

## Convergência lexical entre letras de música e inglês geral: um estudo baseado em *corpus*

Lexical convergence between pop song lyrics and general English: a *corpus*-based study

Patrícia Bértoli\*

---

**RESUMO:** Este trabalho apresenta um estudo de convergência lexical entre letras de música gravadas originalmente em inglês e o inglês geral. O objeto do estudo é um *corpus* composto por cerca de 1 milhão de palavras, advindas de 5.962 letras de músicas diferentes. As letras de música foram consideradas por suas características linguísticas, ou seja, como texto (BÉRTOLI-DUTRA, 2002). Do *corpus* de estudo foram extraídas listas de palavras individuais e de feixes lexicais de três palavras (trigramas). Essas listas foram contrastadas com listas extraídas de dois *corpora* de referência, os quais são demonstrativos do inglês geral; as versões disponibilizadas online do BNC e do ANC. A pesquisa encontra-se fundamentada na Linguística de *Corpus*, que considera a linguagem como um sistema probabilístico, argumentando que o uso de características e traços linguísticos varia de acordo com o contexto imediato (as palavras que a antecedem e sucedem) e com o registro. A análise contrastiva apontou proximidade linguística entre letras de música e o inglês geral, no que diz respeito a ocorrência e frequência de palavras isoladas e trigramas, o que indica proximidade entre a linguagem usada nas letras das músicas analisadas e o inglês coloquial, ou seja, o inglês informal em uso geral (BIBER, 1988)

**PALAVRAS-CHAVE:** Letras de música. Inglês geral. Linguística de *corpus*. Convergência. Frequência.

**ABSTRACT:** This paper presents a study of lexical convergence between pop song lyrics and General English. The corpus is a compilation of roughly 1 million words from 5.962 different pop song lyrics. For the purpose of this study, ‘Lyrics’ is understood as a text type/register. Lists of individual words and lexical bundles of three words (trigrams) were extracted from the corpus. The lists were then contrasted to lists extracted from two reference *corpora* representative of General English: the BNC and the ANC. Grounded on Corpus Linguistics, which views language as a probabilistic system, the study assumes that the use of words—varies according to its immediate context (preceding and following words), as well as according to the register. The contrastive analysis showed strong linguistic similarities between pop song lyrics and General English, respectively to the occurrence and frequency of individual words and trigrams. This result suggests proximity between the language used in pop song lyrics and colloquial English, that is general informal English (BIBER, 1988).

**KEYWORDS:** Song Lyrics. General English. Corpus Linguistics. Convergence. Frequency.

---

---

\* Doutora em Linguística Aplicada e Estudos da Linguagem. Instituto de Letras – ILE – LAG, UERJ.

## 1. Introdução

Na qualidade de registro/tipo textual (BÉRTOLI-DUTRA, 2002; 2012), letras de música têm recebido pouco valor como fonte textual de pesquisa, especialmente, por serem consideradas como um registro não espontâneo e de cunho popular. Grande parte dos estudos sobre letras de música tem como foco suas características poéticas, temáticas ou semióticas (e.g., TATIT, 2001; COSTA, 2002; ROSSATO, 2002; DAMAZO, 2004; CROSSLEY, 2005; MATTE, 2005); ou de suas características sonoras e para o desenvolvimento de atividades de ensino (MURPHEY, 1990; MORA, 2000; ARLEO, 2000; BÉRTOLI-DUTRA, 2002). Este trabalho propõe uma outra abordagem ao investigar similaridades e diferenças de uso lexical entre o discurso de letras de música popular produzida originalmente em língua inglesa e outros registros em inglês, que compõem o que se considera inglês geral. Segundo Berber Sardinha (2004), um *corpus* de geral de uma língua deve “incluir o maior número possível de registros encontrados na língua alvo, e cada registro, por sua vez, deve ter o maior número possível de exemplares.” (BERBER-SARDINHA, 2004, p. 18)

A noção de música popular utilizada nesta pesquisa segue a proposta de Starr e Waterman (2007) que a consideram como toda forma musical produzida e divulgada em massa pelos meios de comunicação, a fim de atingir o maior número de consumidores possíveis. Essa perspectiva foi adotada pois considera-se que esse é o tipo de música que chega rapidamente ao ouvido do consumidor – em especial do consumidor de música internacional que é motivação constante de pesquisa por ser um dos meios de praticar a aprendizagem em língua estrangeira ao considerar suas letras.

O estudo segue também o princípio de que letras de música não podem ser consideradas apenas por suas características poéticas, ou seja, como poesias musicadas, pois embora os dois registros compartilhem características semelhantes como métrica, rima, etc. (COSTA, 2002), também apresentam características diferentes, tanto por seu componente sonoro quanto pelas escolhas lexicais que as caracterizam (BÉRTOLI-DUTRA, 2002, 2012; GRIFFITHS, 2003). Ademais, o foco deste estudo reside exatamente no uso lexical.

A fim de alcançar o objetivo principal deste estudo exploratório a fim de detectar similaridades e diferenças lexicais entre o discurso de letras de música popular escrita em inglês e o inglês geral, o estudo foi orientado especificamente pela Linguística de *Corpus* (BERBER SARDINHA, 2004; BIBER, 1988) e guiado pelos seguintes procedimentos: (a) levantamento de frequências das palavras usadas no *corpus* de estudo; (b) descrição de

padrões léxico-gramaticais do *corpus* de estudo; e (c) análise contrastiva dos padrões encontrados no *corpus* de estudo em oposição aos padrões existentes na língua inglesa, conforme representados pelo *corpus* de referência.

Este artigo está dividido da seguinte forma, primeiramente serão apresentados os princípios teóricos que fundamentaram a pesquisa, em sendo lugar, serão descritos os *corpora* utilizados e os procedimentos metodológicos e analíticos. Em seguida, estarão dispostos os resultados e as conclusões.

## 2. Base Teórica

A seleção de letras de música como fonte de pesquisa encontra sustento teórico na relevância de sua característica social, conforme argumentado por Starr e Waterman (2007) e Frith (1993). Outrossim, o estudo ora descrito partiu do pressuposto de que letras de música podem ser consideradas como representantes de um registro específico, fazendo, portanto, jus a análise linguística (BÉRTOLI-DUTRA, 2002; 2012).

As representações transferidas pela música, sejam sonoras ou referentes às mensagens presentes nas letras ou o conteúdo imagético apresentado por seus intérpretes, integram o conjunto de características com as quais nos identificamos e, conseqüentemente construimos nossas personalidades, nossas identidades particulares e partilhadas. A natureza sociológica da música popular e sua contribuição para a formação da cultura e da identidade dos indivíduos (STARR; WATERMAN, 2007; FRITH, 1993) faz da música um relevante objeto de estudo, ainda que sejam observadas especificamente suas características linguísticas dissociadas de seus componentes musicais, como feito aqui.

