

Estudo contrastivo do uso de Alocutivos em italiano, português e espanhol europeus e português brasileiro

Tommaso Raso*

Flávia Leite**

RESUMO

O artigo estuda a unidade informacional de alocutivo contrastando espanhol europeu (EE), português europeu (PE), italiano (I) e português brasileiro (PB), com base na Teoria da Língua em Ato. O estudo se baseia na análise de 15 textos para cada língua, exceto para o PE em que constam 30 textos (todos paritariamente divididos em monólogos, diálogos e conversações), que possuem em média 1500 palavras retirados de 4 corpora de fala espontânea comparáveis: o C-ORAL-ROM e o C-ORAL-BRASIL. Foram estudados a função e a frequência da unidade; suas características acústicas; sua distribuição; sua variação lexical; e as diferenças nas várias tipologias e no discurso reportado. Os resultados apontam para interessantes diferenças interlinguísticas. A conclusão sugere uma explicação de caráter sociolinguístico para esses resultados.

Palavras-chave: Estrutura informacional; alocutivos; comparação interlinguísticas; português do Brasil; português europeu; espanhol europeu; italiano.

RESUMEN

El artículo estudia la unidad informacional de alocutivo comparando el Español (EE), Portugués Europeos (PE), Italiano (I) y Portugués Brasileño (PB), con base en la Teoría de la Lengua en Acto. El estudio está basado en el análisis de 15 textos para cada lengua, excepto para el PE en el que se cuentan 30 textos (todos igualmente divididos en monólogos, diálogos y conversaciones), esos poseen cerca de 1500 palabras retiradas de 4 corpora de habla espontánea comparables: el C-ORAL-ROM y el C-ORAL-BRASIL. Fueron estudiadas la función y la frecuencia de la unidad; sus características acústicas, su distribución; su variación lexical; y las diferencias entre las varias tipologías y en el discurso reportado. Los resultados apuntan para interesantes diferencias interlinguísticas. La conclusión sugiere una explicación de carácter sociolinguístico para esos resultados.

Palabras llaves: Estructura informacional; alocutivos, comparación interlinguísticas, Portugués de Brasil; Portugués Europeo, Español Europeo; Italiano.

1. INTRODUÇÃO

Este trabalho estuda as unidades informacionais de alocutivos (ALL)¹, com base na Teoria da Língua em Ato (CRESTI 2000), no espanhol europeu (EE, variedade da região de Madrid) e no português europeu (PE, variedade da região de Lisboa), além de comparar os resultados com os dados obtidos no estudo do português brasileiro (PB, variedade de Minas Gerais) e do italiano (I, variedade da Toscana) (RASO-GOULART

* é doutor pela Università di Napoli Federico II

** é graduanda da Faculdade de Letras da UFMG

¹ Ver seção 2.2 por uma definição aprofundada.

2009). Para isso, foram analisados 15 textos do EE e 30 do PE², somando-se os resultados de Raso e Goulart (2009) que trabalharam com 15 textos do PB e 15 do I. Os textos foram extraídos de dois corpora comparáveis: o C-ORAL-ROM (CRESTI-MONEGLIA, 2005; 2007; MONEGLIA, 2000) e o C-ORAL-BRASIL (RASO-MELLO, 2009; 2010). O objetivo é estudar interlinguisticamente as unidades informacionais de alocutivo para: 1- comparar a frequência da unidade; 2- analisar suas características acústicas; 3- analisar a variação lexical; 4- analisar a variação de *tokens* e *types* dentro de cada uma das tipologias interacionais previstas nos corpora comparáveis (diálogo, monólogo e conversação); e 5- verificar a preferência distribucional da unidade nas quatro línguas. As conclusões fornecem dados significativos a serem utilizados para trabalhos de cunho sociolinguístico; entretanto um estudo nessa ótica não entra no escopo do trabalho.

O C-ORAL-ROM é um corpus das quatro principais línguas românicas europeias (Italiano, Francês, Espanhol e Português Europeu), segmentado por enunciados e unidades tonais, e alinhado através do software WinPitch³ (o que permite examinar ao mesmo tempo som, espectrograma e texto). Os subcorpora que compõem o C-ORAL-ROM são formados, cada um, por 300.000 palavras, sendo 150.000 para o âmbito formal e 150.000 para o âmbito informal; no âmbito informal, que é o único levado em consideração nesse estudo, os textos são divididos igualmente nas tipologias interacionais dialógica, monológica e conversacional. O C-ORAL-BRASIL⁴ é a quinta ramificação do C-ORAL-ROM, ao qual é comparável por arquitetura e critérios de segmentação.

2. METODOLOGIA

2.1 A Teoria da Língua em Ato

O marco teórico que embasou a pesquisa foi a Teoria da Língua em Ato (CRESTI, 2000; MONEGLIA, 2005; RASO-MELLO-JESUS-DE DEUS, 2007; VALE, 2009; CRESTI –MONEGLIA, 2008; MONEGLIA, 2005), segundo a qual haveria uma correspondência biunívoca entre a unidade de ação (ilocução) e o enunciado (ato

² A razão pela qual a base de dados do PE é duas vezes aquela das outras línguas é a baixíssima frequência dessa unidade nos textos de PE. Quisemos ter a certeza de que a frequência não fosse devida aos textos escolhidos mas refletisse uma característica geral do corpus. A segunda amostra confirmou a média da primeira.

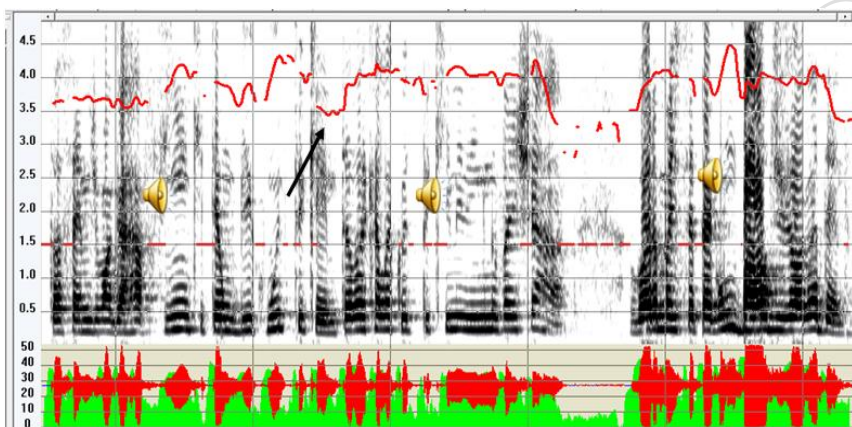
³ www.winpitch.com

⁴ O Projeto C-ORAL-BRASIL é coordenado por Tommaso Raso e Heliana Mello da Universidade Federal de Minas Gerais e é financiado pela FAPEMIG, pelo CNPQ, pela UFMG e pelo Banco Santander.

locutório ou contraparte linguística da unidade de ação) (AUSTIN, 1962), cuja interface seria constituída pela prosódia. De acordo com essa teoria, a fala pode ser segmentada em enunciados, cuja fronteira seria constituída por uma quebra prosódica percebida como terminal (/). O enunciado, por sua vez, pode ser dividido em unidades tonais, cujas fronteiras são constituídas por quebras prosódicas percebidas como não terminais (/). A Fig. 1 mostra uma sequência de 3 enunciados, cuja transcrição etiquetada é a seguinte:

tá saindo de uma garrafinha que tem um bico muito pequeno //=COM= então daquela coisa pequeninim nã vai encher rápido //=COM= agora imagina cê pega um balde e joga dentro //=COM=

FIGURA1 [Som](#) ([som 1](#), [som 2](#), [som 3](#))



É possível notar que entre o primeiro e o segundo enunciado, no ponto indicado pela seta, não há pausa alguma, assim, é a percepção de um perfil terminal que permite a segmentação e a apreciação dos dois enunciados como autônomos. Ao contrário, é muito frequente termos pausas, até longas, dentro de um enunciado, sem que, entretanto, se perceba nem a realização de uma ilocução nem um perfil terminal.

