

***Quer Dizer* e seus usos: análise de dados de escrita**

Cristine Görski Severo (UFSC - PG)*

RESUMO: Com base na metodologia da Lingüística de Corpus, descrevem-se padrões contextuais de uso da expressão *quer dizer* em dados de escrita, mediante: análise dos contextos anterior e posterior à expressão; sistematização das funções desempenhadas pelo *quer dizer*; correlação entre as funções em dados de escrita e de fala e entre os dados de escrita e o dicionário. Foram identificadas cinco diferentes funções para o *quer dizer*, presentes tanto na fala como na escrita, sendo que uma delas – a função corretiva – não é contemplada pelo dicionário Houaiss no verbete *querer*.

Palavras-chave: *Quer dizer*; Função; Contexto; Fala e escrita; Dicionário.

Introdução

Neste trabalho, objetiva-se descrever o funcionamento da expressão *quer dizer* em dados de escrita oriundos do corpus organizado pelo NILC (Núcleo Interinstitucional de Lingüística Computacional). Foram encontradas 1.329 ocorrências de *quer dizer* e selecionados 218 dados, que foram analisados a partir da metodologia oferecida pela Lingüística de Corpus, sendo que para a manipulação dos dados foi utilizado o programa *WordSmith*.

Sucintamente, a abordagem baseada em corpus possibilita o reconhecimento de padrões lingüísticos em contextos reais de uso da língua, através de métodos estatísticos. A quantidade e a qualidade deste reconhecimento se vinculam, grandemente, à velocidade e à habilidade do computador de estocar informações (amostras de textos de língua em uso); por isso, foi especialmente a partir dos anos 90 do século passado que pesquisas em Lingüística de Corpus passaram a contribuir mais efetivamente para a sistematização das regularidades/padrões/trações recursivos da língua.

Primeiramente, para contextualizar o trabalho, são expostos alguns aspectos relevantes da Lingüística de Corpus e do seu método. A análise propriamente dita se orienta pelas seguintes questões gerais: (i) quais as características formais do contexto textual anterior e posterior ao *quer dizer* e que funções essa expressão desempenha nesses contextos? (ii) existe uma correlação entre as funções encontradas nos dados de escrita, de fala¹ e nas acepções dicionarizadas? Por fim, ressalta-se a importância deste tipo de pesquisa para a operacionalização de uma identificação automática das funções com base nos contextos adjacentes; e para o estabelecimento de correlações funcionais possíveis entre os dados de fala e de escrita e entre os dados de escrita e o dicionário.

* crisgorski@hotmail.com

1. Sobre a Lingüística de Corpus (LC) e o Método

A LC não possui status de teoria, mas sim de uma ferramenta metodológica², que visa ao levantamento e análise de dados empíricos, sejam eles fonéticos, morfológicos, sintáticos, semânticos e outros. Por metodologia deve-se entender “um modo típico de aplicar um conjunto de pressupostos de caráter teórico” (SARDINHA, 2000, p. 356), que permite certa observação (análise) do comportamento do objeto de pesquisa e, conseqüentemente, a produção de conhecimento.

A abordagem empirista é o que caracteriza a LC, sendo que a linguagem deve ser vista como um sistema probabilístico e o método aplicado como estatístico. Sobre este método, Mani (2006) comenta que as regras estatísticas deduzidas a partir de exemplos reais operam de forma semelhante àquelas deduzidas a partir da intuição; ademais, o aprendizado geralmente ocorre a partir da experiência, através da indução, ou seja, da aplicação de regularidades estatísticas. Resumidamente, as características essenciais da abordagem baseada em corpus são: o empirismo; o uso de um grupo extenso de textos em linguagem natural (corpus), como base para a análise; o uso de programas de computadores para a análise dos dados; e o uso de técnicas quantitativa e qualitativa de análise (BIBER, CONRAD e REPPEN, 1998). Na fala de Sardinha (2000):

“a Lingüística de Corpus ocupa-se da coleta e exploração de corpora ou conjuntos de dados lingüísticos textuais que foram coletados criteriosamente com o propósito de servirem para a pesquisa de uma língua ou variedade lingüística” (SARDINHA, 2000, p. 325).

Além disso, as diferenças percebidas nas variações dos dados necessariamente associam-se ao contexto, não sendo elas aleatórias. Dessa maneira, é possível falar em padrões de aparecimento – a linguagem não opera de qualquer maneira, mas, de acordo com determinados contextos o uso que se faz da linguagem, esta varia: “a linguagem forma padrões que apresentam regularidade (...) e variação sistemática” (SARDINHA, 2000, p. 351).

Nessa perspectiva, os dados são oriundos de um universo textual determinado previamente segundo os critérios do pesquisador em relação ao objeto de pesquisa, manipulados e analisados a partir de sistemas de computador e mediante um modelo probabilístico adequado³. O fenômeno em estudo deve ser categorizado e classificado segundo os critérios estabelecidos pelo investigador, e “tanto as características de uma anotação de corpus quanto o processo de anotação – e as dificuldades derivadas, por exemplo, da discrepância entre anotações – são objeto de pesquisas intensas” (ROCHA, 2003, p. 192).