É importante esclarecer que a concepção de gênero utilizada neste trabalho, assim como a de registro, segue a perspectiva de Biber e Conrad (BIBER, 1988; BIBER; CONRAD, 2009, p. 2) para quem o nível do gênero demonstra os propósitos comunicativos e situacionais de um texto, enquanto o nível de registro apresenta as características funcionais, as quais são representadas pelas propriedades sintáticas, morfológicas e semânticas de um texto.

A análise de letras de música como texto observa suas características no nível de registro, conseqüentemente, considera-a como pertencente ao gênero oral, tal como sermões e discursos políticos, os quais são preparados antecipadamente a serem apresentados ao ouvinte, mas não envolvem interação direta e imediata entre os interlocutores. Segundo o

sociólogo Frith (1998), as letras fazem parte do gênero coloquialismo conversacional, trazendo consigo inúmeras variações linguísticas, que permitem a exploração de diferentes línguas, modos de falar, de palavras enquanto símbolos sonoros, entre outros, podendo, portanto, ser relacionada a contextos situacionais específicos, que possuem propósitos comunicativos específicos, como se espera de um gênero (BERTOLI-DUTRA, 2012).

Sendo assim, neste trabalho, pressupomos a existência do gênero letras de música, que se caracteriza principalmente por sua peculiaridade sonora, cujos textos compartilham o propósito de terem sido escritos de modo que pudessem ser integrados a estruturas melódicas. Ademais, Frith (1998) argumenta que letras não são simplesmente poesias cantadas, mas uma forma de retórica e oratória, num relacionamento de persuasão entre cantor e ouvinte. Dessa forma, consideraram-se letras de música como fonte de conhecimento linguístico, vez que se entende a transferência das palavras que ouvimos nas canções para outras formas do discurso, como, por exemplo, uma discussão oral sobre música, num ambiente social (MOORE, 2003). Em suma, a presença da música e suas características linguísticas foram fatores preponderantes para a consideração de letras de música enquanto texto, uma forma de registro que merece ser investigado. Sendo assim, o estudo teve como fundamentação teórica principal a Linguística de *Corpus* (BERBER-SARDINHA, 2004; BIBER, 1988)

A Linguística de *Corpus* é uma área que estuda os fenômenos da língua por meio da observação de grandes quantidades de dados empíricos armazenados eletronicamente, por computador. De modo simples, pode-se dizer que um *corpus* é uma coletânea de dados linguísticos autênticos, que podem ser processados por programas computacionais, cuja coleta segue critérios rigorosos, que estejam de acordo com o propósito da pesquisa e que sejam representativos (BERBER SARDINHA, 2004). Segundo Berber Sardinha, as quatro características fundamentais para a coleta dos dados formadores de um *corpus* computadorizado são textos em linguagem natural, autenticidade, critério rigoroso e representatividade. No que diz respeito a representatividade é um pouco mais delicada, enquanto a extensão do *corpus* deve ser considerada, vez que a visão probabilística da língua subjaz que “quanto maior a quantidade de palavras, maior a probabilidade de aparecerem palavras de baixa frequência” (BERBER SARDINHA, 2004, p. 22-23), por outro lado, é preciso considerar-se o quê se pretende representar. Em suma, o tamanho de um *corpus* de estudo é dependente daquilo que se pretende estudar.

O *corpus* coletado para esta pesquisa encaixa-se nas características fundamentais

apontadas por Berber Sardinha, vez que advêm de fonte natural, pois foram produzidos por falantes nativos para outros fins que não o da pesquisa; são autênticos, pois estiveram sujeitos à interação natural com o leitor/ouvinte (são letras de músicas realmente gravadas e veiculadas pela mídia); e são representativos de letras de música, vez que correspondem a uma amostra significativa desse universo linguístico, mais de 1 milhão de palavras, sendo, portanto, classificado como um *corpus* médio-grande.

Vale ressaltar que os linguistas do *corpus* consideram a língua como um sistema probabilístico, o que está enraizado na observação empírica da língua, ou seja, na observação natural da ocorrência de dados linguísticos em situações reais de uso, seguindo as propostas de Halliday, Sinclair e Firth (FIRTH, 1957; HALLIDAY; HASAN, 1989; HALLIDAY; 1991; SINCLAIR; 1991; HALLIDAY; WEBSTER; 2002). Essa visão probabilística assume que, embora muitas escolhas e combinações lexicais sejam possíveis, dentro de um determinado contexto (e na língua em geral), elas não ocorrem da mesma forma nem com a mesma frequência (BERBER SARDINHA, 2004, p. 30). A variação de escolhas linguísticas não ocorre aleatoriamente, ela segue uma certa padronização representativa de cada gênero. Pode-se dizer, portanto, que as combinações lexicais têm maior ou menor probabilidade de ocorrerem, ou são mais ou menos frequentes, de acordo com o contexto. A observação empírica vai permitir uma descrição mais completa e mais precisa da língua, pois evidenciará aquilo que é mais típico (HALLIDAY; WEBSTER, 2002, p. 10) em um determinado contexto. Nesse sentido, vale lembrar que a noção de gênero e registro empregadas para esta pesquisa foge da visão tradicional no que se “refere a diferentes perspectivas de variedade textuais”, onde o registro é analisado pela combinação de características linguísticas funcionais comuns, como pronomes e verbos, com a análise da situação de uso da variedade textual. Já a perspectiva de gênero abarca a descrição dos propósitos comunicativos e situacionais, como a de registro, mas a análise linguística “concentra-se nas estruturas convencionais usadas para construir um texto completo dentro de uma variedade, por exemplo, a maneira convencional como uma carta começa e termina” (BIBER; CONRAD, 2009 p.2).

A Linguística de *Corpus* também considera a análise de características e traços linguísticos (BIBER, 1988; BERBER SARDINHA, 2004) como representativos da língua, vez que sua utilização não obedecem a um critério arbitrário de escolha, ou seja, existe variação, ou seja, os falantes podem utilizar características linguísticas (palavras, estruturas,

etc.) diferentes para dizer as mesmas coisas. A variação linguística pode ser associada às funções comunicativas do texto. Essa ideia permeia o trabalho descrito aqui, uma vez que procura-se observar as semelhanças entre diferentes representações da língua (diferentes registros pertencentes a gêneros variados).

A análise de características linguísticas pode ser feita de diversas maneiras. Para este estudo, consideramos palavras individuais e trigramas (conjunto de três palavras seguidas), por suas ocorrências e frequências, defendendo suas características como marcadores textuais. Um n-grama é um feixe lexical (*lexical bundle*) que contém um número de palavras representado pela letra “n”. Esta pesquisa utilizou trigramas, ou seja, um feixe lexical de três palavras. Esse feixe deve ser analisado por seu sentido ou função como conjunto, pois entende-se que os agrupamentos não acontecem por acaso ou de forma randômica, mas carregam um significado específico quando em bloco. Segundo Sinclair (1991) e Hyland (2008) grande parte das palavras não tem um significado independente e, portanto, a interpretação dos feixes, ou n-gramas, auxilia no reconhecimento de gêneros discursivos.

