

# Flutuação dos modos indicativo e subjuntivo em orações iniciadas pelo advérbio talvez – breves explorações com base em corpus

Juliana da Silva Neto (PUC-Rio)\*1 | Cláudia Freitas (PUC-Rio)\*2

## Resumo

Este trabalho apresenta os resultados preliminares de uma pesquisa em que se buscou verificar a flutuação entre os modos subjuntivo e indicativo nas orações independentes introduzidas pelo advérbio de dúvida *talvez*, com o objetivo de contribuir para a descrição do português tendo em vista o seu ensino como língua não materna. A metodologia baseou-se em exploração de dois *corpora* de língua portuguesa, a saber: a coleção CHAVE (*corpus* majoritariamente de textos jornalístico) e o Museu da Pessoa (*corpus* de transcrições de entrevistas), ambos disponibilizados pelo projeto AC/DC, mantido pela Linguateca. Os principais resultados da pesquisa apontam para a maior flutuação entre os dois modos na linguagem oral e informal, dado que a maior ocorrência do modo indicativo nessas orações se deu no *corpus* Museu da Pessoa. Os resultados também indicam que essas orações são construídas com as formas nominais dos verbos, sobretudo no *corpus* CHAVE.

## 1. Introdução

Este trabalho visa a contribuir para a descrição do português para a área de ensino de português como segunda língua para estrangeiros, tendo a exploração de *corpora* eletrônico como subsídio para o estudo e descrição linguística (Sampson, 2002; Santos, 2008, 2014).

Um fenômeno que oferece desafios para o ensino de português para estrangeiros no Brasil é a flutuação entre os modos subjuntivo e indicativo. Tendo isso em vista, o objetivo deste artigo é apresentar os resultados preliminares de uma pesquisa cujo propósito foi o de analisar a flutuação entre os modos indicativo e subjuntivo em orações iniciadas pelo advérbio de dúvida *talvez*, através da exploração de dois *corpora* de língua portuguesa, tomados como representações das variantes formal/escrita e informal/falada considerando as seguintes perguntas norteadoras: i) Em qual dos *corpora* a flutuação entre indicativo e subjuntivo é maior?; ii) Quais os tempos do subjuntivo são mais usados com talvez?; iii) Quais tempos do indicativo são mais usados com talvez?; iv) Na estrutura talvez+subjuntivo, utilizam-se tempos do modo subjuntivo para os quais não há expectativa?, e v) As orações iniciadas por talvez são realizadas com quais formas nominais do verbo?

A pesquisa foi motivada pela constatação de uma das autoras, em sua prática de professora de português como língua não materna, de que as orações iniciadas

\*1 Mestranda do Programa de Pós-graduação em Estudos da Linguagem do Departamento de Letras da PUC-Rio e bolsista CAPES.

\*2 Professora doutora do mesmo Programa e colaboradora da Linguateca.

por talvez, na modalidade oral, sobretudo, ocorrem não só com o modo subjuntivo, modo recomendado pela norma padrão culta (Bechara, 2009, p. 280), mas com o modo indicativo. Apesar desta exigência da norma padrão, não raro encontram-se enunciados como os listados abaixo:

- (1) *FSP940116-027*: Mas se ela fosse utilizada de forma racional, **talvez não seria preciso** investir nesse setor durante dois ou três anos.
- (2) *FSP950720-112*: Batize seu filho e **talvez ele conseguirá um** parágrafo na Joyce.
- (3) *FSP940129-164*: Por exemplo, ninguém **talvez faz idéia**, e não a faz principalmente o governo, que é entretanto a quem mais interessava saber
- (4) *E144-BR-235*: Vocês **talvez são muito** jovens, não devem lembrar disso.
- (5) *E206-BR-60*: Quando entrei lá -- **talvez eu tinha 15** anos de idade -- pra quem vem de família judaica como eu, depois de fazer a cerimônia bar mitzvah, aos 13 anos eu ia na congregação na qual minha família estava vinculada e tinha duas vezes por semana filosofia judaica e outras questões.

Os enunciados (1), (2) e (3) foram retirados do *corpus* CHAVE (Santos e Rocha, 2005) e os enunciados (4) e (5), do *corpus* Museu da Pessoa (Santos et al. 2004). Nestes enunciados, em que se espera o uso do modo subjuntivo, há, pelo contrário, o uso do modo indicativo.

Os autores de compêndios gramaticais do português (Bechara, 2009; Cunha e Cintra, 2001), ao tecerem suas considerações sobre os modos subjuntivo e indicativo, raramente observam casos em que há interferência do modo indicativo em enunciados em que é previsto o uso do modo subjuntivo. As orações independentes iniciadas por talvez, especificamente, são mencionadas apenas por Bechara (2009). O gramático afirma que o subjuntivo, além de ser empregado em outras orações, nas “independentes optativas”, nas “imperativas negativas” e nas “subordinadas em que o fato é considerado como incerto, duvidoso ou impossível de se realizar”, é empregado nas “dubitativas com o advérbio talvez” (Bechara, 2009, p. 280).