Sobre o corpus, este não se resume a um somatório de textos eletrônicos e tampouco é formado por textos criados por programas de computador. Tomado como artefato para pesquisa, o corpus deve: (a) ser constituído de textos naturais, existentes na linguagem humana; (b) ser formado por fragmentos que sejam representativos de uma língua; (c) ser escolhido criteriosamente de acordo com a finalidade a que se propõe, seguindo as condições de naturalidade e autenticidade; e (d) ser vasto e representativo (cf. SARDINHA, 2000). Sobre este último critério, a representatividade deve associar-se especialmente à extensão,

sendo que, quanto mais palavras, maior a probabilidade de itens com baixa frequência na língua aparecerem; além disso, quanto maior é a amostra, mais confiável, uma vez que não se sabe a dimensão exata da população (da linguagem como um todo). Assim, a probabilidade é um dos aspectos fundamentais na LC, o que permite relações entre traços mais e menos comuns em determinados contextos lingüísticos, incluindo os aspectos sintáticos, estruturais, pragmáticos, discursivos, etc. (cf. SARDINHA, 2000; MANI, 2006).

Por fim, cabe ao pesquisador a coleta dos dados, a sua análise e a interpretação do fenômeno investigado, sendo que esse trabalho é grandemente possibilitado pelo uso de programas computacionais⁴ voltados para dados de texto e da estatística, o que garante maior precisão no trabalho realizado. Tais programas incluem ferramentas computacionais que facilitam a análise dos dados. Dentre as ferramentas utilizadas no Brasil, encontram-se o *Micro Concord*, criado em 1993, e o *WordSmith Tools*, criado em 1995, sendo este o “primeiro a aproveitar os recursos do ambiente Windows para a análise de corpus, divulgar a Lingüística de Corpus entre os usuários de micro computadores” (SARDINHA, 2000, p. 335).

O fenômeno em estudo neste trabalho é a expressão *quer dizer*, cujo funcionamento foi observado em textos escritos, integrantes do corpus do NILC (Núcleo Interinstitucional de Lingüística Computacional), composto por cerca de trinta e cinco milhões de palavras. Os textos do NILC dividem-se em três categorias: (i) os *corrigidos* (que geralmente sofrem revisão), compostos por textos de diversos gêneros – literatura, didáticos, enciclopédias, revistas, jornais etc, totalizando 32.590.000 palavras; (ii) os *semicorrigidos* (que não sofrem revisão muito severa), constituídos por contratos, relatórios, dissertações etc, totalizando 1.115.000 palavras; (iii) os *não corrigidos*, compostos por textos autênticos, como redações, monografias e textos de publicidade, totalizando 738.000 palavras. Neste trabalho, fez-se uso dos dados de textos corrigidos, sendo que as categorias encontradas, por ordem de frequência, foram: livro, literatura, didático, redação, monografia, dissertação, revista e texto jurídico.

Obteve-se uma concordância⁵ da expressão *quer dizer* no programa de manipulação de corpora *WordSmith*, que gerou uma lista de 1.329 ocorrências do fenômeno. O *WordSmith* possui três ferramentas, o *WordList* – que gera listas de frequência das palavras no corpus; o *Concord* – que gera concordância ou lista de ocorrências, no corpus, de palavras escolhidas pelo usuário, acompanhada de texto; e o *KeyWord* – que gera lista de palavras-chave, em corpora escolhidos pelo usuário, comparando “uma lista de palavras de um corpus de estudo com uma lista de palavras de um corpus de referência” (GERBER, 2004, p. 57). Para fins do presente trabalho, foi utilizado o *Concord*, que permite uma análise do comportamento do fenômeno em estudo no contexto.

Dentre as 1.329 ocorrências encontradas, foram analisadas 218. A seleção dos dados da análise foi feita aleatoriamente por meio do programa *Concord*, dentro do software *WordSmith*. Para tanto, o recurso *Horizons* foi utilizado, o que permitiu a seleção do item a cada seis ocorrências, aleatoriamente.

Para facilitar a classificação da expressão em estudo, foram observadas características do contexto imediato de aparecimento de *quer dizer* (a janela de contexto é de cinco palavras à esquerda e cinco à direita), tendo sido encontradas com grande frequência as construções “isso quer dizer” (136 ocorrências no corpus completo e 21 no selecionado), “isto quer dizer” (50 ocorrências no completo e 11 no selecionado), ou seja, a presença do pronome

demonstrativo neutro isso/isto. Na posição à direita, a maior ocorrência fica por conta do “que”, com 508 dados no corpus completo e 81 no selecionado. Além disso, o programa detectou, no corpus de 1.329 dados, 77 ocorrências do item “*, quer dizer,” e 197 do item “*. Quer dizer,*”. Todas essas construções fornecem pistas das funções do *quer dizer*, cujas descrições serão vistas posteriormente.

Por fim, as questões que norteiam a análise dos dados são as seguintes:

- a) Como se caracteriza o contexto anterior e posterior a *quer dizer*?
- b) Que sentidos/ funções são observados nos padrões relativos à expressão *quer dizer*?
- c) Em que medida as funções de *quer dizer* verificadas no corpus escrito correspondem às identificadas em corpus oral?
- d) É possível estabelecer algum tipo de correlação entre as funções desempenhadas por *quer dizer* e as marcas formais do contexto anterior e posterior?
- e) Em que medida as funções de *quer dizer* identificadas no corpus escrito analisado são contempladas no dicionário Houaiss?

