

## CARACTERÍSTICAS DISCURSIVAS EM LÍNGUA ESPANHOLA: UMA CONTRIBUIÇÃO À LINGÜÍSTICA DE *CORPUS*

Barbara Soares da Silva Dias (PUCSP)

**Resumo:** Prover assessoria em uma língua estrangeira é uma necessidade na medida em que a economia se internacionaliza. O presente artigo analisa prestação de serviço de TI inserida aos serviços de gerenciamento de Tecnologia da informação (ITSM) conforme definido pela metodologia ITIL (Biblioteca de Infraestrutura de Tecnologia da Informação) de uma empresa prestadora de serviços. O método da pesquisa consiste dos seguintes passos: um conjunto de ligações foi colhido, criando um *corpus* de 70 textos tendo sido completamente etiquetado com o etiquetador Biber Spanish Tagger. Segundo, os dados foram submetidos ao *software* estatístico SPSS, e uma anotação manual dos níveis de *performance* de qualidade e suas características linguísticas foi executada por meio de taxonomia desenvolvida para essa pesquisa, baseada em Michael Hoey (1983). Justifica-se a relevância do estudo como contribuição à Linguística de Corpus. Os resultados indicaram baixa recorrência de dificuldade de linguagem nas ligações para o que era esperado; embora esta baixa recorrência tenha causado impacto no departamento e justifica que tais resultados possam ser trabalhados em uma rotina de treinamento.

**Palavras-chave:** *Corpus*; Língua Espanhola; Marcadores discursivos.

**Abstract:** Provide assistance in a foreign language is a need as economy becomes international. The present article analyses IT service rendering belonging to ITSM, information technology management service initials that, according to defined methodology, ITIL (Infrastructure Information Technology Library) in a service rendering company. Methodology consists of the following steps: a set of calls has been collected enabling a Corpus made of 70 texts which have been entirely tagged by the Biber Spanish Tagger. Secondly, data has been submitted SPSS statistics software and annotation of performance levels and its linguistic characteristics executed through use of taxonomy developed to this research, based on Michael Hoey (1983).

**Keywords:** Corpus; Spanish Language; Discourse Markers.

## CONSIDERAÇÕES INICIAIS

A pesquisa desenvolvida por ocasião da realização do Mestrado em Linguística na Pontifícia Universidade Católica de São Paulo identificou e estudou as características linguísticas de uma equipe em microinformática ao prestar serviços em língua espanhola a clientes que não estão habituados ao uso da Língua Espanhola como um meio de linguagem neste tipo de serviço. O presente estudo está fundamentado no âmbito da Linguística de *Corpus* (LC) e, portanto, na Linguística Aplicada (LA).

Sendo assim, para dar início, temos que a Linguística de *Corpus* é uma das vertentes da Linguística Aplicada (LA), que, por sua vez, centra-se na resolução de problemas de uso da linguagem, e por esta razão pode ser tida como uma ciência social (SILVA e ALVAREZ, 2008, p.11).

A relevância do estudo reside no fato de que ainda não há uma análise com base na Linguística Aplicada deste teor na literatura, tampouco foram encontradas publicações na área voltando-se em especial ao aspecto aqui explorado.

É importante esclarecer que o suporte em modelo da microinformática está sob os moldes da metodologia ITIL (*Information Technology Infrastructure Library*)<sup>1</sup>. A primeira

1 No final da década de 1980, o governo britânico desenvolveu um padrão para gerenciamento de serviços em tecnologia chamado de Information Technology Infrastructure Library (ITIL), que demonstrou ser útil em todos os setores de gerenciamento de serviços, vindo a ser adotado

situação considerada para o estudo foi o cenário em que os analistas contratados demonstrassem conhecimento avançado em língua espanhola (de acordo ao Quadro Europeu de Referência para Línguas) – sendo que, no momento de sua contratação, pudessem atuar em um nível C2 de proficiência no idioma. Uma informação relevante é que parte destes analistas nasceu em país falante de língua espanhola e outros vivem há muitos anos no Brasil e há também novos analistas brasileiros contratados e com vivência em língua espanhola demonstrando o mesmo nível C2 de proficiência.