No âmbito da pesquisa acadêmica, muitos trabalhos sobre o uso do modo subjuntivo em contextos em que há interferência do modo indicativo estão sendo desenvolvidos com fim de analisar, dentre muitos fatores, aqueles que levam ao uso do modo indicativo num enunciado cujos itens lexicais, gramaticais ou morfo-sintáticos exigem o uso do subjuntivo. Destacam-se, entre eles, os trabalhos de Gonçalves (1995), Medeiros (1996), Gonçalves (2003) e Almeida (2010), mas muitos outros poderiam figurar aqui.

## 1.2. O que é um corpus e qual a sua relação com a descrição da língua portuguesa?

Como o presente trabalho é baseado em exploração de *corpus*, torna-se necessário esclarecer o que se entende por *corpus*. Considera-se que *corpus* é uma coletânea de

textos autênticos, proferidos ou escritos em situações espontâneas de fala. A coletânea que compõe um *corpus* foi reunida com algum propósito, geralmente o de pesquisa em línguas; além disso, hoje em dia, um *corpus* está em formato eletrônico e possui dimensão considerável, o que justifica o seu processamento automático. (Sinclair, 2005; Santos, 2008; Freitas, 2015).

Por se considerar aqui que a descrição de qualquer fenômeno do português – e de qualquer língua – é enriquecida com a análise de *corpus*, que proporciona outra visão sobre os fenômenos linguísticos, é essencial explicitar a relação entre descrição de uma língua e, em especial, da língua portuguesa, e *corpus* – no modo como está sendo entendido aqui. A relação mais óbvia que há entre *corpus* e descrição linguística é que o *corpus* fornece material para tal descrição. O *corpus* seria, então, a melhor fonte de dados para estudar um fenômeno do português e para, a partir daí, descrevê-lo. Robert de Beaugrande (2002) e Geoffrey Sampson (2002) defendem essa posição – e outras assumidas nesse trabalho. Para os autores, o *corpus* pode ser um instrumento de mudança na atividade de descrição e estudo linguístico, porque as tecnologias disponíveis hoje em dia possibilitam a manipulação de grandes volumes de dados linguísticos, permitindo a exploração de grandes quantidades de texto. Esta nova atitude oferece ao linguista novas e mais poderosas possibilidades de generalizar sobre a língua, em estudos respaldados por dados linguísticos autênticos.

### 1.3. Como se relacionam corpus e teoria linguística?

Como afirma Freitas (2015), as diferentes maneiras de aproximar *corpus* e pesquisa, ou de relacionar *corpus* e teorias, costumam ser classificadas como abordagens *corpus-driven* (ou guiada por *corpus*) ou abordagens *corpus-based* (baseada em *corpus*). Ainda que, mais recentemente, essa discussão (e distinção) venha se diluindo, ela sintetiza não apenas diferentes maneiras de relacionar *corpus* e linguagem, mas também diferentes opções epistemológicas.

Segundo Freitas (2015), quando, a partir dos anos 1990, grandes *corpora* eletrônicos começam a participar sistematicamente da prática linguística, tem-se um ambiente propício para pôr em xeque os modelos teóricos vigentes. Nesse contexto, toma corpo boa parte do que se entende por uma abordagem *corpus-driven*, em que é comum a referência a dados ou fenômenos que emergem do *corpus*. Enfatiza-se a obtenção “desinteressada” de dados e a imparcialidade (ou distanciamento) do investigador, responsável por relatar dados ou que, a partir da observação distanciada, irá construir suas hipóteses e generalizações. O *corpus* é o espaço que promove uma observação neutra dos fatos da língua, que, por sua vez, irá promover a criação de hipóteses. Ainda de acordo com Freitas (2015), na abordagem *corpus-based*, o *corpus* assume o espaço de auxiliar na investigação de hipóteses prévias, servindo ora (i) como fonte de exemplos, ora (ii) como fonte de dados capazes de validar, refutar ou especificar uma

dada teoria. Neste caso, seriam as hipóteses, ou questões de pesquisa, as responsáveis por direcionar a observação dos dados, capazes de dar novos contornos a hipóteses e questões previamente formuladas, em um processo dialógico. Frequentemente, a utilização da tecnologia também é vista como garantia de neutralidade e objetividade na abordagem *corpus-based*, mostrando como a polarização entre ambas as perspectivas não é tão clara.

Com Beaugrande (2002), também reconhecemos que não há mais espaço para o abismo que separa, nos estudos atuais da linguagem, dados e teoria..