2. Análise dos dados

Primeiramente, será feito um levantamento, baseado em concordância, dos contextos anterior e posterior à expressão *quer dizer*, salientando a frequência de ocorrência dos aspectos formais que aparecem nos dois contextos. Na segunda parte, serão apreendidos os padrões observados no funcionamento de *quer dizer*, a partir dos diferentes sentidos/funções desempenhados pela expressão. Também serão tecidas considerações sobre a articulação entre as diferentes características formais e os sentidos captados nos diversos contextos.

Note-se que diversas funções assumidas pelo *quer dizer* foram analisadas por trabalhos de cunho funcionalista, dentre os quais cito alguns: Koch (1995) propõe que o *quer dizer* é usado como *esclarecedor* ou *retomador*, frisando uma informação já fornecida; Martelotta (1998) atribui à expressão os papéis de *significação*, *retificação*, *reformulação* e *seqüenciação*; e Dias (2006) sugere que a construção expressa a idéia base de significação, com “extensões de significado”, principalmente de *paráfrase* e *avaliação conclusiva*, podendo também desempenhar os papéis de *ressalva*, *ratificação*, *causalidade*, *especificação* e *contradição*. Ressalto que o foco deste artigo é oferecer uma leitura a partir da Lingüística de Corpus do fenômeno do *quer dizer* presente em dados de escrita, sendo que uma breve correlação será feita entre os resultados aqui encontrados e as funções elencadas, a partir de dados de fala, por Dal Mago e Gorski (2002).

2.1. Contextos de *quer dizer*

O contexto de ocorrência de cada dado foi controlado, tanto o precedente como o seguinte à ocorrência analisada (contexto *quer dizer* contexto). Com relação ao contexto anterior, foi controlada a variável *ausência* ou *presença de sujeito*. Os fatores e suas especificações são caracterizados e exemplificados a seguir.

1) Ausência de sujeito anterior – não há vínculo sintático de *quer dizer* com a construção anterior, podendo haver inserção de ponto, ponto-e-vírgula, vírgula, reticências.

(a) “Eu não fiz faculdade. *Quer dizer*, eu estudei muito, eu estudei na escola” (dado 36).

2) Presença de sujeito anterior – o vínculo sintático de *quer dizer* pode ser estabelecido com um sintagma nominal (SN); um pronome demonstrativo (*isto, isso, o que*) que pode estar explícito ou implícito (neste último caso, retomando uma idéia anterior); um pronome relativo; ou uma sentença-sujeito, conforme respectivamente exemplificado a seguir.

(b) “Conchavo *quer dizer* acordo” (dado 31) – SN.

(c) “Isto *quer dizer* que a poluição ainda pode aumentar” (dado 66) – pronome demonstrativo explícito.

(d) “Por essa teoria, a dispersão inicial dos átomos teria sido absolutamente uniforme. Ø *Quer dizer* que o universo primordial nasceu liso, com uma face muito suave, sem caroços” (dado 23) – pronome demonstrativo implícito.

(e) “Bled El Feith”, que *quer dizer* Terra da Sede” (dado 32) – pronome relativo.

(f) “Elaborar uma nova lei não *quer dizer* que estaremos adquirindo uma mentalidade moderna” (dado 81) – sentença-sujeito.

O resultado estatístico concernente ao aspecto formal do contexto anterior está exposto na tabela a seguir.

CONTEXTO ANTERIOR	FREQÜÊNCIA E PERCENTUAL
Ausência de sujeito	64 (29%)
Presença de sujeito:	
<i>isso, isto, o que</i> – explícitos	59 (27%)
SN	38 (17%)
<i>isso, isto, o que</i> – implícitos	28 (13%)
<i>que</i>	19 (9%)
sentença	10 (5%)
	154 (71%)
Total	218 (100%)

Tabela 1: Distribuição da freqüência de *quer dizer* de acordo com o contexto anterior

Note-se que, enquanto 29% dos dados se apresentam sem vínculo sintático com o contexto anterior, 71% apresentam algum vínculo, ou com pronome demonstrativo explícito

ou implícito (somando 40%), ou com SN (17%), ou ainda em menor escala com pronome relativo (9%) ou com uma sentença (5%).

Além dos antecedentes formais, foram controlados também aspectos formais do contexto posterior ao *quer dizer*. Nesse caso, foi considerada a variável *com* ou *sem vínculo sintático*, cujos fatores e suas especificações são caracterizados e exemplificados a seguir.

1) Com vínculo sintático – foram sistematizados os seguintes elementos formais: presença de SN, de sentença, da conjunção *que*, da expressão *o seguinte*, da construção *com X*, de dois pontos e de ponto de interrogação, conforme respectivamente exemplificado abaixo.

(g) “Goma *quer dizer* casa” (dado 05) – SN.

(h) “... que na língua dos gringos *quer dizer* uma maneira completamente diferente de quebrar o pescoço” (dado 169) – sentença.

(i) “Então *quer dizer* que ele não é mais uma paixão nacional?” (dado 04) – conjunção.

(j) “Isso *quer dizer* o seguinte: o presidente e a inflação baixa vão bem, mas o governo está mal.” (dado 149) – *o seguinte*.

(k) “... o que *quer dizer* com essa falta de inocência?” (dado 189) – *com X*.

(l) “(...) fica proibido de entrar no campo a qualquer momento, em qualquer circunstância, a qualquer pretexto. *Quer dizer*: a CBF quer bancar no campeonato nacional o mesmo rigor que a FIFA usa no mundial” (dado 164) – dois pontos.

