

**“I THINK THIS QUESTION IS VERY RELEVANT NOWADAYS...”:
INVESTIGANDO CATEGORIAS E PADRÕES DE MODALIDADE NA
ESCRITA DE APRENDIZES BRASILEIROS DE LÍNGUA INGLESA**

LEONARDO RECKSI RAMOS*
(Doutorando - UFSC)

RESUMO

Este artigo emprega corpora de aprendizes armazenados em formato eletrônico para comparar a variedade e frequência de algumas expressões modais na escrita de aprendizes universitários de inglês como língua estrangeira (EFL) e na escrita de falantes nativos de inglês. Embora o foco principal da investigação recaia sobre a escrita de aprendizes de EFL brasileiros, o autor freqüentemente utiliza corpora de EFL de aprendizes espanhóis e tchecos para fins comparativos, na tentativa de determinar se certas características da escrita dos aprendizes brasileiros provavelmente advêm de interferências da língua materna, ou são, em grande parte, compartilhadas por aprendizes de EFL de outras nacionalidades. O estudo é baseado em quatro sub-corpus de 33,000 palavras e leva em consideração não apenas verbos auxiliares modais, mas também uma grande variedade de dispositivos modais tais como verbos lexicais e advérbios modais. Os resultados revelam um emprego demasiado de expressões modais por parte dos aprendizes de EFL, tendência esta que pode ser atribuída ao atual estágio de desenvolvimento desses aprendizes, e parcialmente a uma característica da interlíngua. O estudo também mostra evidências de interferências de registro, onde os aprendizes parecem transferir padrões de uso do inglês falado para a sua escrita, e particularmente um alto grau de sensibilidade ao tópico no uso de certos modais. O artigo conclui discutindo a necessidade de oferecer a esses aprendizes uma variedade maior de expressões modais que incluam padrões de orações e frases modais mais extensas.

Palavras-chave: lingüística de corpus, modalidade, escrita acadêmica.

ABSTRACT

This article uses computer learner corpora to compare the variety and frequency of some modal expressions in the writing of university-level EFL students and native speakers. Even though the prime focal point of the investigation is Brazilian EFL writers, the author recurrently relies on comparisons with Spanish and Czech EFL writers in an effort to determine whether certain characteristics of Brazilian EFL writing are likely to stem from mother tongue interference, or are, by and large, shared by EFL writers of different language settings. The study is based on four 33,000-word sub-corpora and takes into account not only modal auxiliary verbs but also a broader variety of modal devices, such as lexical verbs and adverbs with a modal value. Results reveal an overall overuse of modal expressions by all EFL writers, a propensity which may be partially developmental, and partially interlingual. The study also discovers evidence of register interference, where the learners appear to transfer patterns of use from spoken English into their writing, and particularly a high-degree of topic sensitivity in the use of particular modals. It concludes by arguing for the necessity to offer learners a broader variety of modal expressions including larger sentence patterns and modal phrases.

Key-words: corpus linguistics, modality, academic writing.

**Agradecimentos:* Gostaria de agradecer os comentários e sugestões de dois pareceristas anônimos, bem como a minha orientadora e amiga Dra. Viviane Heberle. Quaisquer contradições ou problemas que persistam são de minha inteira responsabilidade.

INTRODUÇÃO

Inúmeros estudos (Horvath, 2001; Leech, 1998; Granger, 1998a) evidenciam que uma das principais contribuições de corpora de aprendizes é a possibilidade que oferecem para a investigação de aspectos específicos de sua linguagem que anteriormente eram difíceis ou impossíveis de serem analisados. Em muitos casos, essas investigações são comparativas: a produção de aprendizes de língua inglesa é comparada à produção de falantes nativos, processo este que se enquadra dentro da análise contrastiva de línguas (cf. Granger, 1998b). A modalidade constitui apenas uma área da linguagem desses aprendizes que ainda não foi substancialmente descrita e onde corpora armazenados em formato eletrônico podem trazer uma contribuição, como este artigo pretende demonstrar.

Diversos estudos mostram que aprendizes de EFL parecem ter dificuldades não só em usar verbos modais apropriadamente, mas também tendem a utilizar certos verbos de forma excessiva ou restrita (DeCarrico, 1986; Hinkel, 1995).

Este artigo procura comparar expressões, significados e usos de modalidade em composições produzidas por aprendizes brasileiros de língua inglesa e falantes nativos de inglês. Os corpora são analisados através do uso de concordâncias, o que permite determinar a frequência e a variedade das expressões modais. Espera-se que a descrição do uso de categorias e padrões de expressões modais exemplificada neste estudo e obtida através de uma abordagem com base em corpora tenha importantes desdobramentos pedagógicos.

1. UMA ABORDAGEM BASEADA EM CORPORA PARA O ESTUDO DE MODALIDADE

1.1. Contrastando interlínguas

Pelo menos dois tipos de comparações podem emergir de uma metodologia baseada na análise contrastiva de línguas: (i) a comparação de uma língua nativa com outras interlínguas; e (ii) a comparação de diferentes interlínguas entre si.

O objetivo do primeiro tipo de comparação é procurar estabelecer algumas das características distintas de uma determinada interlíngua. Essa abordagem possibilita a investigação do uso excessivo ou restrito (não apenas erros ou desvios) de itens lingüísticos, o que, por sua vez, permite revelar certas diferenças entre padrões de distribuição da linguagem desses aprendizes em relação à dos falantes nativos. Esses diferentes padrões de distribuição podem nos ajudar a explicar porque um texto talvez não contenha erros gramaticais ou lexicais explícitos, mas dê a impressão de não ter sido escrito por um falante nativo.