Rajagopalan (2007), cujo posicionamento muito se assemelha ao de Beaugrande (2002), critica a criação de uma teoria linguística a partir de conceitos previamente estabelecidos e o fato de haver seleção de dados para fundamentação desses conceitos:

Entre essas ortodoxias está a ideia de que a análise de uma língua x deve suceder à formulação de uma teoria a respeito do que é a língua, isto é, língua em seu sentido genérico e abstrato. Da mesma forma que línguas x, y, z etc., os dados concretos são convocados apenas para confirmar um conceito universal da língua, para servir de exemplificações de modelos ou hipóteses previamente estabelecidos. (Rajagopalan, 2007, p.24)

Ou seja, nessa visão ortodoxa, os dados têm a função de elucidar conceitos, elaborados sem nenhum respaldo nesses próprios dados. Rajagopalan (2007) menciona estudos feitos por pesquisadores da área da fraseologia (Cowie, 1998 *apud* Rajagopalan, 2007), que mostram que as expressões idiomáticas – consideradas pelos gerativistas um dado irrelevante para sua teoria por terem sentido obscuro e por não obedecerem às regras de composicionalidade estabelecidas nessa teoria – não são exceção à regra e “fazem parte de uma classe enorme de expressões que inclui “provérbios”, “frases feitas”, “expressões congeladas”, “sintagmas pré-fabricados” etc.(...)” (p.38). O autor continua:

ou seja, está-se descobrindo um falante que não está compondo frases novas (em tese, infinitas), a partir de elementos (...) finitos, como prevê a tese gerativista. Na verdade, o falante está utilizando “blocos” inteiros com muito mais frequência do que se imagina. (*ibidem*, p. 38)

No estudo citado pelo autor, prioriza-se a observação dos dados. Além disso, seus resultados estão possibilitando a revisão de conceitos postulados na teoria gerativa. A posição assumida aqui também sustenta que a observação (que nunca é neutra ou desinteressada) dos dados linguísticos é condição fundamental para elaborar/rever uma teoria linguística e para descrever a língua – e, para isso, o *corpus* é indispensável.

#### 1.4. Qual o papel do corpus no âmbito de descrições linguísticas?

Tendo em vista que o *corpus* contribui para a elaboração/reformulação das teorias linguísticas e de suas descrições, cabe avaliar que papéis/papel pode ter o *corpus* no

âmbito de descrições linguísticas. Para Beaugrande (2002), ao descrever uma língua, o linguista deve tentar alcançar o equilíbrio entre quantidade e qualidade; entre abrangência e profundidade; entre uniformidade e diversidade e entre regularidades e irregularidades. A quantidade de dados que um *corpus* oferece contribui positivamente para a qualidade da descrição. A quantidade influencia na abrangência da descrição, na medida em que mais dados aumentam a generalidade de uma descrição, captando, de forma mais abrangente, o fenômeno investigado. Os tipos de dados e os objetivos da descrição contribuem para a profundidade/especificidade da mesma.

No que se refere à uniformidade e à diversidade, a análise linguística com base em *corpus* contribui para o apagamento dessa dicotomia na medida em que revela a diversidade do sistema linguístico, que é atribuída somente ao discurso pelos que defendem o outro posicionamento sobre a linguagem. Além disso, através dos estudos linguísticos baseados em *corpus*, somos constantemente confrontados com as regularidades e as irregularidades da língua, que não devem ser deixadas de fora.

### 1.5. (Des)Vantagens com a exploração de corpus

Resta, no entanto, apontar as desvantagens que existem na exploração de *corpus*. A desvantagem mais evidente está no fato de que a modalidade escrita é a que mais está presente em grandes *corpora* eletrônicos. Portanto, o *corpus* apresenta limitações para pesquisas que precisam de uma grande quantidade de dados da modalidade oral transcritos. Ademais, os *corpora* que não são anotados apresentam limitações no que diz respeito a possibilidades de extração de informações linguísticas. Outra limitação é o fato de que um *corpus* com grande volume de dados, e anotado, ser altamente dependente de uma interface de busca amigável (Santos, 2008), o que frequentemente distancia os linguistas desse tipo de exploração. Embora haja essas limitações, elas não invalidam o trabalho de exploração de *corpus* linguístico.

## 2. Metodologia

Neste trabalho, como referido anteriormente, foram utilizados dois *corpora*, CHAVE (Santos e Rocha, 2005) e Museu da Pessoa (Santos *et al.* 2004), disponíveis na Linguateca. A Linguateca é um centro que fornece recursos para o processamento computacional da língua portuguesa. Ela é responsável pelo projeto/serviço AC/DC (Acesso a *corpus*/Disponibilização de *corpus*) (Santos e Bick, 2000; Santos e Sarmiento, 2003), público e gratuito, cujo objetivo é tornar acessíveis, através da internet, *corpora* de língua portuguesa por meio de uma interface simples (<http://www.linguateca.pt/ACDC>). Todos os *corpora* disponíveis no AC/DC foram processados pelo parser PALAVRAS (Bick, 2000), um sistema que faz, automaticamente, análise morfossintática dos textos, incluindo informações relativas aos lemas das palavras, à

classe de palavras e à função sintática, dentre outras informações. O AC/DC possibilita ao usuário fazer pesquisa através de expressões de busca e os resultados podem ser vistos em concordância ou em distribuição por lemas, classes de palavras, função sintática, por semântica de palavras etc., dependendo das funções por distribuição que estão disponíveis para cada *corpus*.

A coleção CHAVE é um *corpus* com 97.7 milhões de palavras, que é composto, sobretudo, por textos completos referentes aos anos de 1994 e 1995 do jornal português Público e do brasileiro Folha de São Paulo. O *corpus* Museu da Pessoa contém 1.4 milhões de palavras e é composto por entrevistas individuais cujo tema é a vida pessoal do entrevistado. Há 107 entrevistas com cidadãos portugueses e 106 entrevistas com cidadãos brasileiros. Ambos os *corpora* estão disponíveis *on-line* e têm acesso livre (cf. <http://www.linguateca.pt/ACDC/>).