Dados os fundamentos que norteiam esse estudo, passa-se à descrição dos *corpora* e da metodologia adotada.

### 3. *Corpora*

O objetivo do estudo aqui descrito foi extrair, relacionar e contrastar padrões léxico-gramaticais (BÉRTOLI DUTRA, 2002; BERBER SARDINHA, 2004; HUNSTON & FRANCIS, 2000) entre um *corpus* de letras de música gravadas originalmente em inglês coletado especificamente para esta pesquisa e dois *corpora* de inglês geral: as versões disponíveis online do British National Corpus (BNC), com mais de 100 milhões de palavras, e do American National Corpus (ANC) com aproximadamente de 15 milhões de palavras.

A versão online do BNC, também conhecida como BNC World Edition<sup>1</sup>, está dividida em 3.144 textos escritos e 910 textos falados, computando 100.467.090 palavras. Os textos escritos advêm de fontes como: jornais e periódicos voltados para o público de todas as idades e interesses, textos e ensaios acadêmicos, obras ficção popular, cartas, memorandos, entre outros. A porção falada é composta por transcrições de conversas informais e formais (e.g. conversas telefônicas, reuniões de negócios ou governamentais, programas de rádio).

---

<sup>1</sup> Cf. <http://www.natcorp.ox.ac.uk/corpus/index.xml.ID=numbers>

O ANC, no momento de coleta, contava com 22 milhões de palavras<sup>2</sup>. Trata-se de um projeto cujo objetivo é oferecer fonte de pesquisa do inglês americano para fins educacionais, linguísticos, lexicográficos, e para o desenvolvimento tecnológico. Foi utilizado neste estudo por estar disponível online para download na forma de listas de frequência por palavras individuais, ou por associadas (i.e., bigrams e trigramas) .

Para compor o *corpus* de estudo foram coletadas 5.962 letras de música compostas originalmente em inglês, computando 1.078.882 ocorrências (*tokens*), resultantes do uso de 22.136 formas (*types*), ou palavras diferentes. A coleta foi realizada seguindo rigorosamente os seguintes critérios: (1) As letras de música foram coletadas e categorizadas como *subcorpora* por artista (banda ou intérprete) na totalidade de suas músicas gravadas (foram descartadas músicas apenas instrumentais); (2) músicas gravadas por mais de um artista foram consideradas apenas uma vez; (3) os artistas foram selecionados por sua representatividade (marcas de sucesso em determinado período) e pela longevidade (fizeram ou fazem sucesso por muitos anos); (4) um dos *subcorpora* deveria favorecer músicas mais tocadas na época da pesquisa. Tais critérios foram considerados a fim de não se favorecer músicas ou artistas de apenas uma determinada época. Em outras palavras, artistas como Frank Sinatra foram escolhidas por sua presença nas paradas de sucesso por mais de sessenta anos (cf. *Billboard Charts*<sup>3</sup>), mesmo após sua morte, ao mesmo tempo que a banda Greenday, por exemplo, representa o grupo dos que fazem sucesso nos anos 2000 (id.). A versão final do *corpus* de estudo contém 5.962 letras de músicas gravadas por 30 artistas diferentes e pode ser melhor visualizada no quadro a seguir:

Quadro 1: composição do *corpus* de estudo.

ARTISTA	CANÇÕES
Aerosmith	56
Beatles	250
Bon Jovi	195
Creed	69
Def Leppard	143
Elton John	326
Foo Fighters	109
Frank Sinatra	1.224
Greenday	169

<sup>2</sup> Cf. <http://www.americannationalcorpus.org/>

<sup>3</sup> Para ver detalhes sobre paradas de sucesso em diversas épocas e variadas categorias, sugerimos consultar os sites: <http://www.billboard.com/> e <http://www.musicimprint.com/Chart.aspx?id=C000002>

Iron Maiden	215
The Killers	52
Led Zeppelin	79
Lenny Kravitz	125
Madonna	168
My Chemical Romance	61
Metallica	119
Nickelback	79
Nirvana	98
Paramore	50
Pearl Jam	146
Pink Floyd	164
Queen	172
Queensryche	112
Ray Charles	206
Red Hot Chilly Peppers	169
Rolling Stones	432
Simple Plan	49
Teen	202
TOP 100	121
U2	170

Vale ressaltar que o subcorpus TOP 100 é composto pelas músicas mais tocadas pela MTV Brasil e outras rádios brasileiras em agosto de 2008, e o subcorpus TEEN é representado por músicas de sucesso em 2008 voltadas para o público adolescente, sendo composto por letras das trilhas sonoras dos filmes High School Musical I, II e III e de artistas como Hannah Montana, The Jonas Brothers.

Uma vez coletados e definidos os *corpora* a serem utilizados no estudo, passou-se aos procedimentos de análise de convergência lexical entre os mesmos.

#### 4. Análise

A análise contrastiva partiu da obtenção (*download*) de listas de palavras do ANC e da extração de listas de palavras do *corpus* de estudo e do BNC. Essas últimas foram extraídas utilizando-se o software de análise lexical WordSmith tools<sup>4</sup> (SCOTT, 1998) que processa textos apresentando recursos como listas de palavras, de palavras chave e de concordâncias. Em seguida, as 500 palavras mais frequentes no *corpus* de estudo foram investigadas quanto a estarem ou não presentes nos *corpora* de referência. Optou-se pelo contraste das 500 formas

<sup>4</sup> disponível em < <http://www.lexically.net/wordsmith/> >

mais frequentes por se tratar de uma análise manual e por se considerar como uma amostra representativa da totalidade do *corpus* de estudo em relação ao inglês geral. A lista das palavras mais frequentes no *corpus* de estudo pode ser vista no anexo 1.

Todas as 500 palavras aparecem no BNC e no ANC, como era de se esperar. Portanto, passou-se a observar a frequência, ou seja, o número de ocorrência de cada uma delas em cada um dos *corpora*. A fim de que os dados pudessem ser comparados, as ocorrências nos três *corpora* foram normalizadas por 100. A normalização um procedimento que permite que as frequências compartilhem o mesmo parâmetro de comparação e é feita da seguinte maneira, divide-se o número de frequência pelo número de palavras do *corpus* e o resultado é multiplicado por 100. Dessa forma, a palavra “*the*” (a mais frequente nos três *corpora*), por exemplo, aparece 43.428 vezes no *corpus* de estudo, tendo frequência normalizada para 4,025. Já no BNC, a frequência bruta de 6.055.105 é normalizada para 6,027 e no ANC de 1.204.816 para 5,44. Em suma, se cada um dos *corpora* tivesse apenas 100 palavras, o “*the*” teria aparecido cerca de 4 vezes no CE, 6 vezes no BNC e 5 vezes no ANC.