O enunciado é, portanto, considerado a menor unidade da fala autônoma pragmaticamente, ou seja, capaz de veicular uma ilocução. A ilocução veiculada naturalmente não depende do conteúdo locutivo e é identificável essencialmente por parâmetros prosódicos. As Fig. 2-4 mostram o perfil prosódico de 3 ilocuições diferentes com o mesmo material locutivo. Trata-se, respectivamente, de uma ilocução de asserção, de sarcasmo e de pergunta total, com o seguinte conteúdo locutivo: *João vai pro Rio //*

FIGURA. 2 ([som](#))

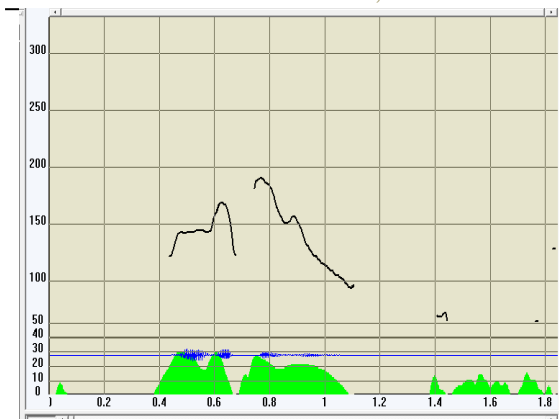


FIGURA 3 ([som](#))

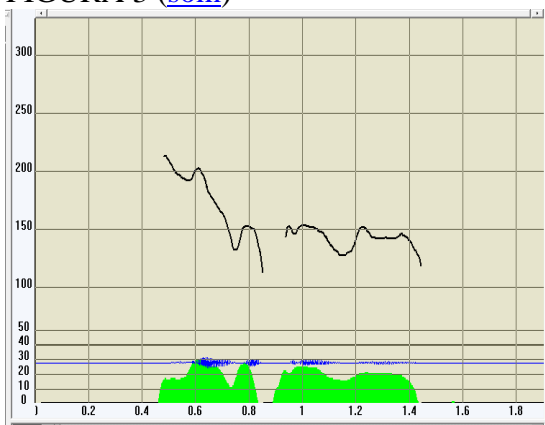
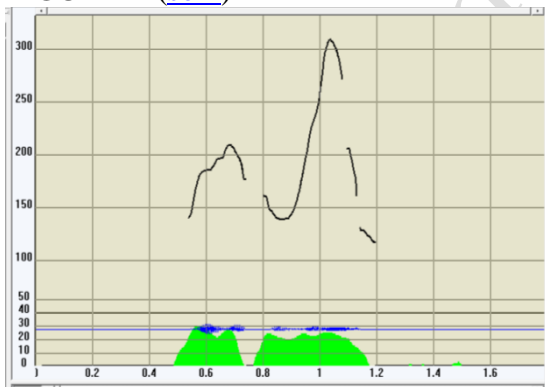


FIGURA 4 ([som](#))



Entre enunciado e ilocução existe, em princípio, uma relação biunívoca, pois, em geral, um mesmo enunciado é capaz de veicular somente uma ilocução⁵. Um enunciado pode ser simples ou complexo. O enunciado é simples quando constituído de uma única unidade tonal (em princípio correspondente a uma unidade informacional). Nesse caso a unidade deve ser a de Comentário, ou seja, aquela que carrega a força ilocucionária e

⁵ Para situações em que o princípio da biunivocidade se perde, veja-se Cresti 2009.

confere de fato autonomia ao enunciado. O enunciado complexo é composto pela unidade de Comentário e uma ou mais outras unidades com funções variadas.

A unidade de Comentário (COM) é, assim, considerada necessária e suficiente para a composição de um enunciado. Seu perfil entonacional é variado, como se viu, dependendo da ilocução veiculada, e sua distribuição é livre. As outras unidades, todas opcionais, são identificáveis com base em função, perfil entonacional e distribuição característicos. Cada unidade possui uma função diferente, um perfil prosódico característico (t'HART-COLLIER-COHEN, 1990) e uma diversa distribuição com relação ao COM. Tópico (TOP) e Apêndice (de Comentário, APC, e de Tópico, APT), junto com o COM, compõem o texto do enunciado propriamente dito. Parentéticos (PAR) e Introdutor locutivo (INT), mesmo não compondo diretamente o texto, são dirigidos a ele, com funções específicas. Outras unidades, os Auxílios Dialógicos (AUX), possuem função dialógica, ou seja, são dirigidas ao interlocutor e não compõem o texto do enunciado propriamente dito nem se referem a ele.

As Fig. 5 e 6 mostram a realização, em duas unidades tonais, do mesmo conteúdo locutivo que nas Fig. 2-4 são realizadas em uma unidade única. Na Fig. 5 o conteúdo é realizado como *João* /=TOP= *foi pro Rio* // =COM=. Na Fig. 6 o conteúdo é realizado como *João* /=COM= *foi pro Rio* // =APC=. É interessante notar que a unidade que confere autonomia ao enunciado é sempre a unidade de COM, independente da parte do conteúdo locutivo que ela realiza. Nas Fig. 5 e 6 as ilocuções realizadas são de asserção, e são, portanto, melhor comparáveis com a Fig. 2.

FIGURA 5 ([som](#))

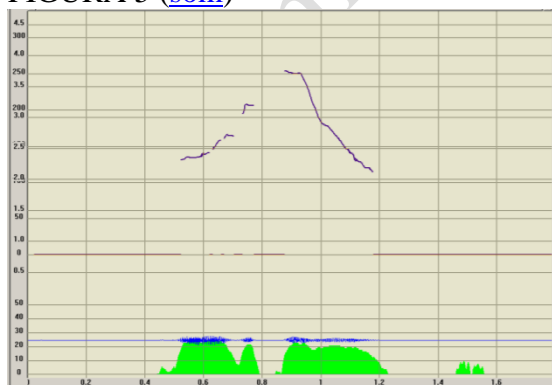
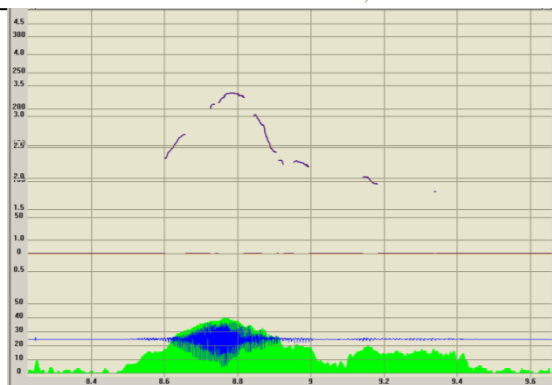


FIGURA 6 ([som](#))



Os enunciados das Fig. 2-6 são realizados em laboratório pela mesma voz e são, portanto, comparáveis. Segue aqui uma rápida exposição das características das diferentes unidades informacionais.