O segundo tipo de comparação, envolvendo diferentes variedades de interlínguas, possibilita a investigação de estratégias compartilhadas por todos os aprendizes, ou por vários grupos de aprendizes, assim como aquelas que parecem ser características de um determinado grupo. Entretanto, é importante que não cheguemos a conclusões apressadas e superficiais ao

argumentarmos sobre a universalidade das estratégias empregadas por esses aprendizes, baseando-nos apenas nas evidências provenientes de um corpus. Na maioria dos casos, os resultados precisam ser investigados através do uso de diferentes tipos de dados e diferentes populações de aprendizes antes que generalizações convincentes possam ser feitas. Quando a análise de um corpus oferece possíveis evidências de transferências da língua materna, então, uma análise contrastiva com base em outros corpora constitui uma maneira importante de complementar as aparentes evidências encontradas em apenas um corpus.

Este estudo emprega corpora não anotados¹, o que envolve uma análise manual em função da necessidade de clarificação de ocorrências que sejam ambíguas (por exemplo, clarificações entre o substantivo e o verbo modal *can*). Análises lingüísticas totalmente automatizadas requerem corpora anotados. Ao anotarmos um corpus é possível, por exemplo, extrair automaticamente todos as ocorrências de verbos modais; evitando desta forma, homônimos. Entretanto, este artigo analisa uma variedade de expressões modais empregadas pelos aprendizes e não apenas verbos modais (Perkins, 1983; Hoyer, 1997). Isto, por sua vez, torna praticamente impossível a automatização de buscas em função da grande diversificação de tais expressões. Modalidade pode, por exemplo, ser expressa através de advérbios equivalentes (*possibly, likely*) e verbos lexicais (*I believe, I think*). O fato de aprendizes terem um amplo repertório de dispositivos para a expressão modalidade ao seu dispor significa que um método envolvendo uma combinação de análises manual e automatizada parece ser mais apropriado.

Devido ao aparente consenso entre lingüistas acerca da centralidade de verbos modais em inglês, esta foi a categoria escolhida como o ponto de partida para minha investigação. Os resultados obtidos foram comparados a informações oferecidas na literatura. Mindt (1995), por exemplo, publicou uma gramática baseada em corpora que oferece informações sobre a freqüência e a distribuição de verbos modais. Esta gramática oferece uma análise detalhada da ocorrência de estruturas modais e descreve as características semânticas de tais expressões em termos quantitativos embasados em vários corpora contemporaneamente disponíveis.

Outros estudos baseados em corpora envolvendo a variedade, freqüência e distribuição contextual de expressões modais enfatizam o fato de que modalidade notoriamente representa dificuldades para os aprendizes. Holmes (1988) investigou a freqüência relativa de 350 itens lexicais utilizados para expressar certeza e incerteza em corpora escritos e falados e empregou os dados para avaliar como a modalidade epistêmica é tratada em diferentes livros de língua inglesa para aprendizes. Os resultados oferecidos por ela indicam que muitos desses livros dedicam grande parte de sua atenção a verbos modais, negligenciando estratégias lingüísticas alternativas para a expressão de certeza e incerteza (Holmes, 1988, p.40).

¹ Corpora não anotados não contêm informações sobre a classe gramatical de cada palavra. Por exemplo, em um corpus sintaticamente anotado através do CLAWS tagger o verbo modal *can* seria anotado como <w VM>*can*, ao passo que o substantivo *can* seria anotado como <w NN1>*can*, onde VM = verb modal e NN1 = substantivo no singular.

Ao compararmos os padrões utilizados por falantes nativos aos dos aprendizes é possível examinar se estes optam por estilos (in)apropriados no emprego de suas expressões modais. Por exemplo, é possível que aprendizes de língua inglesa utilizem expressões modais que se aproximem mais de um registro informal e oral e que falantes nativos sejam mais formais em suas escolhas lingüísticas. O Michigan Corpus of Academic Spoken English (MICASE) (Simpson, et al., 2002) é um ótimo recurso para testarmos hipóteses sobre a possível influência de um registro oral na expressão de modalidade destes aprendizes. O MICASE contém amostras representativas do que pode ser considerada a linguagem de um típico ambiente universitário americano. O corpus contém 1,7 milhões de palavras e cobre uma variedade de tópicos e eventos, além de representar vários graus de formalidade.

Hoye (1997) empregou dados extraídos de corpora de transcrições de fala e escritos em seu estudo sobre combinações entre advérbios e verbos modais em inglês. A fim de comparar falantes nativos de inglês com aprendizes de espanhol, ele conduziu um experimento onde foi solicitado que os falantes nativos completassem um teste inserindo palavras ou frases com valor modal em um texto. Quando o mesmo teste foi empregado com aprendizes de espanhol, Hoye notou que esses consistentemente não perceberam o potencial combinatório de verbos modais e advérbios. Hoye comenta que “existem pontos de contraste e equivalência entre L1 e L2 que podem levar a uma ‘transferência negativa’, ou interferência, e isto prejudica potencialmente o nível de performance dos aprendizes” (Hoye, 1997, p.251, minha tradução).

1.2. Corpora empregados neste estudo

Corpora de aprendizes ainda estão em estágios preliminares, fato surpreendente, haja vista o potencial que oferecem para estudos na área de aquisição de línguas estrangeiras. O corpus utilizado neste artigo é o *International Corpus of Learner English* (ICLE) (para uma descrição detalhada consulte Granger, 1998b). O corpus de aprendizes possui vários sub-corpora representando o inglês produzido por aprendizes de EFL de diferentes nacionalidades. Os sub-corpora empregados neste estudo são o componente brasileiro (Br-ICLE)², que constitui o foco principal de minha análise, e os componentes espanhol (Sp-ICLE) e tcheco (Cze-ICLE). Os sub-corpora espanhol e tcheco contém cerca de 200.000 palavras de textos argumentativos escritos por alunos universitários cursando o terceiro ou quarto anos de instrução no curso de Inglês. O componente brasileiro possui apenas 33754 palavras, mas o grau de escolaridade e o tipo de escrita podem ser considerados análogos ao dos outros dois corpora.

