo, apresentado em 1978 em um simpósio na reunião anual da American Anthropological Association em Los Angeles, foi publicado em 1985 no volume *South American Indian languages: retrospect and prospect* (Klein; Stark, 1985) e recentemente traduzido para o Português e republicado, por iniciativa de meus colegas da UNICAMP, no volume *Saudades da Língua*, comemorativo dos 30 anos do Instituto de Estudos da Linguagem, de que fui co-fundador (Albano et al., 2003, p. 393-410).

Nas duas décadas decorridas desde a primeira publicação poucas reações houve ao artigo. De parte dos lingüistas as duas principais foram as dos especialistas em línguas Karíb Desmond Derbyshire e Spike Gildea, ambas apresentadas no Congresso Internacional de Americanistas de 1991, em New Orleans, ao qual eu não pude comparecer. Derbyshire consubstanciou, com dados das línguas Karíb, cinco semelhanças gramaticais acrescidas às que eu tinha referido (Derbyshire, 1994), e Gildea fez o mesmo com respeito a um caso específico de nominalização (Gildea, 1994). Essas foram, portanto, reações positivas, tendentes a reforçar a hipótese de parentesco genético entre os dois grandes grupos de línguas com base no avanço que se dava no conhecimento da família Karíb. Nenhuma reação chegou a meu conhecimento da parte dos especialistas em línguas Tupí. De parte dos antropólogos houve boa receptividade (Urban, 1992). Outra reação positiva veio recentemente da parte de biólogos geneticistas. Francisco M.

Salzano, Mara H. Hutz, Sabrina P. Salamoni, Paula Rohr e Sídia M. Callegari-Jacques da Universidade Federal do Rio Grande do Sul publicaram em dezembro de 2005, na revista *Current Anthropology*, artigo em que relatam a aplicação de uma metodologia para construir árvores que representem graus de proximidade genética entre populações com base na análise de diversos fatores biológicos e para a qual foram utilizados dados obtidos de índios que falam línguas dos troncos Tupí e Macro-Jê e das famílias Karíb e Aruák, de modo a verificar a compatibilidade das árvores possíveis por eles obtidas com as árvores implícitas nas propostas de classificação lingüística de Cestnir Loukotka (1968), de Joseph H. Greenberg (1987) e de A. D. Rodrigues. A conclusão foi que esta última é a que mostrou melhor compatibilidade com as árvores projetadas pela pesquisa genética biológica (Salzano et al., 2005). Este é, certamente, um novo e valioso meio de verificação para as hipóteses geradas pela pesquisa lingüística histórico-comparativa. Apesar de sempre existir a possibilidade de uma população mudar de língua no curso de seu desenvolvimento histórico, a consideração de dados de vários povos que falam línguas de uma mesma família reduz consideravelmente o efeito de eventuais casos de mudança de língua, de modo que, ainda que não haja – como não há – nenhuma associação direta entre uns e outros, a conservação de fatores genéticos ao longo de várias gerações deve coincidir com a conservação de elementos lingüísticos passados ininterruptamente de geração a geração.

Novas evidências de relações de línguas atribuídas ao tronco Macro-Jê com as línguas da família Karíb e as do tronco Tupí têm sido identificadas em estudos comparativos mais recentes, sobretudo os que temos realizado no Laboratório de Línguas Indígenas da Universidade de Brasília (Cabral; Costa, 2004; Cabral; Rodrigues, 2004; Rodrigues, 1999a, 2000, 2002). Também se continuou trabalhando na constituição interna dos troncos Tupí (Cabral, 2000, 2001a, 2001b, 2002; Cabral; Aragon, 2005; Cabral; Magalhães, 2004; Cabral; Rodrigues, 2005; Rodrigues, 1999b, 2005; Rodrigues; Cabral, 2006; Rodrigues; Dietrich, 1997) e Macro-Jê (Rodrigues, 1997, 1999d, 2007, Rodrigues; Cabral, 2007) e no melhor conhecimento de suas línguas (Tupí: Boufim, 1997, 2003; Braga, 2005; Cunha, 1987; Fargetti, 2001; Franceschini, 1999, 2005; Gabas Júnior, 1999; Galúcio, 2001; Grannier, 2002, 2005; Julião, 2005; Magalhães, 2002, 2005; Praça, 1999, 2001; Reis Rodrigues, 1995, 2001; Rodrigues, 1997, 2001, 2003, 2004; Seki, 2000; Storto, 1999; Macro-Jê: Alves, 1999, 2007; Amado, 1999, 2007; Borges, 1995; Cavalcante, 1987; D'Angelis, 1996, 1998, 2002; Dourado, 2001, 2003, 2005, 2007; Maia, 1998, 2002; Oliveira, 2006; Palácio, 1984, 2004; Ribeiro, 1996, 2004, 2005; Santos, 2007; Seki, 2002; Silva, 1986; Silva, 2005; Silva; Oliveira, 2007; Souza, 1990; Telles, 2007; Telles de Lima, 1995; Viana, 1995, 2004, 2007). Sobre a família Karíb, têm-se destacado os estudos de Spike Gildea (1992, 1994, 1995, 1998, 2001), uma exposição abrangente por Desmond Derbyshire (1999) e diversas teses e outros estudos sobre línguas dessa família (Abott, 1991; Álvarez, 1998; Armellada; Olsa, 1994; Carson, 1982; Cruz, 2005; Franchetto, 1990; Franchetto; Santos, 2001; Hall, 1988; Koehn; Koehn, 1986; Meira, 1999, 2003; Pacheco, 2001, 2005; Sousa Cruz 2002, 2005), além de um estudo comparativo mais recente por Meira e Franchetto (2005).

Com o aumento do conhecimento sobre esses três grandes grupos lingüísticos crescem naturalmente as possibilidades de consolidação ou de reformulação da hipótese formulada e apresentada há um quarto de século.

Por ocasião do IV Encontro Internacional da Associação Brasileira de Linguística fui convidado pela nossa então presidente a participar de uma mesa-redonda, para discutir a hipótese Tupí-Karib, dirigida por Yonne F. Leite e tendo como expositor Sérgio Meira, autor de premiada tese de doutorado sobre a língua Tirió da família Karib. Ao abrir a mesa, Leite anunciou que iria ser discutida a *controvertida* hipótese de Rodrigues. Embora eu não soubesse o que iria dizer Meira, a dirigente da mesa me pediu que desse aos presentes uma idéia da minha hipótese, o que fiz brevemente. Então ela deu a palavra ao expositor, que apresentou uma descrição correta do meu trabalho publicado em 1985, tendo dito que não via como discordar da hipótese baseada na Lista A, a do parentesco genético Tupí-Karib, de modo que desfez a expectativa de que fosse controvertê-la. Quanto à Lista B, com as evidências de empréstimo linguístico entre Tupí-Guaraní e Karib norte-amazônico, Meira também a reconheceu como tal, mas levantou hipótese alternativa quanto ao ponto do contacto, que segundo ele poderia ter sido ao norte do Rio Amazonas. Com respeito às listas C, irrelevantes para as duas hipóteses, limitou-se a pequenos comentários, com o acréscimo de dados de sua própria experiência, e encerrou sua apresentação sugerindo que se mudasse o nome da hipótese Tupí-Karib para "Karib-Tupí" ou "Ka-Tu", já que (!) *katu* significa 'bom' em algumas línguas do tronco Tupí.

Em sua exposição, Meira disse acreditar que seu recente estudo comparativo da família Karib (Meira; Franchetto, 2005) tende a diminuir a possibilidade de que os falantes de Proto-Karib tenham vindo do Brasil central. Entretanto, esse estudo, além de mostrar a existência de dois ramos da família Karib ao sul do Rio Amazonas - Arara-Ikpéng-Bakairí e Kuikúru-Kalapálo-Matipú-Nahukwá -, em contraste com os ramos norte-amazônicos, por razões de método não utilizou os dados lexicais necessários para a verificação da hipótese de empréstimos, a qual continua aguardando exame mais extensivo e mais detalhado por parte dos caribistas. Na verdade, continua fazendo falta a elaboração de dicionários abrangentes para as línguas da família Karib, sobretudo as sul-amazônicas, que documentem o máximo possível do léxico cultural, como procurou fazer W. Ahlbrinck na *Encyclopaedie der Kariben* (1931). Além disso, a idéia apresentada por Meira de um Proto-Karib norte-amazônico é de difícil conciliação com o que até agora se pode admitir como possível localização do Proto-Tupí e do Proto-Tupí-Guaraní. Entendo que o Prof. Sérgio Meira, depois de ter empreendido minuciosa comparação fonológica das línguas da família Karib, com especial atenção para o contraste entre as línguas meridionais e as setentrionais (Meira e Franchetto 2005), quis rever o meu estudo sobre relações históricas entre essa família e o tronco Tupí (Rodrigues 1985). Como fui solicitado a apresentar um resumo do que iria falar aqui, esperei até o último momento dado pela organização do congresso, para saber o que deveria dizer em função das colocações que me seriam apresentadas. Mas tive de projetar um "resumo" sem saber em que direção encaminhar minha fala. Limitei-me, por isso, a dar uma indicação sumária do que apresentei há trinta anos em reunião da American Anthropological Association e que foi publicado num livro de 1985 (Klein e Stark 1985) e recentemente foi traduzido para o Português no livro *Saudades da Língua*, comemorativo dos 25 anos do Instituto de Estudos da Linguagem (IEL) da UNICAMP (Albano et al. 2003). Acrescentei, porém, algumas considerações sobre o que pode ser feito depois de tantos anos, acrescentando ou retificando. Como o foco previsto para esta ocasião era essencialmente Tupí-Karib, não fiz referência às possíveis relações com o tronco Macro-Jê, de que tratei no final

daquele trabalho, o qual por isso passou a ser visto como referente também a um hipotético parentesco pré-histórico enlaçando esses três grandes conjuntos de línguas (v. p. ex. Urban 1992). Agora, antes de concentrar-nos nas relações entre Tupí e Karíb, convém mencionar que, em recente pesquisa, um grupo de geneticistas teve a iniciativa de testar as três principais hipóteses de lingüistas sobre os quatro mais importantes grupos lingüísticos da América do Sul, o Aruák, o Karíb, o Tupí e o Jê, a de Loukotka, a de Greenberg e a de Rodrigues (Salzano et al. 2005). Concluíram que das três a que tem o melhor apoio genético é a de Rodrigues (id., p. S127), ou seja, a que aproxima mais Tupí e Karíb e então Jê, deixando mais afastado o Aruák.

O resumo do Prof. Sérgio Meira foi publicado sendo que ele mostrou claramente os três tipos de relações entre Tupí e Karíb levantados por mim em 1976 (Rodrigues 1985) e se propôs questionar apenas a hipótese sobre a área de um contacto sul-amazônico entre Tupí-Guaraní e Karíb norte-amazônico para explicar a presença de palavras comuns, perguntando se esse contacto não teria sido na costa da Venezuela e da Guiana. Ainda sem ouvir seus argumentos, eu diria, repetindo o que já havia dito em 1976, que o fato de as palavras em questão serem comuns a todas as línguas da família Tupí-Guaraní indica que elas têm nesta a mesma antigüidade do Proto-Tupí-Guaraní, estimável em mais de mil anos; além disso, tanto o fato de não se encontrarem nas línguas das outras famílias do tronco Tupí, como o fato de muitas delas serem fonologicamente atípicas para línguas Tupí-Guaraní, estão a indicar que devem ser empréstimos recebidos pelo ancestral destas. Apesar dos trinta anos já transcorridos desde quando fiz o meu estudo, parece-me que os atuais estudiosos das línguas Karíb ainda não examinaram a situação das mesmas palavras na sua família lingüística, nem a questão da presença de parte dessas palavras também em línguas norte-amazônicas da família Aruák.

Referências

- [1] ABBOTT, M. M. *Handbook of Amazonian languages*, vol. 3. Mouton de Gruyter, Berlin/New York, 1991, pp. 23-160.
- [2] ALBANO, E., ET AL. *Saudades da língua*. Mercado de Letras, Campinas, 2003.
- [3] ALVAREZ, J. Ergatividade escindida y distribución complementaria de fns y afijos pronominales en pemón (caribe). *Opción, Rev. de Ciencias Humanas y Sociales* 25 (1998), 69-94.
- [4] ALVES, F. C. Aspectos fonológicos do apâniekrá (jê). Master's thesis, Universidade de São Paulo, 1999.
- [5] ALVES, F. C. *Línguas e culturas Macro-Jê*. Ed. da UnB, FINATEC, Brasília, 2007, ch. Sistema fonológico do Timbira Apâniekrá (fonemas, sílaba e acento), pp. 45-55.
- [6] ALVES, P. M. *O léxico do Tuparí: proposta de um dicionário bilíngüe*. PhD thesis, Universidade Estadual Paulista, Araraquara, 2004.
- [7] AMADO, R. S. Análise fonológica preliminar do pykobyê. Master's thesis, Universidade de São Paulo, 1999.

- [8] AMADO, R. S. *Línguas e culturas Macro-Jê*. Ed. da UnB; FINATEC, Brasília, 2007. ch. O alongamento vocálico em Pykobyé: motivações prosódicas e morfossintáticas. pp. 77-83.
- [9] BONFIM, F. Análise gramatical das orações da língua tembé. Master's thesis, Universidade de Brasília, 1997.
- [10] BONFIM, F. *Ordem de constituintes e movimento em Tembé: minimalismo e anti-simetria*. PhD thesis, Universidade Federal de Minas Gerais, Belo Horizonte, 2003.
- [11] BORGES, M. N. O. F. Aspectos da morfossintaxe do sintagma nominal na língua kayapó. Master's thesis, Universidade de Brasília, 1995.
- [12] BORGES, M. V. As falas masculina e feminina em karajá. Master's thesis, Universidade Federal de Goiás, 1997.
- [13] BRAGA, A. O. *Aspects morphosyntaxiques de la langue makurap/tupí*. PhD thesis, Université de Toulouse - Le Mirail, 2005.
- [14] CABRAL, A. S. A. C. Flexão relacional na família tupí-guaraní. *Bol. da Ass. Bras. de Ling.* 25 (2000), 233-262.
- [15] CABRAL, A. S. A. C. *Des noms et des verbes en tupi-guarani*. LINCOM, Munique, 2001, ch. Observações sobre a história do morfema - a da família Tupí-Guaraní, pp. 133-162.
- [16] CABRAL, A. S. A. C. *Estudos sobre línguas indígenas*. UFPA, Belém, 2001, ch. O desenvolvimento da marca de objeto de segunda pessoa plural em Tupí-Guaraní, pp. 117-145.
- [17] CABRAL, A. S. A. C. *Ergatividade na Amazônia*, vol. 1. Centre d'études des langues indigènes d'Amérique (CNRS, IRD); Laboratório de Línguas Indígenas (UnB), Brasília, 2002, ch. Natureza e direções das mudanças de alinhamento ocorridas no tronco Tupí, pp. 5-7.
- [18] CABRAL, A. S. A. C., AND ARAGON, C. C. A posição do akuntsú na família lingüística tupari. In *Congresso Internacional da ABRALIN, 4. Anais* (Brasília, 2005), pp. 133-39.
- [19] CABRAL, A. S. A. C., AND COSTA, L. Xikrín e línguas tupí-guaraní: marcas relacionais. *Liames 4* (2004), 7-19.
- [20] CABRAL, A. S. A. C., AND MAGALHÃES, M. Contribuição aos estudos comparativos da família tupí-guaraní. In *Encontro Nacional do Grupo de Estudos Lingüísticos do Centro-Oeste, 2. Atas* (2004), pp. 62-68.
- [21] CABRAL, A. S. A. C., AND RODRIGUES, A. D. *Novos estudos sobre línguas indígenas*. Ed. da UnB, Brasília, 2005, ch. O desenvolvimento do gerúndio e do subjuntivo em Tupí-Guaraní, pp. 47-58.
- [22] CABRAL, A. S. A. C., AND RODRIGUES, A. D. L. C. Notas sobre ergatividade em xikrín. *Liames 4* (2004), 21-28.

- [23] CAVALCANTE, M. P. *Fonologia e morfologia da língua Kaingáng: o dialeto de São Paulo comparado com o do Paraná*. PhD thesis, Universidade Estadual de Campinas, 1987.
- [24] CRUZ, M. O. S. *Línguas indígenas brasileiras: fonologia, gramática e história*, vol. 2. UFPA, Belém, 2002, ch. A posse do nome em Ingaricó (Carib), pp. 13-21.
- [25] CRUZ, M. O. S. *Fonologia e gramática Ingarikó: Kapon - Brasil*. PhD thesis, Vrije Universiteit Amsterdam, 2005.
- [26] CUNHA, P. Análise fonêmica preliminar da língua guajá. Master's thesis, Universidade Estadual de Campinas, 1987.
- [27] D'ANGELIS, W. R. O modelo das geometrias de traços e as línguas do tronco macro-jê. *Letras de Hoje* 104 (1996), 51-56.
- [28] D'ANGELIS, W. R. *Traços de modo e modos de traçar geometrias: línguas Macro-Jê e teoria fonológica*. PhD thesis, Universidade Estadual de Campinas, 1998.
- [29] D'ANGELIS, W. R. *Línguas Jê, estudos vários*. UEL, Londrina, 2002, ch. Gênero em Kaingáng?
- [30] DAS DORES DE OLIVEIRA, M. *Ofayé, a língua do povo do mel: fonologia e gramática*. PhD thesis, Universidade Federal de Alagoas, 2006.
- [31] DE ARMELLADA, C., AND OLZA, J. *Gramática de la lengua pemón (morfossintaxis)*. Universidad Católica del Táchira, San Cristóbal, 1994.
- [32] DE LIMA, S. T. A língua umutína: 'um sopro de vida'. Master's thesis, Universidade Federal de Pernambuco, Recife, 1995.
- [33] DERBYSHIRE, D. C. Clause subordination and nominalization in tupiguaranian and caribbean languages. *Revista Latinoamericana de Estudios Etnolingüísticos* 8 (1994), 179-198.
- [34] DERBYSHIRE, D. C. C. *The Amazonian languages*. CUP, Cambridge, 1999, pp. 23-64.
- [35] DOURADO, L. Estudo preliminar da fonêmica panará. Master's thesis, Universidade de Brasília, 1990.
- [36] DOURADO, L. *Aspectos morfossintáticos da língua Panará (Jê)*. PhD thesis, Universidade de Brasília, 2001.
- [37] DOURADO, L. *Novos estudos sobre línguas indígenas*. Ed. da UnB, Brasília, 2005, ch. Sentenças interrogativas em Panará, pp. 85-98.
- [38] FARGETTI, C. M. *Estudo fonológico e morfossintático da língua Jurúna*. PhD thesis, Universidade Estadual de Campinas, 2001.
- [39] FRANCESCHINI, D. *La langue sateré-mawé: description et analyse morphosyntaxique*. PhD thesis, Université de Paris VII, 1999.

- [40] FRANCESCHINI, D. *Novos estudos sobre línguas indígenas*. Ed. da UnB, Brasília, 2005, ch. Os demonstrativos em Sateré-Mawé (Tupí), pp. 59-68.
- [41] FRANCHETTO, B. *Amazonian linguistics: studies in lowland South American languages*. University of Texas Press, Austin, 1990, ch. Ergativity and nominativity in Kuikúro and other Carib languages, pp. 407-428.
- [42] FRANCHETTO, B., AND SANTOS, M. *Estudos sobre línguas indígenas*. UFPA, Belém, 2001, ch. Estruturas argumentais em Kuikúro (Karib do Alto Xingu), pp. 101-115.
- [43] GALÚCIO, A. V. *The morphosyntax of Mekéns*. PhD thesis, University of Chicago, 2001.
- [44] GILDEA, S. The proto-cariban and tupi-guarani object nominalizing prefix. *Revista Latinoamericana de Estudios Etnolingüísticos* 8 (1994), 163-177.
- [45] GILDEA, S. *On Reconstructing Grammar: Comparative Cariban Morphosyntax*. Oxford University Press, Oxford, 1998.
- [46] GILDEA, S. Pre-Proto-Tupí-Guaraní main clause grammar. In *Encontro Internacional do Grupo de Trabalho de Línguas Indígenas da ANPOLL. 1*, E. da UFPA, Ed. Belém, 2001.
- [47] GRANNIER, D. M. *Aspectos da morfossintaxe do Guaraní Antigo*. PhD thesis, Universidade Federal de Alagoas, 2002.
- [48] GRANNIER, D. M. *Novos estudos sobre línguas indígenas*. Ed. da UnB, Brasília, 2005, ch. A natureza dos prefixos relacionais em Guaraní Antigo, pp. 129-140.
- [49] HALL, K. L. *The morphosyntax of discourse in De'kwana Carib*. PhD thesis, Washington University, 1988.
- [50] HOCK, H. H. *Principles of historical linguistics*. Mouton de Gruyter, Berlin, 1986.
- [51] JULIÃO, M. R. S. *Aspects morphosyntaxiques de l'anambé*. PhD thesis, Université de Toulousou-Le Mirail, 2005.
- [52] JÚNIOR, N. G. *A grammar of Karo, Tupi (Brazil)*. PhD thesis, University of California, Santa Bárbara, 1999.
- [53] KLEIN, H. E. M., AND STARK, L. R. *South American Indian Languages: retrospect and prospect*. University of Texas Press, Austin, 1985.
- [54] KOEHN, E., AND KOEHN, S. S. A. *Handbook of Amazonian languages*, vol. 1. Mouton de Gruyter, Berlin/New York, 1986.
- [55] MAGALHÃES, M. M. S. Aspectos fonológicos e morfossintáticos da língua guajá. Master's thesis, Universidade de Brasília, 2002.
- [56] MAGALHÃES, M. M. S. *Novos estudos sobre línguas indígenas*. Ed. da UnB, Brasília, 2005, ch. Pronomes e prefixos pessoais do Guajá, pp. 141-151.