(m) “Que é que *quer dizer*?” (dado 42) – ponto de interrogação.

2) Ausência de vínculo sintático entre *quer dizer* e aquilo que o segue.

(n) “*Quer dizer*, o sujeito não pode evoluir três estágios à frente do que estão fazendo à sua época” (dado 37)

A tabela 2 mostra o total de ocorrências de *quer dizer* em relação ao contexto seguinte.

CONTEXTO POSTERIOR	FREQÜÊNCIA E PERCENTUAL
Sem vínculo sintático (,)	58 (26,5%)
Com vínculo sintático	
conjunção <i>que</i>	88 (40%)
SN	48 (22%)
dois pontos	10 (4,5%)

ponto de interrogação	04 (2%)
sentença	04 (2%)
<i>o seguinte</i>	03 (1,5%)
<i>com X</i>	03 (1,5%)
	160 (74,5%)
Total	218 (100%)

Tabela 2: Distribuição da freqüência de *quer dizer* de acordo com o contexto posterior

Observa-se que o contexto posterior predominante é com vínculo sintático: a conjunção “que” obteve 40% das ocorrências; em seguida, aparece o contexto sem vínculo sintático (26,5%) e, posteriormente, o contexto mais freqüente é do tipo SN, totalizando 22% das ocorrências. As demais especificidades relativas ao contexto com sujeito mostraram-se pouco freqüentes. Tais resultados serão retomados mais adiante.

2.2. Funções de *quer dizer*

Nesta seção, objetiva-se, a partir de observações dos contextos nos quais o *quer dizer* ocorre, padronizar os sentidos apresentados pela expressão, dado que “a definição de padrão adotada para o estudo inclui a necessidade da descrição dos sentidos associados às colocações recorrentes” (SARDINHA, 2004, p. 229).

Foram identificados, qualitativamente, cinco padrões, que, para fins deste trabalho, defino como funções indicadoras de sentidos específicos: as funções *esclarecedora* (I) e *corretiva* (II), que possuem um caráter mais textual; as funções *de significação* (III) e *modal + verbo dicendi* (IV), que possuem um caráter mais lexical/frasal, ambas atuando com valor verbal; e, por fim, a função *ambígua* (V), que parece sobrepor as funções (III) e (I), com predomínio ora do valor de significação ora do valor esclarecedor.

As funções (I), (II), (III), (IV) podem ser mais bem identificadas substituindo-se, respectivamente, o *quer dizer* por: *ou seja*, *aliás*, *significa* e *deseja dizer*. Também é possível perceber que a função (I) opera como reformuladora de uma idéia ou informação dada previamente para fins de esclarecimento; a (II) como retificadora, corrigindo uma informação prévia; a (III) como introdutora de sinônimo ou de definição; e a (IV) como de intenção de “dizer”. As *funções* desempenhadas pelo *quer dizer* são caracterizadas e exemplificadas a seguir.

(I) **Função esclarecedora** (*ou seja*) – os verbos perdem seu valor semântico individual e não há existência de um sujeito localizado anterior à expressão, sendo que o *quer dizer* assume um caráter relacional. No caso da função esclarecedora, o *quer dizer* pode, por exemplo, retomar uma idéia para enfatizá-la, concluir ou acrescentar alguma informação.

(o) “Eu acho que meio a meio é correto... *Quer dizer*, 100 por 200, qualquer coisa assim” (dado 08).

(II) **Função corretiva** (*aliás*) – similar à função (I), o *quer dizer* assume um caráter relacional, mas, diferente da função anterior, visa corrigir algo que já foi dito.

(p) “Temos um velho cacique do populismo autoritário e um político com formação intelectual e propostas modernizantes. *Quer dizer...* será que temos mesmo?” (dato 103).

(III) **Função de significação** (a *significa* b) – trata-se de uma locução verbal, formada pelo auxiliar “querer” + infinitivo “dizer”, funcionando como uma espécie de introdutor de sinônimo ou de definição, seja num nível menor (o lexical) ou num contexto maior (o frasal). Neste caso, o valor semântico do *quer dizer* encontra-se, em grande parte, no verbo “dizer”. Nessa função, ambos os itens verbais perdem seu significado original ao se fixarem combinados na expressão *quer dizer*, que adquire a acepção de “significa”.⁷

(q) “Gênero *quer dizer* masculino e feminino” (dato 94).

(r) “Essa é a premissa de que ser igual perante a lei *quer dizer* ser desigual na organização da economia, na estrutura da sociedade...” (dato 182)

(IV) **Função modal + dicendi** (*desejar dizer*) – o verbo “querer” opera como modal de intenção e juntamente com outro verbo pleno, no caso, o “dizer”, forma uma seqüência verbal com estatuto categorial de locução. Diferentemente de (III), aqui ambos os verbos mantêm seu significado original.⁸ O sujeito, neste caso, possui os traços semânticos [+ humano] e [+ intencional].

(s) “O senhor *quer dizer* que Fernando Pessoa era homossexual?” (dato 141).

(V) **Função ambígua**, que parece incluir (I) e (III) – o *quer dizer* pode operar tanto com função de significação como esclarecedora. Esta função talvez possa ser compreendida pelas considerações de Dal Mago & Görski (2002) sobre o percurso do *quer dizer* em seu processo de mudança semântica e categorial, que seria: (*aux*) *querer* + *dizer* = *significação* > *articulador textual ratificador* = *esclarecedor*. A função (V) talvez se enquadre neste processo de mudança, podendo assumir tanto características da função (III) como da (I).

