INFERRING CHORD SEQUENCE MEANINGS VIA LYRICS: PROCESS AND EVALUATION

Tom O'Hara[1], **Nico Schüler**[2], **Yijuan Lu**[1], and **Dan E. Tamir**[1]
[1]Dept. of Computer Science and [2]School of Music, Texas State University, San Marcos, TX 78666
{tomohara, nico.schuler, yl12, dan.tamir}@txstate.edu

ABSTRACT

We improve upon our simple approach for learning the "associational meaning" of chord sequences from lyrics based on contingency statistics induced over a set of lyrics with chord annotations. Specifically, we refine this process by using word alignment tools developed for statistical machine translation, and we also use a much larger set of chord annotations. In addition, objective evaluation measures are included. Thus, this work validates a novel application of lexicon induction techniques over parallel corpora to a domain outside of natural language learning. To confirm the associations commonly attributed to major versus minor chords (i.e., happy and sad, respectively), we compare the inferred word associations against synonyms reflecting this dichotomy. To evaluate meanings associated with chord sequences, we check how often tagged chords occur in songs labeled with the same overall meaning.

1. INTRODUCTION

Chords are the foundation of western music, providing the harmony for music and also influencing the melody (given close relation to musical keys). Chords are not simply three or more notes simultaneously played but also involve precise relationships among the notes. For example, the notes in a major chord consist of the root (lowest frequency), a note a third above the root (i.e., two whole steps), and a note a fifth above the root (e.g., three whole steps and a half). An example would be the $CMaj$ chord, which consists of the notes C, E and G. Likewise, chord sequences generally have precise definitions. For example, the popular *12-Bar Blues Progression* commonly uses the following scheme: ⟨I, I, I, I, IV, IV, I, I, V, IV, I, I⟩, where Roman numerals refer to chord intervals [16]. In the key of C, this would be as follows: ⟨C, C, C, C, F, F, C, C, G, F, C, C⟩. Given such precise relationships to musical intervals, meanings typically attached to chord sequences are unlikely to be completely arbitrary. This paper demonstrates how to learn the *associational meaning* [8] of chord sequences (e.g., in terms of word associations).

Parallel text corpora were developed primarily to serve multilingual populations but have proved invaluable for inducing lexicons for machine translation [2, 5]. Similarly, a type of resource intended for musicians can be exploited to associate meaning with music. Guitarists learning new songs often rely upon tablature notation ("tabs") provided by others to show the finger placement for a song measure by measure. Tabs often include lyrics, enabling note sequences to be associated with words. They also might indicate chords as an aid to learning the sequence (as is often done in scores for folk songs). In some cases, the chord annotations for lyrics are sufficient for playing certain songs (e.g., accompaniment by guitar strumming).

There are several web sites with large collections of tabs and chord annotations for songs (e.g., about 250,000 via www.chordie.com). These build upon earlier Usenet-based guitar forums (e.g., alt.guitar.tab). Such repositories provide a practical means to implement unsupervised learning of the meaning of chord sequences from lyrics. As these resources are willingly maintained by thousands of guitarists and other musicians, a system based on them can be readily kept current. This paper investigates how to utilized such resources for associating meaning with chords.

A motivation for this work comes from the context of songwriting. Given lyrics one has written, the challenge is to come up with the structure of the accompaniment, such as chord sequences that might be strummed and/or a series of notes to be played at various points of the song. Although the main consideration is in composing music that sounds good when played, it is often desirable for the music to convey a mood that complements the lyrics. The approach used here could be used to suggest chord sequences that might convey moods suitable for a particular set of lyrics. It could also aid in the reverse direction to aid songwriters who proceed from melody to lyrics, but this would require elaborate natural language generation support [7] to produce coherent lyrics. This is a follow-up to our previous work [15], which presents a simple approach for learning the meaning of chord sequences from associated lyrics. There, co-occurrence statistics are maintained over chord sequences and meaning tokens to determine significant associations. To improve the associations between chords and lyrics, we use tools developed for machine translation, which rely upon word alignments discovered in parallel corpora. This is not to suggest that learning meaning from chords is simply a matter of "translating" chord sequences into text. Our hypothesis is simply that word associations over a large collection of lyrics with chord annotations provide an effective basis for chord meanings. Given the relatively small number of chords used in practice compared to words, this is a *many-to-few* type of association (i.e., course-grained). Text categorization can be used to produce more constrained associations, as done in our previous work [15], which is more suitable for music recommendation.

We first discuss related work (§2). Subsequent sections provide details on the methodology (§3), an overview of the data (§4), and experimentation results (§5). We conclude with a summary and directions for future work (§6).