1. O TOP tem por função a delimitação semântica do COM, ou seja, a definição do âmbito de aplicação da força ilocucionária. Distribucionalmente vem sempre antes do COM. Entonacionalmente é a única unidade, além do Comentário, que possui foco funcional, nesse caso sempre à direita (ALVES DE DEUS, 2008; RASO-MELLO, no prelo; SIGNORINI, 2004.)
2. O Apêndice (de Tópico ou de Comentário, APT e APC, respectivamente) tem como função a integração textual da unidade da qual é apêndice. Distribucionalmente segue sempre a unidade que integra. Entonacionalmente, o APT possui um perfil descendente ou com um pequeno movimento, mas sem foco, enquanto o APC possui perfil descendente ou nivelado, sempre sem movimento (CRESTI- FIRENZUOLI 2002; JESUS, 2008; para o PB RASO-MELLO, no prelo; RASO-ULISSES, 2008; OLIVEIRA, 2009).
3. O Parentético (PAR) possui uma função metalinguística, frequentemente modalizadora, e indica ao interlocutor como interpretar o conteúdo do enunciado. Distribucionalmente pode ocupar qualquer posição, inclusive interrompendo outra unidade textual, com a exceção do começo absoluto do enunciado. Entonacionalmente possui um perfil nivelado, uma frequência mais baixa do que o resto do enunciado, frequentemente com uma maior velocidade de elocução (TENANI, 1996; SCHNEIDER 2007; TUCCI, 2009; MOTA, 2010; VALE, 2010.).
4. O Introdutor Locutivo (INT) tem como função introduzir uma metailocução (como o discurso reportado, uma lista, uma exemplificação, etc.) e marcar, portanto, a suspensão pragmática do enunciado, ou seja, a passagem para um *aqui e agora* diferente daquele

da situação comunicativa (GIANI, 2003 e 2004; CORSI, 2009; MAIA-RASO, em preparação.)

Por outro lado, os AUX (CRESTI, 2000; CRESTI-MONEGLIA, 2008; FROSALI, 2008; MAIA-RASO-ANDRADE, 2008) têm como função ajudar na interação entre os interlocutores, cada um deles com função, perfil e distribuição próprias. Seu papel é desempenhar funções relacionadas à regulação do bom funcionamento da comunicação, como por exemplo, sinalizar, alertar, reforçar ou dirigir-se ao interlocutor de maneira explícita. Dentre as várias unidades informacionais pertencentes a esse grupo encontram-se os ALL, objetivo principal deste trabalho. Além dele fazem parte também dos AUX:

1. os Conativos (CNT), com a função de induzir o interlocutor a fazê-lo desistir de algo; têm perfil descendente, intensidade alta e duração curta, distribuição em começo ou final de enunciado; são pouco frequentes;
2. os Fáticos (PHA), com a função de sinalizar a abertura do canal e o contato com o interlocutor, têm perfil nivelado, intensidade baixa e duração curta, distribuição livre; são muito frequentes;
3. os Incipitários (INP), com a função de iniciar um enunciado marcando um contraste (de natureza afetiva) com o anterior, têm perfil ascendente-descendente, intensidade forte e duração curta; são medianamente frequentes;
4. os Conectores Textuais (DCC), com a função de iniciar um enunciado marcando a continuidade com o anterior, têm perfil modulado, intensidade alta e duração longa; são frequentes;
5. os Expressivos (EXP), com a função de dar suporte emocional ao ato realizado no COM, têm perfil variável, intensidade média a forte, duração média, velocidade alta; são medianamente frequentes.

2.2 Os Alocutivos

Os ALL têm como função se dirigir ao interlocutor, e, portanto, identificá-lo ou marcar o tipo de coesão social com ele existente. Entonacionalmente possuem uma intensidade média ou baixa, com perfil nivelado ou levemente modulado. Diferentemente dos outros AUX, os ALL mantêm uma ligação com o valor semântico dos lexemas utilizados, que,

nos outros AUX são dessemantizados, mantendo unicamente o valor de conectores pragmáticos⁶.

2.2.1. Unidades Informacionais de Alocutivo versus Ilocução de chamamento

Não se deve confundir a ilocução de chamamento com as unidades informacionais de alocutivo, ainda que apresentem semelhanças com relação ao conteúdo locutivo. Esta última possui apenas uma função dialógica, cujo objetivo é manter o bom regulamento da interação, identificando para quem a mensagem está endereçada ou marcando o tipo de coesão social existente; por outro lado, a ilocução de chamamento é um COM. Portanto, se a ilocução de chamamento possui autonomia pragmática, o ALL é apenas uma unidade opcional, pragmaticamente não autônoma.

Nas FIG. 7 e 8 abaixo podem-se perceber claramente as diferenças entonacionais entre essas duas unidades informacionais. Para isso, buscou-se apresentar dois trechos em que o mesmo conteúdo locutivo assume para um mesmo falante a função de Comentário com ilocução de chamamento (FIG. 7) e a função de Alocutivo (FIG. 8)⁷. As diferenças principais mostram a necessidade de uma maior duração e intensidade da unidade de COM, que deve também apresentar um foco funcional para que a ilocução possa ser veiculada. No caso, o conteúdo locutivo (que corresponde ao apelido do interlocutor) é pronunciado pelo mesmo falante e é direcionado ao mesmo interlocutor. Ambas as imagens apresentam uma linha temporal com a mesma duração (7 décimos de segundo). Isso permite visualizar imediatamente as diferenças entre a unidade de COM da FIG. 7 e a unidade de ALL, marcada em preto dentro do enunciado da FIG. 8.

Em termos de duração, a ilocução de chamamento é cinco vezes maior do que o alocutivo, possuindo uma duração de cinco décimos de segundo contra um décimo de segundo de duração do alocutivo. Quanto à variação da curva de F0, o COM de chamamento apresenta uma variação da frequência fundamental de mais de 150hz, iniciando de forma ascendente em 325hz, subindo para mais de 350hz e descendo para 200hz. O foco entonacional desta ilocução está na parte esquerda, como é típico na

⁶ Com uma parcial exceção no caso dos conectores, que mantêm o valor morfológico de conjunções ou locuções conjuncionais.

⁷ O ALL está selecionado em escuro na parte direita da figura.

ilocução de chamamento proximal. A variação da frequência fundamental do alocutivo, registrada em 75hz, é significativamente menor que a do COM de chamamento. A curva entonacional do alocutivo é apenas descendente (não possuindo, portanto, foco entonacional), iniciando-se em 250hz e executando um movimento de descida até 175hz.