(t) “Alice, bastante definida, é o conhecido filme de Walt Disney. Está presente nos seus objetos de cena, nos seus figurinos, talvez também na sua interpretação. *Quer dizer*, é o que se pode pensar quando aparece no palco, por exemplo, uma rainha histórica como aquela de Rodrigo López....” (dato 198).

(u) “... por que os fotógrafos precisavam de uma foto? Exatamente para atestar o boicote, provar que ele existiu. *Quer dizer*: a recusa da fotografia só podia ser dada por uma foto, a ausência pela presença”. (dato 26)

A tabela a seguir mostra os resultados estatísticos relativos às funções.

FUNÇÃO	FREQÜÊNCIA E PERCENTUAL
Significação (<i>significa</i>)	133 (61%)
Esclarecedora (<i>ou seja</i>)	49 (22%)
Modal + Dicendi (<i>deseja dizer</i>)	16 (07%)
Corretiva (<i>aliás</i>)	13 (06%)
Ambígua	07 (03%)
Total	218 (100%)

Tabela 3: Distribuição da freqüência de *quer dizer* de acordo com as funções

Percebe-se que a maior concentração de dados ocorre na função de *significação* (61%), seguida da função *esclarecedora* (22%). A primeira possui um caráter mais lexical/frasal e a segunda um caráter mais textual, funcionando como uma espécie de conector, conforme vimos nos exemplos. Em seguida, em termos de distribuição, estão as funções *modal + dicendi*, *corretiva* e *ambígua*.

Cabe, aqui, uma breve comparação do funcionamento destes dados de *quer dizer* com os oriundos de um corpus de fala (banco de dados do projeto VARSUL). Segundo DAL MAGO & GÖRSKI (2002), as freqüências encontradas em 659 dados de falantes da região sul são as seguintes: 65% das ocorrências são da função de *ratificação* (correspondente à *esclarecedora*, neste trabalho); 21% da função *retificadora* (correspondente à *corretiva*, neste trabalho); 12% de *planejamento verbal* (não encontrada em dados de escrita); e 2% de *significação*.

Comparando-se os resultados dos dois corpora, percebe-se que, em contexto escrito, há um predomínio da função de *significação* (61%), que, em dados de fala, se mostra em último lugar de freqüência (2%) – esta função parece ser mais característica de textos escritos e na fala, talvez, seja desempenhada por uma outra forma. Há também uma diferença acentuada na função *esclarecedora/ ratificadora* (22% na escrita e 65% na fala) com uma significativa diferença de 43 pontos percentuais entre os dois corpora. Percebe-se, nas duas amostras, uma distribuição complementar quanto à freqüência da função predominante de *quer dizer*: na escrita, a função de *significação*; na fala, a função *esclarecedora/ ratificadora*. A função *corretiva/ retificadora* aparece mais em dados de fala (21%) do que na escrita (6%); também é comum na fala a função de *planejamento verbal*, que visa, por exemplo, preencher as pausas no discurso, estando ausente nos dados de escrita.

2.3. Correlação entre contextos e funções de *quer dizer*

Postas as informações referentes às funções desempenhadas pelo *quer dizer*, resta uma análise sobre a correlação existente entre os contextos e as funções. Será possível encontrar certas regularidades contextuais nas funções levantadas? Esses aspectos são discutidos a partir dos resultados expostos nas tabelas 4 e 5.

FUNÇÕES CONT. ANTERIOR	Esclarecedora (ou seja)	Corretiva (aliás)	Significação (significa)	Mod + Dic. (deseja dizer)	Ambígua	Total
Ausência de sujeito	49	13	0	0	2	64
Sujeito SN	0	0	25	13	0	38
Sujeito <i>isso, isto, o que</i>	0	0	58	0	1	59
Sujeito-sentença	0	0	10	0	0	10
Sujeito (<i>isso, isto, o que</i>)	0	0	24	0	4	28
Sujeito <i>que</i> relativo	0	0	16	3	0	19
Total	49	13	133	16	7	218

Tabela 4: Correlação entre o tipo de contexto anterior e as funções de *quer dizer*

Nota-se que o contexto anterior predominante na função de *significação* é o pronome demonstrativo, explícito ou implícito (*isso, isto, o que...*), que retoma, na forma de sujeito, uma ou mais informações dadas previamente. A função de *significação* exige um contexto anterior ao qual, necessariamente, se remete, o que fica evidenciado pela falta de dados no contexto “ausência de sujeito”. Isso já não ocorre na função *esclarecedora*, o que fica claro pelas 49 ocorrências (100%) sem sujeito, categoricamente encontradas nesta função. Este mesmo contexto também equivale a 100% nos dados da função *corretiva*. Assim, as funções *esclarecedora* e *corretiva* se caracterizam por não possuir vínculo sintático à esquerda, atuando muito mais num nível de organização do texto/ discurso, como conectores num sentido amplo, ao contrário do que acontece com a função de *significação*.