A fim de ampliar o escopo da análise, foram também considerados agrupamentos de três palavras, também chamados trigramas (e.g. “*I want you*”; “*I love you*”; “*I don’t know*”), fundamentando-se na premissa de que o que determina o uso de determinada palavra é sua associação com as palavras que a antecedem e as que a sucedem (LAFFERTY; SLEATOR; TEMPERLY, 1992). Além disso, segundo os mesmos autores, trigramas refletem tanto sintaxe, quanto semântica e pragmática, de forma simultânea. Esta análise também considerou as 500 primeiras formas mais frequentes nos três *corpora*. Assim como para as palavras individuais, os dados foram normalizados para se tornarem comparáveis.

De posse das listas das palavras individuais e trigramas mais frequentes nos *corpora* de estudo e de referência, os dados foram contrastados a partir do *corpus* de estudo. Em outras palavras, procurou-se nos *corpora* de referência a ocorrência concomitante das palavras presentes no *corpus* de estudo. Suas frequências foram contrastadas. O mesmo procedimento foi dado à análise de trigramas nos três *corpora*.

## 5. Resultados

Na análise de palavras individuais, foi observado que todas as 500 palavras mais frequentes no *corpus* de estudo também estão presentes no BNC e no ANC. Ademais, houve concomitância entre as 6 primeiras palavras mais frequentes, as quais são exatamente as

mesmas nos três *corpora*, dadas as devidas diferenças de número de ocorrência, ainda que normalizadas. Por exemplo, “*the*”, aparece 4,02 no *corpus* de estudo, 6,02 no BNC e 5,44 no ANC. “*You*” aparece 3,33 no *corpus* de estudo, 0,58 no BNC e 0,80 no ANC. Ou seja, em comparação ao inglês geral, embora “*the*” seja a palavra mais frequente, ela é menos utilizada em letras de música; enquanto que “*you*” é muito mais utilizada em letras de música que no inglês geral. A título de ilustração, as 15 palavras mais frequentes nos três *corpora* estão dispostas na tabela seguinte:

Tabela 1: contraste de frequência das palavras mais frequentes no *corpus* de estudo com o BNC e o ANC.

palavra	frequência <i>Corpus</i> de estudo	frequência BNC	frequência ANC
1. THE	4,02	6,02	5,44
2. YOU	3,33	0,58	0,80
3. I	3,33	0,73	0,85
4. TO	2,36	2,58	2,40
5. AND	2,28	2,61	2,68
6. A	2,14	2,17	2,21
7. ME	1,59	0,13	0,15
8. MY	1,35	0,14	0,24
9. IN	1,29	1,93	1,84
10. IT	1,21	0,91	1,15
11. OF	1,17	3,03	2,73
12. YOUR	0,99	0,13	0,11
13. ON	0,91	0,72	0,63
14. THAT	0,87	1,04	0,76
15. ALL	0,80	0,27	0,23

A observação da tabela indica que embora as palavras estejam presentes nos três *corpora*, há variação de frequência, isto é, de uso. O quadro mostra que nas letras de música presentes no *corpus* de estudo há presença marcante de pronomes pessoais de primeira e segunda pessoas (*I; you*), demonstrando preferência por discurso interpessoal, cujos participantes devem ser a pessoa que canta (eu) e o ouvinte (você), o que corrobora com o achado de Bértoli-Dutra (2002).

Foi considerado como foco de análise particularmente o uso de léxico com funções básicas, como substantivos, verbos, pronomes, etc. O contraste quantitativo entre as frequências dessas palavras nos *corpora* indicou a presença das formas “*m*” e “*are*” sugerindo que acompanhem o uso pronominal mais frequente, de primeira e segunda pessoa, especialmente para “*P*”, cuja frequência apresenta-se muito mais elevada no *corpus* de letras de música. Já “*are*” pode aparecer acompanhando outras palavras, não estando

exclusivamente associada a um pronome, diferente de “’m”, que só aparece quando acompanhado de “I” e, por isso, sua frequência aparece de forma mais equilibrada, quando contrastada aos outros *corpora*.

O uso não flexionado do verbo “be” também aparece de forma equilibrada, assim como “is”, “are” e “we”. Palavras como “love”, “no”, “like”, “do”, “can”, “got”, “if”, “one”, “up”, “time”, “never”, “see” e “baby” apresentam sobreuso no *corpus* de estudo, com destaque para “one” e “baby”, sendo que a última aparece cerca de 40 vezes mais. A tabela a seguir representa uma amostra da análise realizada:

Tabela 2: amostra de contraste de frequência de itens lexicais do *corpus* de estudo com o BNC e o ANC.

palavra	Frequência no <i>corpus</i> de estudo	frequência BNC	frequência ANC
'm	0,82	0,062	0,15
be	0,76	0,64	0,48
is	0,75	0,96	1,02
love	0,70	0,02	0,01
we	0,54	0,29	0,42
know	0,53	0,11	0,21
like	0,44	0,14	0,02
do	0,42	0,17	0,09
can	0,41	0,21	0,21
got	0,41	0,08	0,06
if	0,40	0,25	0,25
one	0,38	0,28	0,0018
get	0,37	0,95	0,10
down	0,36	0,91	0,05
go	0,36	0,86	0,021
up	0,36	0,20	0,017
time	0,35	0,15	0,010
are	0,33	0,45	0,47
never	0,32	0,05	0,059
was	0,32	0,85	0,71
see	0,31	0,11	0,104
baby	0,31	0,0079	0,008

A análise quantitativa de palavras individuais (uma única palavra) indicou que a escolha lexical em letras de música apresenta grande semelhança com a escolha realizada em outros gêneros da língua inglesa, ou seja, em outros registros<sup>5</sup> que não letras de música. Tal constatação sugere que a linguagem usada em letras de música pode ser também classificada

<sup>5</sup> Ver Biber; Conrad, 2009 para concepção de gênero e registro usada neste estudo, já mencionada anteriormente

como coloquial, pois está simultaneamente presente em outros gêneros da língua inglesa, tais como cartas, conversas informais e textos jornalísticos – lembrando que os *corpora* de referência não apresentam letras de música em sua composição.