O *corpus* CHAVE foi escolhido pelo seu tamanho e pela variedade de subgêneros da modalidade escrita que comporta. Já o *corpus* Museu da Pessoa foi selecionado por representar a modalidade oral do português. Os resultados de busca de ambos os *corpora* foram restritos à variante brasileira.

### 3. Análise dos resultados

Nesta seção, serão apresentados os resultados referentes à pesquisa nos dois *corpora*. Foram calculados, para os resultados das buscas em cada *corpus*, os percentuais de ocorrência de todos os tempos verbais dos modos indicativo e subjuntivo; das formas nominais dos verbos; dos tempos de maior frequência dos modos subjuntivo e indicativo e das formas verbais (entendidas aqui como o modo subjuntivo, o modo indicativo e formas nominais). Assim, nas seções a seguir, serão apresentados, primeiramente, com relação a cada *corpus*, a distribuição dos tempos verbais dos dois modos nas orações estudadas; em seguida, o percentual dos tempos mais frequentes do modo subjuntivo e indicativo; logo após, o percentual de ocorrência de cada forma nominal (infinitivo, gerúndio e particípio) e, por fim, os percentuais de ocorrência para cada forma verbal que ocorreu nesse tipo de oração.

#### 3.1. Distribuição dos tempos verbais as orações iniciadas por talvez em ambos os corpora

O objetivo desse trabalho foi analisar a flutuação entre os modos subjuntivo e indicativo nas orações iniciadas por talvez considerando os aspectos mencionados acima. Na Tabela 1, abaixo, apresenta-se a distribuição dos tempos verbais de ambos os modos que mais ocorreram no *corpus* CHAVE. A Tabela 2 é relativa ao *corpus* Museu da Pessoa e apresenta o mesmo tipo de distribuição.

Como se pode observar, nos dois *corpora*, o modo subjuntivo é o que apresenta a maior ocorrência nas orações iniciadas por talvez contidas tanto no CHAVE

quanto no Museu da Pessoa. Na Tabela 1, evidencia-se o uso dos tempos presente do subjuntivo e imperfeito do subjuntivo; do infinitivo, do particípio e do gerúndio; do presente e do futuro do pretérito do indicativo. Além disso, nota-se que é o particípio e o infinitivo que têm as maiores ocorrências depois dos tempos verbais do modo subjuntivo. Na Tabela 2, logo depois dos tempos presente e pretérito imperfeito do subjuntivo, aparecem o infinitivo e o presente do indicativo antes dos tempos compostos do subjuntivo. Desse modo, as duas tabelas tornam patente o uso desse tipo de oração com o modo subjuntivo e com o modo indicativo, bem como com as formas nominais do verbo.

Nas seções subseqüentes, outros percentuais serão analisados.

<b>Tempos</b>	<b>Ocorrências</b>	<b>%</b>
Presente do subjuntivo	1234	66,31%
Pretérito imperfeito do subjuntivo	250	13,43%
Pretérito perfeito composto do subjuntivo	183	9,83%
Particípio passado	49	2,63%
Infinitivo	41	2,20%
Pretérito mais que perfeito composto do subjuntivo	40	2,15%
Gerúndio	21	1,13%
Presente do indicativo	12	0,64%
Infinitivo + acusativo	6	0,32%
Futuro do pretérito	3	0,16%
Futuro do indicativo	3	0,16%
Perfeito ou mais que perfeito do indicativo	3	0,16%
Pretérito mais que perfeito do subjuntivo	2	0,11%
Presente do indicativo na forma progressiva	2	0,11%
Presente do indicativo na voz passiva	2	0,11%
Presente do indicativo na voz passiva	2	0,11%
Pretérito perfeito do indicativo	2	0,11%
Presente do indicativo no futuro perifrástico	2	0,11%
Presente na forma progressiva	1	0,05%
Pretérito Mais que perfeito do indicativo	1	0,05%
Pretérito imperfeito do indicativo na voz passiva	1	0,05%
Gerúndio + acusativo	1	0,05%
Pretérito perfeito do indicativo na voz passiva	1	0,05%
Pretérito imperfeito do indicativo	1	0,05%

**Tabela 1: Distribuição dos tempos verbais com talvez no *corpus* CHAVE**

Tempos	Ocorrências	%
Presente do conjuntivo	31	32,98%
Imperfeito do conjuntivo	19	20,21%
Infinitivo	9	9,57%
Presente do indicativo	9	9,57%
Pretérito perfeito composto do subjuntivo	7	7,45%
Pretérito mais que perfeito composto do subjuntivo	3	3,19%
Gerúndio	3	3,19%
Futuro do pretérito	3	3,19%
Pretérito perfeito do indicativo	3	3,19%
Pretérito imperfeito do indicativo na voz passiva	2	2,13%
Particípio passado	2	2,13%
Presente do indicativo no futuro perifrástico	1	1,06%
Pretérito perfeito ou mais que perfeito	1	1,06%
Pretérito imperfeito do indicativo	1	1,06%