A função *modal + dicendi* necessariamente exige um sujeito, o que pode ser visto pelas 13 ocorrências de contexto com SN (equivalente a 80%) e pelas 3 ocorrências de pronome relativo também funcionando como sujeito. Já a função *ambígua* ora parece retomar uma idéia anterior, com o demonstrativo implícito (4 dados) ou não (1 dado), ora aparece sem vínculo sintático (2 dados). Os dois exemplos abaixo ilustram essa função em dois contextos diferentes: um com vínculo sintático anterior e outro sem.

(v) “... pretende mostrar a trama às 19:45 e reprisá-la às 21:45. Depois, vai optar por apenas um dos horários. Prevalecerá aquele que registrar a maior audiência. Ø *Quer dizer*, então, que se o telespectador sintonizar a emissora às 19:45 não perderá nenhum minuto de novela?” (dado 21). *Contexto anterior implícito*.

(w) “... hoje o mundo é grande porque se pode andar por ele de carro, avião e conhecer tudo, também por uma antena parabólica que pega o mundo todo. *Quer dizer* que a TV manipula as pessoas, não é democrática, manipulada por uma pessoa que submete a empresa a seus interesses” (dado 170). *Sem vínculo aparente com contexto anterior*.

A tabela 5 exibe a correlação entre o contexto posterior e as funções de *quer dizer*.

FUNÇÕES CONT. POSTERIOR	Esclarecedora (ou seja)	Corretiva (aliás)	Significação (significa)	Mod + Dic. (deseja dizer)	Ambígua	Total
Ausência de vínculo (.)	44	12	0	0	2	58
Vínculo: SN	0	0	43	5	0	48

Vínculo: <i>que</i>	0	0	79	6	3	88
Vínculo: <i>o seguinte</i>	0	0	3	0	0	3
Vínculo: <i>com X</i>	0	0	0	3	0	3
Vínculo: sentença	0	1	3	0	4	4
Vínculo: dois pontos	5	0	2	1	2	10
Vínculo: ?	0	0	3	1	0	4
Total	49	13	133	16	7	218

Tabela 5: Correlação entre o tipo de contexto posterior e as funções de *quer dizer*

Verifica-se que o contexto posterior predominante é o com vínculo sintático, com 88 ocorrências da conjunção *que*, presente, em 90% dos casos, na função de *significação*, caracterizando a maioria das ocorrências posteriores nesta função. O *que* também aparece na maioria das ocorrências da função *modal + dicendi*. Observe-se que ambas as funções compartilham predominantemente as mesmas características em termos de contexto seguinte: com vínculo sintático (presença de *que* e de SN).

O segundo contexto posterior mais freqüente é o de ausência de vínculo sintático (58 ocorrências) e ele está presente, em 75% dos casos, na função *esclarecedora*, caracterizando a maioria das ocorrências desta função (44 em 49). Esse contexto também é o mais freqüente na função *corretiva*, com 92% das ocorrências. É possível afirmar que essas duas funções tendem a ter um contexto posterior similar. No que tange à função de *significação*, este contexto é ausente.

O terceiro contexto posterior mais recorrente é com vínculo sintático do tipo SN (48 ocorrências), presente predominantemente na função de *significação*, seguida da função *modal + dicendi*. Os demais contextos apresentam 10 ou menos ocorrências.

Observe-se, em relação à função *ambígua*, que ela compartilha características contextuais tanto da função de *significação (que)* como da função *esclarecedora* (ausência de vínculo; e dois pontos). Seguem exemplificados os dois contextos, um sem vínculo sintático e outro seguido de *que*:

(x)“Alice, bastante definida, é o conhecido filme de Walt Disney. Está presente nos seus objetos de cena, nos seus figurinos, talvez também na sua interpretação. *Quer dizer*, é o que se pode pensar quando aparece no palco, por exemplo, uma rainha histórica como aquela de Rodrigo López....” (dado 198). *Sem vínculo sintático posterior*.

(y)“... pretende mostrar a trama às 19:45 e reprisá-la às 21:45. Depois, vai optar por apenas um dos horários. Prevalecerá aquele que registrar a maior audiência. *Ø Quer dizer*, então, que se o telespectador sintonizar a emissora às 19:45 não perderá nenhum minuto de novela?” (dado 21). *Seguido da conjunção que*.

Vistas sob uma perspectiva de mudança lingüística, via gramaticalização, as funções aqui tratadas podem ser descritas como duas trajetórias distintas, a partir de uma mesma origem, que evidenciam a mudança semântica e categorial pela qual passam os itens lingüísticos combinados *querer e dizer* (cf. GORSKI; ROST; DAL MAGO, 2004):

I – (i) em sua origem, o verbo pleno *querer* (modal) combinado com o verbo *dizer* veiculam o significado de ‘intenção de falar’, requerendo um sujeito [+ humano]; (ii) esses verbos se cristalizam numa locução verbal fixada na 3ª pessoa do singular, passando a funcionar o *querer* apenas como auxiliar de *dizer* (*quer dizer*), com o valor de ‘significa’, perdendo a restrição sobre o tipo de sujeito selecionado; (iii) a expressão *quer dizer* perde seu estatuto verbal e passa a assinalar relações textuais, atuando como um elemento ratificador com valor de ‘ou seja’;

II – (i) (*idem* à anterior); (ii) os verbos *querer* e *dizer* perdem seu estatuto categorial e se cristalizam na expressão ‘quero dizer’, 1ª pessoa do singular, com valor de retificador equivalente a ‘digo’; (iii) a expressão *quero dizer* sofre redução fonética para *quer dizer* e continua funcionando como elemento retificador com valor de ‘aliás’.