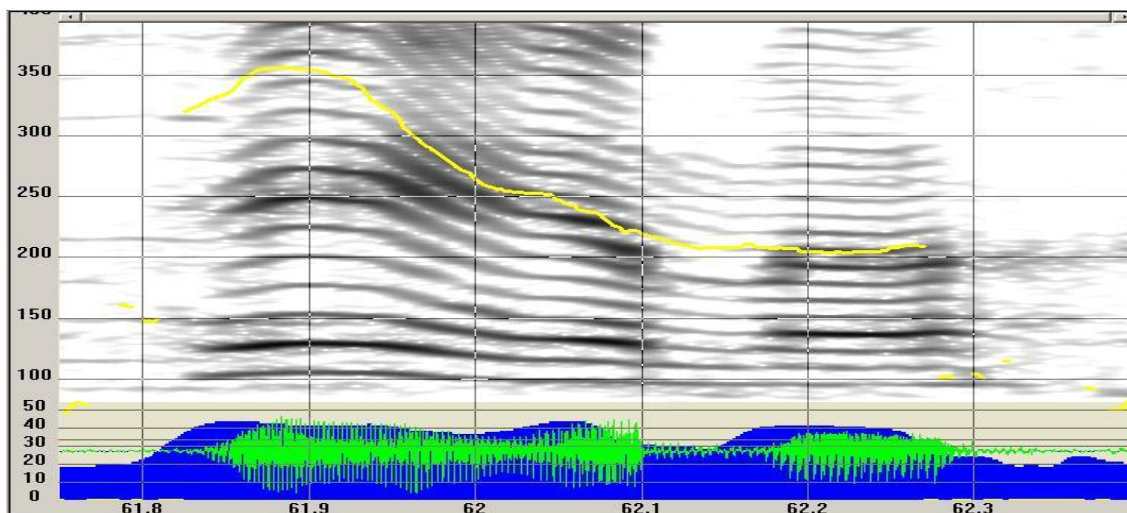
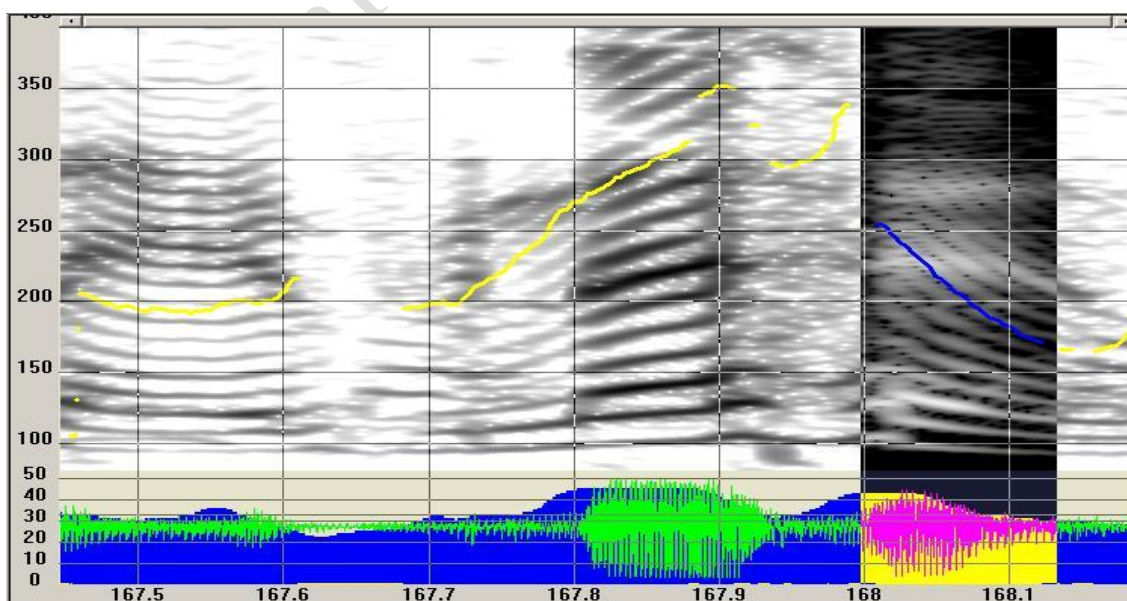


FIG. 7 ([som](#))

*FLA: *Rena* // =COM=

FIG. 8 ([som](#))



*FLA: *é uma neve* / =COM= *né* / =PHA= *Rena* // =ALC=

2.3. Coleta de dados e resultados

Para a análise comparativa dos alocutivos entre EE, I, PB e PE, foram analisados 75 textos de fala espontânea, cada um de cerca de 1500 palavras; desses, 15 para o EE, o I e o PB, e 30 para o PE, pois este último apresentou índices muito inferiores aos das outras línguas românicas. Tais textos foram distribuídos nas três tipologias interacionais previstas: 5 e 10 monólogos, 5 e 10 diálogos, 5 e 10 conversações de EE, I, e PB de um lado e PE de outro, respectivamente. Os dados relativos a I e PB se referem ao trabalho realizado por Raso e Goulart (2009).

Os resultados apresentados nas quatro línguas românicas possuem um grau estatístico de confiabilidade variável: maior quanto à frequência de *tokens*, menor quanto à frequência de *types*, que precisaria de uma confirmação com base em uma amostra maior.

2.3.1. As características dos textos e dos informantes

As características dos informantes e dos textos escolhidos, para cada língua, são as seguintes:

Número de participantes nas conversações: 24 para o EE, 18 para o I, 24 para o PB e 34 para o PE;

Número de participantes nos diálogos: 10 para EE, 12 para o I, 10 para o PB, 21 para o PE;

Número de participantes nos monólogos: 5 para o EE, 11 para o I, 10 para o PB e 19 para o PE;

Total de participantes: 39 para o EE, 41 para o I, 44 para o PB e 74 para o PE. Os números são, portanto, comparáveis, considerando-se que o PE apresenta uma quantidade de textos em dobro comparando-se com as outras línguas e que o EE simplesmente não sinaliza o interlocutor dos monologantes.

Sexo dos participantes: para o sexo feminino temos 18 informantes no EE, 16 no I, 23 no PB e 29 no PE; para o sexo masculino, 20 no EE, 14 no I, 18 no PB e 30 no PE. Em todas as línguas existe, portanto, um bom equilíbrio entre os dois sexos.

Idade dos participantes: 2 menores de idade para o PB e nenhum para as outras línguas; na faixa e 18 a 25 anos, 13 para o EE, 5 para o I, 18 para o PB e 22 para o PE; na faixa dos 26 aos 40 anos, 21 para o EE, 25 para o I, 8 para o PB e 23 para o PE; na faixa dos 41 aos 60 anos, 7 para o EE, 6 para o I, 11 para o PB e 24 para o PE; há também 1 informante com mais de 60 anos no I, 4 no PB e 2 no PE; 2 dos informantes do I e 2 do PE são de idade desconhecida. Isso mostra que todas as faixas etárias são representadas em todas as línguas, e sem grande divergência.