- [57] MAIA, M. *Aspectos tipológicos da língua Javaé*. LINCOM-Europa, Munique, 1998.
- [58] MAIA, M. *Línguas Jê: estudos vários*. Ed. UEL, Londrina, 2002, ch. O mediativo em Karajá, pp. 147-173.
- [59] MEILLET, A. *La méthode comparative en linguistique historique*. Paris/Oslo, 1925.
- [60] MEIRA, S. *A grammar of Tiriyó*. PhD thesis, Rice University, Houston, 1999.
- [61] PACHECO, F. *Morfossintaxe do verbo Ikpeng (Karib)*. PhD thesis, Universidade Estadual de Campinas, Campinas, 2001.
- [62] PACHECO, F. *Novos estudos sobre línguas indígenas*. Ed. da UnB, Brasília, 2005, ch. Funções sintáticas e periféricas em Ikpeng (Karib). pp. 169-176.
- [63] PALÁCIO, A. P. *Guatú, a língua dos índios canociros do rio Paraguai*. PhD thesis, Universidade Estadual de Campinas, 1984.
- [64] PAYNE, T. E. *Amazonian linguistics: studies in lowland South American languages*. Univ. of Texas Press, Austin, 1990, ch. Ergativity and transitivity in Panare (Carib).
- [65] PRAÇA, W. N. *Nomes como predicados na língua tapirapé*. Master's thesis, Universidade de Brasília, 1999.
- [66] PRAÇA, W. N. *Estudos sobre línguas indígenas*. UFPA, Belém, 2001, ch. Sobre o indicativo II no Tapirapé, pp. 167-176.
- [67] RIBEIRO, E. R. *Morfologia do verbo karajá*. Master's thesis, Universidade Federal de Goiás, 1996.
- [68] RIBEIRO, E. R. *Prefixos relacionais em jê e karajá: um estudo histórico-comparativo*. *Mourá 4* (2004), 91-101.
- [69] RIBEIRO, E. R. *Novos estudos sobre línguas indígenas*. Ed. da UnB, Brasília, 2005, ch. Análise morfológica de um texto Karajá, pp. 99-128.
- [70] RODRIGUES, A. D. *Nominal classification in karirí*. *Opción, Rev. de Ciencias Humanas y Sociales 22* (1997), 65-79.
- [71] RODRIGUES, A. D. *The Amazonian languages*. CUP, Cambridge, 1999, ch. Macro-Je, pp. 164-206.
- [72] RODRIGUES, A. D. *Ge-pano-carib x jê-tupí-karib: sobre relaciones lingüísticas prehistóricas en sudamérica*. In *Congreso de Lenguas Indígenas de Sudamérica* (Lima, 1999), vol. 1, Universidad Ricardo Palma, pp. 317-321.
- [73] RODRIGUES, A. D. *Flexão relacional no tronco lingüístico macro-jê*. *Bol. da Ass. Br. de Ling.* 25 (2000), 219-231.

- [74] RODRIGUES, A. D. *Línguas indígenas brasileiras: fonologia, gramática e história*, vol. 1. UFPA, Belém, 2002, ch. Correspondências lexicais e fonológicas entre Tupí-Guaraní e Tuparí, pp. 288-297.
- [75] RODRIGUES, A. D. Ergativité dans le nord-est brésilien. *Faits de Langue, revue de linguistique* 21 (2003), 71-75.
- [76] RODRIGUES, A. D. Silêncio, nasalidade e laringalidade em línguas indígenas brasileiras. *Lecturas de Hoje* 38, 4 (2003), 11-24.
- [77] RODRIGUES, A. D. Sobre a possível origem da diferença fonética entre a fala masculina e a feminina em karajá. *Línguas* 4 (2004), 115-121.
- [78] RODRIGUES, A. D. *Novos estudos sobre línguas indígenas*. Ed. da UnB, Brasília, 2005, ch. As vogais do Proto-Tupí, pp. 35-46.
- [79] RODRIGUES, A. D. *Línguas e culturas Macro-Jê*. Ed. da UnB, FINATEC, Brasília, 2007, ch. O parentesco genético das línguas Umutina e Bororo, pp. 9-18.
- [80] RODRIGUES, A. D., AND CABRAL, A. S. A. C. Sobre o desenvolvimento de padrões absolutivos em famílias orientais do tronco tupí. In *Ergatividade na Amazônia*, F. Queixalós, Ed., vol. 2. Brasília: Centre d'études des langues indigènes d'Amérique (CNRS, IRD); Laboratório de Línguas Indígenas (UnB), 2004, pp. 69-91.
- [81] RODRIGUES, A. D., AND CABRAL, A. S. A. C. Investigando a origem e o desenvolvimento de orações dependentes nas famílias do tronco tupí. *Revista da ABRALIN* 5 (2006), 11-32.
- [82] RODRIGUES, A. D., AND CABRAL, A. S. A. C. *Línguas e culturas Macro-Jê*. Ed. da UnB; FINATEC, Brasília, 2007, ch. Através do léxico Macro-Jê: em busca de cognatos, pp. 175-179.
- [83] RODRIGUES, A. D., AND DIETRICH, W. On the linguistic relationship between mawé and tupí-guaraní. *Diachronica* 14 (1997), 265-304.
- [84] RODRIGUES, C. L. R. *Étude morphosyntaxique de la langue Xipaya*. PhD thesis, Université de Paris VII, Paris, 1995.
- [85] RODRIGUES, C. L. R. *Estudos sobre línguas indígenas*. UFPA, Belém, 2001, ch. Morfemas locativos e direcionais em Xipaya, pp. 177-192.
- [86] SALZANO, F. M., ET AL. Genetic support for proposed patterns of relationship among lowland south american languages. *Current Anthropology* 46, S5 (2005), S121-129. Supplement.
- [87] SANTOS, L. *Línguas e culturas Macro-Jê*. Ed. da UnB; FINATEC, Brasília, 2007, ch. Concordância de número em Kaingáng: um sistema parcialmente ergativo e parcialmente nominativo, pp. 145-152.
- [88] SEKI, L. *Gramática do Kamaiurá, língua tupi-guarani do Alto Xingu*. Ed. da UNICAMP, Campinas, 2000.

- [89] SEKI, L. *Línguas Jê, estudos vários*. Ed. da UEL, Londrina, 2002, ch. O Krenak (Botocudo / Borum) e as línguas Jê, pp. 15-40.
- [90] SILVA, L. J. Aspectos da fonologia e da morfologia da língua rikbáktsa. Master's thesis, Universidade de Brasília, 2005.
- [91] SILVA, L. J., AND DE OLIVEIRA, S. *Línguas e culturas Macro-Jê*. Ed. da UnB; FINATEC, Brasília, 2007, ch. Observações preliminares sobre o sistema pronominal da língua Rikbáktsa, pp. 153-162.
- [92] SILVA, M. A. R., AND SALANOVA, A. P. *Verbo y ergatividade escindida em Mebengôkre*. 1997.
- [93] SILVA, T. C. Descrição fonética e análise de alguns processos fonológicos da língua krenak. Master's thesis, Universidade Federal de Minas Gerais, Belo Horizonte, 1986.
- [94] SOUZA, S. M. O sistema de referência pessoal da língua krahô. Master's thesis, Universidade Federal de Goiás, 1990.
- [95] STORTO, L. R. *Aspects of Karitiana Grammar*. PhD thesis, Massachusetts Institute of Technology, 1999.
- [96] TELLES, S. *Línguas e culturas Macro-Jê*. Ed. da UnB; FINATEC, Brasília, 2007, ch. A flexão nominal em Umutína, pp. 127-137.
- [97] URBAN, G. *História dos Índios no Brasil*. Companhia das Letras, São Paulo, 1992, ch. A história da cultura brasileira segundo as línguas nativas, pp. 87-102. Secretaria Municipal de Cultura; FAPESP.
- [98] VIANA, A. A expressão do atributo na língua karajá. Master's thesis, Universidade de Brasília, 1995.
- [99] VIANA, A. *Línguas e culturas Macro-Jê*. Ed. da UnB; FINATEC, Brasília, 2007, ch. Dissimilação de sonoridade em Boróro: uma abordagem com base no princípio do contorno obrigatório, pp. 57-75.

Language as a biocultural niche and social institution

Chris Sinha
University of Portsmouth

chris.sinha@port.ac.uk

Resumo

Neste artigo a dupla natureza da linguagem como pertencente à cultura (segundo Condillac) ou como uma das faculdades superiores dos humanos (segundo Descartes) será discutida. As questões que se colocam nessa discussão são: em que sentido podemos recolocar nossas idéias sobre a ontologia da linguagem, de forma a que ela seja vista tanto como uma forma sócio-cultural quanto algo intrínseco à biologia humana? Como, essa nova visão "biocultural" da linguagem afeta nossas idéias sobre o seu processamento e sua aquisição?

Grammars... refer to real structures, though not to psychologically real structures in the processing sense... a grammar is a description of our *knowledge* of a *social institution* - the language - and because of this basis in social or institutional reality, rather than in cognitive functioning, grammars and psychological processes have no more than the loose relationships they appear, in fact, to have. The role of grammar during speech programming is analogous to the role of other social institutions during individual behaviour. This role is to define and evaluate the behaviour of individuals. It is not to cause the behaviour. (McNeill, 1979, p. 293).

1 Introduction: language, culture and nature

The place of language in nature and culture is one of the abiding problems of all the language sciences, of which linguistics is but one. Language is at the heart of what it means to be human - indeed it has long been held that language is both *essential* to our humanity and *unique* to our species. Descartes famously argued that language is essentially human *because* it is an expression of uniquely human, universal reason. It is reason, he maintained, that distinguishes humans from animals, which are no more than merely complex machines:

For we can easily understand a machine's being constituted so that it can utter words, and even emit some responses to action on it of a corporeal kind, which brings about a change in its organs... But it never happens that it arranges its speech in various ways, in order to reply appropriately to everything that may be said in its presence... For while reason is a universal instrument that can serve for all contingencies, these organs have need for some special adaptation for every particular action. (Descartes, 1911 [1637], p. 116).

The Cartesian position was the precursor of contemporary theories of the universality and innateness of the human language faculty, although Descartes would perhaps have questioned the conceptualization of this faculty in terms of modular "mental organs" (Chomsky, 1968; Pinker, 1994).

Later Enlightenment philosophers maintained, on the contrary, that language, as an attribute of social association, culture and civilization, was in large part *responsible* for human reason and for what came to be called the "higher mental processes." Condillac, for example, wrote that:

The resemblance between animals and ourselves proves that they have minds, and the difference between them and us proves that their minds are inferior to ours... the mental operations of animals are limited to perception, consciousness, attention, reminiscence and imagination not under their control, while ours include other operations whose origin I am about to make clear... If contemplation consists in preserving perceptions, then before the use of institutional signs, it is merely outside our control; but if it consists in preserving the signs themselves, it has no function at all. So long as imagination, contemplation and memory are unused, or as long as imagination and contemplation operate outside our control, we cannot direct our attention as we please... But when someone begins attaching ideas to signs of his own choosing, we see his memory begin to form... Later, he acquires much greater control over his imagination as he invents more signs, for he has many more ways of using it. (Condillac, 1987 [1746], p. 459).

Condillac rejected the nativism of Decartes, and his ideas not only draw upon the empiricism of Locke and Hume, but also anticipate Vygotsky's notion that human higher mental processes are dependent upon their semiotic, and especially linguistic, mediation (Vygotsky, 1978). Language, in his view, *makes* us human. As Condillac's follower Itard - educator of "Victor", the Wild Boy of Aveyron - wrote, "man is inferior to a large number of animals in the pure state of nature... the moral superiority said to be natural to man is only the result of civilization, which raises him above other animals by a great and powerful force." (apud Lanc, 1977, p. 129).

Descartes and Condillac agreed, then, that language distinguishes humans from other creatures; their disagreement was over whether it *expresses* or *enables* this difference, whether language is to be seen as primarily an expression of an innate faculty of mind, or as primarily a vehicle of social life and social interaction that enables and constitutes the uniqueness of the human mind. In this classical debate, one which defined not only the Enlightenment but Western thought about language down to the present day, we see the alignment of a number of dichotomous categories: human vs non-human; rational *vs* non-rational; culture *vs* nature.

Language, from this point of view, can belong *either* (following Condillac) to culture, making possible the "higher" faculties of human beings, or (following Decartes) to (higher, human) nature-defined in opposition to "lower" animal nature. Either way, the uniqueness of language is assumed to mirror the uniqueness of humans, interpreted in terms either of our unique nature, or, as we now would say, genetic make-up; or of the uniqueness of human culture. With the waning in recent years of the strong nativist program, it has become more

usual to accept that the acquisition of language in children, like other aspects of development, is based upon epigenetic interactions between what is innate and what is available in the environment (Sinha, 1988, 2004). In what sense might it also be possible to recast our ideas about the very ontology of language, in such a way that it comes to be seen as *both* a human socio-cultural form *and* intrinsic to human biology? And how, in its turn, might such a new, synthetic "biocultural" view of language affect our ideas about language processing and language learning?

2 Evolution, ecological niches and animal artefacts

A new synthesis is necessary because the assumptions guiding the thinking of both Cartesian, nativist and Condillacian, culturalist accounts are proving, in the light of 21st century research, to be fundamentally flawed. For the neo-Cartesians, human distinctiveness is to be sought in the genes, from which stem all human attributes, including cultural forms; for the neo-Condillacians, in the existence of human culture, a unique human attribute that, according to some, has rendered the very concept of a "human nature" invalid.¹ Both positions are predicated on an assumption of human uniqueness, of either genes or culture, and this assumption has motivated the traditional paradigm of language and its learning.

In the traditional paradigm, as we have seen, language is seen as part of *either* unique human nature (nativism), *or* unique human culture (environmentalism), and language learning is viewed as the exposure of the learner to an external "input" to be internalized. It is this model that, I argue, must be superseded, as a result of recent advances in biological sciences, advances that confront the traditional paradigm in the human sciences with a striking and challenging paradox.

The paradox is one of discontinuity in continuity. On the one hand, the biological characteristics of the human species display no dramatic discontinuities with those of other species; yet, on the other, human cognitive capacities, and human cultural constructions, appear from our current vantage point to be as exceptional in the living world as they did to Descartes. It can, of course, be argued that the cognitive and cultural discontinuity is merely a symptom of a gap in the available evidence—there are, after all, no living representatives of the human lineage since it diverged from the ancestors of our closest living primate relatives. If there were, the discontinuity would, perhaps, prove to be an illusion. Even so, it is hard to resist the conviction that, however extended the event, or sequence of events in evolutionary time, "something happened" involving language that radically transformed the evolving mind, and this transformation poses a profound and complex problem for both biological and social theory.

¹Malson (1972, p. 9) (in his Introduction to Itard's text) pursued the environmentalist direction of Condillac to its logical, if extreme conclusion, writing that "The idea that man has no nature is now beyond dispute. He has or rather is a history."

To begin with continuity: Darwin's refutation of the idea that the human species is *essentially* different, in biological constitution and evolutionary history, from other species received, in the closing years of the last century, strong confirmation in two very different domains. Succinctly stated, neither genes nor culture, singly, can account for what, if anything, makes humans different from other species.

There is no evidence of dramatic genetic discontinuity between humans and their closest primate relatives, chimpanzees. The two species share, even on the most conservative estimate, about 95% of their genetic material (Britten, 2002). Taken together with initial results of the human genome project, this suggests that whatever cognitive capacities distinguish the human species from other closely related species are unlikely to be attributable to dedicated genetic material available for directly coding such capacities. This does not mean that there is no genetic component of specifically human capacities. It does mean that the ascription of differences between the cognitive capacities of humans and those of non-humans to interspecies genetic differences *alone* is likely to be false. This is bad news for nativist modularity theories.

The news for those who would argue that what is unique about humans is the capacity for culture, a favoured hypothesis for generations past of social anthropologists, is hardly better. Culture can minimally be defined as the existence of intra-species group differences in behavioural patterns and repertoires, which are not directly determined by ecological circumstances (such as the availability of particular resources employed in the differing behavioural repertoires), and which are learned and transmitted across generations. On this definition, there is ample evidence of cultural differences in foraging strategies, tool use, and social behaviours in chimpanzees (Whiten et al., 1999; de Waal, 2001). Such a definition will also qualify, for example, epigenetically learned intra-species dialect differences between songbird communities as cultural and culturally transmitted behaviour (Marler; Peters, 1982). Again, this does not mean that there is *no* cultural foundation for uniquely human cognitive capacities; rather, it suggests that human culture, from an evolutionary and developmental point of view, must be treated as *explicandum* as much as *explicans*.

What is needed, it seems, is a theoretical apparatus capable of integrating culture and biology. One version of such integration, in which culture is analyzed as quasi-heritable units ("memes" or "culturgens") has been argued for by sociobiologists such as Richard Dawkins (1976) and Edward O. Wilson (1998). Such accounts, however, have often been criticized for their reductionism, and recent biological theory suggests that the relation between biology and culture is far more of a two-way street than was ever envisaged by sociobiology. Far from eliminating culture by absorbing it into the genotype, some biologists are increasingly acknowledging the role of culture in shaping the evolutionary process *at the genetic level*, by the construction of new selective environments. Current developments in theoretical biology, amongst which the "niche construction theory" of Laland et al. (2000) is particularly significant, extend and modify the Neo-Darwinian synthesis that dominated 20th century biology by incorporating an ecological dimension that, I shall argue, proves to be particularly important for understanding human cognitive and linguistic evolution.

First, however, I outline (in a simplified fashion) the premises of, and the outstanding problems with, the Neo-Darwinian synthesis unifying Darwin's theory of natural and sexual selection with modern population genetics.² In the Neo-Darwinian synthesis, the *unit of selection* (what is selected) is the gene, or more specifically alternative variants (alleles) of the "same" genes. The *agent of selection* (what *does* the selecting) is the extra-organismic environment, including (a) the inanimate surround, (b) other species (a and b together being the basis of natural selection), and (c) (subpopulations of) genes of the same species (the basis of sexual and kin selection). The relevant *attribute* upon which selection works (what is selected *for*) is any genetically transmitted trait. The mechanism of selection determines the differential reproductive success of the gene (allele) within the population of interacting genes, and thus the frequency distributions of genes and traits in the population. This model, when appropriately formalized, can be extended, as we shall see, by including cultural traits in the environment, that act as "amplifiers" on the selection of genetic variation: this is known as the theory of gene-culture coevolution (Lumsden; Wilson, 1981).

The core issues at the heart of the problems besetting the Neo-Darwinian synthesis can be briefly summarized. First, genes do not come singly, but as combinations (genotypes), packaged in organisms (phenotypes). It is this distinction that Dawkins (1976) recasts as a distinction between the "replicator" (that which is copied), and the "vehicle" (that which embodies the genotypic collection of replicators, and interacts with the environment). For Dawkins, it is only the gene that is actually copied, and therefore he identifies the gene (unit of selection) as the replicator, and the phenotype as a mere "vehicle" for the replicator.