A multifuncionalidade de *quer dizer* observada hoje recobre desde valores lexicais/verbais, que se manifestam nas funções *modal + dicendi* e *significação*, com escopo frasal (núcleo de predicado), até valores relacionais, especialmente na função *ou seja*, com escopo textual (ligando enunciados), e valores mais pragmáticos, como é o caso da função *aliás*, de caráter corretivo.

2.4. Os sentidos dicionarizados de *quer dizer*

Uma consulta ao dicionário Houaiss de Língua Portuguesa mostrou que o funcionamento da expressão *quer dizer* está parcialmente contemplado no dicionário, conforme descrito a seguir. No verbete “querer”, encontra-se:

“1. ter a intenção de dizer; 2. ter o significado de, dar a entender, equivaler a; 3. em frase interrogativa cujo complemento é uma oração subordinada, funciona como pedido de explicação de algo dito ou subentendido; 4. sem sujeito ou complemento, explica melhor ou emenda; isto é, ou seja” (DICIONÁRIO HOUAISS DA LÍNGUA PORTUGUESA, 2001, p. 2355).

Dentre as funções levantadas neste trabalho, três delas são também registradas pelo dicionário, conforme as correlações seguintes: a função *modal + dicendi* (1); a função de *significação* (2) e a função *esclarecedora* (4). A função de *correção* (*aliás*) não é listada pelo dicionário no item lexical “querer”, o que nos levou a supor, num primeiro momento, que não havia uma forma dicionarizada do *quer dizer* correspondente a esse sentido.

Como, no entanto, a expressão em estudo é constituída por dois itens, nos remetemos ao verbete “dizer” e constatamos que o único sentido para *quer dizer* que o dicionário descreve naquele verbete é o de *isto é*. Observe-se que tal acepção corresponde ao significado 4, extraído do item “querer” (“explica melhor ou emenda; isto é, ou seja”), conforme descrito acima. Há, ainda, no dicionário, menção ao seguinte sentido possível para *dizer*: flex. na 1ª p.s. do pres. Ind., possui *valor corretivo* (grifo acrescido) ou explicativo, equivalendo a *isto é, ou seja, a saber, por exemplo*. O dicionário prevê, portanto, o uso de *digo* com valor

corretivo; só que inclui também para essa forma o valor explicativo (que corresponde, em nossa análise, à função *esclarecedora*).

Em resumo, o dicionário não contempla claramente todas as funções encontradas no corpus para *quer dizer*. Algumas observações merecem ser feitas: (i) se, por um lado, considerarmos as funções de *quer dizer* listadas no item verbal “querer”, a função *corretiva* é ignorada; (ii) se, por outro lado, considerarmos as funções de *quer dizer* listadas em “dizer”, encontramos apenas o sentido de *isto é*. Como não há exemplos, infere-se que essa função equivale a que identificamos como *esclarecedora*, cuja forma prototípica é *ou seja*. Uma menção à função *corretiva* aparece apenas associada à forma *digo*, numa das acepções listadas para o verbo “dizer”. Como, segundo o dicionário, *digo* equivale a *isto é*, *ou seja*, etc, pode-se inferir que *digo* também corresponde a *quer dizer*, já que ambas as formas podem ter a acepção de *ou seja* e *isto é*. Fica, portanto, evidente que a função *corretiva* não é satisfatoriamente descrita no dicionário. Considere-se também que em nenhum momento aparece a possibilidade de uso como *aliás*.

Considerações finais

A análise dos dados escritos de *quer dizer* possibilitou a sistematização de cinco funções, assim distribuídas segundo a frequência de uso: de *significação*, *esclarecedora*, *modal + dicendi*, *corretiva* e *ambígua*. Esta última parece compartilhar características da função de *significação* e da função *esclarecedora*, tanto em termos de valor como de aspectos formais do contexto.

A análise das funções elencadas acima e dos contextos nos quais elas aparecem cria condições para a sistematização de uma identificação automática das funções com base nos contextos adjacentes. Assim, tem-se que: a função de *significação* fornece ao *quer dizer* características de uma locução verbal e esta função possui, como predominante no contexto anterior, a presença de sujeito – aparecimento dos pronomes demonstrativos *isto*, *isso*, *o que* – e no contexto posterior a presença de vínculo sintático – alta frequência da conjunção *que*. Já as funções *esclarecedora* e *corretiva* atribuem ao *quer dizer* propriedades de articuladores textuais, sendo que os contextos anterior e posterior que as caracterizam são “sem vínculo sintático”. A função *modal + dicendi* exige, no contexto anterior, um sujeito e, no contexto posterior, a presença da conjunção *que* e de SN.

Além disso, observou-se que o dicionário, no verbete “querer”, deixa de lado apenas uma função do *quer dizer*, que é a *corretiva*; essa mesma função é rapidamente mencionada no verbete “dizer”, porém associada a *digo* e não a *quer dizer*. Embora em número reduzido (a frequência dessa função no corpus analisado foi de 6%), as 13 ocorrências de *quer dizer* corretivo apontam para a pertinência de sua inserção no dicionário, especialmente pelo fato de o resultado ser referente a dados de escrita, que costumam mais lentamente incorporar os usos primeiramente registrados na fala (veja-se que nos dados de fala o percentual para essa função foi de 21% (cf. DAL MAGO, 2002)) Caberia, talvez, uma proposta de trabalho que comparasse os usos corretivos de *digo* e de *quer dizer* para avaliar o quanto a primeira forma tem aparecido em dados de escrita.