A assistência neste seguimento exige o bom desempenho em uma ou mais línguas estrangeiras, entretanto, constatase que boa parte dos profissionais na área não demonstra ser, necessariamente, fluente em outros idiomas.

Identificou-se neste cenário estudado, carência de profissionais com formação em Tecnologia da Informação e simultaneamente proficientes em uma segunda língua, especificamente neste estudo, em Espanhol. Assim, demarca-se que este é considerado um diferencial para a inserção no mercado de trabalho e para o desenvolvimento profissional em corporações de distintas áreas.

---

em diversas empresas desde então. Em meados da década de 1990, o ITIL foi reconhecido mundialmente como um padrão para gerenciamento de serviços para a área de tecnologia da informação. Dentro do ITIL, o Service Desk é uma função que auxilia várias áreas, sendo um meio único de contato para se gerenciar as requisições e solicitações de serviços. Com o crescente aumento da área de Tecnologia da Informação (TI), o Service Desk passou a ser crucial para manter o suporte (PINHEIRO, 2006).

## 1. A LINGUÍSTICA DE *CORPUS*

O referencial da presente investigação é a Linguística de *Corpus*, área da Linguística Aplicada que entre seus contextos e usos, concebe-se de forma clara pelo uso de *corpora* computadorizados. A Linguística de *Corpus* é um campo que se dedica à criação e análise de *corpora* (plural em latim de *corpus*), de conjuntos de textos e também de transcrições de fala armazenadas em arquivos de computador, que vem aprimorando a maneira como se investiga a linguagem nos seus mais diversos níveis, colocando à disposição de quem analisa quantidades de dados antes inacessíveis devido sua quantidade e em exatidão numérica.

Um dos grandes agentes dessa mudança é a informática: sem ela, a Linguística de *Corpus* contemporânea não poderia existir da forma como existe hoje. Assim, observa-se que o linguista de *Corpus* depende de programas de computador para lidar com *corpora*. (BERBER SARDINHA, 2009). A LC pode ser considerada como “a face moderna da linguística empírica” (TEUBERT, 1996, p.06), sendo a linguagem compreendida também como um fenômeno social e analisada a partir de atos concretos de comunicação, isto é, textos reais, buscando o significado em que este é negociado, ou seja, no discurso. A este respeito, José Carlos Paes de Almeida Filho menciona que:

O trabalho (de pesquisa, porque a LA é uma área de pesquisa em Linguagem) a partir da superfície no texto do autor significa ocupar-se das questões manifestas da linguagem na prática social da forma como ela se apresenta a nós no teor de complexidade que possui sempre. Essa é a marca definidora da área teórica de pesquisa aplicada no âmbito da linguagem humana (a LA) há algum tempo também para vários autores da LA brasileira (vide Almeida Filho, 1991, por exemplo). Fenômenos aplicados são aqueles que envolvem justamente o uso real da língua no cotidiano de circulação em que está posta. (ALMEIDA FILHO *Apud* SILVA; ALVAREZ, 2008, p.27)

A Linguística de *Corpus* “trabalha dentro de um quadro conceitual formado por uma abordagem empirista e uma visão da linguagem como sistema probabilístico”, afirma Berber Sardinha (2004, p.30), aproximando-a de uma visão hallidayana, havendo a primazia dos dados provenientes da observação da linguagem reunida no formato de um *Corpus*.

O autor aponta como um dos elementos centrais da conceituação da Linguística de *Corpus* – além do empirismo – a visão probabilística da linguagem, e afirma que “a visão da Linguagem como sistema probabilístico pressupõe que, embora muitos traços linguísticos sejam possíveis teoricamente, não ocorrem com a mesma frequência” (BERBER SARDINHA, 2004, p.31). Ele ainda explica a oposição entre as duas vertentes linguísticas:

Aqui fica mais evidente a oposição entre Halliday e Chomsky. Halliday vê a linguagem como probabilidade, enquanto Chomsky a enxerga como possibilidade. A linguística chomskyana gerativista enfatiza a determinação de quais agrupamentos sintáticos são possíveis (permissíveis) dado o conhecimento que um falante nativo possui de sua língua. Já a linguística hallidayana descreve a probabilidade dos sistemas linguísticos, dados os contextos em que os falantes os empregam. (BERBER SARDINHA, 2004, p.30)

Levando-se em conta a visão da linguagem como sistema probabilístico, bem como o fato de que a experiência única de um analista de atendimento em atuar em um determinado departamento corporativo pode trazer a ele um conhecimento formador de ideia, a métrica da qualidade dos atendimentos em uma operação *Service Desk* é hoje feita por meio dos critérios estabelecidos pela área de qualidade interna da empresa prestadora de serviço e objeto deste estudo.