However, it is *organisms*, not genes, that are subject to *direct* selection pressures in terms of those traits conferring fitness. The organism level of biological organization receives scant attention in population genetics but, even granted that the gene is the *unit* of selection, it is the organism that must be considered as the *site* of selection. Organisms, in most (though not all) cases, can be regarded as morphological individuals. However, the actual *process* of selection by an "agent" occurs in relation to the *functioning, behaving* organism. It was for this reason that Jean Piaget upheld the leading role of behaviour in evolution (Piaget, 1979). In the light of this, it may be (and frequently has been) questioned to what extent it remains legitimate to identify the "replicator" with the genetic unit of selection. Even if the DNA-based biochemical replicator is the gene, the *evolutionary dynamic* of replication-plus-selection should, it can be argued, more profitably be identified with the entire complex of the *site of selection*, which is the *active organism in its ecological niche*.

Ecologists emphasize that species shape, as well as being shaped by, their niches. Organismic behaviours may eventuate in significant *transformations* of the very environment to which the organism must adapt. A simple example (from Sinha, 1988, p. 136) is the following: "A 'path' may... be an unintended consequence of locomotion from one place to another, but it is, nevertheless, a useful

²As will become clear, there is no question here of challenging the overall Darwinian framework of evolutionary theory, but rather of questioning the premises and methodological stances of what is usually referred to as the Neo-Darwinian "modern synthesis".

one... such shaping... can [however] introduce distal consequences—food shortage, erosion, pollution, competition with other species—which are outside the initial circuit of adaptation.” (See also Costall, 2004). In many cases, however, a process of positive feedback will occur in which organism and environment are in a complementary relationship, each shaping the other. An oft-cited example is the hoof of the horse, and its adaptation to the grassland steppe whose ecological characteristics the horse, through its own motion through the landscape, reproduces.

In a subset of such cases, the resulting niche can be seen not merely as a contingent consequence of behaviour, but as an *animal artefact*, inasmuch as phenotypic individuals are genetically, morphologically and behaviourally *adapted* to the *production* of specific niches which are integral to the survival and/or reproduction strategy of the species. Examples of such artefactual niches are the nests of bower birds, and the dams of beavers. The male bower bird builds and decorates an elaborate nest (bower) to attract females., using attractive objects such as flowers, shells and leaves. The bower forms an integral part of the male's mating display, and sexual selection by the female is based upon the bower as much as upon the behavioural display of the male. Beavers construct, through coordinated and collaborative behaviour, dams that serve both as a defence against predators, and as a means to enhance the availability of food. The dams of beavers not only serve as a constructed, artefactual niche for beavers themselves, but also reproduce the wetland ecology in which many other species thrive. As a final example of the significance of animal artefacts, we can mention the termite mound, whose material structure is not only integral to the reproductive strategy of this species of social insect, but also constitutes the morphological structure of the colony as a “group organism”.

In each of these cases, the behavioural repertoire of the species includes behaviours that are specifically adapted to the making of artefactual niches, and these in turn support other behavioural strategies. The artefactual niche in many cases can be regarded as an extension of either a behavioural repertoire (eg male mating display) or of the organism's morphology (eg the bower bird's bower as functionally equivalent, as an indicator of fitness, to the tail of the peacock). Indeed, we can ask if it might be fruitful to consider certain species-specific behavioural repertoires, such as birdsong, to be kinds of animal artefacts, inasmuch the song of the adults provides a niche within which the singing behaviour is learned (see below). It can be argued that the designation of “artefact” should be reserved for more or less enduring, constructed material structures. Even if we accept this, it can still be argued that specialized behavioural repertoires constitute biocultural niches which are functionally analogous to animal artifacts. If so, human natural languages can also be viewed as species-specific biocultural niches.

3 Culture as constructed affordances and the human semiosphere

In the ecological psychology of James J. Gibson (Gibson, 1979), a key role is played by *affordances*, properties of the ecological niche affording or supporting specific kinds of action made possible by the motor system and morphology of

the animal. Such actions are both species-typical (though not necessarily species unique) and adaptive. Because affordances, Gibson maintained, are directly perceived, the phenomenal world of the animal is meaningful, in that it potentiates the activation of perception-action circuits: objects present themselves as edible, climb-able, graspable and so forth.

Gibson neglected, however, to note the crucial importance of the fact that some affordances are *constructed* by the animal itself. Artefactual niches are adaptive precisely because of the behaviours and strategies that they afford: nests are for nesting, and burrows are for burrowing. In such cases, the *site of selection* is no longer just the organism, but the organism in its self-constructed niche: the organism/niche coupling or *organism plus artefact*.

What are the implications of this for the Neo-Darwinian synthesis? A conservative reading would be that the only modification required is that the phenotype, or "vehicle", be extended to incorporate the artefactual niche. This is, indeed, the interpretation favoured by Dawkins (1982), who employs the terminology of the "extended phenotype." Under this interpretation, the "replicator" remains the gene, and only the gene. However, it is not only the gene that is copied or replicated. In fact, the artefactual niche too is both reproduced across generations, and serves as a fundamental precondition for genetic replication. The artefactual niche is thus *both a consequence of and an agent in natural and/or sexual selection*, and must then be seen as a key ingredient of the evolution of the species-typical genotype.

It seems, therefore, that the integration of ecological considerations into evolutionary theory, and specifically the existence of animal artefactual niches, further undermines the hard and fast distinction between germ-line and soma, genotype and phenotype, "replicator" and "vehicle". In fact it makes better sense to say that, even granted that the *unit* of Darwinian selection remains the gene (allele), the "replicator" includes both the artefactual niche, and the niche-adaptive behavioural repertoire of the animal. Such considerations lead us back to Piaget's more general proposition that behaviour is the leading edge and motor of evolution, prompting the conclusion (anticipated above) that the identification by Dawkins of the "replicator" with the *unit* of selection (the gene, or its hypothesized cultural analogue, the "meme") is deeply flawed, and that replication can as well or better be considered as a property of the entire *site* of selection.

At this point, it is useful to make a brief critical detour to re-examine Neo-Darwinist theories of gene-culture co-evolution such as that of Lumsden and Wilson (1981). Such accounts presuppose a functional parallelism between units of biological replication and units of cultural replication ("memes" or "culturgens"); and treat the latter as being structured as human behavioural variants analogous to gene alleles. From this perspective, ethnographic variation is analyzable in terms of aggregate properties of human populations. The Lumsden-Wilson theory has been criticized for making "the reductionist assumption that the characteristics of a society can be understood as simply the sum of the characteristics of the individuals of that society" (Alper; Lange, 1981, p. 3976), and for having no place for *emergent properties* of societies. To lend further force to this critical evaluation, I attempt in a subsequent section to formally specify, for human cultures, such emergent properties in terms of a semiotically defined ontology of the social. For the time being, we can simply note that the Lumsden-Wilson theory *presupposes* an ontological distinction

between gene and meme, nature and culture, without either explaining this distinction, or theoretically motivating the functional parallelism that is proposed to exist between the units of selection in the domains of biology and culture. In summary, the reductionist inadequacy of Neo-Darwinist theories of gene-culture coevolution consists in their recapitulating the failure of Neo-Darwinism to adequately treat the emergent properties of organisms, in their failure to adequately treat the emergent properties of socio-cultural formations.

The critical considerations outlined above have led to a more radical formulation of coevolution, advanced by Laland, Odling-Smee and Feldman (2000).³ A particular role is played in Laland et al.'s (2000, p. 144) theory by genotype/niche combinations labeled by "phenogenotypes." A phenogenotype can be defined as a class of organisms in a bound (though not necessarily genetically determined) relationship with some aspect of a self-constructed environmental niche.

Laland et al. (2000, p. 132) criticize the "human-centred" perspective of previous accounts of gene-culture coevolution, emphasizing that many non-human species behaviourally co-direct genetic evolution through niche construction. This point is important, because it situates the role of culture in human evolution within the wider class of processes, outlined in the previous section, involving adaptation to artefactual niches such as nests, dams, mounds, and burrows. Laland et al.'s model, then, is a general one, not confined to human culture and evolution. They acknowledge, however, that humans are "unique in their extraordinary capacity for culture" (Laland et al., 2000, p. 133). I interpret this to mean primarily that human cultures are unique in some fundamental respect, that is they are different (perhaps discontinuously) from the cultures of other species; and secondarily that the capacity for creating, acquiring, and transmitting cultural forms is uniquely developed (though clearly not unique) in humans.

One evident discontinuity between human and non-human cultures is that human cultures are linguistic; and the capacity for human cultural acquisition and transmission is mediated by the unique human language capacity. The nativist modularist account of this capacity proposes its inscription in the human genotype, a hypothesis vulnerable to many objections, including the difficulty stated above of locating this profound discontinuity in the continuous landscape of the primate genome. An alternative account, along the lines of the co-evolutionary theory of Laland et al. (2000), would view the human language capacity as phenogenotypic. Language, in this account, is an artefactual niche, and the capacity to acquire and use it involves the evolution and replication of a phenogenotypic "biocultural complex" (Laland et al., 2000, p. 144).

Such an account does not require the organism to possess an internal model of the grammar of a language to account for language acquisition, any more than the building of a nest requires a prior internal model of the nest. The grammar of the language is *in the language*, just as the structure of the nest is in the nest. The capacity for language is thus a cognitive-behavioural relationship between language user and the constituents of language, just as the capacity for building a nest is a cognitive-behavioral relationship between the builder and the constituents of the nest; and it is this *relationship* that, in each case, has been

³Here it is important again to emphasize that the theory advanced by Laland et al., while radically departing from Neo-Darwinism, remains Darwinian in the wider sense.

selected for in evolution. This account is thus compatible with usage-based, cognitive functional theories of language and language acquisition (Tomasello, 1998, 2003).

The artefactual niche of language is culturally situated, that is, it is dynamically embedded within the entire semiotic biocultural complex that includes other symbolic and non-symbolic artefacts. This biocultural complex we can, to use the terminology of the Russian semiotician Yuri Lotman (cf. Lotman, 1990), designate as the human *semiosphere*, the constructed, meaningful environment that is reproduced down the human generations along with the human organism itself. It is crucial to appreciate, in this context, that the semiosphere, like other animal artefactual niches, is not merely a constituent of *what* is reproduced, but is also the fundamental mechanism in the *process* of reproduction and transmission.

Because of its pre-eminence in mediating both cultural reproduction and individual cognitive processes, language is the primary and most distinctive constituent of the human semiosphere. The class of organisms with the language capacity (normally developing humans) can thus be theorized as a phenogenotypic replicator, systemically associated with a wider biocultural complex of symbolic and constructive cognitive capacities, also of a phenogenotypic nature; and individual language acquisition and use is situated in the contexts of actuation of these inter-related capacities. This account accords with the view that what makes humans unique is not an innate language acquisition device plus a variety of other species-specific innate cognitive modules, but a generalized semiotic or symbolic capacity (Piaget, 1945; Deacon, 1998; Zlatev et al., 2006); epigenetically developed from a suite of cognitive capacities largely shared with other species, but attaining higher levels of organization in humans.

4 The evolution of complexity: emergence and epigenesis

It was noted above that one of the criticisms made by Alper and Lange (1981) of the Lunsden-Wilson theory of gene-culture coevolution was its inability to account for emergent properties of human society. The term *emergence* is commonly used to mean the evolution and development of new properties and/or levels of organization of behavioural and cognitive systems as a consequence of the operation or cooperation of simpler processes. Emergence is a consequence of the evolutionary and developmental process of elaboration, involving an increase in the complexity of organism, behaviour and cognition. The adaptation of organismic behaviour to constructed, artefactual niches offers clear examples of phylogenetic elaboration leading to emergence. For example, the female bower bird's behavioural repertoire for evaluating the fitness of a prospective mate has extended (elaborated) its scope from evaluating the ritualistic behaviour of the male suitor to evaluating the bower that he constructs, entraining a more complex organismic "level" with emergent properties (the bower as a constituent of the phenogenotypic replicator).

Emergence as an evolutionary process can be conceptualized as "locking" elaboration in new adaptive circuits, in a way similar to the "ratchet effect" discussed for human cultural evolution by Tomasello (1999). Indeed, from the bi-cultural perspective, Tomasello's ratchet effect is but one, albeit exceptionally dramatic, case of a wider, not specifically human, phenomenon of emergence. If it is elaboration that lends directionality (through complexification) to "local" evolutionary processes, it can be hypothesized that it is emergence that underlies the global trend of evolution towards greater complexity. Although it is correct to reject teleological *explanations* for Darwinian evolution, a kind of teleology of process (as Piaget recognized) is a consequence of the locking and reproduction of elaboration through emergence.

Emergence is also characteristic of ontogenetic development, including cognitive development, and has been advanced as an alternative to nativist accounts of language development (MacWhinney, 1999). The developmental biological (and psychobiological) mechanism underlying ontogenetic emergence is *epigenesis*.

Contemporary theories of epigenesis in biological and psychological development build upon the pioneering accounts of Waddington (1975) and Piaget (1979). Epigenetic naturalism (Sinha, 1988) proposes a constructivist account of the interaction between the genotype and its somatic and extra-somatic environment in organismic development. The claim that such an interaction exists is, as such, trivial and undisputed, since everyone agrees that phenotype is co-determined by genes and environment. There are, however, two particularly important characteristics of the theory of epigenesis that I wish to highlight.

The first is that the role of the environmental factors is *constructive* in addition to being *selective*. Nativist approaches to the developmental interaction between genotype and environment stress the role of specific input either in permitting a developmental process to unfold, or in parametrically selecting a particular variant of development. An example of the former would be phenomena such as "imprinting", where an innate and fully endogenous process of development is "triggered" by an environmental event during a critical developmental window. An example of the latter would be the role hypothesized by generative linguists to be played by typological characteristics of target languages in setting parameters and thereby permitting the child non-inductively to acquire the grammar of the target language (Chomsky, 2000). In neither of these cases does the environmental information add any emergent *level of organization* to the genetically coded information. That is to say, the alternate pathways along which the behaviour develops, and its terminal structural complexity, are assumed already to be directly encoded in genes.⁴

⁴It is noteworthy that the Lumsden-Wilson account of gene-culture coevolution, though it employs a terminology of epigenesis, does so in a way that is more reminiscent of the Chomskian notion of parameterization. The Lumsden-Wilson theory envisages two (or, in principle, more) alternate traits distributed in a population, with the possibility of epigenetic "switching" between these traits, which could therefore be alternate expressions of the same gene(s). They hypothesize that an interaction of "epigenetic rule"-encoding genes, and environment (in particular, existing trait frequencies in the population), is responsible for the development in any given individual of one or other trait. Alper and Lange (1981, p. 3976), whose critical assessment of the Lumsden-Wilson theory I have already cited, claim that there is "absolutely no evidence that any genes of this type exist." This author is not competent to pronounce on the biological facts of this disputed point. However, it should be noted that (a)

By contrast, in epigenesis the developmental pathway and final structure of the behaviour that develops are a consequence as much of the environmental information as of the genetically encoded information. For example, the development of birdsong seems to involve reproduction by imitative epigenetic learning, rather than selection from amongst pre-established alternatives (Marler; Peters, 1982). Fledglings not exposed to a model do develop birdsong, but it is impoverished or unelaborated relative to that of those individuals developing in a normal environment in which models are available.

The second key characteristic of epigenesis is, accordingly, that a genetically specified developmental envelope or window specifies an initial behavioural (or perceptual) repertoire that is subsequently *elaborated* through experience of a relevant environment. This process of elaboration is directional (see below), and once it has taken place the initial plasticity of the embryonic, or unelaborated, repertoire is lost. A typical example is the development in human infancy of speech sound perception, in which the "universal" initial processor is transformed into a "language-specific" processor in a process that is probably analogous with that of the development of birdsong. We can note here that an epigenetic account of this process differs from a nativist, parameter-setting process inasmuch as no assumption is made that the infant brain is innately equipped with an inventory of all possible natural language phonemes. Equally, however, it differs from a classical learning account, inasmuch as epigenesis depends upon the elaboration of an initial repertoire which itself is not learned, in a process which cannot be re-run—the initial, unelaborated capacity cannot be re-accessed after the epigenetic developmental process has taken place, as all second language learners come rapidly to realise. In other words, the process of developmental elaboration implies in epigenetic development a transition from relative plasticity and informational openness to relative rigidity and informational closure.

Epigenesis is a developmental process whose genotypic distribution can itself be selected, through the standard Darwinian mechanisms of natural and sexual selection. The onset and closure of periods of plasticity and informational openness is under the control of regulatory genes, which are responsible for the timing and sequencing of all developmental processes from embryonic to mature organismic stages. Although all mammals display epigenetic features in various domains of development, especially during embryogenesis, the "weighting" of behavioural and cognitive development towards epigenetic processes seems to be markedly greater in more complex organisms than less complex ones. As an example, we can cite the well-known difference between the capacities of human-enculturated vs non-enculturated apes for symbol learning. The enhanced learning capacity of the enculturated apes must be due to their developmental environment, but the developmental environment would not make a difference if its relevant features were not available for assimilation into epigenetic construction processes (and indeed in lower non-human primate species no such differences have been observed). Epigenesis, then, seems to be a key

Lumsden and Wilson's "code-switching probability" genes are substantially different from the regulatory genes discussed below, and (b) their model continues to fix the stochastic space of possible developmental outcomes in the genes, rather than this itself being co-determined by genes and environment. Given these considerations, it could be said that the Lumsden-Wilson gene-culture coevolution model (intended as a substantial revision of standard sociobiological theory) severely circumscribes the role of epigenesis in such a way as to call into question whether it is really epigenetic at all.

mechanism in enabling individual organisms to acquire and exploit emergent complexity in phenogenotypic couplings. It is epigenesis, and in particular selectively *augmented* epigenesis, that serves as the crucial bridge between parallel processes of emergence in phylogeny and ontogeny. Epigenesis, I would suggest, is the key stabilizing process through which cultural and cognitive elaboration are emergently, dynamically and concertedly locked.⁵

Epigenesis may be analyzed at different levels, from the cellular to the cognitive. The neurobiological implementation of epigenetic processes at the cognitive and behavioural level is "Neural Darwinism", the selective stabilization of waves of periodically proliferating synaptic connections during ontogenesis (Changeux, 1985), which is, as we might expect, especially marked in the developing human organism, extending through adolescence and young adulthood. Why should humans be, more than any other, the "epigenetic species", and how is this related to humans being the "symbolic species" (Deacon, 1998)?

The answer to this question, I suggest, is that augmented epigenesis is advantageous for organisms in which phenogenotypic organism-niche couplings are both frequent and variable, which is a good enough general description of the human cultural organism. Regulatory genes augmenting epigenetic openness can therefore be expected to have been phenogenotypically selected for in the human genome, permitting further adaptive selection for domain-specific learning in the semiotic biocultural complex, in particular for language. Note, however, that in an epigenetic perspective, any developmental predisposition for learning language is unlikely either to involve direct coding of, or to be dedicated exclusively to, linguistic structure (Mueller, 1996).

The account I have offered revolves around the proposition that the evolutionary elaboration and epigenetic stabilization of the phenogenotypic semiosphere introduced the discontinuity characterizing both human culture and human cognition. Signs are both transformative cognitive tools, and constitutive of specifically human cultural ecologies. The semiotic capacity is hypothesized to have triggered transformative effects across all or most cognitive domains, thereby potentiating human symbolic cultures, which constitute the biocultural niche complexes in which human cultural innovation and transmission occur. The semiotic capacity is the explanatory link binding what is unique to human cognition with what is unique to human culture, bridging the biological with the social and human sciences in the evolutionary and developmental science of human cognition and language. It is to the social and semiotic ontology of language that I turn in the next section.

5 Language as a social fact and social institution

I begin by summarizing two theories of social ontology, classical and modern, separated in time by a century. These are the theories of, respectively, the sociologist Emile Durkheim (1895) and the philosopher John Searle (1995). Probably

⁵This is something of an oversimplification, since the stabilizing role accorded here to epigenesis also involves the canalization of phylogenesis through "Baldwin effects" and genetic assimilation (Sinha, 1988, p. 137-138). However, this detail is not crucial to the account presented here, and the precise status and nature of "Baldwin effects" is a strongly contested issue (Deacon, 2004).

not by coincidence, the theories employ the same terminology of social facts and social institutions, although Searle nowhere cites Durkheim.

Durkheim, a founding father of social theory, attempted a theoretical and methodological clarification of social science and its object. This object he stipulated to be the domain of *social facts*, which he described as "a category of facts which present very special characteristics: they consist of manners of acting, thinking, and feeling external to the individual, which are invested with a coercive power by virtue of which they exercise control over him." (Durkheim, 1982 [1895]).