A fim de não restringir tal constatação e a análise ao uso de palavras individuais, o mesmo procedimento de análise foi dado à ocorrência e frequência de trigramas presentes nas letras de música de forma concomitante aos *corpora* de inglês geral. A extração de trigramas resultou na ocorrência de 129.117 trigramas diferentes para o *corpus* de estudo, 5.431.734 para o BNC e as listas do ANC apresentaram 1.453.050 para sua porção falada e 4.236.030 para sua porção escrita, cujos 30 mais frequentes podem ser vistos no quadro a seguir:

Quadro 2: 30 primeiros trigramas no *corpus* de estudo contrastados com trigramas dos *corpora* de referência.

trigramas CE	trigramas BNC	trigramas ANC falado	trigramas ANC escrito
1. I WANT TO 540	ONE OF THE 20760	I I DO 997	THE UNITED STATES 2600
2. I LOVE YOU 505	THE END OF 11604	AND I AM 993	THE NEW YORK 2477
3. I DON'T KNOW 371	AS WELL AS 11502	YOU THINK OF 99	NEW YORK TIMES 2427
4. LA LA LA 333	PART OF THE 10182	YEAH YEAH OH 99	ONE OF THE 2217
5. I DON'T WANT 294	THERE IS A 9655	YEAH BUT UM 99	THE THE THE 2141
6. DON'T WANT TO 285	OUT OF THE 9162	WAS NICE TALKING 99	THE OF THE 1635
7. OH OH OH 256	SOME OF THE 9142	UM THERE IS 99	IT IS A 1554
8. YEAH YEAH YEAH 254	EQUO HE SAID 8922	TIME AND I 99	THE IN THE 1505
9. I KNOW THAT 239	A NUMBER OF 8358	THAT IS SO 99	I DO NOT 1313
10. I DON'T WANNA 228	END OF THE 7931	TAKING CARE OF 99	IT IS NOT 1213
11. I GOT A 214	THERE WAS A 7741	SIT DOWN AND 99	THE THE A 1156
12. IN LOVE WITH 207	IT WAS A 7685	RIGHT BUT UH 99	IN THE THE 1124
13. I WANT YOU 205	THE FACT THAT 7673	NOT KNOW THE 99	THE A THE 1112
14. NA NA NA 202	THERE IS NO 7457	IT THAT WAY 99	A THE THE 1061
15. WANT TO BE 200	BE ABLE TO 7298	IT FOR A 99	A IN THE 1056
16. YOU AND ME 195	IN ORDER TO 7110	IT A LITTLE 99	OF THE NEW 1049
17. YOU WANT TO 191	TO BE A 6992	IS I AM 99	THERE IS A 1006
18. YOU AND I 187	IT IS NOT 6893	I UNDERSTAND THAT 99	AS WELL AS 980
19. IN YOUR EYES 184	PER CENT OF 6841	I THINK ABOUT 99	A LOT OF 958
20. I WANNA BE 183	A LOT OF 6574	I REALLY I 99	THE WHITE HOUSE 948
21. I DON'T NEED 181	IT IS A 6377	I LISTEN TO 99	THE END OF 934
22. IF YOU WANT 180	EQUO BQUO I 6291	I I UH 99	A A A 879
23. AND I DON'T 178	HE SAID BQUO 5929	I HAD THE 99	THERE IS NO 871
24. IN MY HEART 177	IN TERMS OF 5816	I GUESS THERE 99	THE THE IN 865
25. AND I KNOW 176	IT WOULD BE 5764	I AM LOOKING 99	THE THE OF 818
26. THIS IS THE 175	AT THE END 5608	HUM UM I 99	OF THE THE 815
27. CAN'T YOU SEE 174	MOST OF THE 5403	HUH AND UM 99	THE A A 812
28. TO BE A 169	ON THE OTHER 5391	GOOD THAT IS 99	SOME OF THE 795
29. YOU KNOW THAT 167	THAT IT IS 5389	AND UH MY 99	PART OF THE 782
30. ON AND ON 166	THE NUMBER OF 5372	AND THERE ARE 99	THE TO THE 775

O quadro anterior representa apenas uma amostra dos trigramas analisados, vez que foram considerados os 500 trigramas mais frequentes do *corpus* de estudo, os quais foram manualmente observamos por sua presença nos *corpora* de referência. Esse primeiro

levantamento considerou simplesmente a ocorrência concomitante dos trigramas. Dessa forma, constatou-se que apenas 12 dos 500 trigramas mais frequentes no *corpus* de letras de música não estão presentes em um ou outro dos *corpora* de referência. Os 222 primeiros trigramas mais frequentes no *corpus* de estudo estão também presentes no BNC e ANC. A partir do 223º trigrama (*ooh ooh ooh*), os outros onze começam a aparecer exclusivamente no *corpus* de letras de música, sendo: “*c'mon c'mon c'mon*”; “*oooh oooh oooh*”; “*oo oo oo*”; “*good to ya*”; “*aah aah aah*”; “*wanna be your*”; “*need your love*”; “*love love love*”; “*all I wanna*”; “*ah ah ah*”; “*be your man*”.

A análise desses onze trigramas demonstra que sete dessas expressões apresentam variações da escrita em relação à forma culta (i.e: fogem dos padrões gramaticais estabelecidos), como a repetição de letras (e.g. “*aah*”; “*oooh*”) que é muito usada em letras de música, normalmente indicando o prolongamento da emissão do som (no cantar); da mesma forma, apresenta formas contraídas como “*wanna*” e “*ya*” (para *you*); e a repetição de palavras (*c'mon c'mon*).

O trigrama “*need your love*”, surpreendentemente não aparece nos *corpora* de referência, onde aparecem expressões como “*need your advice/passport/help/expertise/attention*”. Tal fato pode se justificar pelo sobreuso da palavra “*love*”, conforme observado na análise de palavras individuais. O mesmo aconteceu com o trigrama “*be your man*”, que não foi encontrado nos *corpora* de referência, os quais apresentaram variações como “*be your dream/friend/guide/home/idea/you*”. Pode-se inferir que essa expressão parece ser de uso mais comum em letras de música por também apresentar referência a amor e sexualidade.

No que diz respeito aos 22 trigramas que estão presentes nos três *corpora*, o trigrama “*I want to*”, que ocorre 540 vezes no *corpus* de estudo, sendo o mais frequente, refere-se a 0,5% das vezes em contraste ao BNC, onde aparece 2.110 vezes, equivalendo a 0,02%. Embora pareça uma grande diferença, a presença do trigrama já confirma seu uso na língua inglesa em geral, não apenas nas músicas. Interessantemente, expressões como “*la la la*” aparecem nos três *corpora*. Isto é, embora possam parecer de uso exclusivo da música, também são usadas no inglês geral.

## 6. Conclusão

O estudo exploratório aqui apresentando buscou traços de convergência lexical entre

um *corpus* de letras de música e dois *corpora* de inglês geral, o BNC (de inglês britânico) e o ANC (de inglês americano). A constatação de que as palavras individuais mais frequentes no *corpus* de estudo também estão entre as mais frequentes nos *corpora* de referência, seja americano ou britânico, confirma o identificado pela pesquisa de Bértoli-Dutra (2002), que ressalta a proximidade entre a língua utilizada em letras de música e o inglês coloquial. Por outro lado, este estudo também observou a mesma tendência de uso quando consideradas três palavras juntas, ou trigramas.