Nível de escolaridade dos participantes: sem estudo ou até primeiro grau completo 5 no EE, 9 no I, 10 no PB, 1 no PE; com segundo grau completo 16 no EE, 1 no I, 15 no PB, 3 no PE; com terceiro grau completo 18 no EE, 29 no I, 19 no PB, 69 no PE; para 2 informantes do I não é conhecido o grau de escolaridade. Nesse caso nota-se um certo desequilíbrio: o PE e, em medida muito menor o I, tendem a representar em medida privilegiada as faixas de escolaridade mais altas.

As situações comunicativas gravadas: nos textos dialógicos, seja entre duas ou mais pessoas, o EE representa um conversa entre amigos em uma festa, um diálogo entre parentes em uma loja, uma conversa entre amigos tomando café e 7 conversas entre amigos ou parentes sem ação nenhuma, ou em casa ou na universidade. O I representa uma conversa entre vendedor e clientes, uma entre mãe e filha, uma entrevista a um amigo durante almoço, uma entrevista a uma empregada durante a limpeza de casa, uma entrevista a um artesão, uma conversa entre amigos cozinhando, uma explicação de um jogo entre amigos, a discussão de um projeto entre alunos e professores e um bate-papo entre amigos. O PB representa 3 conversas entre amigos, duas amigas conversando enquanto guardam as compras, um engenheiro e um pedreiro em uma obra, três conversas durante jantar ou lanche, um grupo de faxineiras de um prédio conversando durante a pausa do almoço. O PE representa somente conversas entre amigos ou conhecidos em casa, ou, na maioria dos casos, na universidade, sem acionalidade alguma. O quadro comparativo mostra que se PB e I representam uma variação diafásica significativa, o EE, e principalmente o PE, representam situações repetitivas.

Posteriormente analisaremos os possíveis efeitos das características dos textos e dos informantes sobre os dados.

2.3.2. Tokens

As tabelas 1-4 mostram que o PB é a língua que mais se utiliza de alocutivos, apresentando uma quantidade de *tokens* superior àquela das demais línguas, seguido pelo EE, pelo I e por último pelo PE. Note-se que os dados do PE se referem a uma base de dados duas vezes superior às das outras línguas. Estatisticamente, portanto, seus valores, para os fins de comparabilidade com as demais línguas, deveriam ser divididos por dois.

Tabela 1

Conversaço

	Português Brasileiro	Espanhol Europeu	Italiano	Português Europeu
<i>Tokens</i>	81	43	26	8

Tabela 2

Diálogo

	Português Brasileiro	Espanhol Europeu	Italiano	Português Europeu
<i>Tokens</i>	38	12	16	6

Tabela 3

Monólogo

	Português Brasileiro	Espanhol Europeu	Italiano	Português Europeu
<i>Tokens</i>	14	7	1	6

Tabela 4

Total de Alocutivos

	Português Brasileiro	Espanhol Europeu	Italiano	Português Europeu
<i>Tokens</i>	133	62	43	20

Cabe justificar que o número alto de alocutivos no *corpus* do PE no quadro dos monólogos não pode ser considerado um fator de grande importância, pois dos 6 ALL, 5 são utilizados pelo mesmo falante em um único texto.

Proporcionalmente falando, a frequência de ALL encontrados em PB é mais de 2 vezes superior à do EE e mais de 3 vezes a do I. Mas o que impressiona é a distância com o PE, que apresenta uma frequência 13 vezes menor que a do PB, 6 vezes menor que a do EE e 4 vezes menor que a do I.

2.3.3. Types

As tabelas 5, 6 e 7 mostram os lexemas encontrados na amostra e suas quantidades, respectivamente para as interações conversacionais, dialógicas e monológicas. A tabela 8 oferece os dados totais.

Tabela 5

Conversação

	Português Brasileiro	Espanhol Europeu	Italiano	Português Europeu
Total de lexemas	9	10	4	5
Tipos de lexema	Nome próprio: 56	Nome próprio: 15	Nome próprio: 21	Ó / nome: 1
	Minha filha: 8	Hija mia: 2	Ragazzi: 3	Ó senhor doutor: 2
	Menina: 6	Rica: 1	Bella: 1	Ó senhora doutora: 2
	Senhora: 3	Hombre: 7	Nonna: 1	Senhor doutor: 2
	Mãe: 2	Mamá: 2		Ó senhor engenheiro: 1
	Gente: 2	Tia: 9		
	Bobo: 1	Cariño: 1		
	Moço: 1	Mi niña: 2		
	Filho: 2	Hija: 3		
	Tú: 1			

Observa-se que nas interações conversacionais houve a ocorrência de 81 *tokens* em PB, 43 em EE, 26 em I e 8 em PE. Proporcionalmente o PB apresenta uma ocorrência de 1,8 vezes a mais que o EE, 3,1 vezes maior que o I e 20,2 vezes maior que o PE. Quanto a variação de *types*, nessa mesma tipologia encontramos 9 possibilidades no PB, 10 no EE, 4 no I e 5 no PE (porém no dobro de textos). Os nomes próprios são considerados em conjunto como um *type* único.

É importante ressaltar, quanto aos dados do PE apresentados na tabela 5, que cinco dos seis diferentes tipos de lexemas foram enunciados pelo mesmo falante, na mesma gravação e em discurso reportado.

As interações conversacionais são aquelas que apresentam o maior número de alocutivos em todas as línguas, e este fato pode ser explicado pela constante necessidade de o interlocutor identificar para quem está passando a mensagem, já que dada a presença de mais interlocutores isso não é sempre situacionalmente dado. Nas interações conversacionais aparecem, portanto, ambas as funções do ALL: identificar o interlocutor e sinalizar a coesão social.

Tabela 6

Diálogo

	Português Brasileiro	Espanhol Europeu	Italiano	Português Europeu
Total de lexemas:	6	4	2	6
Tipos de lexema:	Nome próprio: 31	Nomes próprios: 1	Nomes próprios: 7	Nome próprio: 1
	Minha filha: 2	Tía: 3	Mamma: 6	Ó amigo: 1
	Filha: 2	Hija: 1	Te: 1	Ó / nome próprio: 2
	Gente: 1	Hombre: 7		Ó minha senhora: 1
	Moço: 2			Ó dona Laura: 1

Nas interações dialógicas o PB apresentou 38 *tokens*, o EE 12, o I 13 e o PE 6. Proporcionalmente o PB apresenta uma ocorrência 3 vezes maior do que o EE, 3 do que o I e 12,6 vezes maior do que o PE.

Nessa mesma tipologia foram encontrados 6 *types* no PB, 4 no EE, 2 no I e 6 no PE. Proporcionalmente o PB apresenta uma ocorrência 1,5 vezes maior que o EE, 3 vezes maior que o I e 2 vezes maior que a de PE.

Nos diálogos, o número de ALL encontrados é significativamente menor que na conversação, pois, neste caso, não há a necessidade de diferenciar para quem se está direcionando a mensagem. Como essa função em princípio não é necessária, os ALL servem para desempenhar uma função afetiva entre participantes da interação que, através deles, reiteram certo grau de coesão na interação.