Social facts, for Durkheim, are not merely aggregates of the individual cognitive representations of them by the subjects that are regulated or "coerced" by the social facts, since for each individual subject the social fact presents itself as a part of an out-there, already given *objective* reality. The objectivity of social facts consists, for Durkheim, in the fact they are independent of any single individual's thoughts or will. As Jones (1986, p. 61) puts it, "it is precisely this property of resistance to the action of individual wills which characterizes social facts. The most basic rule of all sociological method, Durkheim thus concluded, is to treat social facts as things." Durkheim's treatment of social facts consists therefore in, first, an ontological proposition, that social facts are irreducible to biological or psychological facts (or structures or processes); coupled with, second, an epistemological and methodological proposition regarding their treatment: as *objects* of a particular kind, whose determinate nature consists in their "coercion" of conduct.

Durkheim has often been criticized for the breadth and vagueness of his notion of "social fact". A particularly problematic aspect of his theory is that, in counterposing "social facts" to "individual conscience" (or mind), he sometimes identified the former with "states of the collective conscience". Some social psychologists (cf. Moscovici, 2000) have followed this direction in constructing a theory of "social representations", but critics have claimed that Durkheim sympathized with a view of society as a kind of super-organic "collective personality".

Whether Durkheim believed in a "collective mind" or not, such a notion is not only scientifically untenable, it is unnecessary. I propose that a social fact can most simply be defined as something regulating an aspect of conduct which requires the *participation* (Goodwin; Goodwin, 2004) of more than one individual. This "something" may be a codified law, a norm, an institution, a rule in the Wittgensteinian sense, or a canon of interpretation. A natural language, therefore, qualifies as a social fact (or institution) under this reading of Durkheim's theory. Social facts, for mature human beings, are objects of *common knowledge*; language is a prime example of this (Lewis, 1969; Ikonen, 1983; Clark, 1996). However, the social fact itself is not the sum, average or common denominator of all the individual beliefs of participants (since it is, indeed, the *object* of these beliefs). Social facts, in this sense, are in some way prior to individual cognitions about them. Yet it cannot be claimed that social facts are *independent* of cognitions, in this case of social cognitions, since their normative status is dependent upon agreement in cognition. We shall return to this paradoxical problem in discussing Searle's theory of social facts.

Social facts, for Durkheim, are constitutive of the domain of human social theory. Given that non-human species also display social behaviours, should we regard social facts as being uniquely human? Ethologists have pointed to

the evolutionary roots of norms, rules and conventions in the ritualized displays that many species exhibit in, for example, mating and agonistic displays. Ritualization, in turn, can be regarded as falling under the definition of a biocultural niche as discussed above. If so, we could argue that social facts are biocultural niches regulating and sustaining, supporting and constraining, the participatory behaviour of more than one individual. This definition is entirely consonant with Durkheim's view that social facts "consist of manners of acting, thinking, and feeling external to the individual, which are invested with a coercive power by virtue of which they exercise control over him."

Under this interpretation, social facts would be seen as no more unique to humans than culture. Yet there is clearly something unique about human social facts. This uniqueness consists, surely, in the way in which social facts are cognitively constructed as *objects* of intersubjective common knowledge (and common emotional investment), so that they can be *known* in the way in which the rules of football, the laws of the land, or a family history may be known.

To bring some order into the definitions employed here, I will stipulate that the concept of "social fact" pertains to those biocultural niches which are of a fully normative nature; that is, those which not only regulate behaviour, but are known to do so, and knowledge of which (whether explicit or tacit) is essential to their regulative status. Social facts, on this definition, can only be *constructed* by human beings with a certain level of cognitive development, although the institutions that they construct may be participated in by animals which lack this cognitive status (eg prelinguistic infants in language practices, racehorses in horse races etc). Social facts, then, constitute an emergent ontological level within the wider category of biocultural niches, and one which is uniquely human.

Searle (1995, p. 1-2) situates knowledge and belief at the heart of his account of social facts: "There are things that exist only because we believe them to exist. I am thinking of things like money, property, government, and marriages... [such] Institutional facts are so called because they depend upon human institutions for their existence." In an unfortunate inconsistency of terminology, Searle (1995, p. 26) regards "institutional facts", which seem to be more less equivalent to Durkheim's social facts, as a subset of what he (Searle) calls "social facts", which are basically all activities which involve participation in joint action: "I will henceforth use the expression 'social fact' to refer to any fact involving collective intentionality. So, for example, the fact that two people are going for a walk together is a social fact. A special subclass of social facts are institutional facts... for example, the fact that this piece of paper is a twenty dollar bill is an institutional fact."

Searle's account of social or institutional facts (such as money) is that they depend upon collective agreement and knowledge that, under determinate rules, something *counts* as an instance of a social object. Hence, the general form of such rules is:

- (1) "X counts as Y in context C" (Searle, 1995, p. 28).

Although he never uses the term, Searle's definition is in effect a semiotic one, in that the "counting as" relationship is one of *meaning* or *signification*. The twenty dollar bill, for example, signifies a certain monetary value or equivalence. However, the relationship between the bill and its monetary value is not a fully-fledged *sign relationship*. The bill does not *represent* or *stand for* twenty dollars:

it simply is twenty dollars, it is self-identical to its monetary exchange value. To clarify this difference, we can point out that the numeral 20 printed on the bill *stands for* (represents) the number twenty, but the bill itself does not represent, for example, twenty one dollar bills, but rather is *equivalent* to them in the value that it possesses, or counts as having.

Sinha (1988, p. 37) defines the pragmatic and semiotic conditions on *representation* as follows: "To represent something... is to cause something else to stand for it, in such a way that *both* the relationship of 'standing for', *and* that which is intended to be represented, can be recognized." (italics added). It must be emphasized that built into the conditions on representation is a duality of cognition, paralleling the duality of sign structure (the conventional unity of signifying substance and its signification). Two cognitions are necessitated: the recognition of the sign relationship, and the recognition of what is signified. The "counting as" relationship, by contrast, has no such duality: to know that something counts as a particular object, however abstract or complex that object may be, it is necessary only to recognize it as a token of that category of objects.

What is necessary to grasp the "counting as" relationship is knowledge of the rules and norms that constitute the category (for example money, or a language). In one fundamental (if limited) sense, then, knowledge of a language is definitionally knowledge of what counts as a token of the language, and in order to know this, the knowing subject must necessarily know (in some way and to some degree) the rules of the language. It is this level of knowledge that is considered to be primary in generativist and other formalist theories of language, which attempt to elucidate the rules that constitute the full range of tokens for which it is the case that:

(2) X counts as (a sentence) S in L (a language)

This definition does not, however, encompass in any way the *representational* function of language, its capacity to represent things (situations, events, actions, objects) outside of the formal context of L, that is, the world outside language. The knowledge constituting this *semantic* domain is governed, not by the "counting as" relationship and its conditions, but by the "standing for" relationship" and its conditions. This "standing for" relationship can be noted, in a way parallel with Searle's notation of the "counting as" relationship, as follows:

(3) S (a sign) stands for M (a message) in context C

However, the duality inherent in the conditions on representation (above) requires that this preliminary notation be expanded, to include knowledge on the part of the subject that S *counts* as a sign, or, more accurately, that a particular object counts as a signifier. This expansion yields:

(4) [X counts as S and S stands for M] in C
Where X is a token of the class of signifiers in C

(4) is sufficiently general to cover all cases of sign use, including highly idiosyncratic and context bound cases, such as non-conventional gestures. We can now undertake a further expansion to specify cases in which a given sign is part of a sign *system*, shared by a particular community of users:

(5) [X counts as S and S stands for M in C_s] for C_u

Where:

C_s = sign system

C_u = community of users

In the specific case of language, we can reduce the notion of a sign system shared by a community of users to the simple term L, language, thus:

(6) $L = C_s$ for C_u

Now any grammatical and meaningful instance of language use can be annotated:

(7) [X counts as S and S stands for M] in L

Note that, consistently with the approach of Cognitive Grammar (Langacker, 1987), S (the signifier) is an expression at any level, sub-lexical, lexical or constructional; grammatical assemblies of signs are also signs.

The definition offered in (7) can thus be considered to be the notational reduction of the broader theoretical approach to language taken by cognitive and functional semantically based theories, and indeed by all linguistic theories that include representational meaning in the linguistic theory. It is clearly a more inclusive definition than the formal-sentential definition (2), reproduced here:

(2) X counts as (a sentence) S in L

Definition (7) is also, quite simply, more psychologically complete than (2): what we usually mean by "knowing a language" is the knowledge of *both* what counts as a token of the language, *and* what it means. In the concluding section, I will attempt to elucidate further just what is, and is not, necessary for such knowledge.

Before doing so, I pursue this formal-notational exercise further by exploring how the conjoint definitions of "counting as" and "standing for" can be employed to define the sub-systems of language as traditionally employed in linguistic theory.

Grammar (in the wide, cognitive grammar sense, including lexical form and phonology) can be defined as:

(8) X counts as S in L

X is an instance of S, and S is a grammatical expression in L. The distinction between X and S is the distinction between, for example, phonetics and phonology.

Presupposing (8), semantics can be defined as:

(9) S stands for M in L

This is the relation between, for example, word form and lexical entry or concept; or, more generally, between linguistic expression and linguist conceptualization.

Presupposing (9), pragmatics can be defined as:

(10) S counts as A_s for Participants_(2...n) in C_d

Where:

A_s = Speech act (including reference)

C_d = Discourse context

Under this description, pragmatics is the closest of the linguistic subsystems to the "counting as" relationship. This accords with the intuition that pragmatics is not "systematic" in quite the same way as grammar and semantics: that speech acts are specifically linguistic instances of more general communicative acts (such as "threats" and "invitations" in both human and non-human species); and that their interpretation is strongly dependent on gesture, prosody, posture, physical and linguistic context.

Having employed the notational formalism to distinguish the subsystems of language one from another, we can now re-assemble them to analyze the structure of particular utterances in their context.

(11) [X counts as S & S stands for M] in L & S counts as A_s for Participants_(2...n) in C_d

Such a re-assembly does not yet account for the *interaction* between semantics, pragmatics, extra-linguistic context and shared world knowledge in actual utterances. For example, if the utterance is "You really did well this time!", and it is clear from the context that the speech act is one of ironic praise, the contextual meaning is "You did very badly". Or, if the utterance is "The road meanders up the hill", the contextual meaning is that the road has a winding path, not that the road is itself in motion (Talmy, 1996). How can we capture such facts of language?

It seems impossible to do so without appealing to psychological processes such inference, default and prototypic reasoning, subjectivization and perspectivization. This is the fundamental insight which drives cognitive linguistics. If we wish to formalize this, it would look something like this:

(12) [X counts as S & S stands for M] in L & S counts as A_s in C_d
⇒ S counts as (having) M_c for H in C_d

Where:

M_c = Contextual meaning

H = Hearer

This brings us back, in an intriguing hermeneutic circle, to Scale's original definition of a social fact, and emphasizes the truism that, in the end, all meaning is contextual and situated. This does not, however, mean the same as saying that there are no institutionalized, relatively stable, relatively autonomous and systematic social facts; indeed, it is precisely this very relative stability and autonomy which constitutes the objectivity of social facts emphasized by Durkheim.

This objectivity is not to be confused with the *objectivism* of formal, truth conditional semantics. Amongst the advantages of the simple notational definitions developed here are:

1. The account of semantic meaning is underdetermined by this formulation. The semantic theory need not be truth-functional, but is (necessarily) conventional and normative (as indeed are all the subsystems).
2. Semantics is distinguished from pragmatics without necessitating a truth functional semantics.
3. Contextual dependence characterizes all subsystems, as well as the interactions between them, but does not erase the distinctions between them.
4. Language as a social institution has its own proper structure which necessitates, but is irreducible to, the intentionality of its users. Language, like all social facts, is an objectification of intersubjectivity, with an emergent structure relatively autonomous from the intentional states (such as mutual knowledge of the language) which are possessed by its users and "subjects". It is in this fact, and this fact alone, that the objectivity of language inheres.

6 Conclusion: restricting the need to know

Biocultural niches are integral to the evolution of many species, including the human species. Radical nativist hypotheses of strong discontinuity between human and non-human genome are neither necessary nor plausible, if we view biocultural niches as constituents of phenogenotypic sites of Darwinian selection. The human semiosphere can be viewed as a species-specific biocultural niche, whose distinguishing feature consists in the elaboration and emergence of the semiotic function. This function, in turn, is constituted by the interplay and developmental interlacing of its two constituent semiotic relations, "counting as" and "standing for". It can be hypothesized that these two semiotic relations are evolutionary derivatives of, on the one hand, ritualization and, on the other, the evolution of symbols from signals (Sinha, 2004). In both of them, the conventionalization of intersubjective participation in niche-regulated activities plays a central role (Sinha; Rodríguez, in press). Current evolutionary biological theory, including niche-dependent evolution and epigenesis, accounts for the continuity between human and non-human culture and cognition. The evolution of the human semiosphere, in which language as a biocultural niche is developmentally and processually interdependent with other artefactual supports for human social interaction and social practice (Sinha, 2005), is what accounts for the discontinuity dividing human from non-human cognition and culture, and the evolutionary emergence of human social facts and social institutions. This discontinuity has been amplified by the consolidation, through language, of human culture as a fundamentally symbolic order.

From a biocultural perspective, the human language capacity, although it is almost certainly supported by genetic adaptations to maximize exploitation of

the human biocultural niche, is not innate, but epigenetically developed. Language as a social institution comes to be known by language-acquiring infants, but the knowledge required is not that of a grammar as a formal object divorced from its semiotic function. Although there can be no scientific objection to the study of language as a purely formal system, insistence on the disciplinary autonomy and full explanatory adequacy of formal theories leads to a distorted picture of the human language capacity, and to unnecessarily constrained theories of language acquisition. If "knowledge of language" is restricted to knowledge of what counts as a grammatical sentence, not only is language itself as a semiotic system truncated and reduced, but the process of its acquisition is rendered incomprehensible. To fill this conceptual vacuum, innate knowledge of Universal Grammar is invoked.

The biocultural theory of language and its acquisition restores, quite literally, life to language, for far too long reduced to formal structures and operations. It suggests a picture of "knowledge of language" that is both richer, in one sense, and poorer, in another, than that to which we have become accustomed from generative linguistics. It is richer because it incorporates meaning and context, the fundamental pillars supporting both language acquisition and language use. It is poorer because there is no longer a compelling reason to attribute a knowledge equivalent to the results of formal analysis to the learners and users of language. Simply stated, in the biocultural theory, *there is no mental grammar* isomorphic with autonomous grammar. Rather, grammar is in *language*, as a biocultural niche and social institution, just as the structure of the bowerbird's nest is in the nest as a biocultural niche. The learner need not internalize a formal description of the structure in order to acquire the ability to *act* in it. Language is not an "input" to a processor or device, but a structured niche affording complex and semiotically mediated communication and cognition. Grammar is a social institution, normatively regulating linguistic practice, and it is the practical ability to adhere to its constraints and supports that is acquired by the language learner.

7 Acknowledgments

I wish to thank my colleagues Alan Costall and Jörg Zinken for their valuable comments on parts of this article. I am also grateful for the stimulating environment provided by my colleagues on SEDSU (Stages in the Evolution and Development of Sign Use), a collaborative project supported by the European Union 6th Framework programme "What it Means to be Human" (Zlatev et al. 2006).

References

- [1] ALPER, J., AND LANGE, R. Lumsden-wilson theory of gene-culture cocvolution. In *Proceedings of the National Academy of Sciences* (1981), vol. 78, pp. 3976–3979.
- [2] BRITTEN, R. Divergence between samples of chimpanzee and human dna sequences is 5%, counting indels. In *Proceedings of the National Academy of Sciences* (2002), vol. 21, pp. 13633–13635.

- [3] CHANGEUX, J.-P. *Neuronal man: The biology of mind*. Oxford University Press, Oxford, 1985.
- [4] CHOMSKY, N. *Language and mind*. Brace & World, New York, Harcourt, 1968.
- [5] CHOMSKY, N. *The architecture of language*. Oxford University Press, New Delhi, 2000.
- [6] CLARK, H. H. *Using language*. Cambridge University Press, Cambridge, 1996.
- [7] COSTALL, A. From darwin to watson (and cognitivism) and back again: the principle of animal-environment mutuality. *Behavior & Philosophy*, 32 (2004), 179-195.
- [8] DAWKINS, R. *The selfish gene*. Oxford University Press, Oxford, 1976.
- [9] DE CONDILLAC, A., AND BONNOT, E. *Philosophical writings*. Lawrence Earlbaum, NJ, 1987. Translated by F. Philip. Hillsdale.
- [10] DE WAAL, F. *The ape and the sushi master*. Allen Lane, London, 2001.
- [11] DEACON, T. *The symbolic species: The co-evolution of language and the human brain*. Penguin, Harmondsworth, 1998.
- [12] DEACON, T. Culture and mind. In *Plenary lecture to the 1st International Conference on Language* (jul 2004), University of Portsmouth.
- [13] DESCARTES, R. *The philosophical works of Descartes*. Cambridge University Press, Cambridge, 1911, ch. Discourse on the method of rightly conducting the reason.
- [14] DURKHEIM, E. *The rules of sociological method and selected texts on Sociology and its method*. Macmillan, London and Basingstoke, 1982, ch. The rules of sociological method, pp. 29-163.
- [15] GIBSON, J. *The ecological approach to visual perception*. Houghton Mifflin, Boston, 1979.
- [16] GOODWIN, C., AND GOODWIN, M. H. *A companion to linguistic anthropology*. Blackwell, Oxford, 2004, ch. Participation, pp. 222-244.
- [17] ITKONEN, E. *Causality in Linguistic Theory*. Croom Helm, London, 1983.
- [18] JONES, R. A., AND DURKHEIM, E. *An introduction to four major works*. Sage Publications, Beverly Hills, CA, 1986.
- [19] LALAND, K. N., ODLING-SMEE, J., AND FELDMAN, M. W. Niche construction, biological evolution, and cultural change. *Behavioral and Brain Sciences* 23 (2000), 131-175.
- [20] LANE, H. *The wild boy of Aveyron*. Allen & Unwin, London, 1990.
- [21] LANGACKER, R. W. *Foundations of Cognitive Grammar*, vol. 1. Stanford University Press, Stanford, 1987. Theoretical prerequisites.

- [22] LEWIS, D. K. *Convention: A philosophical study*. Harvard University Press, Cambridge, MA, 1969.
- [23] LOTMAN, Y. *Universe of the mind: A semiotic theory of culture*. I.B. Tauris & Co. Ltd, London & New York, 1990.
- [24] LUMSDEN, C. J., AND WILSON, E. O. *Genes, mind and culture: The coevolutionary process*. Harvard University Press, Cambridge, MA, 1981.
- [25] MACWHINNEY, B. *The emergence of language*. Lawrence Earlbaum, Mahwah, NJ, 1999.
- [26] MALSON, L. *Wolf children and the wild boy of Aveyron*. NLB, London, 1982.
- [27] MARLER, P., AND PETERS, S. Developmental overproduction and selective attrition: New processes in the epigenesis of birdsong. *Developmental Psychobiology* 15 (1982), 369-378.
- [28] MCNEILL, D. *The conceptual basis of language*. Lawrence Earlbaum, Hillsdale, NJ, 1979.
- [29] MOSCOVICI, S. *Social representations*. Polity Press, Cambridge, 2000.
- [30] MUELLER, R. A. Innateness, autonomy, universality? neurobiological approaches to language. *Behavioral and Brain Sciences* 29 (1996), 611-675.
- [31] PIAGET, J. *La formation du symbole chez l'enfant*. Delachaux & Niestlé, Neuchatel, 1945.
- [32] PIAGET, J. *Behaviour and evolution*. Routledge & Kegan Paul, London, 1979.
- [33] PINKER, S. *The language instinct: How the mind creates language*. W. Morrow, New York, 1994.
- [34] SEARLE, J. *The construction of social reality*. Allen Lane, London, 1995.
- [35] SINHA, C. *Language and representation: A socio-naturalistic approach to human development*. Harvester-Wheatsheaf, Hemel Hempstead, 1988.
- [36] SINHA, C. *Evolution of communication systems: A comparative approach*. MIT Press, Cambridge, MA, 2004, ch. The evolution of language: From signals to symbols to system, pp. 217-235.
- [37] SINHA, C. Bleeding out of the background: Play, props and staging in the material world. *Journal of Pragmatics* 37 (2005), 1537-1554.
- [38] SINHA, C., AND RODRÍGUEZ, C. *The shared mind: Perspectives on intersubjectivity*. John Benjamins, Amsterdam, ch. Language and the signifying object. in press.
- [39] TALMY, L. *Language and Space*. MIT Press, Cambridge, MA, 1996, ch. Fictive motion in language and perception.