## **2. A INTERAÇÃO COMUNICACIONAL E A ANÁLISE LINGUÍSTICA PARA O PRESENTE ESTUDO**

Por meio da escuta diária de ligações em Língua Espanhola por cerca de 18 (dezoito) meses, a percepção da necessidade de novos desenhos para a análise linguística em suportes telefônicos dentro do cenário estudado foi latente e recorrente para que a realização desta pesquisa fosse concretizada e encarada como fundamental. A medição da qualidade e monitoração constante dos níveis de

atendimento e medidas de correção de desvios da língua foi altamente relevante na comunicação entre quem demandava informações e de quem as recepcionava. Tognini-Bonelli (2001) explica a diferença de trabalhar com evidências empíricas identificadas por meio de *corpora* eletrônicos:

Um texto existe em um contexto único de comunicação como um evento de linguagem unificada mediada entre dois participantes: o *Corpus*, por outro lado, reúne muitos textos distintos e não pode ser identificado como um evento comunicativo único e coerente; as citações em um corpus expandido do formato *KWIC* (*key words in context*, palavras-chave dentro de um contexto) para incluir o número de palavras “n” remanescentes de fragmentos textuais e elementos representativos em um *Corpus* são os padrões de repetição e co-seleção. A diferença implica uma diferente leitura de ambos: o texto deve ser lido horizontalmente da esquerda para a direita, dando atenção às fronteiras entre unidades maiores tais como cláusulas, frases e parágrafos. Um *Corpus*, examinado primeiramente em um formato de *KWIC* (*key words in context*, palavras-chave dentro de um contexto) com o nódulo da palavra alinhado no centro, é lido verticalmente, varrendo os padrões repetidos atuais no contexto do nódulo.<sup>2</sup> (2001, p.2-3)

---

2 Tradução livre da autora. No original: “(...) a text exists in a unique communicative context as a single, unified language event mediated between two (sets of) participants; the corpus, on the other hand, brings together many different texts and therefore cannot be identified with a unique and coherent communicative event; the citations in a corpus expandable from the KWIC format to include n number of words remain fragments of texts and the significant elements in a corpus are the patterns of repetition and patterns of co-selection. This difference entails a different reading of the two: the text is to be read horizontally, from left to right, paying attention to the boundaries between larger units such as clauses, sentences and paragraphs. A corpus, examined first in KWIC format with the node word aligned in the center, is read vertically, scanning for the repeated patterns present in the co-text of the node.”

Ainda abordando a linguagem na literatura, com relação *corpus-driven approach* (abordagem direcionada pelo *Corpus*), esta em adaptação às leituras da área, é mencionada como crítica às pesquisas baseadas em *Corpus* (*corpus-based*), isto é, aquelas em que os *corpora* (e as ferramentas informatizadas de exploração) são utilizados apenas para confirmar teorias pré-concebidas (RÖMER, U, 2004). No presente estudo foi utilizado o *software Statistical Package for Social Science for Windows* (SPSS) licenciado pela IBM<sup>3</sup>, *International Business Machines Corporation* que gera análise estatística de dados. Permite realizar cálculos complexos além de visualizar resultados úteis para a pesquisa aqui desenvolvida.

Dessa forma, para o desenvolvimento da pesquisa foram coletados os diálogos que ocorreram entre os analistas de uma instituição americana líder global em Tecnologia da Informação, chamada aqui de empresa A; prestadora de serviços *Service Desk* e os clientes, funcionários de uma empresa multinacional brasileira, líder mundial em mineração; chamada aqui de empresa B.