Como a interpretação dos significados e padrões de expressões modais nos textos requer uma análise muito detalhada e o tamanho do sub-corpus brasileiro é limitado, este estudo é baseado em amostras de aproximadamente 33000 palavras extraídas dos outros dois corpora e do componente brasileiro em sua totalidade (vide Tabela 1). Os corpora de

² Para informações mais detalhadas sobre o projeto Br-ICLE consulte Sardinha (2000) ou acesse www.lrecski.hpg.com.br.

aprendizes foram contrastados com o *Louvain Corpus of Native English Essays* (LOCNESS)³. Como o LOCNESS possui apenas escrita não profissional, comparações também serão feitas utilizando-se outro corpus padrão de escrita de falantes nativos – o Freiburg-Brown Corpus of American English (FROWN)⁴ – mais especificamente as sessões de *Academic Prose* (Sessão J) e *Editorial* (Sessão B). Adicionalmente, alguns resultados serão comparados com os resultados oferecidos no *Longman Grammar of Written and Spoken Language* (LSWE) (Biber et al., 1999).

É necessário que se faça um comentário sobre os tópicos das composições desses aprendizes. A variedade de tópicos argumentativos oferecidos aos aprendizes brasileiros envolve temas tais como a capacidade do sistema penitenciário brasileiro de reformular a conduta de seus carcerários, a capacidade das universidades brasileiras de preparar seus alunos para as demandas do mercado de trabalho e a televisão como sendo o ‘ópio’ das massas. Os alunos universitários britânicos e americanos, por outro lado, escreveram sobre uma variedade de tópicos literários e argumentativos. Os tópicos literários envolvem resenhas críticas às obras de Sartre, Victor Hugo, Voltaire, Camus etc. Os tópicos argumentativos, que constituem cerca de um quarto do corpus de falantes nativos (Granger, comunicação pessoal), envolvem tópicos relacionados à situação econômica, política e social de diversos países da Europa.

Tabela 1. Tamanho dos sub-corpora utilizados para a análise de expressões modais

Corpora	Número de palavras
Br-ICLE	33754
Sp-ICLE	33519
Cze-ICLE	33356
LOCNESS	33361

Embora as amostras da produção escrita de falantes nativos e dos aprendizes de língua inglesa possam ser consideradas suficientemente similares para serem utilizadas na análise da variedade de formas empregadas pelos aprendizes na expressão de modalidade, os diferentes tópicos e, acima de tudo, os diferentes tipos de textos (argumentativos versus literários), requerem que os resultados quantitativos sejam interpretados com certa cautela, como sugere Dagneaux (1995, citada em Granger, 1998a), em um estudo mais detalhado sobre o assunto. Em sua investigação sobre emprego de expressões modais por aprendizes franceses, Dagneaux descobriu que escritores nativos utilizaram uma quantidade maior de verbos modais para expressar probabilidade e possibilidade do que aprendizes franceses

³ O ICLE contém composições argumentativas escritas por aprendizes de EFL de 14 diferentes nacionalidades. O LOCNESS é formado por composições que são tanto argumentativas quanto literárias escritas por alunos universitários britânicos e americanos. Vide Granger (1998b, p.13).

⁴ O FROWN corpus foi compilado para ser equivalente ao BROWN e LOB corpora com a diferença de conter amostras de inglês americano representativo do início da década de 90. Para maiores informações acesse o site <http://khnt.hit.uib.no/icame/manuals/frown/INDEX.HTM>.

quando o tópico era argumentativo (Dagneaux, 1995, citada em Granger, 1998a). Em outro estudo, Hinkel (1995) demonstrou a influência de tópicos na distribuição de frequência de certas expressões modais. Em composições escritas na área da educação, Hinkel descobriu que aprendizes de EFL japoneses, chineses, vietnamitas, coreanos e indonésios empregam demasiadamente os modais *must* e *should*, sendo que *must* não foi observado nas composições de falantes nativos cujo tópico era relacionado a questões políticas.

O tópico é provavelmente apenas um, dentre uma variedade de fatores que podem explicar diferenças na expressão de modalidade. O uso excessivo ou restrito de uma certa estrutura ou palavra nos sub-corpora de aprendizes pode estar relacionado a transferências de padrões de suas línguas maternas, ou pode constituir estratégias geralmente empregadas por eles. Os problemas encontrados pelos aprendizes na escolha do(s) item(ns) apropriado(s) podem também estar relacionados a diferentes distribuições de expressões modais na escrita e na fala (Biber, 1988; Biber e Finegan, 1989). Outra explicação pode ser atribuída ao fato de professores de EFL enfatizarem mais expressões modais que são freqüentemente salientadas em livros texto de EFL.

Pode haver outros fatores bem mais complexos dos que os sugeridos acima. Por exemplo, os padrões de modalidade encontrados em certos textos podem ser um reflexo da transferência de certas normas culturais específicas de uma língua; um reflexo de diferentes tradições ou métodos retóricos no ensino de escrita acadêmica; ou até mesmo um reflexo de diferentes atitudes pessoais em relação ao processo da escrita.

2. ANÁLISE E RESULTADOS

2.1. O emprego de verbos modais

Na análise que se sucede, as composições dos falantes nativos constituem o parâmetro utilizado para a comparação e avaliação das características do uso de expressões e padrões de modalidade utilizados pelos aprendizes de EFL. Quando estruturas modais são muito mais freqüentes nos sub-corpora de EFL do que no LOCNESS, isto é caracterizado como uso *excessivo* ou *demasiado*.

A Tabela 2 mostra a distribuição geral de frequência de verbos modais no sub-corpus brasileiro e no corpus de falantes nativos. Dentre todos os modais, *can*, *will* e *could* e o par de modais *should* e *have to* são os que particularmente mais se destacam.

Poderia este uso excessivo de *can* e *will* por parte dos aprendizes brasileiros estar relacionado a transferências de práticas conversacionais de gêneros acadêmicos, ou à dificuldade destes aprendizes de distinguir entre o registro escrito e o oral? Uma comparação entre *can* e *will* em um corpus de inglês falado (MICASE) e um corpus de inglês escrito (FROWN) indica que ambos ocorrem mais freqüentemente na fala do que na escrita (resultado este semelhante àquele obtido por Kennedy, 1998, p.130).