ABSTRACT: Based on a Corpus Linguistics methodology, patterns of the expression *quer dizer* in written data are described, through: the analysis of the previous and posterior contexts of the expression; the systematization of the functions played by *quer dizer*; the correlation between the functions in written and spoken data and between written data and the dictionary. Five different functions, located both in spoken and written data, were identified for *quer dizer*; one of these functions – the corrective one – is not considered by the Houaiss dictionary in the word *querer*.

Keywords: *Quer dizer*; Function; Spoken and written context; Dictionary.

Notas

- 1 Dal Mago (2002) analisou dados de fala oriundos do Projeto VARSUL (Variação Lingüística Urbana na Região Sul do Brasil).
- 2 Para Sardinha (2000), “A Lingüística de Corpus não é uma disciplina tal qual a psicolingüística, sociolingüística, semântica etc., pois seu objeto de pesquisa não é delimitado como em outras áreas” (SARDINHA, 2000, p. 355).
- 3 Segundo Rocha (2003), “Modelos probabilísticos podem definir quais os elementos – tais como contextos lingüísticos de qualquer natureza, fatores sociais ou faixa etária – que influenciam na probabilidade de ocorrência de uma ou outra estrutura, especificando, inclusive, o momento em que uma determinada estrutura é descartada, e uma outra se torna estável” (ROCHA, 2003, p. 205).
- 4 Nota-se, a partir dos anos 90, um movimento da lingüística computacional em direção à utilização de métodos baseados em análise estatística de corpus, motivado pelo aumento substancial da velocidade dos computadores, pela possibilidade de acesso a uma ampla variedade de dados lingüísticos *on line* e pelo amplo desenvolvimento dos métodos estatísticos; a partir das anotações oriundas dos estudos da lingüística de corpus, muitos padrões de uso da língua passaram a ser sistematizados e transformados em programas computacionais (MANI, 2006).
- 5 “Uma concordância é uma lista de exemplos de uma determinada palavra, expressão ou morfema, apresentados no contexto em que ocorrem em um corpus” (ROCHA, 2003, p. 214).
- 6 Apesar de haver presença de ponto, infere-se pelo contexto a retomada da idéia antecedente.
- 7 Observe-se que, nessa acepção, *quer dizer* tem seu uso cristalizado na terceira pessoa do singular.
- 8 Nesse caso, a locução formada por *querer* + *dizer* pode ser flexionada em todas as pessoas do discurso (eu quero dizer, tu queres dizer...).

Referências bibliográficas

- BIBER, Douglas; CONRAD, Susan; REPPEN, Randi. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press, 1998.
- DAL MAGO, Diane; GÖRSKI, Edair. *Quer dizer: Um elemento lingüístico com múltiplas funções*. In: VANDRESEN, Paulino (Org.). *Variação e Mudança no português falado da região Sul*. Pelotas/RS: EDUCAT, 2002.
- DIAS, Nilza Barrozo. A gramaticalização de “quer dizer”. In: GONÇALVEZ, Sebastião C. L.; CASSEB-GALVÃO, Vânia C.; LIMA-HERNANDES, Maria Célia (Orgs.). *Introdução à Gramaticalização*. São Paulo: Parábola, 2007.

- GERBER, Regina Márcia. *VIDE BULA: auxiliares modais em textos de bulas de medicamentos em francês e em português*. Florianópolis, 2004. Dissertação (Mestrado em Lingüística) – Centro de Comunicação e Expressão, Universidade Federal de Santa Catarina.
- GORSKI, Edair; ROST, Cláudia DAL MAGO, Diane. Aspectos pragmáticos da mudança via gramaticalização. In: CHRISTIANO, M. E. et al. (Orgs.). *Funcionalismo e gramaticalização: teoria, análise, ensino*. João Pessoa: Idéia, 2004.
- HOUAISS, Antonio; VILLA, Mauro de Sales; FRANCO, Francisco Manoel de Mello (Orgs.). *Dicionário Houaiss da Língua Portuguesa*. Rio de Janeiro: Editora Objetiva, 2001.
- KOCH, Ingedore. *A inter – ação pela linguagem*. São Paulo: Contexto, 1995.
- MANI, Inderjeet. Computational Linguistics. In: FASOLD, Ralph W. e CONNOR-LINTON, Jeff (Eds.). *An Introduction to Language and Linguistics*. Cambridge: Cambridge University Press, 2006. p. 465-491.
- MARTELOTTA, Mário E. Marcadores discursivos e operadores argumentativos. In: VOTRE, Sebastião; MARTELOTTA, Mário E. (org.) *Trajetórias de gramaticalização e discursivização*. Rio de Janeiro, 1998.
- ROCHA, Marco. Aspectos Cognitivos e Computacionais do Tratamento da polissemia através de métodos estatísticos. *Fórum Lingüístico*, v.3, n.2: 191-237, 2003.
- SARDINHA, Tony Berber. *Lingüística de Corpus*. Barueri, São Paulo: Manole, 2004.
- _____. Lingüística de Corpus: histórico e problemática. *D.E.L.T.A.*, vol. 16, no 2: 323-67, 2000.