Tabela 7

Monólogo

	Português Brasileiro	Espanhol Europeu	Italiano	Português Europeu
Total de lexemas:	4	2	1	3
Tipos de lexema:	Nome próprio: 10	Nomes próprios: 4	Nomes próprios: 1	Man: 4
	Minha filha: 1	Hombre: 3		Minha querida senhora: 1
	Senhora: 2			Ó mãe: 1
	Meu filho: 1			

Nas interações monológicas foram encontrados 14 *tokens* no *corpus* do PB, 7 no EE, 1 no I e 6 no PE. No caso do PE, 5 dos 6 ALL foram encontrados no mesmo texto.

O número total de *types* lexicais, somando-se os textos das três diferentes tipologias, é de 9 para o PB, 10 para o EE, 5 para o I e 10 para o PE. No caso do PE, como já foi dito, a base de cálculo é duas vezes aquela das outras línguas. No caso dos *types*, contrariamente aos *tokens*, não podemos estabelecer uma proporção dividindo por dois os resultados obtidos. O mais justo parece considerar como proporção algum ponto dentro de um intervalo entre 5 e 10 *types*. É interessante notar que, enquanto a ocorrência de *tokens* no PB é muito mais alta do que nas outras línguas (respectivamente 2, 3 e 13 vezes maior), o mesmo não acontece ao compararmos a quantidade de *types*: em particular, o EE apresenta a mesma variação de *types* do PB, e o PE, que apresentava uma distância tão grande das outras línguas quanto à ocorrência dos ALL, mostra em números absolutos a mesma variação do que o PB. Além disso, a variação de *types* em PE seria ainda maior se considerássemos como *types* diferentes os casos em que o alocutivo é precedido por *ó*. Mesmo considerando que a base de cálculo é diferente, o resultado do PE é surpreendente, e uma explicação deve ser tentada.

Note-se também que em todas as línguas a ocorrência de ALL de nomes próprios é alta em relação aos outros lexemas utilizados, porém esse número é mais destacado no PB e no I.

Tabela 8

Total de Lexemas							
PB		EE		I		PE	
Nomes	97	Nomes	20	Nomes	29	Nomes	1
Minha filha /meu filho	12	Hija mía	2	Ragazzi	3	Ó senhor doutor/ senhora doutora	4
Menina/o	6	Tía	12	Bella	1	Ó / nomes	3
Senhora	5	Hombre	17	Nonna	1	Minha querida senhora	1
Mãe	2	Mamá	2	Mamma	6	Ó mãe	1
Gente	3	Rica	1			Man	4
Filho/a	4	Hija	4			Ó amigo	1
Bobo	1	mi niña	2			Senhor doutor	2
Moço	3	Cariño	1			Ó senhor engenheiro	1
		Tú	1				
Total	133	Total	62	Total	40	Total	18

2.4.O discurso reportado

Uma ajuda para a compreensão da situação apresentada até agora pode vir da análise dos ALL realizados em discurso reportado, onde o uso do ALL é importante para situar o interlocutor dentro da situação fictícia que é narrada, fornecendo, através do ALL, a informação do interlocutor. A tabela 9 mostra o quadro de ocorrências:

Tabela 9

Alocutivos em discurso reportado

	Conversação	Diálogo	Monólogo	Total
PB	0	6	4	10

EE	2	0	3	5
I	0	6	0	6
PE	7	6	1	14

Proporcionalmente falando, o PB apresentou 2 vezes mais ALL em discurso reportado do que o EE, 1,6 vezes do que o I e 1,4 vezes do que o PE. Portanto, comparando os dados gerais com aqueles relativos ao discurso reportado, se confirma a distância com o EE, que era a menor, diminui aquela com o I, e se reduz drasticamente a distância entre PB e PE. Se em geral o PB apresenta 13 vezes mais ALL do que o PE, em discurso reportado essa distância se reduz a apenas 1,4 vezes.

2.5. Posição dos alocutivos nos enunciados

Sabe-se que os ALL podem ser encontrados em qualquer posição nos enunciados, mas nunca se estudou se línguas diferentes preferem posições diferentes. A tabela 10 mostra os dados gerais, enquanto as tabelas 11-13 mostram os dados diferenciados por tipologia interacional.

Tabela 10
Posição no enunciado

Quadro geral										
	PB		EE		I		PE		Total	
Início	7	5%	14	22%	11	27,5%	6	30%	38	15%
Meio	44	33%	22	35%	10	25%	3	15%	79	31%
Fim	82	61%	26	42%	19	47,5%	11	55%	138	54%

A posição final é, portanto, a preferida em todas as línguas, mas principalmente nas duas variedades de português. O PB, em particular, parece evitar de maneira muito marcada a posição inicial. Ao contrário, EE e I parecem diferenciar menos entre as posições.

Os dados das tabelas 11-13 mostram que essa variação pode estar relacionada, pelo menos parcialmente, com o tipo de interação em que os participantes estão envolvidos.

Nas interações conversacionais, é mais comum que os ALL apareçam no final dos enunciados nas quatro línguas analisadas. Já nas interações dialógicas, somente o PB mostra uma clara preferência pela posição final. Nas interações monológicas, a posição interna de enunciado parece adquirir um peso maior.

Tabela 11

Posição no enunciado

Conversaão										
	PB		EE		I		PE		Total	
Início	3	4%	7	17%	7	26%	1	12%	18	11%
Meio	28	35%	16	37%	5	20%	1	12%	50	32%
Fim	49	61%	20	46%	14	53%	6	75%	89	57%

Nas conversações os ALL aparecem no final dos enunciados em mais de 60% dos casos em PB, mais de 45% em EE, quase 55% dos casos em I e 75% no PE. Somando-se as três línguas, são encontrados 57% dos ALL no final dos enunciados, 32% no meio e apenas 11% no início.

Tabela 12

Posição no enunciado

Diálogo										
	PB		EE		I		PE		Total	
Início	4	11%	6	50%	4	31%	4	66%	18	25%
Meio	8	19%	1	9%	4	31%	0	0%	13	18%
Fim	29	70%	5	41%	5	38%	2	34%	41	57%

Nas interações dialógicas os resultados não são tão homogêneos como no caso das conversações. Nos diálogos, os ALL aparecem em final dos enunciados em 70% dos casos no PB e em 38% no I. Já no EE e no PE, a maioria dos ALL aparece no início dos enunciados: 50% dos casos do EE e 66% dos casos no PE. Mas os números, com exceção do PB, são pequenos demais para se chegar a conclusões.

Tabela 13

Posição no enunciado

Monólogo										
	PB		EE		I		PE		Total	
Início	0	0%	1	15%	0	0%	1	17%	2	8%
Meio	8	66%	5	71%	1	100%	2	33%	16	62%
Fim	4	44%	1	15%	0	0%	3	50%	8	30%

Quanto aos monólogos, a posição interna de enunciado parece ser a preferida, pelo menos em PB e EE. Os números de I e PE não permitem chegar a uma conclusão.