Tabela 2. Distribuição total de verbos modais encontrados no Br-ICLE e no LOCNESS

Modal	Br-ICLE	LOCNESS	Chi-quadrado
can	177**	83	32,9
will	105*	66	8,4
should	71**	26	20,3
would	69	55	1,4
could	56*	34	5,1
have to	49**	15	17,7
must	33	20	3,0
may	24	22	0,1
might	5	7	0,4
shall	3	0	3,0
ought to	1	0	1,0
Total	593**	328	73,2

Esta desigualdade pode refletir também as diferenças entre os tipos de textos (o componente brasileiro é composto em sua totalidade por textos argumentativos). É interessante relatar que na Sessão B (editoriais) do FROWN corpus, *will* é mais freqüente do que *would*, ao passo que no corpus como um todo a ordem é invertida (um resultado idêntico é relatado por Krogvig e Johansson (1981, p.42) em relação ao Lancaster Oslo-Bergen (LOB) corpus).

Uma possível explicação para o uso excessivo de certos modais pode estar relacionada com a transferência da língua materna. Ao contrastarmos os três sub-corpora de inglês de aprendizes com o LOCNESS, podemos não apenas investigar a possível influência da língua materna, mas também averiguar se diferentes línguas maternas deixam diferentes ‘marcas lingüísticas’ nos textos dos aprendizes. Na Tabela 3, as composições dos falantes nativos são comparadas às dos aprendizes espanhóis, tchecos e brasileiros.

Tabela 3. Comparação do uso de modais de possibilidade entre aprendizes espanhóis, tchecos brasileiros e nativos

Modal	Sp-ICLE	Cze-ICLE	Br-ICLE	LOCNESS
Can	136**	195**	177**	83
Could	52*	53*	56*	34
May	20	37*	24	22
Might	3	11	5	7
Total	211**	296**	262**	146

A Tabela 3 mostra que todos os grupos de aprendizes usam excessivamente os verbos modais *can* e *could* e que os três grupos empregam a categoria possibilidade de forma demasiada se comparados aos falantes nativos.

Qual é, então, a razão para esse excesso generalizado do uso dos verbos modais *can* e *could*? Uma possível explicação para tal excesso é que ele reflita a incerteza dos aprendizes

ao argumentar em inglês. No exemplo a seguir, extraído do Br-ICLE, a possibilidade expressa por *could* para convencer o leitor de que a TV pode, de fato, ser considerada um veículo de mídia com desdobramentos positivos, parece ser muito remota para ser utilizada como um mecanismo de persuasão, haja vista que o aprendiz em questão constrói todo o seu argumento salientando, ao longo de sua composição, o lado perverso da TV:

1. Finally, TV is a kind of bad and useless thing. In order to make it better, it should be added in its programming some programs for instructing not only children but also adolescents and adults. Since, it is important to know how to behave in our real life, how important it is to be a respectable and solidary person, for example. As I have said, TV, nowadays, is, really, the opium of masses. So, it could be a good thing, but it has been used in a wrong way and has been dangerous to the body and the mind of those who do not know how to watch it. (Br-ICLE)

É possível também que diferentes culturas tenham normas pré-estabelecidas com relação a graus de certeza e ao quão assertivos aprendizes possam ser ao escreverem academicamente. Esta hipótese é corroborada por estudos que identificaram variações no grau de certeza expressos nos argumentos em diferentes línguas (Hinkel, 1995; Recski, no prelo; Hyland e Milton, 1997). De acordo com esses estudos, finlandeses, japoneses e chineses parecem preferir um estilo mais indireto em sua escrita acadêmica do que, por exemplo, brasileiros, alemães e tchecos. Aprendizes muitas vezes utilizam *might* e *may* onde provisoriedade não parece ser o tom apropriado na tarefa de convencer o leitor sobre o ponto de vista que está sendo argumentado:

2. So, these standards are imposed on us since we are children, but I sincerely believe that it is still certainly possible to live without following them, but if each one of us teach our children that no one needs to be thin in order to be happy or a good friend, these concepts might change, and in the future we may have a better society. (Br-ICLE).

Uma outra possível explicação para o uso excessivo de alguns verbos modais de possibilidade pode estar relacionada à influência de um registro oral (note a estrutura retórica semelhante à fala nos exemplos (1) e (2) acima). Holmes (1988, p.29) demonstra, por exemplo, que expressões de possibilidade ocorrem mais em transcrições de fala do que na escrita. A frequência relativamente alta de verbos modais de possibilidade por parte dos aprendizes pode, também, refletir o fato de não estarem empregando outros dispositivos epistêmicos (e.g. adjetivos, substantivos, etc), preferidos pelos falantes nativos, para expressar possibilidade.

2.2. Categorias de significados nas composições dos aprendizes e dos falantes nativos

Normalmente a escolha de dispositivos que expressam modalidade está diretamente relacionada aos tipos de significados criados e/ou expressos. Verbos modais tais como *might*, *may*, *should* e *must* possuem pelo menos dois significados prototípicos (epistêmico e deôntico). Ambas modalidades epistêmica e deôntica são importantes em escrita argumentativa embora pertençam a diferentes estilos argumentativos (Kennedy, 1998;

Heberle, 1997). Verbos modais expressam graus de probabilidade (o sistema de modalidade epistêmica), ou graus de obrigação, necessidade, permissão e vontade (o sistema de modalidade deôntica). Para alguns desses modais, o significado deôntico é mais freqüente, enquanto para outros, o uso epistêmico prevalece. *Must*, *ought to*, *have to* e *should* geralmente expressam significados deônticos. Recski (2002) analisou um corpus de língua inglesa de 1 milhão de palavras subdividido em cinco registros (escrita acadêmica, economia, ficção, ciência e transcrições de fala) e descobriu que das 653 ocorrências analisadas, 70,6% (ou 461 ocorrências) foram utilizadas para expressar significados deônticos. Recski (2002) nota que em registros mais formais tais como escrita acadêmica, economia - onde o grau de subjetividade dos argumentos parece ser evitado por motivos de clareza, autoridade e credibilidade - encontra-se um número maior de ocorrências do modal *must* empregado deonticamente. A Tabela 4 mostra a distribuição de significados de *must* no Br-ICLE e LOCNESS. É possível observar claramente que nenhum dos grupos emprega *must* epistemicamente de forma significativa.