O sobreuso de certas palavras, ou seja, a maior ocorrência dessas palavras no *corpus* de estudo do que nos *corpora* de inglês geral, sugere que exista uma espécie de conversa romântica<sup>6</sup> (entre pessoas que se amam) nas letras de música, de forma muito mais acentuada do que no inglês geral<sup>7</sup>, dadas as frequências de palavras como “*baby*” e “*love*”, ou seja, evidencia um tipo de “fala à pessoa amada”. O que também pode ser confirmado pelo uso exclusivo dos trigramas “*be your man*” e “*need your love*”. Deve-se considerar, contudo, a possibilidade de que parte das ocorrências de “*baby*” nos *corpora* de referência refira-se a bebê, e não a tratamento de afeto, o qual é a forma de uso mais frequentemente no *corpus* de estudo.

Este estudo indicou também que se uma análise puramente quantitativa de frequências normalizadas pode apontar convergências com o inglês geral, ao mesmo tempo não parece ser suficiente para apontar as características das letras de música como um todo (o que não era objeto deste estudo, mas que parece ser bastante interessante). Um dos fatores que contribuiria para uma análise mais profunda das letras de música seria a etiquetagem do *corpus*. Sendo assim, uma das limitações do estudo refere-se a, por exemplo, não se poder separar os usos de palavras como “*love*”, “*like*” ou “*know*”, bastante frequentes no *corpus* de estudo, em relação ao seu uso como morfossintático, o que, poderia nos dar uma ideia melhor de como as letras se comportam. Isto é, se as letras usam a forma “*know*”, por exemplo, para indicar o que se sabe ou conhece da mesma forma que no inglês geral ou se está sendo usada mais frequentemente para preencher espaços conversacionais, com expressões como “*you know*”.

A análise de trigramas possibilitou uma nova perspectiva de aproximação do discurso geral (utilizado em variados registros) com o das letras de música e, portanto, foi considerada

---

<sup>6</sup> Os temas das canções são variados, há canções de amor, Guerra, morte, paz, justiça, religião, etc. (ver Bértoli-Dutra, 2014)

<sup>7</sup> Entende-se por inglês geral a língua utilizada nos mais variados registros falados e escritos

de forma satisfatória para justificar o uso de letras de música em sala de aula para fins de ensino de línguas, não simplesmente para práticas auditivas.

As duas esferas de análise revelaram que letras de música estão muito próximas do inglês geral, ao mesmo tempo que provocaram a necessidade estudos futuros que venham a descrever mais detalhadamente o discurso da música.

## Referências

ARLEO, A. Music, song and foreign language teaching. **Les Cahiers de l'APLUIT**, vol. XIX, no. 4. 2004 p. 5-19.

BERBER SARDINHA, A. P. **Linguística de Corpus**. Barueri: Manole, 2004.

BÉRTOLI-DUTRA, P. **Explorando a linguística de *corpus* e letras de música na produção de atividades pedagógicas**. Dissertação Inédita (Mestrado em LAEL), PUC-SP, 2002.

\_\_\_\_\_. Song Lyrics and Speech: similarities, differences and multidimensional analysis of Song Lyrics from 1940 to 2009. In: MELLO, H.; PETTORINO, M.; RASO, T. **Proceedings of the VII GSCP International Conference: Speech and Corpora**. Firenze: Firenze University Press, 2012.

\_\_\_\_\_. Multi-Dimensional analysis of pop songs In: BERBER SARDINHA, T; VEIRANO PINTO, M. **Multi-Dimensional Analysis, 25 years on: A tribute to Douglas Biber**. 1 ed. Amsterdam : John Benjamins, 2014, p. 151-177.

BIBER, D. **Variation across speech and writing**. Cambridge: Cambridge University Press, 1988. **crossref** <http://dx.doi.org/10.1017/CBO9780511621024>

COSTA, N. B. da. As letras e a letra: o gênero canção na mídia literária. In: DIONÍSIO, A. P.; MACHADO, A. R.; BEZERRA, M. A.. (orgs.). **Gêneros textuais e ensino**. Rio de Janeiro: Lucerna, 2002.

CROSSLEY, S. A. Metaphorical conceptions in hip-hop music. **African American Review**, vol. 39, no. 4, Winter, 2005. p. 501-512.

DAMAZO, F.A.F.T. **O canto do povo de um lugar: uma leitura das canções de João do Vale**. Tese de doutorado apresentada ao Programa de Estudos Literários da Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP: SJ Rio Preto, 2004.

FIRTH, A. The discursive accomplishment of normality on 'língua franca' English and conversation analysis. **Journal of Pragmatics** 26. 1996 (237-259). **crossref** [http://dx.doi.org/10.1016/0378-2166\(96\)00014-8](http://dx.doi.org/10.1016/0378-2166(96)00014-8)

FRITH, S. Music and identity. In: HALL, S.; Du GAY, P. (Eds). **Questions of Cultural Identity**, London, UK: Sage publications, 1993. p. 108-127.

\_\_\_\_\_. Performin Rites. **On the value of popular music**. Cambridge, Massachusetts, USA: Harvard University Press, 1998.

GRIFFITHS, D. From Lyrics to anti-lyric: analyzing the words in pop songs. In: A. F. MOORE. (ed). **Analysing Popular Music**. Cambridge: Cambridge University Press, 2003. (p. 39-59). **crossref** <http://dx.doi.org/10.1017/CBO9780511482014.003>

HALLIDAY, M. A. K. Corpus studies and probabilistic grammar. In: K. Aijmer; B. Altenberg (org.) **English Corpus Linguistics: Studies in honour of Jan Svartvik**. (30-43). London: Longman, 1991.

HALLIDAY, M.A.K; HASAN ; R. **Language, context, and text: aspects of language in a social-semiotic perspective**. 2nd edition. Deakin University Press/Oxford University Press, 1989.

HALLIDAY, M.A.K; WEBSTER, J. (ed.). **On grammar: By Michael Alexander Kirkwood Halliday**. New York: Continuum, 2002.

HUNSTON, S.; FRANCIS, G. **Pattern grammar: a corpus-driven approach to the lexical grammar of English**. Amsterdam/Phildelphia: John Benjamins, 2000. **crossref** <http://dx.doi.org/10.1075/scl.4>

HYLAND, K. As can be seen: lexical bundles and disciplinary variatuon. **ESP**. Vol. 27, 2008, p. 4-21.

LAFFERTY, J.; SLEATOR, D.; TEMPERLEY, D. Grammatical trigrams: A probabilistic model of link grammar, in **Proceedings of the AAAI Fall Symposium on Probabilistic Approaches to Natural Language**, Cambridge, MA, October 1992.