- [40] TOMASELLO, M. *The new Psychology of Language: Cognitive and functional approaches to language structure*. Erlbaum, Mahwah, 1998.
- [41] TOMASELLO, M. *The cultural origins of human cognition*. Harvard University Press, Cambridge, MA, 1999.
- [42] TOMASELLO, M. *Constructing a language: A usage-based theory of language acquisition*. Harvard University Press, Cambridge, MA, 2003.
- [43] VYGOTSKY, L. S. *Mind in society: the development of higher mental processes*. Harvard University Press, Cambridge, MA, 1978.
- [44] WADDINGTON, C. H. *The evolution of an evolutionist*. Edinburgh University Press, Edinburgh, 1975.
- [45] WHITEN, A., GOODALL, J., MCGREW, W. C., NISHIDA, T., REYNOLDS, V., SUGIYAMA, Y., TUTIN, C. E., WRANGHA, R. W., AND BOESCH, C. Cultures and chimpanzees. *Nature* 399 (1999), 682-685.
- [46] WILSON, E. O. *Consilience: The unity of knowledge*. Knopf, New York, 1998.
- [47] ZLATEV, J. The sedsu project. stages in the evolution and development of sign use (sedsu). In *6th International Conference on the Evolution of Language* (Rome, apr. 2006).

A Dança das linguagens na web: critérios para a definição de hipertexto

Antonio Carlos Xavier
Universidade Federal de Pernambuco/ Nehte

tonaz@uol.com.br

Abstract

The internet is perhaps the most revolutionary invention among the avalanche of technological innovations from the last thirty years. Sophisticated navigation programs allow a mix of languages (verbal, visual and sonorous) presented simultaneously to the users of the web. The objectives of this paper are two, in fact: to discuss a set of parameters to define hypertext and, thus, avoid confusion of terms, such as: assume hypertextuality as intertextuality, and consider the lack of linearity as an exclusive characteristic of the hypertext. The other objective is to point out the needs of the users to understand relatively well the rules of language games that makes it viable their productive incursion in the web. The hybrid enunciative nature of the hypertext permits a unique experience to the user that, before so many languages that dance on the computer screen, needs to learn how to play with such languages in order to avoid falling into the hypertextual labyrinth of the web.

1 Introdução

Talvez hoje não seja mais tão necessário convencer os acadêmicos e estudiosos das ciências humanas sobre a necessidade de desenvolver mais pesquisas que tenham como tema os novos comportamentos lingüísticos, cognitivos e sociais manifestos pelos usuários da Internet. Isso porque, depois de 10 anos da chegada da grande rede ao Brasil, a persistência heróica de alguns pesquisadores para levar adiante projetos sobre o tema e o aumento vertiginoso de usuários dessa mídia se tornaram fortes argumentos para superar a resistência de muitos intelectuais. O mal-estar que o tema causava em alguns estudiosos, inclusive lingüistas, diminuiu notadamente. O aumento da adesão de pesquisadores ao tema tem permitido um avanço considerável na produção de conhecimento sobre o que é e como tem sido a prática linguageira mediada pela Internet.

Não se pode mais negar o fato de que a Internet não é só mais uma nova mídia figurando entre as demais já criadas ao longo da história da inteligência humana. Ela é a mais revolucionária invenção dentro da avalanche de inovações tecnológicas das últimas três décadas, com reflexos mundiais.

Um estudo da ONU, intitulado *Digital Life 2006*, visando observar o impacto das tecnologias no comportamento das sociedades, constatou que a Web já é a mídia mais consumida entre usuários com até 54 anos de idade em todo o mundo.

Os dados foram divulgados em dezembro de 2006 pela União Internacional de Telecomunicações¹.

De acordo com o estudo, o aumento no consumo da mídia digital reflete a ampliação das conexões por banda larga, que quintuplicaram em apenas três anos, com destaque para o impulso da banda larga móvel em 2004. Já em 2006, o total de conexões rápidas em todo o mundo chegou a 277 milhões, com 166 países dispondo do serviço, sendo a maior concentração na Ásia (41,2%)² e na América (32,2%). No Brasil, a entrada da banda larga aumentou para 20 horas semanais o tempo médio de navegação dos internautas brasileiros.

A pesquisa revelou ainda que houve mudanças significativas no perfil de uso de computadores – ao invés dos antigos computadores centrais, a partir dos quais diversos usuários usavam a mesma máquina, o setor de Tecnologia de Informação atualmente chega a registrar diversos microcomputadores para apenas um usuário. Já são mais de 100 bilhões de páginas indexadas com mais de 170 países em rede.

Todavia, o objetivo deste ensaio não é o de convencer e estimular lingüistas e educadores em geral a pensarem em projetos de pesquisa sobre o tema linguagem e Internet. Isso é fato, é preciso fazer tais projetos o quanto antes, pois ignorá-los é correr o risco de cair na obsolescência tecnocientífica. Este trabalho busca, pois, apresentar e discutir parâmetros para definir hipertexto, a fim de percebê-lo como tecnologia enunciativa fundamental pelo qual passam as diversas linguagens humanas criadas ao longo do tempo, bem como pretendemos alertar a comunidade científica para as implicações cognitivas, sociais e pedagógicas do modo de enunciação digital que emerge no hipertexto.

É também nossa pretensão refutar as imprecisões terminológicas que têm surgido ultimamente em alguns trabalhos publicados em revistas e livros brasileiros. Alguns equívocos têm se instaurado, em geral, por razões teóricas bem específicas. São elas: 1) a confusão que alguns acadêmicos têm feito entre o conceito de hipertextualidade com o de intertextualidade; 2) o superdimensionamento do fenômeno da não-linearidade ou deslinearidade conferida ao hipertexto por alguns estudiosos mais entusiastas com o tema, como se essa fosse uma de suas características exclusivas e não uma estratégia intrínseca ao processamento cognitivo da leitura em qualquer suporte. 3) Há também quem postule que todo texto é um hipertexto em oposição aos que defendem que todo hipertexto pode ser textualizado, mas nem todo texto é um hipertexto. 4) Hipertexto seria, na opinião de alguns, apenas uma nova forma de circulação de textos; e por último, 5) encontra-se quem diga que a maior vantagem do hipertexto ainda é obter um melhor conhecimento do que seja um texto.

As razões que sustentam essas posições serão aqui debatidas. Antes, faz-se necessário explicitar as bases técnicas e epistemológicas que dão suporte aos posicionamentos adotados neste trabalho.

¹<http://computerworld.uol.com.br/comunicacoes/2006/12/04/idgnoticia.2006-12-04.2620009640/IDGNoticia.view>

²Em Hong Kong é possível acessar a rede a 1 gigabyte por segundo (1,04 milhão de Kbps), o suficiente para baixar um dvd em 30 segundos. A banda larga poderá ser substituída por fibra ótica. Ela comprime bits e bytes em microtubos de vidro, na forma de ondas de luz. O sistema de conexões óticas já tem 3 milhões de usuários em 7 países (EUA, França, Itália, Holanda, Suécia, Japão e Coréia). A velocidade, por enquanto, fica em 100 Mbps (12 vezes mais que a melhor conexão no Brasil que é de 8 Mbps).

2 As Bases Epistemológicas

Como se sabe, as línguas naturais são heterogêneas, variadas, variáveis, indeterminadas, sociais, históricas, situadas e interativas. Essas características tornam-nas fundamentais na constituição das sociedades, pois os indivíduos delas dependem para compreender sua condição de ser no mundo e estabelecer sua relação com o outro de modo satisfatório para todos os envolvidos no processo interacional. Enquanto ferramentas de mediação entre interlocutores, as línguas naturais permitem que a verbalização de idéias bem como sua socialização no espaço e no tempo sejam realizadas na modalidade oral ou ancoradas graficamente em suportes de papiro, pergaminho, papel ou página digital.

No entanto, a construção de sentido, nas línguas naturais, não se dá de forma direta ou transparente, mas de modo complexo. Esse laborioso processo leva em consideração uma atmosfera situacional permeada de fatores, tais como, tempo e lugar, bem como outros constituintes envolvidos diretamente na produção e interpretação do dito que dizem respeito a quem diz, o que diz, como diz e para que diz o que quis dizer.

2.1 Do jogo da linguagem para “o jogo com as linguagens”

Na segunda fase de suas reflexões filosóficas sobre a natureza da linguagem, Wittgenstein passou a postular a função constitutiva da linguagem natural no raciocínio humano e nas formas de sua existência e organização social. Para Wittgenstein (1953), o mais fascinante da linguagem é sua capacidade plástica de mudar de sentido conforme o uso que dela se fizer. Ele afirma que: “As palavras só adquirem significado no fluxo da vida: o signo, considerado separadamente do fluxo de suas aplicações, parece morto, sendo o uso seu sopro vital” (1953, p. 43). Segundo ele, o “significado” de um signo é a soma das regras que determinam seus “lances” possíveis. Defende que não cabe aos filósofos da linguagem indagar sobre os significados das palavras, mas sobre suas funções práticas que são muitas e variadas.

Assim, diante dessa pluralidade de funções que a linguagem pode assumir na vida cotidiana, pareceu razoável ao filósofo vienense conceber o funcionamento da linguagem comparativamente ao dos jogos. O ponto em comum para a analogia é a existência de regras que guiam ambas as atividades. A linguagem seria, então, constituída por uma série de regras como há nos jogos. A exemplo do jogo de xadrez, na linguagem, cada lance é calculado pelo falante e executado com as palavras. A expressão “jogos de linguagem” foi, em geral, empregada como equivalente a “cálculo”, que serve para o locutor realizar várias ações humanas (como perguntar, asseverar, consolar, indignar-se, descrever, etc) a serem percebidas enquanto tal pelos interlocutores.

O autor de *Investigações Filosóficas* postula que a práxis da linguagem de um dado locutor, que se expressa também pelo tom (entoação) em que uma frase é dita, informa ao interlocutor como ela quer e deve ser entendida. As inúmeras espécies de signos existentes indicam a pluralidade e dinamicidade da linguagem que a cada dia ganha novos tipos e formas por serem diferentes os usos e propósitos que se fazem dela.

Enfim, com a idéia de “jogos de linguagem”, Wittgenstein procurou sublinhar a natureza heterogênea das línguas naturais, percepção que também pode ser estendida às outras linguagens, pois os processos semióticos (signicos ou

simbólicos) exigem que os interlocutores estejam sempre calculando os lances realizados no processo comunicacional, seja em que forma de linguagem for.

Na esteira dessa heterogeneidade lingüística e mesmo consciente de que o termo wittgensteiniano estava remetido ao âmbito da linguagem verbal, o conceito de "jogo de linguagem" é aqui ampliado para o "jogo com as linguagens", a fim de dar conta do emaranho de semioses implicado no hipertexto. Da forma como se apresenta, o hipertexto joga com toda a riqueza da linguagem verbal e, ao mesmo tempo, agrega a capacidade signíca das outras linguagens paralelamente. É na perspectiva da flexibilidade das linguagens em geral para a produção de efeitos de sentidos que se deve ver o hipertexto e seu funcionamento integrador de diversas semioses que nele convergem.

Se a linguagem verbal possui a adequada maleabilidade para produzir significação quando posta em uso real, em situação viva, ela não deve se comportar de modo diferente de quando acompanhada por outras linguagens, como imagens e efeitos sonoros. Pelo contrário, o verbal nesse espaço de co-ocorrência de linguagens deve flexibilizar-se ainda mais, pois não é o único modo enunciativo responsável pela significação, mas divide essa tarefa com as demais linguagens que se lhe interpõem.

Assim, o ambiente e as condições de possibilidade para a ocorrência dos "jogos de linguagem" de forma muito mais ampla e diversa vêm ao encontro da arquitetura organizacional do hipertexto, pois a fusão de linguagens, garantida pelos recursos tecnológicos hipermediáticos, multiplica as possibilidades de lances intencionais dotados de significação, por permitirem aos sujeitos jogar com as linguagens para produzir e acessar as informações discursivizadas na tela do computador.

Portanto, com a perspectiva wittgensteiniana, segundo a qual a linguagem se constitui por lances realizados pelos jogadores usuários da língua e para quem a significação é o uso da linguagem em contexto, o presente trabalho propõe que se conceba o hipertexto como uma tecnologia enunciativa em que não apenas a linguagem verbal produz sentido mas as linguagens visual e sonora também contribuem para tal. Cabe ao usuário jogar com as diversas linguagens, para delas extrair significação que satisfaça seu objetivo.

3 As Bases Tecnológicas

Como rede de computadores interligados, a Internet esteve restrita por três décadas a alguns poucos privilegiados que dominavam o funcionamento das máquinas, não tão amistosas como as que usamos hoje. O trânsito de informações era pequeno em comparação ao atual, e o serviço mais utilizado era o correio eletrônico entre centros acadêmicos e militares de pesquisa. Os *experts* em informática trocavam mensagens e compartilhavam informações de modo muito particular. Só a partir da década de 1990 é que foram criados os protocolos de marcação e transferência de dados, de modo a tornar seu uso mais fácil a não-iniciados em telemática.

Os programas de navegação permitiram que pudesse vir à tona toda a potencialidade comunicativa e interacional da rede, a partir de então, disponível a um grande número de pessoas. Entretanto, tais programas de navegação, utilizando

códigos e algoritmos próprios, geraram a possibilidade de mesclar as três principais formas de expressão desenvolvidas pela humanidade para manifestar-se, a saber, a verbal, a visual e a sonora.

Na tela do computador, é possível ter acesso não apenas a textos verbais, mas também a informações em imagens estáticas e dinâmicas, e ouvir, por meio de caixas de som acopladas, a sonoridade que foi ancorada em uma "página digital". Acessam-se agora os três modos enunciativos simultaneamente. É verdade que a junção de formas enunciativas não é uma novidade, pois os suportes em celulose sempre uniram texto verbal a imagens. O novo está em adicionar o som e as imagens com movimento tanto por meio de animação de desenhos quanto por meio de vídeos inteiros.

A "página digital" conhecida hoje nasceu da idéia inicial de construir uma máquina capaz de armazenar todo o conhecimento produzido pela humanidade desde a invenção da escrita até o presente momento. Esse equipamento, chamado de *Memex*, além de armazenar conhecimento, também deveria permitir o compartilhamento desses saberes de forma democrática para uma quantidade infinita de habitantes. Quem primeiro imaginou a criação dessa supermáquina foi o americano Vannevar Bush no ano de 1945, mas não logrou êxito em razão de uma série de limitações técnicas daquele momento.

A idéia foi reativada na década de 1960 por Theodore Nelson no intitulado *Projeto Xanadu*. Foi ele quem cunhou oportunamente o termo HIPERTEXTO. Com ele, Nelson buscava exprimir a superdimensão de *locus* onde textos poderiam ser reunidos e lidos por quaisquer pessoas, as quais poderiam adicionar outros e vinculá-los entre si de forma que esse ambiente crescesse *ad infinitum*. Todo o *thesaurus* de saberes já produzido e ainda por produzir poderia ficar à mão dos homens, condição que poderia os tornar mais informados e fazê-los vivenciar o sentido da cidadania do modo mais amplo.

Nas palavras de Nelson (1993, p. 57):

Nosso objetivo no Projeto Xanadu não foi suprir a necessidade da indústria ou fazer coisas acontecerem um pouco mais rápido ou de modo mais eficiente. Temos apenas um objetivo justo: fazer um novo mundo... A abertura do hipertexto manifesta publicamente o destino de uma sociedade livre, que é desejável, poderosa e que está a caminho.

A pretensão expressa nas palavras de Nelson reflete o sonho da civilização pós-Renascimento, já "iluminada", que acreditava na possibilidade de todos os homens terem acesso irrestrito ao mundo das idéias e à riqueza cultural e científica produzida na Terra. Esse desejo contou com a militância aguerrida dos enciclopedistas Diderot e D'Alambert, dos filósofos Kant e Condorcet e outros intelectuais pouco mais de três séculos atrás.

Mas foi com Douglas Engelbart que o projeto da rede mundial de computadores decolou efetivamente, pois criou as atuais janelas de interface e idealizou o mouse, periférico que nos permite gerenciar o interior da página eletrônica. Engelbart desenvolveu o NLS (oN Line System), um sistema baseado em hipertexto, no qual textos, imagens e vídeos eram apresentados de forma conjunta e interativamente. Foi esse sistema que permitiu ao usuário acessar todos os elementos semióticos de modo não-linear e colaborativo. Sua comercialização, no entanto, só começou a acontecer no início dos anos 1980.

Em geral, cientistas e engenheiros de computação têm como objetivo fundamental construir linguagens e algoritmos que possam viabilizar comportamentos computacionais. Acreditam que a união e o compartilhamento de informações entre os indivíduos via rede de computadores podem suprir a incompletude de conhecimentos dos indivíduos, e assim solucionar problemas com base em volumes maciços de informação distribuída em rede.

4 Definindo Hipertexto

Apresentaremos o conceito de hipertexto sob dois aspectos: técnico-informático e enunciativo, pois é preciso refazer seu percurso desde sua concepção original imaginada pelo cunhador do termo e outros pesquisadores antes e depois dele para, dessa maneira, compreendê-lo como tecnologia enunciativa que abriga o *modo de enunciação digital*.

4.1 Ponto de vista técnico-informático

Sob essa perspectiva, é interessante notar que o hipertexto é produto da combinação de programas computacionais baseados em um sistema complexo de codificação de algoritmos, os quais obedecem a uma rígida lógica procedural. Por meio dessa programação complexa, foi possível dispor à percepção humana textos, imagens estáticas e dinâmicas e sons simultaneamente, o que gerou a demanda por máquinas ordenadoras com capacidade para suportar várias mídias. Por essa razão, essas poderosas máquinas passaram a ser denominados equipamentos *multimídias*. Tais aparelhos trabalham com estratégias de combinação de bytes e agrega-os em seqüência digital, utilizando para tal apenas os números 0 e 1.

A intrincada sintaxe dos programas hipertextuais camufla-se aos olhos do usuário comum e, ao mesmo tempo, permite que se realizem explorações a partir da tela do computador ou de outro aparelho multimídia, de uma maneira totalmente inusitada em relação as outras superfícies de exploração perceptual, como a página analógica de um livro impresso, por exemplo. Com base nessa criação tecnológica, podemos afirmar que o hipertexto é um *constructo pluri-enunciativo produzido e processado prioritariamente na tela de um aparelho multimídia*.

Theodore Nelson (1993) apresenta o conceito de hipertexto sob dois aspectos:

1. Trata-se de um conceito unificado de idéias e de dados interconectados de maneira que podem ser editados no computador.
2. Uma instância com a qual se podem religar as idéias e os dados, evidenciando a dupla vocação do hipertexto: um sistema de organização de dados e um modo de pensar.

Em outras palavras, Nelson defende, de um lado, a capacidade e funcionalidade de um sistema organizador que arquiva e distribui dados e, de outro, um sistema cuja vocação seja a de interligar as idéias articuladamente, a exemplo de como funcionam os neurônios e suas sinapses na mente humana. A articulação das idéias é feita pelos links que as "amarram" e indicam ao usuário outros lugares na rede onde aquela palavra ou até mesmo um tema inteiro também está em discussão.

No início dos anos 1990, o físico britânico Tim Banners-Lee criou no CERN (Laboratório Europeu de Física de Partícula) a interface gráfica da Internet. Ele desenvolveu a WWW (World Wide Web) ou teia de alcance mundial. Esse programa facilitou o acesso à rede dos usuários em geral. A web é, na verdade, uma aplicação dos recursos da hipermídia à Internet. *Off-line*, a hipermídia permite a mixagem de diferentes mídias, tais como, arquivos de texto, de imagens, de sons e de vídeos. Computadores e aparelhos de telefones celulares são os equipamentos digitais mais conhecidos em cuja tela são vistas as aplicações da hipermídia. Banners-Lee conseguiu elaborar um programa que viabilizou a navegação por várias mídias vinculadas não apenas à mesma máquina, mas a toda a rede, desde que os equipamentos estejam *on-line*.