Tabela 4. Distribuição de usos deônticos e epistêmicos do modal *must*

Uso deôntico		Uso epistêmico	
Br-ICLE	LOCNESS	Br-ICLE	LOCNESS
28	16	5	4

Por outro lado, *may* foi empregado epistemicamente em 100% das ocorrências (24) no sub-corpus dos aprendizes brasileiros. Possibilidade deôntica pode ser parafraseada como “é possível que” ou “circunstâncias permitem que”. Este tipo de modalidade foi encontrada no LOCNESS em exemplos tais como:

3. Despite this the terrorists' moral approach invites admiration. Kaliyev is presented in a more favourable light than Stepan - the latter is seen to kick Kaliyev when he is down while the rest of the group tend to back Kaliyev. This may suggest that Camus is endorsing a moral approach to achieving political aims. (LOCNESS)

May deôntico é formal, o que talvez explique o seu freqüente uso com voz passiva impessoal:

4. Les Mains sales presents aspects of the political world and moral values which may be considered in various ways, depending on the political and ethical morals of the audience. When it was first produced, as a play in 1948 and later as a film, it provoked great indignation from Marxist intellectuals, who felt that their position had been severely attacked by Sartre. (LOCNESS)

Outra aparente função de certos verbos modais nas composições é a de influenciar crenças e atitudes dos leitores através de argumentos. Ao usar de maneira demasiada modais e semi-modais de obrigação e necessidade tais como *should* e *have to*, os aprendizes brasileiros parecem estar adotando um estilo direto e enfático de persuasão, que em algumas composições parece estar relacionado ao tópico em questão. De acordo com Biber (1988, p.150), “dispositivos tais como modais de necessidade [...] podem marcar a tentativa de

persuadir ouvintes/leitores de que certas circunstâncias são desejáveis ou prováveis” (minha tradução). Assim, muitas ocorrências de *should* e *must* foram observadas nas composições dos aprendizes brasileiros cujos tópicos eram extremamente polêmicos. *Should* e *must* demonstram que algum tipo de ação é desejada, recomendável, ou obrigatória. *Should*, particularmente, refere-se a normas culturais de comportamento e/ou ética ou códigos morais que podem ser específicos para os aprendizes, como os exemplos (5) e (6) relacionados ao sistema penitenciário brasileiro atestam:

5. The basic objective of any intelligent prison system should be the rehabilitation of criminals. These men and women should be recovered to society through work and education. (Br-ICLE)

6. In my opinion, criminals (as human beings they are) should be taught, well treated, informed, given (psychological, social, philosophical) support, and given a second (third, fourth, if necessary, why not?) chance always. Common sense is, most of the times, unworthy, because it can be anything, mainly contradictory. It is part of the common sense to say that a person should pay for his/her crime equally (an eye for an eye, a tooth for a tooth), and the same common sense says that (and this is Christianity) vengeance is a bad thing and that we should do like Jesus did: offer the other face. Moreover, common sense says that everybody deserves a second chance. Isn't prison revenge? (Br-ICLE)

O mesmo estilo didático, onde *should* e *must* são freqüentemente utilizados para indicar que tipos de ações são necessárias ou desejadas, também foi observado em composições cujo tópico envolvia o sistema de ensino:

7. No matter what you decide to study, to be, to reach, you must feel the need inside you; you must be strong, face all the problems, struggle, in order to achieve your aim. (Br-ICLE)

8. Willing to learn something takes a lot of hard work and time; of course short-term aims should be set in order to find easier to manage the learning and see the progress steps, then, comes the long term aims. Time should be planned properly and well in advance, selecting what activities will be performed to search a specific objective. It is important to set one objective at a time. After reaching the first one, then it is possible to set the second, and the next one and then next one and so on, until reaching the whole and feel happy with the results. In all the objectives there are priorities which should be considered in terms of importance to oneself. Thus, always the relevant things should come first and the others can be postponed. (Br-ICLE)

É importante ressaltar que de um total de 71 ocorrências de *should*, 29 (40,8%) foram encontradas nas composições cujo tópico era relacionado ao sistema penitenciário brasileiro (o que representa 12,3% do corpus (4151 palavras). Por outro lado, em outro grupo de composições representando 31,8% do corpus (10733 palavras) foram encontradas apenas 10 instâncias de *should* (14,1%). Este resultado ressalta o às vezes impressionante efeito que um tópico pode exercer sobre o uso de expressões modais.

2.3. A expressão de modalidade epistêmica

Os principais verbos modais empregados epistemicamente foram *could*, *may*, *might*, *would* e *will*. *Will* é utilizado para prever o que ocorrerá no futuro com algum grau de

certeza (persuasão predictiva). *May* e *might* expressam probabilidades e/ou possibilidades (o que pode ou não ocorrer).

9. I am a student of the course in Letras - Licenciatura, and I will probably graduate next year. However, so far, all my classes have been theoretical. The practical experience will happen only in the last semester. I believe theory is also necessary, but theory alone is not enough. All the theoretical courses I attended (Psychology of Education, Didactics, Fundamentals of Education, and so many others), will not be very helpful when, after graduation, I have to give a class for the first time. Only theory is not enough to prepare professionals (teachers, lawyers, doctors, etc.) to daily experience.

10. The criminal, also, may be simply crazy, may have some mind pathology, etc.; I am not a doctor, but we know it is possible. Nobody cares about it, nobody evaluates the cases properly and cautiously and, because of that, I believe, many people that could recover and have a normal social life get worse and worse into the jails.