MATTE, A. C. F. Gostar de música: percurso de uma paixão. **Significação: revista brasileira de semiótica**. 23. São Paulo: ANNABLUME. Junho, 2005. p. 72-92.

MOORE, A. F. (ed). **Analysing Popular Music**. Cambridge: Cambridge University Press, 2003.

MORA, C. F. Foreign language acquisition and melody singing. **ELT Journal**, vol. 54, no. 2, 2000, p. 146-152. **crossref** <http://dx.doi.org/10.1093/elt/54.2.146>

MURPHEY, T. The song stuck in my head phenomenon: a melodic din in the lad? **System**, vol. 18, vol. 1, 1990, p. 53-64.

ROSSATO, E. V. A brasilidade na poesia contemporânea de Caetano Veloso. In: III SEMINÁRIO DE ESTUDOS SOBRE LINGUAGEM E SIGNIFICAÇÃO - **SELISIGNO: DISCURSO E REPRESENTAÇÃO** E IV SIMPÓSIO DE LEITURA DA UEL, 2002, Londrina. Caderno de Resumos do IV Simpósio de Leitura da UEL: III Selisigno: Discurso e Representação, 2002. v. 1. p. 93-93.

SCOTT, M. **WordSmith Tools**. Version 3. Oxford University Press: England, 1998.

SINCLAIR, J. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press, 1991.

STARR, L.; WATERMAN, C.. **American Popular Music**. From minstrelsy to MP3. 2nd. ed. New York: OUP, 2007.

TATIT, L. **O cancionista: composição de canções no Brasil**. São Paulo: EDUSP, 1996.

## ANEXO 1

### LISTA DAS 500 PALAVRAS MAIS FREQUENTES NO CORPUS DE ESTUDO

1 THE 43.428	126 THEN 1.419	251 MYSELF 598	376 DEAR 388
2 YOU 35.995	127 HOME 1.402	252 REMEMBER 596	377 ITS 387
3 I 35.964	128 ALWAYS 1.359	253 THOUGHT 596	378 UNTIL 386
4 TO 25.485	129 ABOUT 1.356	254 WAIT 589	379 LATE 382
5 AND 24.662	130 STILL 1.340	255 SOUL 587	380 FINE 381
6 A 23.138	131 HAD 1.333	256 LIVING 584	381 MORNING 379
7 ME 17.209	132 EVERY 1.327	257 COLD 581	382 YEARS 379
8 MY 14.635	133 KEEP 1.321	258 FEELING 580	383 ANYTHING 374
9 IN 13.966	134 WOULD 1.305	259 ONCE 575	384 TRYING 374
10 IT 13.118	135 OVER 1.287	260 SIDE 575	385 DOES 373
11 OF 12.685	136 US 1.261	261 PAIN 574	386 MISS 373
12 YOUR 10.694	137 EVER 1.255	262 THOSE 574	387 PRETTY 370
13 ON 9.951	138 HEAR 1.233	263 HANDS 566	388 TOUCH 370
14 THAT 9.471	139 AN 1.225	264 CHORUS 562	389 LAY 369
15 I'M 8.891	140 WERE 1.216	265 DONE 562	390 MONEY 369
16 ALL 8.729	141 YOU'LL 1.209	266 TILL 561	391 SUCH 366
17 BE 8.205	142 AM 1.185	267 HELP 559	392 NOBODY 365
18 IS 8.199	143 MUCH 1.146	268 LOOKING 559	393 MEET 364
19 FOR 7.759	144 MIND 1.145	269 DAYS 558	394 REASON 359
20 LOVE 7.606	145 BETTER 1.089	270 TOWN 553	395 RUNNING 358
21 DONT 7.217	146 BELIEVE 1.087	271 WHAT'S 552	396 EASY 357
22 SO 6.492	147 OLD 1.085	272 BRING 550	397 GROUND 357
23 IT'S 6.220	148 NOTHING 1.083	273 GOODBYE 550	398 GUESS 355
24 BUT 6.157	149 INTO 1.074	274 NA 543	399 SOON 353
25 WE 5.865	150 THINGS 1.065	275 REAL 541	400 CHILD 351
26 JUST 5.815	151 GONE 1.057	276 FOUND 538	401 BYE 348
27 KNOW 5.712	152 HOLD 1.037	277 FORGET 535	402 MATTER 347
28 NO 5.488	153 LIVE 1.018	278 THEY'RE 531	403 LISTEN 346
29 WITH 5.459	154 ALONE 1.011	279 LOSE 530	404 SOUND 346
30 WHEN 5.307	155 TRY 991	280 GOES 520	405 WENT 343
31 WHAT 4.956	156 I'D 988	281 KNEW 518	406 SAD 341
32 OH 4.905	157 GOTTA 964	282 ALRIGHT 517	407 UNDERSTAND 339