**Tabela 2: Distribuição dos tempos verbais com talvez no *corpus* Museu da Pessoa**

### 3.2. Tempos do subjuntivo utilizados com talvez em ambos os corpora

Como este trabalho visa a contribuir com o ensino de português como língua não materna, o detalhamento dos tempos utilizados nas orações em questão é importante, visto que a ocorrência dessa estrutura não se dá com qualquer tempo do modo do subjuntivo e nem com qualquer tempo do modo indicativo. Nesta seção, o Gráfico 1 e o Gráfico 2 esclarecem essa afirmação. O Gráfico 1 mostra o percentual de ocorrência de cada tempo do subjuntivo no *corpus* CHAVE e o Gráfico 2 apresenta o percentual dessas ocorrências no outro *corpus* utilizado, o Museu da Pessoa. Os maiores percentuais de ocorrência foram para os tempos presente do subjuntivo, com mais de 72% no CHAVE e mais de 51% no Museu da Pessoa; pretérito imperfeito do subjuntivo, com mais 14% no CHAVE e mais de 31% no Museu da Pessoa; pretérito perfeito do subjuntivo, com 10,72% no CHAVE e 11,67% das ocorrências no Museu da Pessoa; pretérito mais que perfeito do subjuntivo, com 2, 34% das ocorrências no CHAVE e 5% no Museu da Pessoa.

O fato de os maiores percentuais de ocorrência terem sido para o tempo presente no CHAVE e para o tempo passado no Museu da Pessoa parece ter explicação. O Museu da Pessoa é um *corpus* composto por entrevistas em que as pessoas contam suas histórias de vida. Ou seja, nessas entrevistas, a fala das pessoas centra-se no passado e, por isso, usam muito mais os tempos do passado. Já texto jornalístico parece que se centra muito mais nas questões do presente. De fato, Berber Sardinha (2005), depois de estudo feito com *corpus* composto por textos do mesmo jornal

utilizado no CHAVE, Folha de São Paulo, concluiu que, por serem muito frequentes expressões relativas ao tempo presente, o jornal “passa a ser um instrumento de reflexão sobre o presente e sobre tendências futuras, mais do que uma crônica sobre eventos passados” (p. 240).

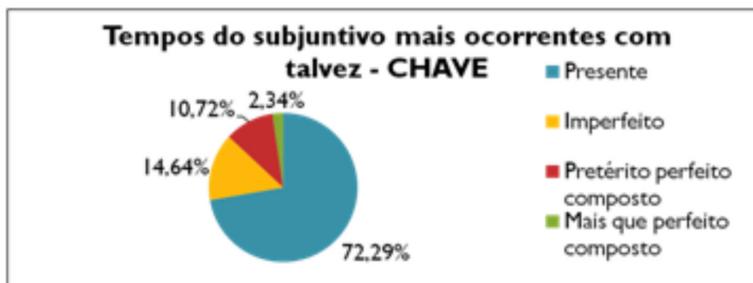


Gráfico 1: Percentuais relativos aos tempos do modo subjuntivo nas orações com talvez no CHAVE

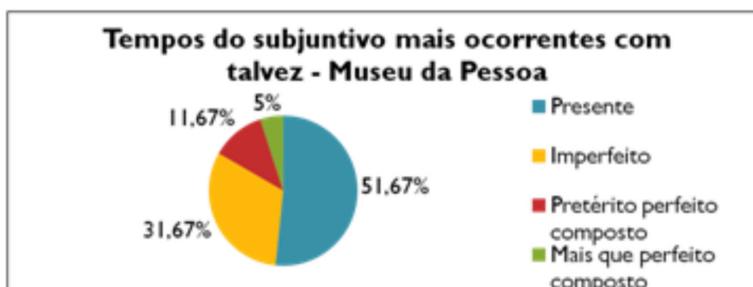


Gráfico 2: Percentuais relativos aos tempos do modo subjuntivo nas orações com talvez no Museu da Pessoa

### 3.3. Tempos do indicativo utilizados com talvez em ambos os corpora

Se a estrutura que está sendo analisada aqui não permite o uso de qualquer tempo do modo subjuntivo, tampouco permite o uso de qualquer tempo do modo indicativo. Para este modo, os maiores percentuais de ocorrência foram para os tempos presente, com mais de 46% de ocorrências para o CHAVE e 50% para o Museu da Pessoa, e futuro do pretérito, com 13,33% para o CHAVE e 15% para o Museu da Pessoa. O futuro do presente do indicativo representou 10% das ocorrências no *corpus* CHAVE. O percentual de ocorrência para esse mesmo tempo não foi significativo no Museu da Pessoa. Outros tempos, como o pretérito perfeito e imperfeito do indicativo, tiveram percentuais significativos de ocorrência nos dois *corpora*, como pode ser visto no Gráfico 3 e no Gráfico 4, abaixo.

Esses resultados podem indicar que, nas orações independentes introduzidas pelo advérbio talvez, não se usa apenas o modo subjuntivo, mas o modo indicativo. Embora outras explorações tenham que ser feitas, os percentuais relativos aos tempos deste modo e do modo subjuntivo podem evidenciar a flutuação entre o presente do subjuntivo e o presente do indicativo e entre o pretérito imperfeito do subjuntivo e o futuro do pretérito do indicativo.