As dificuldades encontradas por aprendizes de EFL de diferentes culturas para controlar o grau de certeza e/ou incerteza contido em suas proposições são bem relatadas na literatura (Holmes, 1988; Recski, no prelo; Aijmer, 1999).

Modalidade epistêmica é uma área complexa envolvendo uma grande variedade de dispositivos lingüísticos para a expressão de certeza e incerteza. Verbos modais epistêmicos possuem muitos equivalentes lexicais e gramaticais. Alguns significados e estruturas são mais prototípicos do que outros. Por exemplo, verbos modais representam a maneira mais prototípica ou 'gramaticalizada' para expressar modalidade em inglês. Equivalentes advérbiais de *may* são ilustrados nos exemplos inventados (ii) e (iii):

- (i) Hamilton may be rich
- (ii) Hamilton is probably rich
- (iii) Perhaps Hamilton is rich

Estas não são apenas variações de estilo; elas representam padrões alternativos de modalidade. É preciso que consideremos quando e com que frequência elas são empregadas e se exercem a mesma função de verbos modais.

Como foi demonstrado anteriormente, os aprendizes apresentam uma tendência a empregar verbos modais de maneira excessiva, tendência esta que pode ser um reflexo de certas práticas de ensino. Como Holmes (1988, p.40) aponta, muitos livros didáticos dedicam uma atenção muitas vezes demasiada a verbos modais, deixando de enfatizar, desta forma, estratégias alternativas.

Quando o uso de advérbios utilizados para expressar certeza ou incerteza foi investigado nos corpora (cf. Recski, no prelo), pôde-se observar que a maioria deles foram empregados excessivamente pelos aprendizes brasileiros se comparados ao LOCNESS (Tabela 5). A maior diferenças entre os dois corpora está relacionada às frequências dos advérbios *usually*, *maybe*, *sometimes*, *of course* e *always*. Cerca da metade dos advérbios epistêmicos empregados pelos aprendizes brasileiros são advérbios de certeza, utilizados

para ampliar a convicção, às vezes, subjetiva dos aprendizes. É interessante ressaltar que em um recente estudo sobre modalidade epistêmica na escrita de aprendizes brasileiros, Recski (no prelo) notou que eles freqüentemente encontram problemas ao dosar o grau de assertividade de suas proposições, o que, por sua vez, resulta em argumentos expressos com graus de convicção inapropriados.

Tabela 5. Distribuição de advérbios epistêmicos.

Modalidade	Br-ICLE	LOCNESS	Chi-quadrado
INCERTEZA			
probably	13	5	3,4
quite	12	8	0,7
usually	18**	1	15,0
perhaps	2	17	12,0
sometimes	24**	0	23,7
possibly	1	0	0,9
maybe	24**	2	18,4
CERTEZA			
certainly	8	3	2,2
indeed	7	9	0,3
in fact	15	21	1,1
of course	15*	4	6,2
surely/ for sure	7	2	2,7
always	31**	12	8,2
never	12	17	0,9
actually	14	18	0,5
Total	203	119	

Pode-se argumentar que o emprego excessivo de certas expressões enfáticas pode estar relacionado a uma habilidade ainda insuficiente e preliminar da escrita desses aprendizes assim como a diferentes princípios retóricos que governam o seu emprego (cf. os resultados encontrados por Recski, no prelo e Lorenz, 1998, p.64). Segundo Barbaresi (1987, p.21, minha tradução), *of course* tem pelo menos duas funções retóricas: (i) a de “fazer com que o leitor aceite mais facilmente um argumento”; e (ii) a de “reduzir um grande argumento a doses aceitáveis”. O uso de *of course* no excerto (11), por exemplo, demonstra como esta aprendiz procura persuadir o leitor a aceitar a sua argumentação ao reduzir a sua amplitude:

11. Television in Brazil became a synonymous of sex in the last decade. It's funny isn't it? This is exactly where the money is. Remember, showing the people what they want to see, even if it's not the right thing to do, is a good way to fill up your pocket. Showing this kind of program, like sex scenes or insinuations, can change children mind's, giving them a wrong idea, or showing them a

world that's not real. It's time for a break on this exploitation. The competent organizations should do more about that, not only keep saying that it's wrong, but act quickly. If we want to make better country for everyone, this is one of the things that has to be changed. *Of course*, if the problems of Brazil were summarized to this, things probably would be better. But that's another story.

Modalidade epistêmica também pode ser observada através do uso de verbos lexicais e outros tipos de expressões fixas. Quando dois dos verbos lexicais epistêmicos mais freqüentes (*I think* e *I believe*) assim como a expressão *in my opinion* foram investigados, observou-se que ambos foram quase que exclusivamente encontrados na escrita dos aprendizes brasileiros (Tabela 6).

Tabela 6. Distribuição de *I think*, *I believe* e *in my opinion*

	Br-ICLE	LOCNESS
<i>I think</i>	17	0
<i>I believe</i>	6	2
<i>In my opinion</i>	15	0
Total	38	2

O emprego excessivo desses verbos, assim como da expressão epistêmica *in my opinion*, indicam o alto grau de envolvimento dos aprendizes em relação ao que escrevem. Baseada em quatro sub-corpora (alemão, francês, sueco e finlandês) e um corpus comparável de composições escritas por falantes nativos, Petch-Tyson (1998 p.110) verificou que aprendizes tendem a empregar excessivamente expressões que refletem o seu envolvimento pessoal. Vários outros estudos apontam para o uso excessivo de *I think* por parte dos aprendizes. Por exemplo, Granger (1998c) relata que aprendizes preferem estruturas discursivas ativas ao invés de passivas (cf. Altenberg, 1998; Ringbom 1998). É possível também argumentar que esse uso demasiado de expressões de envolvimento pessoal seja um reflexo da influência de um registro oral. *I think* é um dos bigramas mais freqüentes no MICASE, ocorrendo 43,7 vezes a cada 10000 palavras, resultado este semelhante ao encontrado por Aijmer (1996, p.9) em sua análise do London-Lund Corpus (51 ocorrências a cada 10000 palavras).