---

3 O Statistical Package for Social Science for Windows (SPSS) é um software licenciado pela IBM, International Business Machines Corporation utilizado para análise estatística de dados e permite realizar cálculos complexos além de visualizar resultados de forma didática. É um pacote que inclui aplicação analítica, data mining, ou seja, mineração de dados, Text Mining, ou seja, a minúcia e mineração de dados em texto, entre outras aplicabilidades.



### 3. MÉTODO

O método para executar a pesquisa consistiu dos seguintes passos: um conjunto de ligações foi colhido, criando um *Corpus* de 70 textos tendo sido completamente etiquetado com o programa *Biber Spanish Tagger*. Em seguida, os dados foram inseridos em formatação adequada para o *software* estatístico SPSS, e uma anotação manual dos níveis de *performance* de qualidade foi executada por meio de taxonomia desenvolvida para a pesquisa, baseada em Michael Hoey (1983). Dentro das dependências da empresa A, mediante autorização e cumprimento dos procedimentos devidos, cerca de 1.000 (mil) áudios de atendimento *Service Desk* foram ouvidos e 70 (setenta) destes foram recortados e selecionados para análise linguística. O critério para seleção foi a escolha de interações que tiveram tempo médio de 09 a 12 minutos.

#### 3.1. Coleta do *corpus*

Os textos formam o *Corpus* de estudo (chamado de *Customer Service Corpus* ou CSC) contendo 13.322 *tokens* ou ocorrências. Cada texto refere-se a uma ligação e todas as ligações foram realizadas em Língua Espanhola. A partir da coleta do *Corpus* CSC (*Customer Service Corpus*) de atendimentos telefônicos entre usuários finais e analistas

de atendimento de *Service Desk* – em que o *Corpus* é apresentado inteiramente analisado, transcrito e submetido às ferramentas necessárias para análise -, o objetivo é: (1) identificar a relação de qualidade existente na comunicação dos atendimentos, (2) em seguida, identificar seus padrões de desempenho por grupos de baixa, média e alta *performance*, (3) efetuar uma anotação manual de todos os atendimentos segundo critérios de qualidade corporativos e segundo base metodológica estabelecida em Hoey (1983), (4) submetê-los à etiquetagem segundo o processador *Spanish Biber Tagger*, (5) anotar manualmente seus marcadores discursivos em espanhol, e (6) efetuar comparação entre os marcadores discursivos encontrados no *Corpus* CSC e os gerais encontrados na língua espanhola e verificar sua distribuição - mais frequente e menos frequentes.

Com base na distribuição dos problemas de atendimento, são propostas as seguintes questões de pesquisa:

A) Qual é a avaliação dos atendimentos segundo os critérios de qualidade da empresa? Como esses atendimentos podem ser divididos em grupos de *performance*?

B) Quais são os marcadores conversacionais usados com base em Samuel Gili Gaya (1943) e Martín-Zorraquino & Montolío-Duran (1988)? Há diferenças de frequências entre os grupos de atendimentos de baixa, média e alta *performance*?

C) Quais são as instâncias de ocorrência de *code mixing* nos turnos dos atendentes? Há diferenças destas frequências entre os grupos de atendimentos de baixa, média e alta *performance*?

D) Quais são os termos da microinformática (*Service Desk*) empregados pelos atendentes e pelos usuários? Há diferenças de frequências entre os grupos de (1) atendimentos de baixa, média e alta *performance*?

Temos então que os traços linguísticos não ocorrem de forma aleatória, sendo possível evidenciar e quantificar regularidades (padrões). É comum assim, sugerir que a linguagem é padronizada (*patterned*), isto é, há correlação entre os traços linguísticos e os contextos situacionais de uso da linguagem.

### 3.2. Análise dos dados

Para responder à pergunta de pesquisa 1: *Qual é a avaliação dos atendimentos segundo os critérios de qualidade da empresa? Como esses atendimentos podem ser divididos em grupos de performance?* A tabela (consta no apêndice A) foi submetida a um cálculo simples de média de avaliações. Para cada áudio escutado, foi colhida a resposta sobre o critério que um analista cumpriu e o critério que o analista não cumpriu na tabela. Por exemplo: O analista demonstrou conhecimento técnico durante atendimento? Se a resposta fosse que **sim**, no cálculo da média o analista ganharia 02

pontos; se a resposta fosse **não**, ele somaria 1 ponto. O peso 2 significa o dobro do peso 1. O peso 2 vale 2 e o peso 1 vale 1. Se a soma das características de qualidade é igual a 9, a média ponderada deve ser:

Após a média simples criada para os atendimentos

$$\text{Média} = \frac{P(2) \times 9 + P(1) \times 9}{P(2) + P(1)}$$

colhidos na pesquisa, estes foram submetidos à análise de *cluster*, por meio do *software SPSS*. Tudo isto gerou 03 grupos de *performance* dos analistas, que os classificou em desempenhos bons, regulares ou ruins e suas notas atribuídas. Tal análise permitiu responder à segunda parte da pergunta 1 da pesquisa. A anotação manual dos marcadores discursivos em Espanhol foi feita da seguinte forma para responder à pergunta 2: *Quais são os marcadores conversacionais usados com base em Samuel Gili Gaya (1943) e Martín-Zorraquino & Montolío-Duran (1988)?* Foi compilada uma lista dos marcadores mais frequentes encontrados no espanhol falado a partir de Samuel Gili Gaya (1943) e Martín-Zorraquino & Montolío-Duran (1988).

a) Foi feita busca por estes marcadores no *corpus* CSC, ao coletar a amostra dos marcadores conversacionais mais comuns encontrados na língua espanhola. Há classificações, tais como: os estruturadores de informação, os ordenadores,

os digressores, os conectores, os reformuladores, os operadores argumentativos, os marcadores conversacionais e os meta-discursivos conversacionais.

b) Em seguida, tendo como referência os critérios da tabela, os padrões foram comparados para responder à segunda parte da pergunta 2.

Para responder à pergunta 3: *quais são as instâncias de ocorrência de “code mixing” nos turnos dos atendentes? Há diferenças destas frequências entre os grupos de (1) atendimentos de baixa, média e alta performance?*, foi necessário efetuar uma anotação manual em tabelas (*file maker*) em que as dificuldades de comunicação em espanhol (aqui reconhecidas como *code mixing*) foram detalhadas, as expressões típicas e frequentes de um atendimento de suporte telefônico em âmbito de *Service Desk* foram identificadas, além das expressões inadequadas em espanhol que um brasileiro pode cometer ou que um falante desde o nascimento de língua espanhola porém vivendo por longa data vivendo no Brasil, como é este o caso, possa misturar sua habilidade oral com palavras do português.

Para responder à pergunta 4: *quais são os termos da microinformática (Service Desk) empregados pelos atendentes e pelos usuários? Há diferenças destas frequências entre os*

*grupos de atendimentos de baixa, média e alta performance?* foi necessário, inicialmente, listar manualmente e efetuar uma busca manual no *Corpus* sobre os termos frequentes da microinformática. São entendidas como termos da microinformática todas as palavras e expressões frequentes utilizadas por uma ilha de atendimento específica, neste caso, o *Service Desk*, e entendidas segundo normas da empresa A como termos técnicos ao invés de usos rotineiros de uma conversação fora deste ambiente. Em seguida, para responder à segunda parte da pergunta 4 da pesquisa, foi necessário efetuar uma comparação com a tabela e a análise de *cluster*) conforme foi feito também para as perguntas anteriores.

### **3.3. Dos resultados da pesquisa**

Os resultados obtidos na pesquisa permitem em primeiro lugar estabelecer os grupos de *performance* e identificar a qualidade de atendimento nas ligações em língua espanhola para o departamento de *Service Desk*.

#### **3.3.1. Pergunta 1**

Qual é a avaliação dos atendimentos segundo os critérios de qualidade da empresa? Como esses atendimentos podem ser divididos em grupos de *performance*? Segue amostra para observação na íntegra localizada no Quadro 1:

| Áudio | Conhecimento Técnico | Dificuldades com idioma | Script Técnico |
|-------|----------------------|-------------------------|----------------|
| 2     | Não                  | Não                     | Sim            |
| 3     | Não                  | Não                     | Não            |
| 4     | Sim                  | Não                     | Sim            |
| 5     | Não                  | Não                     | Não            |

**Quadro 1:** Coleta inicial (Fonte: elaborado pela Autora)

Em seguida, para compor a mescla foi utilizado quadro contendo as dificuldades de comunicação em espanhol, conforme segue:

**Quadro 2:** Dificuldades de comunicação de atendimento (Fonte: elaborado pela Autora)

| Dificuldades de comunicação no atendimento | Qtde |
|--|------|
| <i>me puede pasar sí o sí</i>              | 3    |
| <i>me sale error ahorita</i>               | 2    |
| <i>voy a chequear Okey</i>                 | 4    |

Foram calculadas as médias e os pesos por atendimento, além de terem sido identificados os grupos de baixa, média e alta *performance*.

| Gravação | Nota | Análise de <i>Cluster</i> | 3 grupos<br>(Análise de <i>Cluster</i> ) |
|----------|------|---------------------------|--|
| 2        | 10   | Bom                       | 1,0 até 5,0 Ruim                         |
| 4        | 5    | Regular                   | 5,0 até 8,0 Regular                      |
| 5        | 10   | Bom                       | 8,0 até 13,0 Bom                         |

**Quadro 3:** Identificação de grupos de *performance* (Fonte: elaborado pela Autora)

Para responder à segunda parte da pergunta 1, o grupo de baixa *performance* foi, então, reconhecido, segundo estatística do *software* SPSS, com notas de ranqueamento

entre 1,0 até 5,0; o nome dado para esse tipo de desempenho foi “baixo”. O grupo de média *performance* teve notas de 5,0 até 8,0 e é reconhecido como “médio”. O grupo de alta *performance* teve notas de 8,0 até 13,0. Segundo análise de *cluster*, os primeiros resultados observáveis a partir dos grupos identificados foram os seguintes: 18,8% refere-se à quantidade de atendimentos de baixa *performance*, 31,9% refere-se aos atendimentos de média *performance* e 49,3%, aos de alta *performance*, reconhecidos no quadro como “*baixo, regular e bom*”.

### 3.3.2. Pergunta 2

Quais são os marcadores conversacionais usados com base em Samuel Gili Gaya (1943) e Martín-Zorraquino & Montolío-Duran (1988)? Há diferenças de frequências entre os grupos de atendimentos de baixa, média e alta *performance*?

Entre as classificações estabelecidas e as comparações feitas com o uso da língua espanhola geral, somente os marcadores **meta-discursivos conversacionais** tiveram representatividade suficiente para traçar sugestões de uso real na língua conforme amostra do quadro a seguir:



| Áudio | Marcadores Meta-discursivos |
|-------|-----------------------------|
| 6     | <i>Ah</i>                   |
| 7     | <i>Mire</i>                 |
| 11    | <i>Sabes decir</i>          |
| 17    | <i>Bueno</i>                |

**Quadro 4:** Marcadores (Fonte: elaborado pela Autora)

Os meta-discursivos conversacionais referem-se aos vários tipos de *tokens*, ou *ocorrências*, linguísticas que um autor emprega no seu texto para guiar ou direcionar o leitor para o modo como deve entendê-lo, e o seu posicionamento em relação a ele (MAO, 1993, p.265). É possível observar que pelo fato justamente de o *Corpus* ser um *Corpus* falado, os **marcadores meta-discursivos conversacionais** foram os mais expressivos na colheita de amostras na tiragem do *Corpus* de estudo.

### 3.3.3. Pergunta 3

Conforme mencionado anteriormente, para responder a pergunta 3 (*quais são as instâncias de ocorrência de “code mixing” nos turnos dos atendentes? Há diferenças destas frequências entre os grupos de (1) atendimentos de baixa, média e alta performance?*) da pesquisa, por meio da *análise de cluster*, foram feitas as contagens do *code mixing*, interpretado aqui como sendo as *dificuldades de comunicação em espanhol que o analista lida ao interagir com o usuário*. Segue quadro em que consta atendimento de *code mixing* identificado:

| Áudio | Uso de <i>CodeMixing</i>        |
|-------|---------------------------------|
| 2     | <i>Nuevamente</i>               |
| 7     | <i>tiene error</i>              |
| 8     | <i>voy a pedir que la anote</i> |