Para o criador da web, ela alterou a maneira como os cientistas se comunicam, colaboram uns com os outros e como transferem seus conhecimentos. Mais de quinze anos depois de criada, a web se expandiu exponencialmente, de modo a levar seu criador e outros pesquisadores do MIT (Instituto de Tecnologia de Massachusetts) e das Universidades de Southampton e de Maryland a sugerirem a criação de uma nova ciência que nomearam de **Ciência da Web**. O artigo com a proposta foi publicado na edição de maio de 2006 da revista "Science" (www.sciencemag.org).

O objetivo desta nova vertente científica, interdisciplinar por natureza, é entender e dar suporte ao crescimento da web. Na proposta de criação da Ciência da Web, os cientistas focalizam o estudo da arquitetura da rede a partir da complexidade de sua interface. Querem se certificar de que a web "apóia os valores sociais básicos de segurança, confiabilidade, privacidade e respeito aos limites sociais".

Um outro projeto importante que está contemplado no interior dos estudos da Ciência da Web é a *Semântica da Web*, também proposta por Banners-Lee e colaboradores. Em linhas gerais, eles esperam com ela produzir mecanismos que facilitem o cruzamento de informações arquivadas de diferentes maneiras na rede, independentemente do tipo de sistema de programação em que os dados estejam originalmente armazenados.

Para atingir esse objetivo, os pesquisadores esperam desenvolver técnicas científicas e equações matemáticas poderosas ligadas a várias disciplinas, considerando as características macro e microscópicas da web e suas conseqüências nas interações humanas em rede.

Enfim, podemos verificar que o avanço tecnológico possibilitou a utilização de programas computacionais diversos, cuja síntese é a babel semiótica atualizável na tela de equipamentos multimidiáticos da nossa contemporaneidade. Além do mais, há muitos outros recursos telemáticos em vias de desenvolvimento e implementação que poderão promover mudanças nas formas de processamento lingüístico-cognitivo da informação fartamente disponível no espaço digital.

4.2 Ponto de vista enunciativo

Constatados os dispositivos que governam a estrutura informacional capazes de mesclar os diversos recursos semióticos que constituem o hipertexto, interessamos identificar e compreender como esse mosaico de linguagens funciona e quais suas implicações para a construção de sentido dos discursos ali digitalizados.

Como já dissemos, o hipertexto permite a fusão de várias linguagens no mesmo espaço de percepção visual e sonora. Essa superposição de linguagens

torna sua abordagem muito mais desafiadora, pois o usuário terá que processar formas enunciativas diferentes a cada abordagem do mesmo hipertexto que, paradoxalmente, nunca é o mesmo.

Portanto, o hipertexto on-line é a "tecnologia enunciativa que viabiliza a surgimento do modo de enunciação digital, uma nova forma de produzir, acessar e interpretar informações" (2002, p. 97) expostas no *suporte* de percepção, a tela ou monitor de um equipamento multimídia.

Estamos denominando *modos de enunciação* todas as formas de semiotização de signos originalmente elaboradas para expressar idéias, manifestar desejos e intenções latentes à mente humana. Referimo-nos especificamente às linguagens naturais e artificiais, como sistemas de representação de conhecimentos, utilizadas há séculos nas sociedades civilizadas com propósitos sócio-comunicativos. As línguas naturais baseadas no *logos*, na palavra oral e escrita, as baseadas em pictogramas, símbolos e ícones já consagrados e as imagens fixas (desenho, pintura, figura, foto) ou movimentadas (animação gráfica, vídeos), assim como sons e efeitos sonoros intencionais, são aqui chamados de *modos de enunciação*. Em síntese, são as várias "linguagens criadas, socialmente convencionalizadas e pragmaticamente reproduzidas em contextos situacionais adequados nas diferentes esferas sociais." (idem)

Sendo assim, definimos enunciativamente o hipertexto como um espaço virtual singular que apresenta, reapresenta e articula os recursos lingüísticos e semióticos já em circulação centrados num só lugar de acesso perceptual. Não se trata de um novo gênero de discurso, mas de uma forma outra de dispor e compor entrelaçadamente as informações expostas em diferentes linguagens. Cada linguagem que se ancora no hipertexto guarda suas peculiaridades sógnicas, mas ao mesmo tempo cede a primazia de significação para que possa cooperar com o propósito principal que é a construção do sentido pretendido pelo sujeito-enunciador do espaço virtual.

O hipertexto, como aplicação da hipermídia ou multimídia on-line, cria as condições de possibilidade para equilibrar, no mesmo locus enunciativo, os diversos modos de enunciação que co-atuam sinergicamente para a constituição da inteligibilidade dos discursos dispostos em rede. Nessa perspectiva, o hipertexto postula a seguinte equação enunciativa (Xavier, 2002, p. 110):

TEXTO + IMAGEM + SOM = HIPERTEXTO → MODO DE ENUNCIÇÃO DIGITAL

5 A Dança das linguagens na web

Eis, portanto, a razão para a idéia de que as linguagens estejam bailando na web, ou seja, que os diversos modos de enunciação estejam dançando multisensorialmente no hipertexto on-line. O espaço que ocupa na tela paralelamente às outras, a forma perceptual, cores, traços e nuances de sonoridade com os quais se apresenta dão o "tom" e a relevância que cada modo enunciativo possui para a construção dos efeitos de sentido pretendidos na elaboração de um hipertexto, embora não garantam que sejam exatamente recuperados pelo usuário durante a leitura-navegação. Por princípio, o hipertexto é uma tecnologia enunciativa aberta, que oferece ao usuário escolhas de caminhos a trilhar conforme suas vontades e necessidades individuais.

Essa simultaneidade de linguagens sem posição fixa na página *web*, uma vez que cabe ao produtor-enunciador fazer as modificações que desejar nos sítios e links criados a qualquer tempo sem ter que justificá-las para seus consumidores, confere a noção da necessidade da oscilação e balanço que tais linguagens sofrem no espaço virtual. Saber jogar com elas para cá e para lá na superfície hipertextual é um desafio constante a ser enfrentado pelo enunciador, pois tem como objetivo fazer o hiperleitor entrar na dança virtual proposta na arquitetura da página *web* inicialmente planejada.

Já que dança é jogo de sedução, o enunciador joga com as diversas linguagens a fim de seduzir seu co-enunciador a dançar com ele virtualmente, durante a navegação pela página. A dança das linguagens na *web* é, portanto, um constante jogar do produtor do hipertexto que sugere certo ritmo ao dançante, usuário do site. Este, por sua vez, não tem obrigação de segui-lo, pois diferentemente das danças em geral, que têm uma seqüência mais ou menos fixa de movimentos corporais a serem executados, na *web*, o dançante é quem dá o ritmo, é quem escolhe o andamento e os passos de acordo com a coreografia que melhor lhe aprouver pelos links que clicar.

6 Critérios para definição de hipertexto

Os passos dessa dança não estão pré-determinados por nenhum coreógrafo específico; estão temporariamente propostos na arquitetura organizacional do hipertexto pelo autor da página. Cabe ao dançante (hiperleitor) perceber a coreografia indicada pelo enunciador e também criar outra, na medida em que desejar entrar no jogo com as linguagens que constituem o hipertexto.

Os dispositivos digitais que dão *imaterialidade*, *ubiquidade*, (*hiper*) *inter textualidade* e *multisemiótica* ao hipertexto permitem uma grande mobilidade ao usuário que poderá acessá-los de qualquer ponto da página, sem compromisso com qualquer linearidade que lhe obrigue a abordá-lo de cima para baixo ou da esquerda para direita, por exemplo. Ele tem a liberdade de sair virtualmente da página *web* inicial e chegar à outra, armazenada em um servidor distante fisicamente milhares de quilômetros de onde partiu. São essas idas e vindas, esses movimentos sem ritmo determinado que permitem ao hiperleitor mover-se livremente pela página *web*, assim como dançam as diversas linguagens que compõem o espaço digital.

Essa dança de linguagens na *web* só é possível graças às características do hipertexto supracitadas. Assim, a afirmação que assegura que “– pelo menos do ponto de vista da recepção – todo texto é um hipertexto” (Koch, 2002, p. 61) não se sustenta integralmente, quando cotejada com algumas das propriedades do hipertexto *on-line*. Observemos, a título de exemplo, duas dessas características: *imaterialidade* e *ubiquidade*. A primeira característica só se aplica parcialmente, enquanto a outra não se aplica sob hipótese alguma ao texto. Este nasce na dependência de um suporte material que lhe dê forma e concretude. Recordemos das tábuas de argila, dos papiros e pergaminhos que antecederam o papel derivado das árvores; todos esses são espaços de textualização que materializam as idéias verbalizadas oralmente mais suscetíveis ao vento e à fragilidade da memória humana.

Um texto produzido em um processador eletrônico apresenta flexibilidade de edição, ou seja, pode-se copiar, cortar, colar porções de um texto e inserir outras com muito mais facilidade e agilidade que se escrito à mão ou em máquinas de escrever, mecânicas ou elétricas. Entretanto, o texto estará na eminência de ser impresso, de ser manuseado, tocado. Logo, podemos dizer que texto escrito no computador *off-line* pode ser denominado de *texto eletrônico*. Todavia, nem todo texto eletrônico é um hipertexto, mas todo hipertexto é um texto eletrônico, i.e., todo hipertexto precisa ser escrito na tela do computador e deve obedecer às regras de programação em *html* ou qualquer outro programa de construção de páginas para a web. Nem todo texto eletrônico é indexado à rede, nem sempre compartilha a produção do sentido com outras mídias (imagens com movimento e sons). Há, portanto, uma notória diferença entre texto verbal produzido no computador e hipertexto que mescla modos enunciativos. O primeiro é parte do todo que é o segundo.

Por outro lado, um texto, ainda que eletrônico, se não estiver *on-line*, jamais poderá gozar das vantagens da *ubiquidade*, ou seja, não poderá ser acessado por diferentes usuários em vários lugares do planeta onde houver um pc conectado. Essa propriedade potencializa o acesso a um certo documento de modo simultâneo a várias pessoas, algo impensável às edições impressas de um livro.

A ressalva entre travessões que antecede a afirmação citada de que *todo texto é um hipertexto*, qual seja, *pelo menos do ponto de vista da recepção*, modaliza estrategicamente o dito conferindo-lhe uma verdade parcial. É fato que a leitura de um discurso em qualquer suporte, inclusive na tela do computador, pode ser feita a partir de qualquer ponto da superfície de acesso. Um livro, embora produzido com certa ordenação em suas partes cujo autor achou importantes, pode ser consumido do fim para o começo, do meio para suas extremidades, bem como suas notas de rodapé ou índices remissivos podem nunca ser utilizados. Fazê-lo é uma decisão exclusiva do leitor. O raciocínio deslinear é próprio do leitor e não do modo enunciativo em que o autor escolhe para expressar-se.

A novidade nesse aspecto está em transformar a não-linearidade, a falta de um ponto de partida obrigatório para o início da leitura, em princípio básico da construção do próprio hipertexto. Dizendo de uma outra maneira, o diferencial proporcionado por essa tecnologia enunciativa é o de assumir a deslinearidade já na sua arquitetura inicial e produzi-lo com vários links que podem levar seus usuários a percorrer diferentes estradas, de forma que um dado hipertexto jamais será navegado duas vezes da mesma maneira ainda que feito pelo mesmo hiperleitor.

Que fique claro, pois, que a deslinearidade é reivindicada como novidade enquanto princípio de produção do hipertexto e não de sua recepção. Barthes, bem antes da chegada dessa tecnologia enunciativa, já anunciava isso para a abordagem pelo leitor dos livros impressos.

Gostaríamos de explicitar um outro equívoco comum que tem andado em artigos científicos de alguns linguistas: tomar intertextualidade por hipertextualidade. De acordo com a definição de hipertexto que temos assumido neste trabalho, a intertextualidade é uma das características constitutivas da idéia inicial de um sistema interligado de textos acessíveis a todos os cidadãos do mundo. Certamente esse princípio deriva da concepção de texto como rede interna de enunciados que formam um tecido único, embora eventualmente remeta-se a outros fora dele. Mas essa remissão é textual, os trechos citados explicitamente não aparecem na tela após o toque do dedo do leitor sobre o fragmento entre aspas.

No hipertexto isto é possível, a intertextualidade é executada “concretamente” pelo leitor, pois a remissão é feita de modo direto e instantâneo, bastando apenas um clicar de mouse sobre o link desejado para se ter acesso ao hipertexto indexado à web.

A tecnologia enunciativa do hipertexto acentua a função da intertextualidade na construção e até na desconstrução de sentido de um discurso *on-line*, pois os links dispostos na página web abrem a possibilidade para o encadeamento do que chamamos de *hipertextualidade* que enovela o hiperleitor até o infinito dos documentos já indexados.

Convém salientar que a hipertextualidade é uma tecnologia enunciativa que funciona principalmente com base no princípio da intertextualidade, mas que se distingue desta, posto que é mais abrangente por absorver e articular outras tecnologias enunciativas, que dialogam diretamente entre si para fazer funcionar o modo de enunciação digital.

Essa *hipertextualidade* intrínseca ao hipertexto (Xavier, 2003) representa um indiscutível ganho para as sociedades, pois faz circular um número maior de discursos. Entretanto, não concordamos com a afirmação reducionista de que o hipertexto seja apenas “uma nova forma de circulação de textos” (Possenti, 2002, p. 221).

A afirmação acima é reducionista por duas razões; por um lado, ignora as vantagens da ubiqüidade de discursos indexados à rede, condição técnica que contribui inquestionavelmente para a democratização do conhecimento, fazendo com que muito mais pessoas leiam, e “que, por exemplo, em países de censura férrea, a liberdade maior de ler pode produzir não só novos sentidos, mas também mais liberdade política e de costumes” (idem); e, por outro, desconsidera os comprovados efeitos cognitivos, quando da recepção de informações *on-line*, provocados pela convergência (dança) das linguagens, viabilizada pelos recursos multimídias que compõem o hipertexto.

Dizer que “a maior vantagem do hipertexto ainda é nos fazer entender melhor o que é um texto” (idem, p. 217) demonstra um grande desconhecimento sobre o potencial enunciativo, intelectual e social desta tecnologia digital, além de deixar nas entrelinhas a crença no texto verbal escrito como o limite para a expressividade dos sujeitos de discursos. Considerando o notável salto tecnológico das últimas décadas, afirmações assim soam resistentemente nostálgicas, exceto por aqueles que teimam em colocar a cabeça no buraco para fugir à realidade contemporânea. É provável que o advérbio “ainda” que aparece na citação signifique o quão pouco o autor conhece aquilo sobre o que se arrisca a comentar. Talvez, aprendendo um pouco mais sobre a natureza multimidiática dos hipertextos, ele perceba mais claramente o quanto eles se diferem dos textos e os ampliam.

7 Considerações Semifinais

Toda essa reflexão pretende, em última análise, defender a natureza enunciativa híbrida e, por isso, inédita do hipertexto *on-line*. Certamente, a bricolagem de linguagens na mesma superfície de percepção – a tela do computador ou o aparelho celular – exige de seu usuário um outro comportamento cognitivo, pede-lhe estratégias perceptuais um pouco diferentes das que são normalmente realizadas durante o processamento da leitura em textos impressos.

Uma grande quantidade de pesquisas em diferentes países e nas mais diversas ciências vem sendo feita com o objetivo de compreender como acontece o processamento em hipertextos, quais as táticas cognitivas dos hiperleitores e que fatores auxiliam ou determinam as decisões que são tomadas durante o contato com essa tecnologia enunciativa.

Estamos convencidos de que a chave para abordar as diversas semioses que se espraiam pelo hipertexto é, sem dúvida, a apropriação de pelo menos parte das “regras” que regem o “jogo de linguagem” referido por Wittgenstein, claro que ampliada para lidar não só com a verbal, mas também com a visual e a sonora. Quais são essas regras e como as usamos efetivamente ainda não foi possível determinar.

As táticas para a construção de sentido no hipertexto são desenvolvidas à medida que os sujeitos lidam com ele, que dele se tornam íntimos, descobrindo-lhes os elos e atalhos. Diante de tantas possibilidades de lugares a explorar na esfera digital, o estranhamento inicial é esperado; a familiarização com as regras desse *jogo com linguagens* nessa nova tecnologia enunciativa virá paulatinamente. No salão virtual da dança das linguagens, é importante aprender a “dançar” e a jogar com as diferentes linguagens para não patinar no labirinto hipertextual da web.

Referências

- [1] KOCH, I. *Desvendando os segredos do texto*. São Paulo: Cortez, 2002.
- [2] NELSON, T. *Literary Machines 93.1*. Sausalito: Mindful Press, 1992.
- [3] POSSENTI, S. *Os Limites do discurso*. Edições Criar, Curitiba, 2002.
- [4] WITTGENSTEIN, L. *Investigações Filosóficas*. Nova Cultural, São Paulo, 1999. (Os Pensadores).
- [5] XAVIER, A. C. *O Hipertexto na sociedade da informação: a constituição de um modo de enunciação digital*. PhD thesis, Campinas, 2002. Inédito.
- [6] XAVIER, A. C. *Cadernos de Estudos Linguísticos*. Campinas, 2003, ch. Hipertexto e intertextualidade, pp. 293-290.

C-ORAL-ROM: Comparing romance languages in spontaneous speech corpora

Emanuela Cresti & Massimo Moneglia
LABLITA, University of Florence

elcresti@unifi.it

Resumo

Este artigo apresenta o corpus C-ORAL-ROM (Corpora de Referência Integrada para as Línguas Românicas Faladas) e sua arquitetura, que tem como objetivo representar os principais fatores que afetam a variabilidade da língua falada. O artigo enfoca as condições que possibilitam a comparabilidade de corpora para a análise linguística, ao nível da arquitetura de corpora e de sua anotação. O valor da anotação das quebras prosódicas para a determinação de enunciados e fronteiras é discutido em face de informação sintática.

1 The C-ORAL-ROM Corpus

C-ORAL-ROM has accomplished a set of corpora of spontaneous speech for the main romance languages that are spoken in Europe (Italian, French, Portuguese and Spanish)¹. This multilingual resource consists of 772 different texts corresponding to 300,000 words for each Language. The entire corpus (1,200,000 words) recorded 123:27:35 hours of speech produced by 1,427 different speakers².

The main objective of C-ORAL-ROM is to allow cross-linguistics studies of romance languages based on spontaneous speech data which represent everyday use. This task leads to three connected requirements that the resource meet: a) the corpus design must ensure general representativeness and internal balance (Sinclair, 2005); b) if we want to use corpora for comparative studies the corpus design of each language resource must be comparable with the others. c) Corpora must be annotated with regards to the unit of reference ranking above the word level in the same manner; d) the linguistic information conveyed by the

¹The resource is the result of the C-ORAL-ROM project co-ordinated by the University of Florence and funded within the Fifth framework program of the European Union. Language corpora have been provided respectively by: University of Florence (LABLITA, Laboratorio linguistico del Dipartimento di italianistica); Université de Provence (DELIC, Description Linguistique Informatisée sur Corpus); Centro de Linguística da Universidade de Lisboa (CLUL); Universidad Autónoma de Madrid (Departamento de linguística, Laboratorio de Linguística Informática).

²C-ORAL-ROM is delivered in two multimedia editions: 1) for speech laboratories, through the ELDA catalogue, in 9 DVDs in non-compressed and non-encrypted files <http://www.elda.org/catalogue/en/speech/S0172.html>; 2) for personal use, in one DVD, through Benjaminus Publishing company, in compressed and encrypted format (Cresti; Moneglia, 2005) <http://www.benjaminus.com/cgi-bin/t.bookview.cgi?bookid=SCL2015>.

speech signal must be correlated to transcripts in a meaningful way. In section 2 the paper will discuss the first two issues presenting results that validate the C-ORAL-ROM choices with regards to the corpus design of the resource. In section 3 the synchronization issue will be discussed in connection with challenging problems presented by spontaneous speech analysis. Then a conclusion is presented.

2 Spoken language variability and resource comparison in the spontaneous speech domain

2.1 Corpus design

The problem of representation is particularly sensitive in the spoken domain due to the huge variability of the spoken performance. For instance, a story told to a child and a row between husband and wife vary in language register, dialogue structure and of course in many language properties at various levels of the linguistic representation. There are many types of such variations in the spoken universe. For example comparing a business transaction with a lesson at school other core variations may occur (again dialogical character, programming, style, task, etc.) On the other hand, as sociolinguistic studies have shown during half a century of research, also difference in education, geographical origin, age, occupation, etc. have strong consequences on the language performance of the population. (Berruto, 1987; Biber, 1988; De Mauro et al., 1993; Gadet, 1996; Labov, 1966).