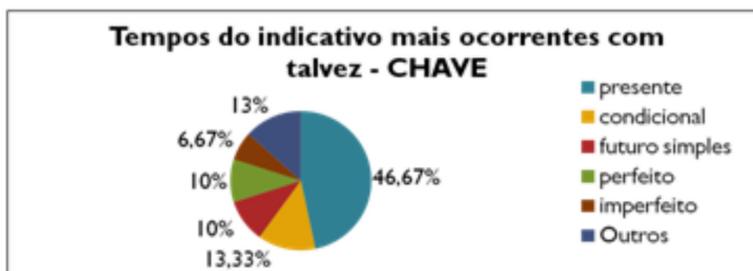


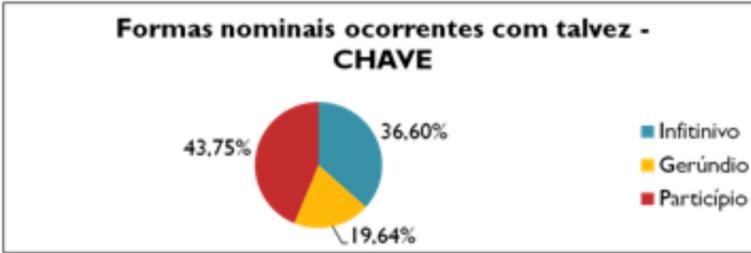
Gráfico 3: Percentuais relativos aos tempos do modo indicativo nas orações com talvez no CHAVE



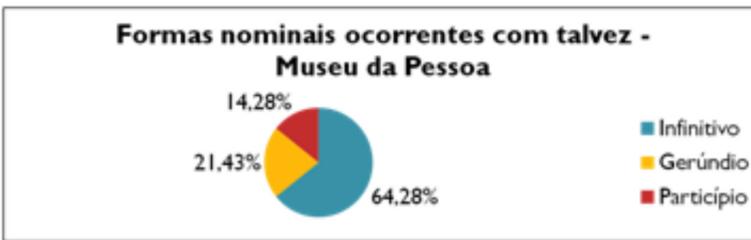
Gráfico 4: Percentuais relativos aos tempos do modo indicativo nas orações com talvez no Museu da Pessoa

### 3.4. Formas nominais com talvez em ambos os corpora

Esta pesquisa mostrou que, com as orações estudadas, não apenas o modo subjuntivo e o modo indicativo ocorrem, mas as formas nominais dos verbos. O Gráfico 5 e o Gráfico 6 abaixo mostram os percentuais de ocorrência do infinitivo, gerúndio e particípio. No *corpus* CHAVE, a maior ocorrência foi do particípio, com 43,75% das ocorrências. O infinitivo constituiu 36,60% das ocorrências e o gerúndio, 19,64%. No *corpus* Museu da Pessoa, a distribuição dos percentuais foi um pouco diferente. O infinitivo teve 64,28% das ocorrências; o particípio, 21,43% e o gerúndio, 14,28%.



**Gráfico 5:** Percentuais relativos às formas nominais do verbo usadas nas orações com talvez no CHAVE



**Gráfico 6:** Percentuais relativos às formas nominais do verbo usadas nas orações com talvez no Museu da Pessoa

### 3.5. Formas verbais usadas com talvez em ambos os corpora

Com o fim de comparar as formas verbais que mais ocorrem nas orações iniciadas com talvez, foram calculados os percentuais relativos à ocorrência do modo subjuntivo, do modo indicativo e das formas nominais dos verbos. O Gráfico 7, a seguir, mostra que, no *corpus* CHAVE, mais de 91% das ocorrências foram de orações iniciadas com talvez junto a verbos no modo subjuntivo. Esse resultado já era esperado, dado que, nos textos jornalísticos, busca-se o uso da norma padrão. Entretanto, houve enunciados em que outros tipos de estrutura foram utilizados. No mesmo gráfico, mostra-se que 1,77% foram de orações iniciadas com talvez junto a verbos do modo indicativo e que 6,34% foram de orações com as formas nominais dos verbos.

O Gráfico 8 apresenta percentuais diferentes, pois há uma maior distribuição de ocorrência dos dois modos e das formas nominais. O modo subjuntivo ocorreu em 63,38% dos enunciados; o indicativo, em 21,28%, e as formas nominais, em 14,89%. Este gráfico mostra ser incontestável a flutuação entre os dois modos. No *corpus* Museu da Pessoa, estão presentes a modalidade oral e o registro informal, nos quais o monitoramento da fala diminui e os desvios aparecem com mais nitidez e em maior quantidade.