**Quadro 5:** Fragmento *Code mixing* (Fonte: elaborado pela Autora)

É importante mencionar que a frequência de *code mixing* encontrada foi baixa. No entanto, ainda que a incidência seja baixa e por sua vez a reincidência também, sugere-se que o impacto operacional seja alto devido à seriedade do uso inadequado. Para responder à 2ª parte da pergunta 3, segue quadro onde constam as ocorrências de *code mixing* relacionadas aos grupos de *performance*:

| Áudio | Ranqueamento | Média por GRUPO | Ocorrência de <i>Code Mixing</i> |
|-------|--------------|-----------------|----------------------------------|
| 8     | Bom          | 13              | <i>ha percebido otras</i>        |
| 9     | Bom          | 9               | <i>Estamos con problemas</i>     |
| 10    | Regular      | 4               | <i>me puedes pasá</i>            |

**Quadro 6:** Fragmentos *code mixing* x grupos de *performance* (Fonte: elaborado pela Autora)

Observa-se que mesmo os grupos de alta *performance* com médias entre 8 e 13, que seriam as notas máximas segundo critério da análise por *cluster*, também há incidências de *code mixing*; tal critério não afeta, necessariamente, os padrões estabelecidos dentro do quadro inicial.

### 3.3.4. Pergunta 4

Quais são os termos da microinformática (*Service Desk*) empregados pelos atendentes e pelos usuários? Há diferenças

de frequências entre os grupos de (1) atendimentos de baixa, média e alta *performance*? Conforme mencionado anteriormente, para responder à 2ª parte da pergunta 4 foi efetuada comparação com a tabela de dados conforme amostra:

| Áudio | Termos da microinformática         |
|-------|------------------------------------|
| 02    | Cuando me llegue el correo         |
| 03    | Quiere que yo haga un nuevo ticket |
| 04    | Se trata de un tema de nivel local |
| 05    | Voy a generar un ticket            |

**Quadro 7:** Amostra tabela da dados (Fonte: elaborado pela Autora)

A partir dos grupos identificados de *performance* segundo a análise, foram feitas as contagens dos termos frequentes da microinformática em um *Service Desk*. A contagem foi feita manualmente e a identificação foi reconhecida segundo os termos usuais de um ambiente desta área em espanhol, que não se assemelha necessariamente a uma análise de conversação regular em língua espanhola. Segundo Alison Wray, por *chunks* entende-se aqui nesta pesquisa como unidades lexicais que são formadas por duas palavras ou mais (2005, p.85). O resultado obtido para encontrar os *chunks* ou palavras isoladas foi baixo e sugere-se que os termos podem ser identificados com maior recorrência no ambiente da microinformática. Para responder a 2ª parte da pergunta 4, “há diferenças de frequências entre os grupos de (1) atendimentos de baixa, média e alta *performance*?”

foi efetuada comparação com o resultado da análise de *cluster*, onde foram identificados 3 grupos de *performance* e uma mescla de dados conforme amostra do quadro a seguir:

| Escuta | Termos da microinformática                | Notas por análise de <i>cluster</i> |
|--------|---|-------------------------------------|
| 02     | <i>Cuando me llegue el correo</i>         | Bom                                 |
| 03     | <i>Quiere que yo haga un nuevo ticket</i> | Ruim                                |
| 04     | <i>Se trata de un tema de nivel local</i> | Regular                             |
| 05     | <i>Voy a generar un ticket</i>            | Regular                             |

**Quadro 8:** Mescla termos da microinformática (Fonte: elaborado pela autora)

Para melhor interpretação do quadro acima, as notas foram extraídas da identificação de grupos por áudio. Com o intuito de aferir melhor as informações, cabe revisar método da análise da 2ª parte da pergunta 4: foi estabelecida a comparação com alguns dos critérios encontrados nas classificações da tabela. Os critérios colhidos nesta pesquisa e que se assemelham ao uso dos termos da microinformática diretamente são os seguintes: a) O analista segue o *script* técnico sugerido pela empresa A? b) O analista demonstra uso de linguagem técnica durante atendimento?