According to this tradition a significant sampling of the spontaneous speech universe should in principle represent:

- diachronic variations (time)
- diatopic variations (geography)
- diaphasic variation (social relation)
- diastratic variation (socio-economic characters of population)

Given the multilingual nature of this resource and the cost of spoken corpora, C-ORAL-ROM, as with other current spoken resources³, focuses mainly on the diaphasic parameter that, as we will see in this paper, has in principle the larger impact specifically on the structuring aspects of spoken language variability (Moneglia, 2006). Therefore C-ORAL-ROM is representative of the diaphasic variation, and it is unbalanced from the point of view of diastratic and diatopic variation.⁴ Figure 1 shows the number of speakers recorded in each language resource and their regional origin.

³The DUTCH CORPUS is an example of context-driven corpus design

⁴Diachronic variation is out of present possibilities in the spontaneous speech domain. Only a few spontaneous speech recordings were collected before the seventies. See Moneglia et al. (to appear) for recent diachronic studies on spoken Italian and the NURC collection for Brazilian Portuguese.

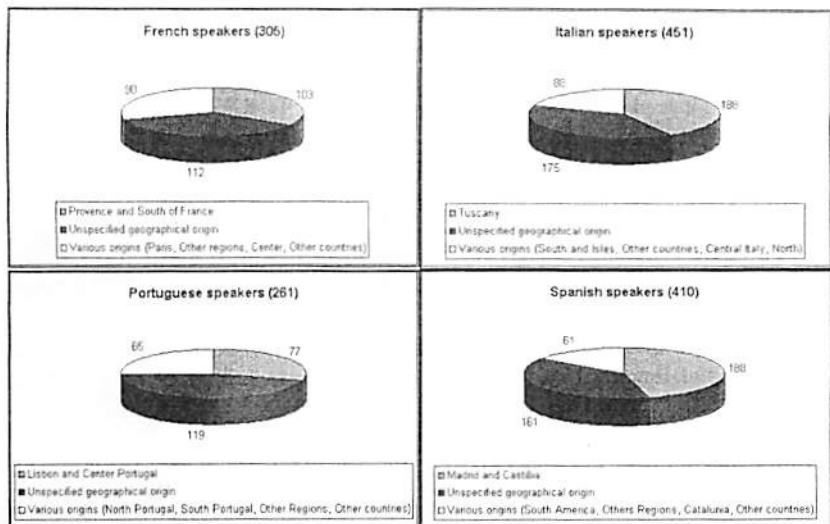


Fig. 1: Speakers in the C-ORAL-ROM corpora

These data demonstrate that indeed all C-ORAL-ROM corpora embody some geographical variation and also record a huge number of speakers, that may belong to various diastatic sectors of the population.⁵ Therefore the variety of each corpus is ensured also from the diatopic and the diastatic point of view, but the four collections are crucially unbalanced in those respects and the diastatic and diatopic variation of each corpus do not strictly represent the universe in the same manner.⁶ On the contrary C-ORAL-ROM corpora demonstrate in a uniform manner the diaphasic variation and are therefore comparable in this respect.

The diaphasic variation is demonstrated in C-ORAL-ROM through a set of contextual parameters that define the diaphasic properties of each text:

- Register parameter*: sessions showing a formal language register vs. sessions demonstrating an informal language use;
- Channel parameter*: face-to-face interactions, broadcasting, interactions through telephone;
- Dialogical structure parameter*: speech events having a dialogue or a multi-dialogue structure vs. monologues;
- Social context parameter*: interactions belonging to family and private life vs. interactions taking place in public;

⁵Education, Age, Profession, and geographical origin of each speaker is recorded in the metadata of each session.

⁶The spoken part of the British National Corpus (BNC) provided a significant sampling of the British population. More recently CoSIH (Izre'el et al., 2001) designed a frame which applies this strategy to the Israeli population. In CoSIH informants are captured in recordings while they go through all the contexts of one day, so ensuring speech data that are balanced both at sociological and contextual variation level.

- e. *Domain of use parameter*: domains such as law, business, research, teaching, church, etc. are represented

On the basis of the previous variation parameters C-ORAL-ROM adopts two different sampling strategies for formal contexts and informal contexts. As the following table shows, the Formal is sampled according to Channel and Domain of use, while Informal is sampled according to the Sociological context in which the event take place and to its Communication Structure.

This design strategy is pretty new in the spontaneous speech corpus tradition and it tries to capture the use conditions of the formal and the informal registers. Indeed, at least in principle, the set of contexts in which the formal register is preferred in a given society is substantially defined (closed list). Therefore the formal use of language can be represented by listing the typical contexts of its use. The same does not hold for the universe of informal speech whose contexts of use are in principle an open set. In other words no context is more typical than another as an instance of informal speech. Therefore the C-ORAL-ROM matrix demonstrate the formal use of language through the genre and the domain of application of the sessions in a closed list of typical domains of use, while for informal contexts, the domain and the text genre is left random and the variability is recommended.

Language type	Sociological context	Structure of the communication event
INFORMAL	Family / Private Public	Monologues Dialogues Multi-dialogues
Language type	Channel	Typical domain of use
FORMAL	Format in natural context	political speech; political debate; preaching; teaching; professional explanation; conference; business; law;
FORMAL	Media	news; sport; interviews; science; weather; forecast; scientific press; reportage; talk show
INFORMAL	Telephone	Private conversation; Human-Machine interactions

Tab. 1: The C-ORAL-ROM Corpus Design

Given the above schema each romance corpus represents the universe in a comparable way, as far as it complies with the following matrix, which specifies the proportion of words and samples in each field.

Section	Context	Domain	Target Words	Italian Words	French Words	Spanish Words	Port. Words
Informal	Family-private		150000	154967	152185	168868	165436
		Monologue	42000	45212	47702	42082	45937
		Dialogue-Conversation	82500	83404	77144	88973	86950
	Public	Monologue	25500	26201	27499	37812	32540
		Dialogue-Conversation	6000	6050	6900	6116	7696
			19500	20241	20539	31696	24853
Formal			150000	156002	143418	164614	152480
	Natural context	(see the above list)	65000	68328	57319	72268	66140
	Media	(see the above list)	60000	61759	57143	62739	62018
	Telephone	(see the above list)	25000	25015	28956	28607	24322
Total			300000	310969	295803	333482	317916

Tab. 2: The matrix of the C-ORAL-ROM corpora

Sample length is determined in a uniform manner in the four corpora. Informal samples record about 1500 words while formal samples record around 3000. Therefore, given the above figures, corpora record also a comparable number of samples.

2.2 Cross-linguistic comparison and contextual variations

Research based on the C-ORAL-ROM corpora has shown that this strategy is successful in capturing crucial properties of spoken language variation, that turns out strictly correlated to the diaphasic variation, rather than to the diastatic properties of the speakers.⁷

In the following we will show that the contextual variation recorded in the C-ORAL-ROM corpus design is correlated to the variation of high level language properties that characterize the overall nature of the spoken performance. Among the various correlations that have been found, we will focus here on two main examples that characterize the performance at the level of the lexical choice and at the level of the constructive principles.

2.2.1 Nouns and Verbs

The lexicon is a core area of the language and huge differences have been recorded in the spoken and written channel. It is well known that Nouns are the prevalent part of speech in written language (Biber, 1988; Biber et al., 1999; Giordano; Voghera, 2002; Halliday, 1989). This is confirmed also by our measurements of tokens in the BNC whose lexicon records around 26% of Nouns and 17% of verbs.⁸ On the contrary in spoken language the use of verbs is much more frequent and the ratio change in favor of verbs (Halliday, 1989).

As Figure 3 shows the proportions of verbs and nouns is cross linguistically consistent in C-ORAL-ROM and verbs always record more tokens than Nouns in the four corpora. The 'verbal' lexical strategy of speech noticed by Halliday is therefore confirmed at cross-linguistic level in the Romance languages and it is therefore a genuine differential feature of the spoken performance.

⁷ See in particular the speed, the Mid-length of the utterance, the mid-length of the dialogic turn, the mid length of the tone unit, the relation between simple verbless and compound verbal utterance, the main connectives. See Moneglia (2004), Moneglia (2006), Cresti (2005) for detailed discussions regarding different kinds of language variations.

⁸ This measurement is based on a LABLITA internal tagging and will be verified with the new tagged release of BNC that will be available shortly.

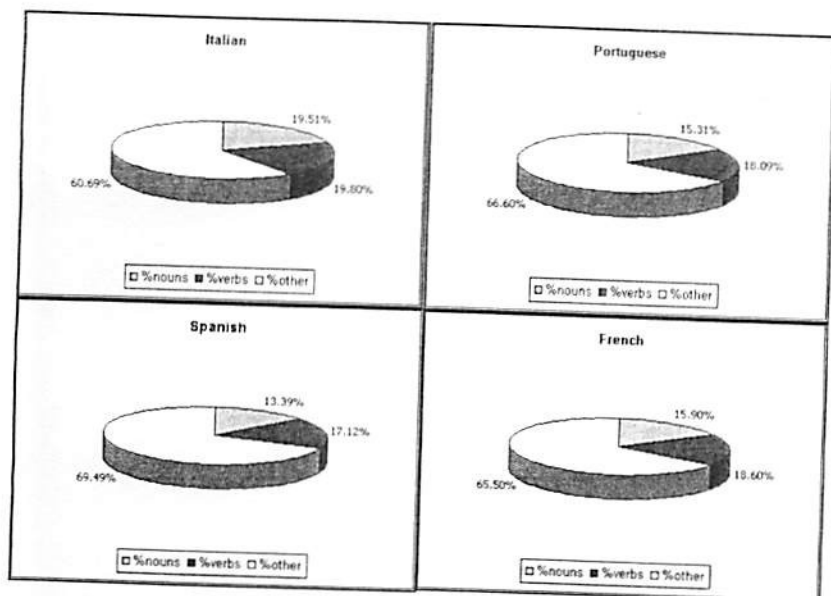


Fig. 2: Nouns and Verbs in the C-ORAL-ROM corpora

However the Verb/Nouns ratio in spontaneous speech is even more significant if values are considered in the contextual variation of the corpus design. The line diagrams in Figure 4⁹ show a regular increase in nouns from Informal Dialogues to Media and Formal Monologues, with a very marked drop in the Telephone node, and, in a complementary way, a decrease in verbs from Informal Dialogues to Formal Monologues. The variation of the ratio between Nouns and Verbs in the romance corpus is therefore regular at cross-linguistic level and moreover strictly corresponds to contextual variation.

The feature of Formality required by the context seems to be the factor most responsible for the progressive increase of Nouns through the corpus structure and the proportional decrease of Verbs. In summary the ratio of Nouns and Verbs, that is the main feature of the lexical strategy in speech, changes as a function of context variation.

2.2.2 Verb-less utterances

Specific contextual variations strictly correlate also with systematic variations in the constructive strategies of the utterance, that is in the core level of the spoken performance. The primary index of the utterance structure is of course the presence of a finite verbal form. C-ORAL-ROM confirms that verb-less utterances in speech have a remarkable frequency, when compared to the total utterance. This feature of spontaneous speech has been already noticed for

⁹The diagrams here and in the figures below list in the horizontal axis respectively: inf-dial (informal dialogues); inf_mon (informal monologues); form_dial (formal dialogues); form_mon (formal monologues); media, telephone

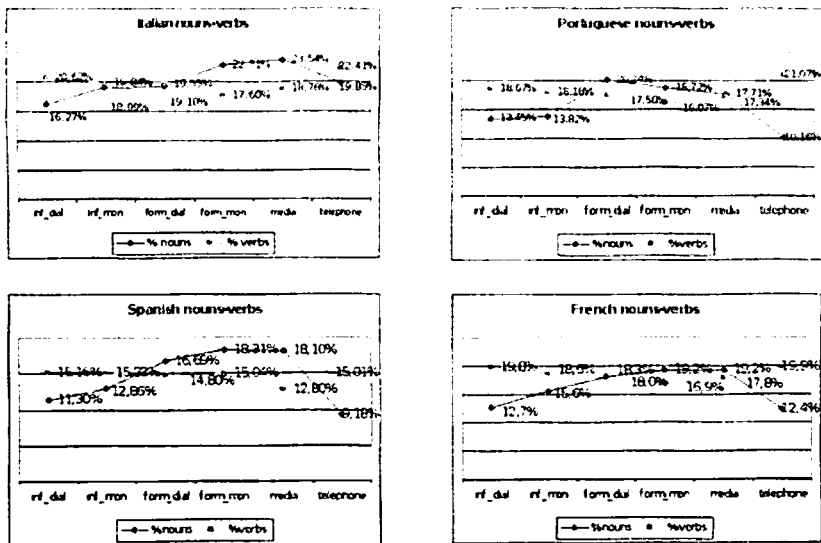


Fig. 3: The variation of Nouns and Verbs through the corpus design nodes

English in the Longman Grammar (38% of verb-less recorded in spoken English) and, as the following figure shows, the percentage of verb-less utterances is similar also in romance corpora, with a significant difference in French.¹⁰

However also the ratio between verb-less and verbal utterances strongly varies in accordance to the contexts recorded in the corpus and this variation is paralleled in the four romance languages.

As the line diagrams in Figure 6 show the 'verbal structuring strategy' is positively influenced in the four romance languages by the monologic structure of the communicative event. According to the relation "the more the event is monologic the more utterances are verbal, while the distribution of verb-less utterances is complementary". The variation in percentage of verb-less utterances in connection with the dialogical structure of the linguistic event occurs at cross-linguistic level and looks like a genuine variation parameter of spoken language.

In summary, if we want to compare languages in the spontaneous speech domain the corpus design of each resource must comprise a huge contextual variation, as important language characters of language performance are driven by contextual variation. Moreover all resources must represent this variation in a uniform manner. C-ORAL-ROM is designed to meet both requirements.

¹⁰See Deolofeu and Blanche-Benveniste (2006) for a discussion.

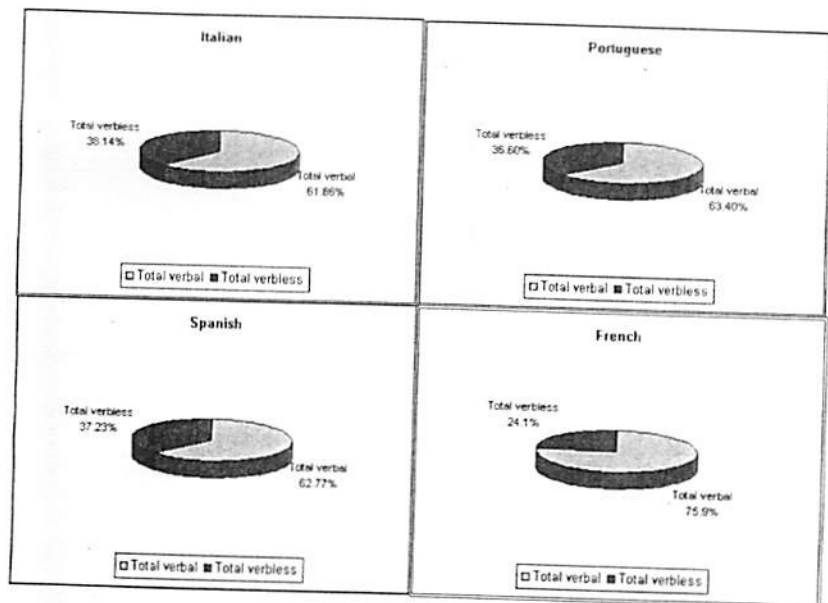


Fig. 4: Distribution of verbal and verbless utterances in the C-ORAL-ROM corpora

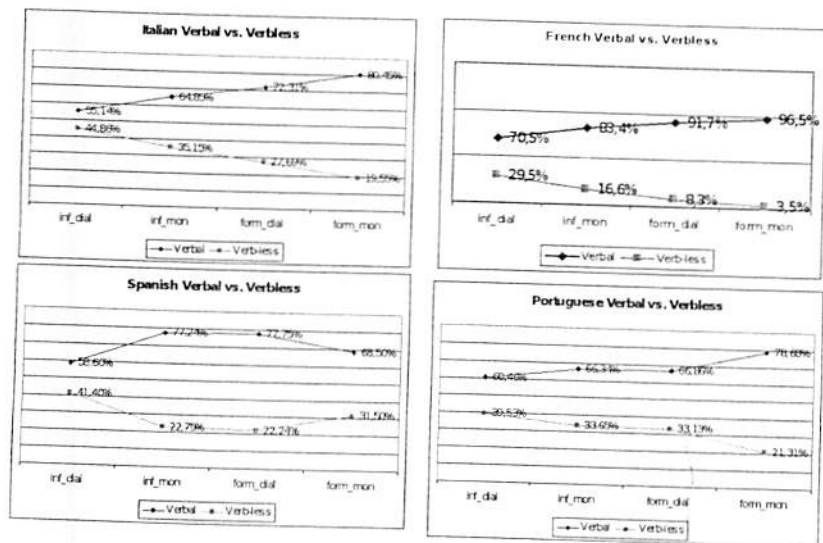


Fig. 5: Variation of the distribution of verbal and verbless utterances through main contexts of the design

3 Text to speech alignment and the “reference unit” of spontaneous speech

The collection of files of the C-ORAL-ROM resource is also conceived to allow maximum and easy exploitation of the linguistic information. Each recorded session of the C-ORAL-ROM corpus is stored in acoustic files¹¹ and is delivered with the following main annotations:

- a. Session metadata (in CHAT and IMDI format)
- b. The orthographic transcription in CHAT format (MacWhinney, 1994) enriched by the tagging of terminal and non-terminal prosodic breaks.

However, as it is widely recognized, the exploitation of spoken resources requires a link between the linguistic annotation and the acoustic source (Llisterri, 1996; Oostdijk et al., 2004; Sinclair, 1994). To this end C-ORAL-ROM is integrated by the Win Pitch Corpus speech software (©Pitch France) that allows the direct and simultaneous exploitation of both the acoustic and textual information.¹² In connection with this feature C-ORAL-ROM also delivers:

- c. The text-to-speech synchronization, based on the alignment to the acoustic source of each transcribed utterance, in .xml files.

For text to speech synchronization, special attention is devoted to the annotation of the ‘reference units’ ranking above the word level and to the link of this unit with prosody. This is because the linguistic analysis of each corpus and the comparison among different language corpora strictly require that the segmentation of speech into linguistic units is done always in the same manner. The previous case is quite clear in this respect: if we want to measure the occurrence of verb-less utterances in the four romance languages, the utterance must be identified in all corpora in the same manner. At the same time the necessity of a text to speech synchronization also requires the speech event in the continuum that is suitable for both language perception and linguistic analysis.

In the following we will briefly show that the identification of the reference unit is the key information which allows a meaningful synchronization of the transcripts to the acoustic source, and a solid basis for the linguistic analysis in the spontaneous speech domain.

3.1 The utterance in the C-ORAL-ROM annotation

The product of the speech performance is the ‘utterance’ (Biber et al., 1999), but the definition of this entity is a complex matter. The utterance can hardly be identified through the same syntactic and semantic cues used for the description of written language (Biber et al., 1999; Blanche-Benveniste, 1997; Cresti, 2000;

¹¹In the ELDA edition acoustic files are .wav Windows PCM, 22.050 Hz, 16 bit, while in the Benjamins edition the sound is stored in MP3 files.

¹²Textual resource with the annotation of Part of Speech (PoS) and lemma of each form is also delivered. The Benjamins edition is also integrated with a concordance software (Contextes ©Jean Veronis).

Izre'el, 2005; Miller; Weinert, 1998). As the previous paragraph made clear, almost 1/3 of utterances does not have a verb and therefore do not show a clear syntactic structure: this means that configurations that are not syntactic clauses may be reference units in the speech flow.

In the C-ORAL-ROM approach, following the pragmatic tradition (Austin, 1962), the utterance is defined as the linguistic entity that copes with a *speech act*. An utterance is a stretch of the flow of speech that is 'concluded' and 'autonomous' from a pragmatic point of view (Quirk et al., 1985). More specifically, in our definition, the utterance is the *minimal linguistic entity that can be pragmatically interpreted*, and it is 'autonomous' and 'concluded' for this reason (Cresti, 2000).