3. CONSIDERAÇÕES FINAIS

Este estudo permite, a nosso ver, chegarmos a algumas conclusões e formularmos algumas hipóteses que devem ser averiguadas analisando-se uma base de dados maior. Parece certo que os ALL estão presentes em línguas/culturas diferentes com uma frequência que pode variar muito. O PB apresenta, sem dúvida, uma ocorrência muito mais alta de ALL do que as outras línguas e, principalmente, muito mais alta do que a variedade portuguesa europeia. Este resultado é extremamente significativo do quanto o contexto cultural e sociolinguístico influencia essa unidade informacional específica. Os dados mostrados podem, portanto, ser utilizados para estudos de ordem sociolinguística.

De fato a unidade de ALL possui duas funções diferentes: identificar o interlocutor e marcar a coesão social com ele, através de títulos, adjetivos, epítetos ou simplesmente o nome próprio. É lógico pensar que as diferenças entre as línguas não se manifestem tanto na primeira das duas funções. É provável que quando existe a necessidade de se identificar o interlocutor, o ALL, que parece o recurso natural para isso, além da ilocução de chamamento, seja utilizado com uma frequência parecida nas quatro línguas. Isso parece confirmado também observando-se os dados relativos ao discurso reportado, onde a função identificativa adquire um peso muito maior. Nesse contexto as proporções gerais mudam, e o PE não parece se comportar de maneira diferente das outras línguas. Essa consideração deve ser verificada com um estudo específico sobre os ALL em discurso reportado. Nesse estudo os dados sobre o discurso reportado são pequenos para se obter confiabilidade estatística, e não foi verificado que os discursos reportados nas diferentes línguas fossem comparáveis quantitativamente.

Outras considerações podem ser feitas a partir dos dados sobre a variação lexical. Se comparada com a diferença de frequência dos *tokens*, a variação de *types* não parece grande. Não se registra alguma variação em número de *types* entre PB e EE, e mesmo a distância com o I, e especialmente o PE, se reduz muito. Isso nos deixa hipotetizar que se algumas línguas usam mais ou menos o ALL, isso não se reflete imediatamente na riqueza de recursos que cada uma delas possui para realizá-lo. Isso é especialmente evidente na comparação entre PB e PE. O PE parece ter substancialmente a mesma variação de *types* que o PB, apesar de usar muito menos o ALL. Parece, então, que o código pesa mais quanto aos recursos lexicais para realizar a função e, a cultura, mais quanto à frequência com que a função é concretamente realizada. Mas o tipo de lexema preferido reflete também diferenças culturais interessantes: boa parte da superioridade quantitativa do PB é devida à enorme presença de nomes próprios, quase completamente ausentes em PE. Os nomes próprios parecem ter um papel muito grande também em italiano, onde constituem quase $\frac{3}{4}$ de todos os ALL. A impressão de uma maior formalidade nas escolhas lexicais do PE deve ser verificada com uma base de dados maior, pois pode ser devida aos efeitos de dois textos específicos.

Mais considerações podem ser retiradas observando-se os resultados relativos à posição do ALL no enunciado. Todas as línguas parecem preferir a posição final, mas, além disso, subsistem claras diferenças. Podemos afirmar que o PB é mais rígido na escolha da posição, que é final em 61% dos casos e inicial somente 5%. O PE também prefere claramente a posição final (55% dos casos) mas parece evitar, principalmente, a posição interna (somente 15% dos casos). Ao contrário, o EE e o I parecem ter maior liberdade de posição.

Finalmente, as diferenças relativas às diversas tipologias nos sugerem outras considerações: não deve surpreender que, em todas as línguas, os ALL ocorram mais nas conversações. É somente nessa tipologia, de fato, que a função identificativa se manifesta (a não ser nos casos de discurso reportado), já que o interlocutor é automaticamente dado em diálogos e monólogos. Analogamente, não surpreende que os monólogos apresentem menos ALL, dada a forte redução da interatividade nessa tipologia.

Antes de concluir o trabalho é necessário discutir se os dados diferentes para as várias línguas podem refletir os aspectos dos corpora que não são imediatamente comparáveis, observados na seção 2.3.1. As diferenças entre os corpora são basicamente duas: a pouca variação diafásica representada nos corpora de EE e PE, e o fato de que os

corpora de I e PE privilegiam (principalmente o PE) a representatividade de falantes com escolaridade alta. Quanto à primeira diferença, não parece ser relevante para os nossos fins, já que o fato de o EE ser diafasicamente pouco diferenciado (e de ser nesse aspecto muito parecido com o PE) não impede que o EE e o PE mostrem comportamentos muito diferentes quanto à ocorrência da unidade de ALL. A baixíssima presença de ALL em PE não pode, portanto, ser imputada à escolha situacional; uma vez que, se assim fosse, não poderíamos explicar a alta frequência em EE. Quanto ao segundo aspecto, poder-se-ia dizer que I e PE apresentam uma baixa ocorrência de ALL devido ao fato de que ambos os corpora privilegiam falantes de alta escolaridade. Essa hipótese, contudo, não se sustenta se consideramos que quando aumenta a porcentagem de escolaridade alta do PE e no I (nas conversações) não há diminuição na porcentagem de ALL, e que nos monólogos, onde os falantes de PE são praticamente todos de escolaridade alta, o PE consegue a maior ocorrência de ALL em comparação às outras línguas, enquanto o I, com um número de falantes de baixa escolaridade superior às outras línguas, registra uma ocorrência de ALL baixíssima. Além disso, note-se que nos diálogos onde o PB apresenta uma porcentagem de falantes com alta escolaridade superior às outras tipologias, o número de ALL é 3 vezes aquele do EE, muito superior à média. A porcentagem de ALL não parece depender do nível de escolaridade. No caso do EE, por exemplo, registra-se a maior porcentagem de ALL no nível de escolaridade média (55% dos ALL em 37% dos falantes).

Concluindo, podemos afirmar que este estudo mostra claramente que o ALL é um excelente indicador sociolinguístico a ser explorado em pesquisas desse cunho. Através dele é possível apreciar diferenças culturais no uso funcional de um recurso comum a diferentes línguas, mas também a códigos iguais em contextos culturais diferentes, como é o caso do PB e do PE. Além disso, o fato de que a única língua românica usada em uma cultura não europeia, entre as quatro aqui estudadas, mostre as características mais distantes daquelas das outras línguas (e em muitos aspectos opostas à variedade europeia do mesmo código), reforça o peso do contexto sociolinguístico para o uso dessa unidade informacional.

REFERÊNCIAS BIBLIOGRÁFICAS

ALVES DE DEUS, L. *A Unidade Informacional de Tópico no Português do Brasil*. 2008. 230f. Dissertação (Mestrado em Estudos Linguísticos) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2008.

AUSTIN, J. *How to do things with words*. London: Oxford University Press, 1962.

CORSI, G. 2009. *L'Introduttore Locutivo: una ricerca corpus-based di Italiano parlato informale*. 2009. Monografia (Laurea Triennale) – Facoltà di Lettere e Filosofia, Università degli studi di Firenze, Firenze, 2009.