Segue amostra contendo comparação com o seguimento de *script* técnico pelo analista e notas de *performance*:

| Escuta | Seguiu <i>script</i> técnico? | Notas |
|--------|-------------------------------|-------|
| 02     | Sim                           | Bom   |
| 03     | Não                           | Ruim  |

**Quadro 9:** *Script* técnico (Fonte: quadro elaborado pela autora)

Com relação ao uso dos termos da microinformática e uso de linguagem técnica, nome dado pelo critério da tabela elaborada segundo padrão da empresa A para atingir a excelência em qualidade, segue amostra de quadro comparativo de dados:

| Escuta | Uso Linguagem Técnica? | Notas |
|--------|------------------------|-------|
| 04     | Sim                    | Bom   |
| 06     | Sim                    | Bom   |

**Quadro 10:** Linguagem Técnica (Fonte: Quadro elaborado pela Autora)

Observa-se que, o que a empresa A nomeia como linguagem técnica pode ser interpretado como o uso de termos da microinformática. Desta forma, há uma relação direta com os analistas que atingiram o critério de uso da linguagem técnica e os analistas que atingiram o critério para uso de termos da microinformática.

### **CONSIDERAÇÕES FINAIS**

O grupo de desempenho regular detém 49,3% dos atendimentos e, segundo a análise estatística há menos uso de marcadores discursivos, no caso, o uso escasso de meta-discursivos, porém também menos uso de português onde deveria haver discurso em espanhol e os termos da microinformática aparecem como lineares. O grupo de desempenho baixo detém 31,9% dos atendimentos

e, segundo cálculo demonstra baixo uso de marcadores discursivos. Por outro lado, há baixo uso do português e termos da microinformática.

Ainda que alguns dos analistas contratados tenham demonstrado conhecimento avançado na língua espanhola após contratação e período de curva de aprendizado (nível C2 de proficiência no idioma), conclui-se que não necessariamente eles obtiveram níveis de qualidade de atendimento satisfatórios em relação à prestação de serviços. Foi possível ainda observar a presença de marcadores meta-discursivos juntamente com a fala do analista que soluciona o problema do usuário. É por meio justamente deles que os analistas seguem com o atendimento procurando usar-se de poder articulatório ou em outros momentos não demonstrando argumentação e isso pode ser observado tanto na tabela quanto na porcentagem de utilização de marcadores meta-discursivos.

## REFERÊNCIAS

- ALMEIDA FILHO, José Carlos Paes de (1993). *Dimensões comunicativas no ensino de línguas*. Campinas, SP: Pontes.
- BERBER SARDINHA, Tony (2000). “Linguística de corpus: Histórico e problemática”. *DELTA*, São Paulo, 16(2), 323-367.
- \_\_\_\_\_. (2004). *Linguística de Corpus*. Barueri, SP: Editora Manole.
- \_\_\_\_\_. (2009). *Pesquisa em Linguística de Corpus com WordSmith Tools*. Campinas: Mercado de Letras, p.272.
- HOEY, Michael (1983). *On the Surface of Discourse*. London: GeorgeAllen and Unwin, p.211.

PINHEIRO, F. A (2006). *Implantação de um Service Desk*: Um estudo de caso aplicando conceitos do Itil e do Pmbok. BPM Brasil, São Paulo.

RÖMER, U. (2004) “A corpus-driven approach to modal auxiliaries and their didactics”. In: SINCLAIR, John McHardy. *How to use corpora in language teaching*. Amsterdam: John Benjamins. p.185-199.

SILVA, K. A.; ALVAREZ, M. L.(2008) *Perspectivas de investigação em linguística aplicada*. Campinas: Pontes.

TEUBERT, W. (1996). “Comparable or parallel corpora?” *International Journal of Lexicography*, 9 (3), 238-264.

TOGNINI-BONELLI, E. (2001). *Corpus Linguistics at Work*. Amsterdam & Philadelphia: Benjamins.

**Barbara Soares da Silva Dias** é Doutoranda em Linguística Aplicada pela PUCSP, Mestre em Linguística Aplicada pela PUCSP, Licenciatura em Letras-Português pelo Instituto Claretiano e Bacharel em Secretariado Executivo Bilíngue pela PUCSP. Professora de Ensino Superior no Centro Paula Souza, FATEC BAIXADA SANTISTA, Professora de Ensino Superior no Instituto Sumaré e tradutora na Casa da Tradução.

*Recebido em 22 de maio de 2018.  
Aprovado em 27 de agosto de 2018.*