33 THIS 4.879	158 DREAM 963	283 MAKES 517	408 MUSIC 338
34 LIKE 4.842	159 THEM 963	284 BEST 515	409 NEXT 338
35 DO 4.550	160 LEAVE 946	285 START 512	410 RED 338
36 CAN 4.520	161 FACE 942	286 DEAD 507	411 SHAKE 338
37 GOT 4.447	162 RUN 940	287 WOMAN 505	412 READY 337
38 IF 4.386	163 YOU'VE 927	288 HEAVEN 504	413 PART 336
39 NOW 4.379	164 ANOTHER 925	289 LONELY 504	414 SEA 336
40 YOU'RE 4.253	165 LIGHT 915	290 ROCK 503	415 WORK 336
41 OUT 4.194	166 PLEASE 915	291 RAIN 500	416 FEAR 335
42 ONE 4.122	167 DID 911	292 DOOR 497	417 FUN 334
43 GET 4.024	168 SUN 907	293 CLOSE 496	418 BURNING 331
44 YEAH 4.024	169 HAS 906	294 WISH 496	419 TRUTH 331
45 DOWN 3.906	170 BEFORE 905	295 BECAUSE 495	420 ABOVE 330
46 GO 3.889	171 GOING 904	296 SMILE 492	421 FEET 330
47 UP 3.883	172 NEW 902	297 SEEN 490	422 GIRLS 330
48 TIME 3.820	173 SOMETHING 902	298 BLACK 489	423 LADY 330
49 ARE 3.613	174 REALLY 898	299 MAYBE 488	424 TOMORROW 330
50 NEVER 3.546	175 CALL 891	300 MEAN 485	425 WONDER 330
51 WAS 3.485	176 HEAD 884	301 KIND 484	426 BOYS 329
52 SEE 3.438	177 YES 882	302 TALK 484	427 EVERYONE 327
53 BABY 3.387	178 HAND 879	303 EVERYBODY 478	428 WANTS 327
54 SHE 3.387	179 INSIDE 876	304 CRAZY 477	429 TRIED 326
55 WILL 3.374	180 TONIGHT 871	305 HAPPY 475	430 WHITE 324
56 CAN'T 3.270	181 LAST 860	306 ROLL 475	431 MAMA 321
57 HAVE 3.260	182 PLACE 842	307 FLY 474	432 BROKEN 320
58 COME 3.249	183 STAY 840	308 ROUND 471	433 STRONG 318
59 NOT 3.194	184 OFF 838	309 OPEN 469	434 BABE 316
60 I'LL 3.181	185 THAN 837	310 TOLD 466	435 LIGHTS 316
61 WANT 3.103	186 THING 837	311 YOURSELF 465	436 FALLING 314
62 SAY 3.070	187 HIM 804	312 AH 462	437 SAYS 314
63 FROM 3.041	188 STOP 798	313 SOMEBODY 461	438 SET 312
64 TAKE 2.973	189 THEIR 791	314 SURE 461	439 WANTED 312
65 HER 2.933	190 LET'S 788	315 MANY 460	440 DARK 311
66 WAY 2.851	191 DIE 784	316 USED 460	441 C'MON 306
67 THEY 2.822	192 WALK 780	317 CHANCE 459	442 BORN 305
68 BACK 2.734	193 WE'LL 780	318 CAME 457	443 COULDN'T 305
69 GONNA 2.733	194 MINE 778	319 ENOUGH 457	444 ASK 304
70 AS 2.731	195 FALL 777	320 LINE 457	445 MOTHER 302
71 MAKE 2.729	196 CRY 774	321 FRIEND 456	446 WHOA 302
72 AWAY 2.689	197 MADE 774	322 TEARS 454	447 LAND 299
73 THERE 2.661	198 OOH 774	323 MOON 453	448 WHO'S 298
74 AT 2.642	199 TWO 770	324 SOMETIMES 452	449 WILD 298
75 LET 2.575	200 TRUE 768	325 FRIENDS 451	450 LORD 297
76 I'VE 2.425	201 TURN 768	326 FOREVER 450	451 SEEM 297
77 DAY 2.382	202 SHOULD 767	327 FIGHT 448	452 GAME 296

78 HOW 2.320	203 SOMEONE 766	328 ARMS 444	453 KILL 292
79 NIGHT 2.306	204 WRONG 760	329 RIDE 444	454 LOW 291
80 HEART 2.260	205 SWEET 758	330 WORDS 444	455 BEAUTIFUL 290
81 HE 2.223	206 LA 757	331 LIE 443	456 STANDING 290
82 LIFE 2.221	207 HARD 756	332 GETTING 441	457 HATE 288
83 FEEL 2.178	208 WITHOUT 756	333 TIMES 440	458 WHOLE 288
84 TELL 2.079	209 HE'S 751	334 STAR 439	459 CITY 287
85 RIGHT 2.061	210 MUST 737	335 KNOWS 438	460 TAKES 286
86 COULD 2.055	211 DANCE 736	336 TODAY 438	461 SHINE 284
87 WELL 2.051	212 LOST 733	337 YOUNG 438	462 WAR 283
88 HERE 2.046	213 EVERYTHING 729	338 ANY 437	463 HEARTS 282
89 CAUSE 2.041	214 LEFT 726	339 HOPE 437	464 HOT 282
90 THERE'S 1.938	215 COMING 725	340 ALIVE 435	465 HIT 281
91 TOO 1.937	216 FREE 725	341 HIDE 434	466 LIPS 279
92 MORE 1.932	217 SONG 720	342 HELL 432	467 WALKING 279
93 WHERE 1.871	218 COMES 718	343 BLOOD 430	468 DEATH 278
94 BY 1.845	219 END 710	344 STARS 429	469 MEN 277
95 MAN 1.836	220 SING 710	345 EACH 428	470 PARTY 277
96 LITTLE 1.831	221 PEOPLE 709	346 SLEEP 426	471 GUN 276
97 GOOD 1.823	222 SHOW 703	347 MIGHT 424	472 HA 275
98 NEED 1.820	223 CARE 699	348 AFTER 422	473 DOESN'T 274
99 BEEN 1.814	224 WHILE 698	349 DIDN'T 420	474 GETS 274
100 HEY 1.775	225 THESE 695	350 LIES 420	475 SINCE 273
101 OR 1.727	226 BAD 689	351 MOVE 420	476 BURN 272
102 GIVE 1.694	227 HIGH 683	352 ALONG 418	477 GAVE 272
103 THAT'S 1.663	228 SAME 679	353 WIND 414	478 HURT 272
104 THINK 1.655	229 WAITING 676	354 SEEMS 411	479 AIR 270
105 WORLD 1.652	230 OWN 670	355 BEAT 409	480 LIVES 267
106 OUR 1.643	231 PLAY 668	356 TOOK 409	481 SOMEWHERE 267
107 EYES 1.639	232 SKY 659	357 SAVE 408	482 TIRED 267
108 AGAIN 1.632	233 BREAK 658	358 HONEY 407	483 LEARN 266
109 ONLY 1.628	234 KISS 658	359 STREET 402	484 SOMEDAY 266
110 THROUGH 1.614	235 PUT 657	360 THOUGH 402	485 BED 263
111 WHO 1.604	236 NAME 649	361 YOU'D 402	486 BIT 259
112 WANNA 1.553	237 CHANGE 647	362 WATCH 401	487 PAST 259
113 SAID 1.547	238 DREAMS 645	363 UNDER 400	488 UPON 258
114 LONG 1.545	239 BOY 639	364 FOOL 397	489 TIGHT 256
115 WHY 1.542	240 BLUE 633	365 VERY 397	490 DARLING 255
116 GIRL 1.534	241 GOD 619	366 DEEP 396	491 FULL 255
117 WE'RE 1.533	242 MAY 618	367 BEHIND 395	492 BLUES 254
118 AROUND 1.517	243 STAND 618	368 ROOM 395	493 CONTROL 252
119 SOME 1.517	244 BIG 614	369 WORD 395	494 LOVED 251
120 HIS 1.496	245 TOGETHER 606	370 SAW 394	495 USE 251
121 SHE'S 1.494	246 FAR 604	371 WE'VE 391	496 STEP 250
122 AIN'T 1.478	247 OTHER 604	372 ELSE 390	497 MOMENT 249

123 FIND 1.471	248 EVEN 602	373 FIRST 390	498 EYE 248
124 WON'T 1.471	249 YA 602	374 HEARD 389	499 HELLO 248
125 LOOK 1.437	250 FIRE 600	375 ROAD 389	500 LOVING 248

Artigo recebido em: 15.10.2014

Artigo aprovado em: 29.11.2014

Letras & Letras