Gráfico 7: Percentuais relativos às formas verbais usadas nas orações com talvez no CHAVE



Gráfico 8: Percentuais relativos às formas verbais usadas nas orações com talvez no Museu da Pessoa

#### 4. Considerações finais

Neste trabalho, buscou-se identificar as formas verbais que ocorrem na estrutura mencionada, tendo em vista as perguntas norteadoras citadas na primeira seção. Cabe aqui retomá-las e respondê-las, tornando possível, assim, analisar a flutuação entre essas formas verbais, considerando que a forma verbal predominante é o modo subjuntivo. A primeira pergunta norteadora do trabalho foi a seguinte: em qual dos *corpora* a flutuação entre indicativo e subjuntivo é maior? Os resultados quantitativos expostos acima mostram que a flutuação entre os dois modos é maior no *corpus* Museu da Pessoa. Acredita-se que a tendência é o indicativo junto ao advérbio talvez aparecer num *corpus* de transcrição de entrevistas, com predominância da linguagem informal oral.

As perguntas um e dois foram estas: quais os tempos do subjuntivo são mais usados com talvez? e quais tempos do indicativo são mais usados com talvez? Os resultados mostram i) que os tempos do subjuntivo que mais ocorreram em ambos os *corpora* foram presente, imperfeito, pretérito perfeito composto e mais que perfeito composto e ii) que os tempos do indicativo que mais aparecem nessas orações são presente e futuro do pretérito. A alta frequência desses tempos verbais pode indicar que as orações estudadas não permitem o uso aleatório dos tempos verbais, mas apenas de determinados tempos. Considera-se, da mesma forma, que a ocorrência

desses tempos verbais nos *corpora* mostra o uso que comumente é feito pelos falantes do português brasileiro, sendo possível afirmar, portanto, que os falantes fazem uso de determinadas formas verbais nessas estruturas provavelmente por fatores discursivos, que não foram analisados aqui.

A quarta pergunta do trabalho foi a seguinte: na estrutura talvez+subjuntivo, utilizam-se tempos do modo subjuntivo para os quais não há expectativa? Na Tabela 1 e na Tabela 2, em que aparece a distribuição dos tempos verbais presentes nos dois *corpora*, não há ocorrência dos tempos futuro e futuro composto desse modo. Esta pergunta foi considerada importante no trabalho porque o objetivo é contribuir para a descrição do português tendo em vista seu ensino como língua não materna. A possível construção de orações com o futuro e com o futuro composto teria de ser, portanto, desmotivada nesse processo de ensino. A questão do ensino ou não da construção desses enunciados com os tempos do modo indicativo é uma questão em aberto que não pretende ser respondida nesse trabalho. No entanto, acredita-se que a construção dessas orações com o modo indicativo deve ser desmotivada em situações em que o aprendiz esteja produzindo textos formais, orais ou escritos.

Quanto à quinta pergunta norteadora (as orações iniciadas por talvez são realizadas com quais formas nominais do verbo?), nota-se que as orações em questão também ocorrem com as formas nominais do verbo e que essa ocorrência é maior no *corpus* jornalístico e menor no *corpus* de fala.

Os resultados preliminares deste trabalho, de cunho quantitativo, suscitam outras questões e necessitam de análises qualitativas, para verificar, por exemplo, uma possível função discursiva que pode haver no uso de outras formas verbais, diferentes do subjuntivo, nas orações iniciadas pelo advérbio talvez. Bechara (2009) afirma, sobre a ocorrência do indicativo nessas orações, o seguinte:

Às vezes ocorre o indicativo com talvez: “Magistrado ou guerreiro de justo ou generoso se gaba: - e as turbas talvez o *aplaudem* e *celebram* seu nome” [AH.2,180] Parece que o indicativo deixa antever melhor a certeza de que o de que se duvida se pode bem realizar. (p. 281 – grifos do autor)

O gramático registra sua impressão de que o uso do indicativo nessas orações tem uma função discursiva, qual seja: a de indicar que a dúvida é menos evidente e que o fato expresso pelo verbo que está no modo indicativo pode ser mais certo do que aquele expresso pelo modo subjuntivo.

Contudo, os resultados expostos aqui dão subsídios para outros trabalhos e para os professores de português como língua não materna que já se depararam com a questão abordada aqui e que não encontraram respostas satisfatórias ou para aqueles que se interessarem pelo tema.