3. CONSIDERAÇÕES FINAIS

Esta investigação preliminar sobre modalidade na escrita de aprendizes brasileiros de língua inglesa revelou um excesso generalizado no uso de uma grande parte das categorias de modalidade examinadas: verbos modais, advérbios epistêmicos e verbos lexicais. Entretanto, é importante ressaltar que embora este estudo tenha considerado uma variedade de verbos e expressões modais, ele não pode ser considerado exaustivo porque as formas disponíveis para a expressão de modalidade são tão numerosas e variadas que isto requereria uma análise manual, baseada em um corpus manualmente anotado para cada tipo de

expressão modal, antes que generalizações mais definitivas sobre o emprego de modalidade na escrita destes aprendizes pudessem ser postuladas. Por exemplo, é possível que enquanto os aprendizes empregam de forma excessiva as categorias investigadas neste estudo, eles talvez utilizem de forma restrita outras categorias que não foram investigadas (e.g., adjetivos e substantivos empregados modalmente). Entretanto, mesmo em um estudo não-exaustivo como este, os resultados parecem ser interessantes, não apenas pelo que revelam sobre certas características do emprego de modalidade na escrita destes aprendizes, mas também pelas promissoras possibilidades que se abrem de novos caminhos para pesquisa na aquisição de línguas estrangeiras.

Em primeiro lugar, há um excesso generalizado das categorias de modalidade examinadas nos corpora dos aprendizes. Isto pode ser, em parte, o resultado do fato desses aprendizes adotarem um estilo mais *falado* em sua escrita se comparados aos falantes nativos representados no LOCNESS⁵. Este excesso também pode estar relacionado à tendência apresentada por aprendizes de empregar expressões pleonásticas em sua escrita.

Pelo menos no que diz respeito aos aprendizes brasileiros, parece ser possível que esse excesso seja reforçado por fatores relacionados às duas línguas. Ao passo que em inglês, expressões modais são prototipicamente expressas através do uso de verbos modais, em português, expressões de modalidade são geralmente expressas não apenas por intermédio de verbos modais, mas também através de advérbios e outras expressões. A transferência da língua materna pode, assim, reforçar a tendência já existente na interlíngua de expressar modalidade pleonasticamente.

A distribuição de expressões de modalidade certamente contribui de forma importante para que se possa estabelecer o tom e o estilo apropriados em qualquer tipo de escrita. Certas expressões modais são mais freqüentes na fala do que na escrita e, se elas são utilizadas na escrita com freqüências tipicamente associadas à fala, isto necessariamente será um fator que contribuirá para um estilo mais *falado* do que *escrito*. O fato disto ser o que realmente acontece na escrita destes aprendizes não deve ser surpreendente para aqueles familiarizados com outros estudos relacionados ao projeto ICLE. Muitos destes estudos vêm revelando que vários aspectos da linguagem desses aprendizes refletem um estilo mais *falado* do que o encontrado na escrita dos falantes nativos (Granger e Rayson 1998; Altenberg e Tapper, 1998; Recski, no prelo). Em relação ao tom, tanto o emprego significativo de modais para a expressão de provisoriedade e certeza quanto o uso alternado de estratégias contraditórias encontradas nos textos aqui analisados (por exemplo, a combinação de provisoriedade com expressões de certeza) parecem influenciar o efeito retórico atingido pelo texto, dando a impressão de que o texto não foi escrito por um falante nativo, mas sim por um aprendiz de língua estrangeira.

⁵ Uma análise contrastiva entre MICASE e FROWN revelou que *can*, *could*, *will*, *would* e *have to* são todos preponderantes em inglês falado. O número de ocorrências encontradas nos dois corpora a cada 10000 palavras foi o seguinte: (a) MICASE: *can* (52,1), *could* (18,3), *will* (21,5), *would* (32,8), e *have to* (13,2); e (b) FROWN: *can* (21,7), *could* (16,5), *will* (19,9), *would* (25,5), e *have to* (3,3). Vide também Biber et al. (1999: § 6.6) para maiores informações sobre a preponderância de verbos modais em registros orais.

Este estudo também revelou que o tópico pode ser um fator de extrema importância para a escolha de verbos modais por parte dos aprendizes. Além disso, parece ser bem provável, embora não seja comprovado aqui, que exista uma grande influência de práticas e materiais pedagógicos na escrita desses aprendizes. Claramente, não há apenas um fator que possa explicar certos padrões de modalidade apresentados pelos aprendizes.

Várias implicações pedagógicas podem ser apontadas. Por exemplo, discussões sobre modalidade freqüentemente tendem a focar verbos modais em detrimento de outras categorias. É necessário que definições mais amplas de modalidade sejam adotadas, definições que levem em consideração a grande variedade de categorias estruturais utilizadas para a expressão de modalidade. Os resultados aqui apresentados apontam também para a necessidade de situar o estudo de modalidade dentro de uma perspectiva discursiva, uma perspectiva que leve em consideração não apenas aspectos como tom e registro, mas também certos aspectos semânticos e culturais.

Acredito que aprendizes se beneficiariam, se nós professores, no papel de facilitadores, salientássemos o fato de que certas maneiras de descrever o mundo, ou de expressar opiniões provavelmente variam de uma língua para outra e que certos dispositivos modais funcionam regidos por diferentes sistemas de valores, normas e expectativas culturais.

Em estudos futuros que envolvam questões pertinentes à modalidade na escrita de aprendizes de língua estrangeira seria desejável exercer um controle mais acirrado em relação ao tópico. Isso se faz necessário porque o tópico parece ser um fator crucial no uso de certos modais (cf. o emprego de modais deônticos discutido em 3.2). Seria igualmente proveitoso estudar uma variedade maior de corpora de aprendizes: os resultados deste estudo sugerem que algumas tendências no emprego de modalidade, tais como o uso pleonástico de algumas expressões epistêmicas, são compartilhadas por outros grupos. No entanto, antes que seja possível falar de tendências universais, outros grupos de aprendizes, de diferentes níveis de proficiência, precisariam ser investigados. Modalidade é uma área de estudo complexa e promissora; entretanto, é necessário que avancemos além de sua superfície para que possamos melhor entender o seu papel na expressão de vários significados.