CRESTI, E. *Corpus di italiano parlato*. Firenze: Accademia della Crusca, 2000.

CRESTI, E.; FIRENZUOLI, V. L'articolazione informativa topic-comment e comment-appendice: correlati intonativi. In: REGNICOLI, A. (Ed.). *La fonetica acustica come strumento di analisi della variazione linguistica in Italia*. Atti delle XII giornate del Gruppo di Fonetica Sperimentale (GFS). Roma: Il Calamo, 2002, p. 153-166.

CRESTI, E. La Stanza: un'unità di costruzione testuale del parlato. In: *Atti del X Congresso della Società Internazionale di Linguistica e Filologia Italiana, SILFI 2008*. Basilea, 30.06-03.07 2008. 2009, p. 1-25.

CRESTI, E.; MONEGLIA, M. *C-ORAL-ROM: Integrated Reference Corpora for Spoken Romance Languages*. Amsterdam-Philadelphia: John Benjamins, 2005.

CRESTI, E.; MONEGLIA, M. *Transcrição e análise da fala com base na Teoria da Língua em Ato*. Belo Horizonte: Faculdade de Letras da UFMG, 2008. Workshop.

FIRENZUOLI, V.; TUCCI, I. Una strategia di costruzione del testo parlato: l'introduttore locutivo. In: *Atti del Convegno "Il parlato italiano"*, no prelo.

FROSALI, F. Le unità di informazione di ausilio dialogico: valori percentuali, caratteri intonativi, lessicali e morfo-sintattici in un corpus di italiano parlato (C-ORAL-ROM). In: CRESTI, E. (Org.). *Prospettive nello studio del lessico italiano*. Firenze University Press, 2008, p. 417-424.

GIANI, D. Le discours directe rapporté dans l'italien arlé e écrit. Em: SCARANO: Antonietta (Org.). *Macrosyntaxe et Pragmatique: l'analyse de l'oral*. Roma: Bulzoni, 2003: 203-213.

GIANI, D. Una strategia di costruzione del testo parlato: l'introduttore locutivo. Em: ALBANO LEONI, F. (org). *Atti del congresso "Il parlato italiano"*. Napoli: D'Auria, 2004: 84-97.

JESUS, A. C. U. *A unidade de apêndice no português do Brasil*. 2008. 230 f. Dissertação (Mestrado em Estudos Linguísticos) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2008.

MAIA ROCHA, B.; RASO, T. *Análise da unidade informacional de introdutor locutivo no português brasileiro*, em preparação.

MAIA ROCHA, B.; RASO, T.; ANDRADE, M. I. Alguns auxílios dialógicos em italiano, português do Brasil e em italianos cultos em contato prolongado com o português do Brasil. *Fragmentos*,(UFSC), v.35, 2009. p.205-218.

MONEGLIA, M.; CRESTI, E. L'intonazione e i criteri di trascrizione del parlato adulto e infantile. In: BORTOLINI, U., PIZZUTO, E. *Il Progetto CHILDES Italia*. Pisa: Del Cerro, 1997, p. 57-90.

MOTA, B. A. *A unidade informacional de Parentético na fala espontânea do espanhol*. 2010. 92 f. Dissertação (Mestrado em Estudos Linguísticos Neolatinos), Faculdade de Letras – Universidade Federal do Rio de Janeiro.

OLIVEIRA, C. J. F. Apêndice de Comentário e Comentários Ligados: uma distinção à luz da Teoria da Língua em Ato. *Revista Eletrônica de Divulgação Científica em Língua Portuguesa, Linguística e Literatura*, ano 5, n. 10, 2009.

RASO, T.; GOULART, L. Estudo contrastivo do uso de alocutivos em português brasileiro e italiano. *Fragments*, (UFSC), 2009.

RASO, T.; MELLO, H., Parâmetros de compilação de um corpus oral: o caso do C-ORAL-BRASIL, Em: Veredas, 2009, p. 20-35.

RASO, T.-MELLO, H. The C-ORAL-BRASIL corpus. In: Moneglia, M.-Panunzi, A., (orgs.) *Bootstrapping Information from Corpora in a Cross Linguistic Perspective*. Firenze University Press, 2010.

RASO, T.; MELLO, H. As especificidades da unidade de tópico em PB e possíveis efeitos do contato linguístico. In: SARAIVA, E.; CHAVES MARINHO, J. *Estudos da língua em uso: da gramática ao texto*. Belo Horizonte: Faculdade de Letras, UFMG, no prelo.

RASO, T. - ULISSES, A. Tópico e Apêndice no português do Brasil: algumas considerações. In: *Revista de estudos da linguagem*. 2008.

SIGNORINI, S. Il Topic: criteri di identificazione e correlati morfosintattici in un corpus di italiano parlato. In: LEONI, F. A.; CUTUGNO, F.; PETTORINO, M.; SAVY, R. (Eds.). *Atti del Convegno Nazionale "Il parlato italiano"*. Napoli: M. D'Auria, 2004a, p. 15-39.

SIGNORINI, S. L'unità di topic: caratteristiche e frequenza in un corpus di italiano parlato. Il topic complesso. In: D'ACHILLE, P. (Ed.). *Generi, architetture e forme testuali*. Atti del VII Convegno Internazionale SILFI. Firenze: Franco Cesati, 2004b, p. 227-238.

SCHNEIDER, S. *Reduced Parenthetical Clauses as Mitigators: a corpus study of spoken French, Italian and Spanish*. Amsterdam-Philadelphia: John Benjamins, 2007.

TENANI, L. E. *Análise Prosódica das Inserções Parentéticas no Corpus do Projeto da Gramática do Português Falado*. 1996. Dissertação de Mestrado – Universidade Estadual de Campinas, Campinas, 1996a.

TENANI, L. E. Marcas prosódicas das inserções parentéticas. *Estudos Linguísticos*. Taubaté, 1996b. v. 25, p. 803-808.

TENANI, L. E. O fenômeno de parentetização na fala. *Estudos Linguísticos*. São Paulo, 1997. v. 26, p. 787-793.

t'HART, J. – COHEN, A. – COLLIER, R. *A perceptual study on intonation: an experimental approach to speech melody*. Cambridge: Cambridge University Press, 1990.

TUCCI, I. "Obiter dictum": la funzione informativa delle unità parentetiche. In: *Atti del Convegno Internazionale G.S.C.P. "La comunicazione parlata"*. Napoli, 23-25.02 2009.

TUCCI, I. L'inciso: caratteristiche morfosintattiche e intonative in un corpus di riferimento. In: LEONI, F. A. *et al.* (Orgs.). *Atti del Convegno "Il parlato italiano"*. Napoli, 13-15.02 2003. Napoli: M. D'Auria, 2004, p. 1-14.

TUCCI, I. *L'espressione lessicale della modalità nel parlato spontaneo: dati dal corpus C-ORAL-ROM italiano*. PhD diss. Università di Firenze, Dipartimento di Linguistica, 2007.

VALE, H. P. *A unidade informacional de Parentético no português do Brasil: uma análise baseada em corpus*. 2010. 149 f. (Mestrado em Estudos Linguísticos) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2009.