## Referências

- ALMEIDA, E. S. de. *Variação de uso do subjuntivo em estruturas subordinadas: do século XIII ao XX*. Tese de Doutorado. Rio de Janeiro: Universidade Federal do Rio de Janeiro, 2010.
- BEAUGRANDE, Robert de. Descriptive linguistics at the millennium: corpus data as authentic language. In: *Journal of language and linguistics* 1(2), 2002. pp. 91-131.
- BECHARA, Evanildo. Cap. 6 - Verbo. In: \_\_\_\_\_. *Moderna gramática portuguesa*. Rio de Janeiro: Nova Fronteira, 2009. 37ª ed. pp. 209-286.
- BERBER SARDINHA, Tony. A língua portuguesa nos CD-Roms da Folha de S. Paulo. In:\_\_\_\_\_ (Org.). *A língua portuguesa no computador*. Campinas, SP: Mercado das Letras, 2005. pp. 215-248.
- BICK, Eckhard. *The Parsing System "Palavras": Automatic Grammatical Analysis of Portuguese in a Constraint Grammar Framework*. Aarhus, Denmark: Aarhus University Press, 2000.
- BRANDÃO PEREIRA, A. L. *O uso do modo subjuntivo na fala de jovens da cidade de Juiz de Fora/MG*. Dissertação de Mestrado. Rio de Janeiro: Pontifícia Universidade Católica do Rio de Janeiro, 1995.
- CUNHA, Celso; CINTRA, Lindley. *Nova Gramática do português contemporâneo*. Rio de Janeiro: Nova Frontera, 2001.
- FREITAS, Cláudia. "Os corpos da Linguateca na prática". Encontro de Linguística de Corpos (ELC2011) (Minas Gerais, 11-12 de Novembro de 2011), 2011.
- FREITAS, Cláudia. Corpus, Linguística Computacional e as Humanidades Digitais. In: LEITE, M. & GABRIEL, C. T. *Linguagem, Discurso, Pesquisa e Educação*. Rio de Janeiro: DP et ali, 2014. pp 18-46.
- GONÇALVES, Jussara Regina. *Considerações sobre a flutuação no emprego do subjuntivo em contextos orais do Português do Brasil*. Dissertação de Mestrado. Rio de Janeiro: Pontifícia Universidade Católica do Rio de Janeiro, 2003.
- MEDEIROS, Vanise Gomes de. *Princípios norteadores da ocorrência do subjuntivo em orações relativas de SN em posição de sujeito*. Dissertação de Mestrado. Rio de Janeiro: Pontifícia Universidade Católica do Rio de Janeiro, 1996.
- OKSEFJELL, Signe; SANTOS, Diana. 'Breve panorâmica dos recursos de português mencionados na Web', in Vera Lúcia Strube de Lima (ed.), III Encontro para o Processamento Computacional do Português Escrito e Falado (PROPOR'98) (Porto Alegre, RS, 3-4 de Novembro de 1998), 1998. pp. 38-47.

RAJAGOPALAN, Kanavillil. A linguística de corpus no tempo de no espaço: visão reflexiva. In: Gerber, R. M. e Vasilévski, V. (Eds.). *Um percurso para pesquisas com base em corpus*. Florianópolis, SC: Ed. da UFCS, 2007. pp. 23-44.

SAMPSON, Geoffrey. *Empirical linguistics*. London and New York: Continuum, 2002.

SANTOS, Diana; BICK, Eckhard. “Providing Internet access to Portuguese corpora: the AC/DC project”. In: Gavriladou, Maria; Carayannis, George. Markantonatou, Stella, Stelios Piperidis; Stainhaouer, Gregory (eds.). *Proceedings of the Second International Conference on Language Resources and Evaluation, LREC2000*. ELRA, 2000. pp. 205-210.

SANTOS, Diana. Corporizando algumas questões. In: STELLA, E. O.; TAGNIN, S.; OTO, A. V. (Orgs.). *Avanços da linguística de corpus no Brasil*. Editora Humanitas: São Paulo, 2008. pp. 41-66.

SANTOS, Diana. Podemos contar com as contas? In: ALUÍSIO, S.; TAGNIN, S. (Orgs.). *11th Corpus Linguistics Conference*. Cambridge (UK): Cambridge Scholars Publishing, 2014.

SANTOS, Diana. “Gramateca: corpus-based grammar of Portuguese”, in Jorge Baptista, Nuno Mamede, Sara Candeias, Ivandré Paraboni, Thiago A.S. Pardo & Maria das Graças Volpe Nunes (eds.), International Conference on Computational Processing of Portuguese (PROPOR’2014) (São Carlos, Outubro de 2014), Springer, pp. 214—219.

SANTOS, Diana; ROCHA, Paulo. The key to the first CLEF in portuguese: topics, questions and answers in CHAVE. In: Peters, Carol; Clough, Paul; Gonzalo, Julio; Jones, Gareth J. F.; Kluck, Michael; Magnini, Bernardo (eds.). *Multilingual information access for text*. Speech and Images, 5th Workshop of the Cross-Language Evaluation Forum. CLEF 2004, Bath, UK, September 15-17, 2005.

SANTOS, Diana; SARMENTO, Luís. O projecto AC/DC: acesso a corpora/ disponibilização de corpora. In: Mendes, Amália; Freitas, Tiago (eds.). *Actas do XVIII Encontro Nacional da Associação Portuguesa de Linguística* (APL 2002). Porto, Portugal, 2-4 de Outubro de 2003. pp. 705-717.

SANTOS, Diana; SIMÕES, Alberto; FRANKENBERG-GARCIA. “Linguateca: um centro de recursos distribuído para o processamento computacional da língua portuguesa”. In: Luna, Guillermo De Ita; Chávez, Olac Fuentes; Galindo, Mauricio Osorio (eds.). *Proceedings of the international workshop “Taller de Herramientas y Recursos Lingüísticos para el Español y el Portugués”*, IX Iberoamerican Conference on Artificial Intelligence (IBERAMIA 2004) (Puebla, México, Novembro de 2004), pp. 147-154.

SINCLAIR, J. 2005. “*Corpus and Text - Basic Principles*” in *Developing Linguistic Corpora: a Guide to Good Practice*, ed. M. Wynne. Oxford: Oxbow Books: 1-16.