However, despite this pragmatic definition, the annotation procedure adopted in C-ORAL-ROM to parse the speech continuum exploits the prosodic correlations of the utterance, that are the more evident to perception and are therefore suitable for an annotation schema.¹³ It is assumed that each utterance has a profile of *terminal intonation* (Crystal, 1975; Karcevsky, 1931) and therefore the presence of *terminal breaks* in a string is considered the cue that allows the detection of utterance boundaries. The utterance is identified through a heuristic that relies on the annotation of *prosodic breaks* in the speech flow by competent speakers (Cresti, 2000; Cresti; Firenzuoli, 1999).

Prosodic tagging is accomplished on the basis of the sole perceptual evidence¹⁴: each perceptively relevant prosodic variation in the speech continuum is such that it causes the parsing into discrete prosodic units and will be reported by the annotator onto the transcription. Prosodic breaks are distinguished in accordance with two main qualities: *terminal vs. non-terminal*.

The speech flow and its transcription are so divided into reference units which rank above the word level taking into account a single prosodic feature: *each prosodic unit ending with a terminal break is considered an utterance*. This heuristic corresponds to a strong perceptual evidence and therefore leads to an high inter-annotator agreement. It was already known that competent speakers have a strong perception of prosodic boundaries (Buhmann et al., 2002; Sorianello, to appear). The C-ORAL-ROM annotation and its subsequent validation show that, at least with regards romance languages, competent speakers can also coherently discriminate prosodic boundaries that have a terminal value from those boundaries that indicate that the utterance continues (Danieli et al., 2004).¹⁵

3.2 Speech performance and the reference unit of spontaneous speech

The idea to specify the reference unit for linguistic analysis on the basis of prosodic cues may sound strange to many scholars. The fact is that syntax does not provide enough evidence for the identification of the linguistic unit ranking

¹³See Carletta et al. (1997) for the problems encountered in speech acts annotation.

¹⁴For a general framework of the relationship between perceptual and prosodic structures see Hart et al. (1992).

¹⁵The K value recorded for Italian, Spanish and Portuguese corpora in the inter-annotator agreement evaluation is over 8; the K value for the French corpus is over 6.

above word level and to define this unit is absolutely necessary for exploiting spoken resources for linguistic purposes.¹⁶

The detection of terminal prosodic breaks is relevant to this task for two main reasons: 1) the reference units are in one-to-one correspondence with speech acts; 2) the information conveyed by terminal breaks is necessary to determine the syntactic domains, and therefore the analysis of speech performance syntax depends on prosody and not vice versa. The following paragraphs will present the C-ORAL-ROM multimedia format while dealing with both issues.

3.2.1

The conception that the language performance “read” a predetermined syntactic structure cannot be applied to the analysis of spoken language performance. The examples below present two typical spontaneous speech turns of the C-ORAL-ROM corpus (Portuguese) that explain why.¹⁷

* JOS: já chegou para algums

* HEL: estive no Chiado ha pouco tempo

The wording of the turns will bring the analyst to figure out one speech event, in both cases, with the same syntactic structure. In this hypothesis both turns will be analyzed as one speech event compounded by one sentence with an adverbial phrase. But this analysis does not correspond to the actual status of both stretches of speech. Given that in the C-ORAL-ROM format the double slashes “//” indicate the terminal breaks, which identify the utterance limit, the single slash “/” refers to the non-terminal prosodic breaks, the following are the data transcribed:

* NOE: [48] estive no Chiado / há pouco tempo //
(pfamd114)

* JOS: [147] já chegou // [148] para algums //
(pfamd21)

The first turn is one sequence of two units ending with only one terminal break and corresponds to a single assertion. This is clear as soon as the speech performance is analyzed. The first envelope cannot be judged as an autonomous speech act, as it is suspended, and cannot be interpreted in isolation. Therefore the turn is one and only one autonomous utterance with an internal chunking. The reader can verify this by accessing the audio information in C-ORAL-ROM. This is done by playing the audio and visualizing the following analysis of the speech data by just clicking on each transcribed utterance.

On the contrary the second turn is a sequence of two envelopes each one ending with a terminal break. Both units are judged autonomous and concluded in itself and both can be interpreted in isolation as a single assertion. Each one is therefore an independent speech act, and the turn is compound by two distinct utterances.

¹⁶See Moneglia and Cresti (2006) for a full discussion of the issue.

¹⁷The turn is introduced by the speaker's label in capitals, preceded by “*” and followed by “:”. The sequence of speech events in the same turn should be ordered horizontally by the analyst. A number in square brackets indicate the rank of the utterance in the recording.

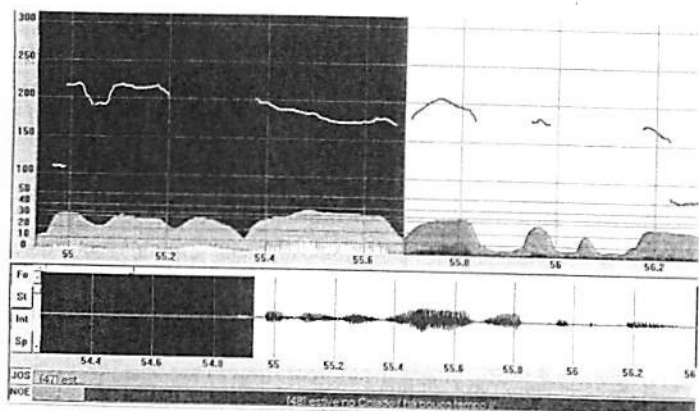


Fig. 6: WinPitch analysis

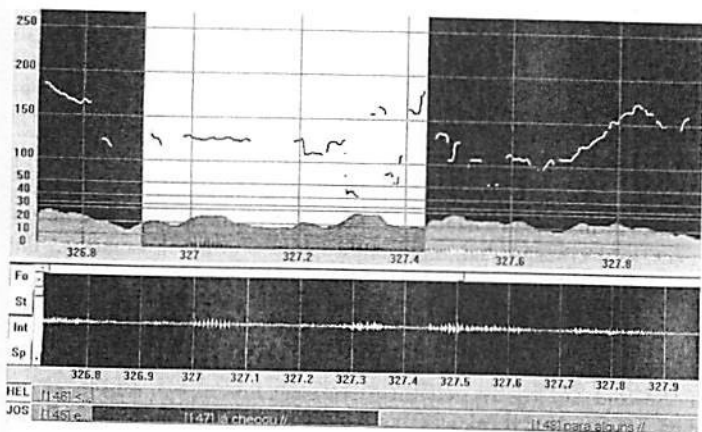


Fig. 7: WinPitch analysis

Given that from both the syntactic and semantic point of view the two sequences could have the same structure the distinction is specifically conveyed by prosody.

From this point of view the syntactic structure is a function of the prosodic information and not vice-versa, and to each utterance there corresponds an independent syntactic domain.

C-ORAL-ROM only marks the break type and does not identify the specific prosodic cues that lead to the previous judgment. However, marking this minimal property of the utterance, it specifies the reference unit ranking over the word level in an uniform manner for all utterances of the four romance corpora (roughly 135,000 utterances in C-ORAL-ROM) and therefore allows cross-linguistics measurements of the actual structures of speech performance.

For example, in the case under consideration, the fact that the second turn is about two independent utterances also specifies that the adverbial phrase is a verbless utterance and not the adverbial phrase of a verbal utterance. This information is radically underdetermined from a sole syntactic point of view.

Through the multimedia format of the C-ORAL-ROM resource the user also has the possibility to verify (real time) the segmentation of the speech flow provided in the corpus annotation and can therefore immediately study the acoustic parameters, that are simultaneously displayed while the textual information is played.

3.2.2

One could say that the performance should not enter into decisions regarding the syntactic structure of a given text. From this point of view both turns could still have the same structure that, in the second case, has accidentally been performed by two independent acts.

Despite the fact that many linguists may take this position, it must be stressed that this objection is not consistent with data in the domain of spontaneous speech and brings dramatic consequences for the linguistic analysis of spontaneous speech corpora. This is because the syntactic structure of whatever stretch of speech is not pre-theoretically determined, but must be decided upon on the basis of positive evidence. To ignore the information provided by the speech act performance leads to a radical under-determinacy of the spoken language structure. This under-determinacy does not occur in reality due to the fact that competent speakers do assign a value to the boundaries of speech events.

The following dialogic turn, also taken from C-ORAL-ROM Portuguese, shows this point. The turn is about the place that people of a village used in the past for washing things. The transcription corresponds to the word sequence uttered by the speaker:

* EMI: [...] no inverno lavámos ali no ribeiro ali ali ao pé do
pontão da linha¹⁸

From a theoretical point of view the syntactic structures that can be assigned to the previous sequence are many. The reference units ranking over the word level are therefore strongly underdetermined from the syntactic point of view. Among the more perspicuous syntactic grouping the reader can consider:

¹⁸The reader should not consider a fragment that occurs in the turn, marked in bold.

1. [no inverno lavámos ali] [no ribeiro ali ao pé do pontão da linha]
2. [no inverno lavámos ali] [no ribeiro ali] [ao pé do pontão da linha]
3. [no inverno lavámos] [ali no ribeiro] [ali ao pé do pontão da linha]
4. [no inverno lavámos ali no ribeiro] [ali] [ao pé do pontão da linha]
5. [no inverno lavámos ali no ribeiro] [ali ao pé do pontão da linha]

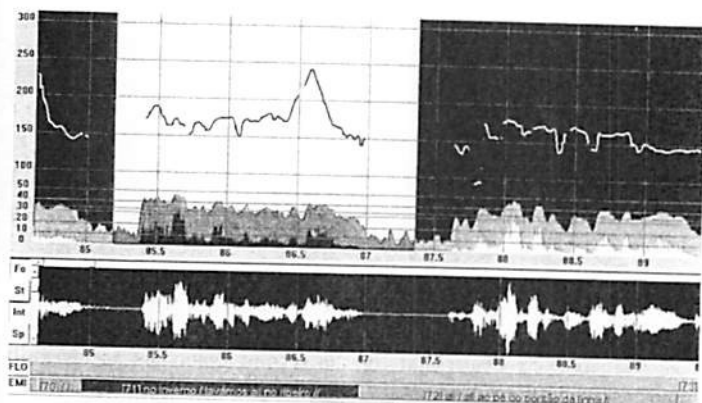


Fig. 8: WinPitch analysis

All the previous five alternatives are adequate from a syntactic, a semantic and also a contextual point of view. So a generic reference to the "context" does not help for the selection of the alternatives. However the stretch of speech is not ambiguous to the listener in reality and only one structure is compatible with the input data. The following is the data transcribed in C-ORAL-ROM¹⁹:

* EMI: ... [71] no inverno / lavámos ali no ribeiro // [72] ali [/]
ali ao pé do pontão da linha //
(pfamd103)

As specified in the transcription, the hearer perceives that in the turn two, speech acts have been accomplished, each one concluded by a terminal prosodic break. Therefore only the syntactic structure in 5 is compatible with the input data; that is a verbal utterance with an adverbial locative modification, followed by a nominal utterance:

5 [no inverno lavámos ali no ribeiro] [ali ao pé do pontão da linha]

¹⁹ "[/]" refers to the occurrence of a retracting generating a fragment

Each utterance is aligned to the acoustic counterpart. The following is the result of the text sound alignment as it is showed in the Win Pitch Corpus graphical interface:

To ignore speech acts in spontaneous speech analysis is to make underdetermined a domain that is on the contrary completely determined.

Of course one could say that in these cases the information provided by the speaker is, in practice, sufficient to avoid the ambiguity. But why should the prosodic information, that is considered relevant in this case, be irrelevant in the previous one? If prosodic indexes are needed to determine the syntactic boundaries, this role must be recognized if we want the analysis to be consistent with the data.

4 Conclusion

Without the prosodic information concerning the utterance boundaries the syntactic structure is underdetermined. The definition of the utterance boundaries through prosodic properties allows a proper analysis of the four spoken romance corpora and consequently permits us to carry out their general comparison from a linguistic point of view. More specifically, the prosodic tagging in C-ORAL-ROM provides the term of reference for the statistic evaluation of the main properties of spoken language performance in the multilingual corpus and a linguistically meaningful criterion for the text-to-speech alignment of large corpora. This result is achieved with low effort, trusting the direct perception of prosodic boundaries.

References

- [1] AUSTIN, L. J. *How to Do Things with Words*. Oxford University Press, Oxford, 1962.
- [2] BERRUTO, G. *Sociolinguistica dell'Italiano Contemporaneo*. La Nuova Italia Scientifica, Roma, 1987.
- [3] BIBER, D. *Variation Across Speech and Writing*. Cambridge University Press, Cambridge, 1988.
- [4] BIBER, D. *The Longman Grammar of Spoken and Written English*. Longman, London, New York, 1999.
- [5] BLANCHE-BENVENISTE, C. *Approches de la Langue Parlée en Français*. Ophrys, Paris, 1997.
- [6] BRITISHNATIONALCORPUS. British national corpus. Online; Available at <http://www.natcorp.ox.ac.uk/>.
- [7] BUHMANN, J., ET AL. Annotation of prominent words, prosodic boundaries and segmental lengthening by no-expert transcribers in the spoken dutch corpus. *2nd. Proceedings of the International Conference on Language Resources and Evaluation (LREC 2002) (2002)*, 779-785.

- [8] CARLETTA, J., ET AL. The reliability of a dialogue structure coding scheme. *Computational Linguistics* 1, 23 (1997), 13-31.
- [9] CHAT. Chat. Online; Available at <http://childes.psy.cmu.edu/manuals/CHAT.pdf>.
- [10] CONTEXTES. Contextes. Online; Available at <http://www.up.univ-mrs.fr/veronis/logiciels/Contextes/index-en.html>.
- [11] CoSIHCORPUS. Cosih corpus. Online; Available at <http://www.tau.ac.il/humanities/semitic/cosih.html>.
- [12] CRESTI, E. *Corpus di Italiano Parlato*, vol. 1, 2. Accademia della Crusca, Firenze, 2000. CD-Rom.
- [13] CRESTI, E. *C-ORAL-ROM: Integrated Reference Corpora for Spoken Romance Languages*. Benjamins, Amsterdam, 2005, ch. Notes on Lexical strategy, structural strategies and surface clause indexes in the C-ORAL-ROM spoken corpora, pp. 209-256.
- [14] CRESTI, E., AND FIRENZUOLI, V. Illocution et profils intonatifs de l'italien. *Revue Française de Linguistique Appliquée* 4, 2 (1999), 77-98.
- [15] CRESTI, E., AND MONEGLIA, M. *C-ORAL-ROM: Integrated Reference Corpora for Spoken Romance Languages*. Benjamins, Amsterdam, 2005.
- [16] CRYSTAL, D. *The English Tone of Voice*. Edward Arnold, London, 1975.
- [17] DANIELI, M., ET AL. Evaluation of consensus on the annotation of prosodic breaks in the romance corpus of spontaneous speech "c-oral-rom". *4th. Proceedings of the LREC Conference 4* (2004), 1513-1516.
- [18] DEULOFEU, J., AND BLANCHE-BENVENISTE, C. *Spoken Language Corpus and Linguistics Informatics*. Benjamins, Amsterdam, 2006, ch. C-Oral-Rom - French Corpus, pp. 181-198.
- [19] DUTCHCORPUS. Dutch corpus. Online, Available at http://lands.let.kun.nl/cgu/doc_English/topics/project/pro_info.htm.
- [20] GADET, F. Variabilité, variation, variété. *Journal of French Language Studies*, 1 (1996), 75-98.
- [21] GIORDANO, R., AND VOGHERA, M. Verb system and verb usage in spoken and written italian. *International Conference on the Statistical Analysis of Textual Data (JADT 2002)*, 6th. *Proceedings* (2002), 289-299.
- [22] HALLIDAY, M. A. K. *Spoken and Written Languages*. Oxford University Press, Oxford, 1989.
- [23] HART, J., ET AL. *A Perceptual Study on Intonation: an Experimental Approach to Speech Melody*. Cambridge University Press, Cambridge, 1990.
- [24] IMDI. Imdi. Online, Available at <http://www.mpi.nl/IMDI/>.

- [25] ISARD, A., AND CARLETTA, J. *Empirical Methods in Discourse Interpretation and Generation: Working Notes of the AAAI Spring Symposium Series*. Stanford University, Stanford, CA, 1995, ch. Replicability of transaction and action coding in the Map Task corpus, pp. 60-66.
- [26] IZRE'EL, S. Intonation units and the structure of spontaneous spoken language: a view from hebrew. *IDP05 International Symposium on Discourse-Prosody Interfaces, Proceedings* (2005). Available at <http://aune.lpl.univ-aix.fr/prodige/idp05/actes/izreel.pdf>.
- [27] IZRE'EL, S., ET AL. Designing cosih: the corpus of spoken israeli hebrew. *International Journal of Corpus Linguistics*, 6 (2001), 171-197.
- [28] KARCEVSKY, S. Sur la phonologie de la phrase. *Travaux du Cercle Linguistique de Prague*, 4 (1931), 188-228.
- [29] LABLITA. Lablita. Online; Available: <http://lablita.dit.unifi.it>.
- [30] LABOV, W. *The Social Stratification of English in New York City*. Center for Applied Linguistics, Washington, DC, 1966.
- [31] LLISTERRI, J. Eagles: preliminary recommendations on spoken texts, 1996. Available at <http://www.ilc.cnr.it/EAGLES96/spokentx/spokentx.html>.
- [32] MACWHINNEY, B. *The CHILDES Project: Tools for Analyzing Talk*. Lawrence Erlbaum Associates, Hillsdale, NJ, 1994.
- [33] MAURO, T. D., ET AL. *Lessico di Frequenza dell'Italiano Parlato*. ETAS, Milano, 1993.
- [34] MILLER, J., AND WEINERT, R. *Spontaneous Spoken Language*. Clarendon Press, Oxford, 1998.
- [35] MONEGLIA, M. Measurements of spoken language variability in a multilingual corpus: Predictable aspects. *LREC Conference, 4th. Proceedings 4* (2004), 419-422.
- [36] MONEGLIA, M. *Spoken Language Corpus and Linguistics Informatics*. Benjamins, Amsterdam, 2006, ch. Units of Analysis of Spontaneous Speech and Speech Variation in a Cross-linguistic Perspective, pp. 153-179.
- [37] MONEGLIA, M., AND CRESTI, E. *Spoken Language Corpus and Linguistics Informatics*. Benjamins, Amsterdam, 2006, ch. C-ORAL-ROM Prosodic boundaries for spontaneous speech analysis, pp. 89-114.
- [38] MONEGLIA, M., ET AL. L'incidenza del lessico fiorentino nella lingua d'uso a firenze: un confronto tra il corpus stammerjohann del 1965 e un corpus di parlato contemporaneo. In *Prospettive nello studio del lessico Italiano* (Firenze:FUP, Giugno 2006), E. Cresti, Ed., vol. 9, IX Congresso internazionale della Società di Linguistica e Filologia Italiana. To appear.
- [39] NURC. Nurc. Online; Available at <http://www.lettras.ufjf.br/nurc-rj/>.
- [40] OOSTDIJK, N., ET AL. Compiling and processing spoken language corpora. *LREC Conference, 4th. Proceedings* (2004), 563-566.

- [41] QUIRK, R., ET AL. *A Comprehensive Grammar of the English Language*. Longman, London, 1985.
- [42] SINCLAIR, J. The history of a corpus. *International Summer School of Corpus Linguistics* (2005). Unpublished.
- [43] SINCLAIR, J. M. Spoken language, phonetic/phonemic and prosodic annotation. Tech. rep., NERC, 1994. NERC-1 Network of European Reference Corpora: Final Report.
- [44] SORIANELLO, P. Per una definizione fonetica e fonologica dei confine prosodici. *Congresso Internazionale La Comunicazione Parlata* (2006). To appear.
- [45] WINPITCH. Win pitch. Available at <http://www.winpitch.com/>.

este livro foi dado pelo
GRUPO DE PÓS-GRADUAÇÃO
em Linguistics
por R\$ 30,00
Em 17 / 09 / 09

ABRALIN
Associação Brasileira de Lingüística
www.abralin.org

Realização



Universidade
Federal de
Minas Gerais



Patrocínio



Conselho Nacional de Desenvolvimento
Científico e Tecnológico



PETROBRAS



C A P E S

FAPEMIG

Fundação de Amparo à Pesquisa do
Estado de Minas Gerais

Apoio



editora**contexto**



CEDECOM - UFMG
CENTRO DE COMUNICAÇÃO



A água de Minas



BH, SÓ PENSO EM VOCÊ.



PREFEITURA BH
TRABALHO PELA VIDA

ISBN: 978-85-7758-018-7