REFERÊNCIAS BIBLIOGRÁFICAS

- AIJMER, K. (1996) *I think* – an English modal particle. In: T. SWAN & O. J. WESTVIK (Eds.) *Modality in Germanic languages. Historical and comparative perspectives*. Berlin: Mouton de Gruyter, p. 1-47.
- _____. (1999) Epistemic possibility in an English-Swedish contrastive perspective. In: H. HASSELGÅRD & S. OKSEFJELL (Eds.) *Out of corpora*. Amsterdam & Atlanta, GA: Rodopi, p. 301-323.
- ALTENBERG, B. (1998) On the phraseology of spoken English: the evidence of recurrent word combinations. In: A. P. COWIE (Ed.) *Phraseology: theory, analysis, and applications*. Oxford: Oxford University Press, p. 101-122.
- ALTENBERG, B. & TAPPER, M. (1998) The use of adverbial connectors in advanced Swedish learner's written English. In: S. GRANGER (Ed.) *Learner English on Computer*. London & New York: Longman, p. 80-93.

RECSKI - I think this question is very relevant nowadays...

- BARBARESI, L. (1987) 'Obviously' and 'certainly': Two different functions in argumentative discourse. *Folia Linguistica*, 21, p. 3-24.
- BIBER, D. (1988) *Variations across speech and writing*. Cambridge University Press.
- BIBER, D. & FINEGAN, E. (1988) Adverbial stance types in English. *Discourse Processes*, 11, p.1-34.
- BIBER, D.; JOHANSSON, S.; LEECH, G.; CONRAD, S. & FINEGAN, E. (1999) *Longman Grammar of Spoken and Written English*. Essex: Longman.
- COATES, J. (1983) *The Semantics of the Modal Auxiliaries*. Beckenham: Croom Helm.
- DAGNEAUX, E. (1995) *Expressions of epistemic modality in native and non-native essay-writing*. Unpublished MA dissertation. Département d'Etudes Germaniques, Université Catholique de Louvain.
- DeCARRICO, J. (1986) Tense, aspect and time in the English modality system. *TESOL Quarterly*, 20, p.665-682.
- GRANGER, S. (Ed.). (1998a) *Learner English on Computer*. London & New York: Longman.
- _____. (1998b) The computer learner corpus: a versatile new source of data for SLA research. In: GRANGER, S. (Ed.) *Learner English on Computer*. London & New York: Longman, p. 3-18.
- _____. (1998c) Prefabricated patterns in advanced EFL writing: Collocations and formulae. In: COWIE, A. P. (Ed.) *Phraseology: theory, analysis, and applications*. Oxford: Oxford University Press, p. 145-160.
- GRANGER, S. & RAYSON, P. (1998) Automatic profiling of learner texts. In: GRANGER, S. (Ed.) *Learner English on Computer*. London & New York: Longman, p. 119-131.
- HEBERLE, V. (1997). *An investigation of the textual and contextual parameters of women's magazines*. Tese de Doutorado. Programa de Pós-graduação em Letras/Inglês e Literatura Correspondente. UFSC.
- HINKEL, E. (1995) The use of modal verbs as a reflection of cultural values. *TESOL Quarterly*, 29(2), p. 325-343.
- HOLMES, J. (1988) Doubt and certainty in ESL textbooks. *Applied Linguistics*, 9, p. 21-44.
- HORVATH, J. (2001) *Advanced Writing in English as a Foreign Language: A Corpus-Based Study of Processes and Products*. Pecs: Lingua Franca Csoport.
- HOYE, L. (1997) *Adverbs and Modality in English*. London & New York: Longman.
- HYLAND, K. & MILTON, J. (1997) Qualifications and certainty in L1 and L2 students' writing. *Journal of Second Language Writing*, 6(2), p.183-205.
- KENNEDY, G. (1998) *An Introduction to Corpus Linguistics*. London: Longman.
- KROGVIG, I. & JOHANSSON, S. (1981) *Shall, will, should and would* in British and American English. *ICAME News*, 5, p.32-56.
- LEECH, G. (1998) Learner corpora: what they are and what can be done with them. In: GRANGER, S. (Ed.) *Learner English on Computer*. London & New York: Longman, p. xiv-xx.
- LORENZ, G. (1998) Overstatement in advanced learners' writing: stylistic aspects of adjective intensification. In: GRANGER, S. (Ed.) *Learner English on Computer*. London & New York: Longman, p. 53-66.

- MINDT, D. (1995) *An Empirical Grammar of the English Verb: Modal Verbs*. Berlin: Cornelsen.
- PALMER, F. (1990) *Modality and the English modal*. London & New York: Longman.
- PERKINS, M. (1983) *Modal Expression in English*. London: Pinter.
- PETCH-TYSON, S. (1998) Writer/reader visibility in EFL written discourse. In: GRANGER, S. (Ed.) *Learner English on Computer*. London & New York: Longman, p. 107-118.
- RECKSKI, L. (2002) The English Modal Auxiliary *Must*: A Corpus-Based Semantic-Syntactic Account. *Revista da ABRALIN*, 1(2), p.99-122.
- _____. (no prelo) Negociando certeza e incerteza em escrita acadêmica em inglês.
- RINGBOM, H. (1998) Vocabulary frequencies in advanced learner English: a cross-linguistic approach. In: GRANGER, S. (Ed.) *Learner English on Computer*. London & New York: Longman, p. 41-52.
- SARDINHA, T. (2000) O corpus de aprendiz BR-ICLE. *DIRECT Papers* 40. São Paulo: LAEL – PUC/SP.