2. BACKGROUND

There has been a variety of work in music information retrieval on learning the meaning of music. Most approaches have used supervised classification in which user tags serve as ground truth for machine learning algorithms. A few have inferred the labels based on existing resources. The approaches differ mainly on the types of features used. Whitman and Ellis [20] combine audio features based on signal processing with features based on significant terms extracted from reviews for the album in question, thus an unsupervised approach relying only upon metadata about songs (e.g., author and title). Turnbull et al. [19] use similar types of audio features, but they incorporate tagged data describing the song in terms of genre, instrumentality, mood, and other attributes. Hu et al. [6] combine word-level lyrics and audio features, using tags derived from social media, filtered based on degree of affect, and then revised by humans (i.e., partly supervised). McKay et al. [11] combine class-level lyric features (e.g., part of speech frequencies and readability level) with ones extracted from user tags from social media, specifically via Last.fm. [1] They also include features for general term co-occurrence via web searches for the task of genre classification.

There has been other recent work in analyzing symbolic chord annotations. Macrae and Dixon [9] extract online chord annotations and show how they can be ranked according to sequence similarity to help filter bad annotations. McVicar et al. [13] use chord sequences from online sources to augment the task of chord recognition from audio via Hidden Markov Models (HMM's). Barthet et al. [1] extract chord annotations to augment a guitar tutor program (e.g., to illustrate chord fingering).

Lastly, there are a few approaches addressing the relations between lyrics and audio, rather than using them as separate features. Torres et al. [18] use a correlation-based approach referred to as *Canonical Correlation Analysis* (CCA) to associate lyrics with audio features. Under the CCA methodology, songs are represented in two feature spaces: a semantic annotation feature space and an audio feature space. For each space, the CCA identifies a one dimensional projection that maximizes the correlation between the projected data. The identified projections are used to construct and refine a musically meaningful vocabulary applied to assigning meaning to music. In addition, they present an approach to infer the projections under the assumption that the vector spaces are sparse. More recently, McVicar et al. [12] apply CCA to assess the correlation between lyrics and audio features as a part of an unsupervised system for quantifying mood. The system exploits a special dictionary on affect, specifically with ratings for valence (e.g., 'pleased' vs. 'frustrated') and arousal (e.g., 'excited' vs. 'sleepy'). Both approaches deal with meaning at the song level, but we address the issue of assigning meaning to smaller units. Furthermore, rather than audio features, we assign meanings to musical units more commonly used in music theory (e.g., chord progressions), making the results more accessible to musicians.

[1] See http://www.last.fm.

Overall process

1. Obtain large collection of lyrics with chord annotations
2. Extract lyrics proper with annotations from dataset
3. Convert into tab-delimited chord annotation data format
4. Determine best chord-word associations

Simple approach

4a. Fill contingency table: chord(s)/word co-occurrences
4b. Determine significant chord(s)/word associations

Preferred approach

4a. Invoke GIZA to produce chord(s)/word alignments
4b. Filter extraneous alignments

Figure 1. Process in learning meanings for chord sequences. The meanings are via individual words; and, $chord(s)$ is a single chord or a four-chord sequence.

Parallel corpora are vital for machine translation. Gale and Church [5] show how translation lexicons can be induced via co-occurrence statistics over contingency tables derived from such corpora. Parallel corpora have also been exploited to develop statistical machine translation systems, following pioneering work by IBM [2]. This incorporates sophisticated statistical models to account not only for co-occurrence, but also word order and degree to which alignment with multiple words are allowed (i.e., "fertility", which can account for phrasal alignments). Och and Ney [14] show that these models outperform other approaches for alignment (using GIZA, their implementation of them).

3. METHODOLOGY

Figure 1 lists the steps involved in the overall process for learning the meaning of chord sequences. First, a website for guitar instruction is downloaded to obtain a large sample of lyrics with chord annotations. The resulting data then is passed through a filter to remove extraneous text associated with the lyrics (e.g., transcriber notes). Next, the data is converted into a tabular format reflecting the chord/lyrics correspondences.

There are two approaches for obtaining the chord/word

Alternating lines:

```
C                            F
They're gonna put me in the movies
C                                     G
They're gonna make a big star out of me
      C
We'll make a film about a man that's sad
        F
    and lonely
    G7                       C
And all I have to do is act naturally
```

In-line chords:

```
[C] They're gonna put me in the [F] movies
[C] They're gonna make a big star out of [G]
    me
We'll [C] make a film about a man that's sad
    and [F] lonely
And [G7] all I have to do is act
    [C] naturally
```

Figure 2. **Chord annotation sample.** Lyrics are from "Act Naturally" by Johnny Russell, with chord annotations for the song as recorded by Buck Owens.

associations. In the simple approach, the data is converted into contingency tables from which co-occurrence statistics [10] are computed (e.g., Dice and mutual information). In the preferred approach (i.e., current NLP "best practice"), the data is formatted as a parallel corpus file and fed into a statistical word alignment system, such as GIZA. Afterwards, extraneous alignments are filtered.

3.1 Lyric Chord Annotation Data

The most critical resource required is a large set of lyrics with chord annotations. These annotations are often specified with alternative lines for chords and for the lyrics. They can also be specified with chords in-line with the lyrics. Figure 2 shows some examples of both formats. The popular website Chordie is used to obtain the data. [2] The website is crawled, and all the songs in the *chord.pere* directory are extracted (other directories are for user songbooks, etc.). There are over 65,000 files, but preprocessing complications reduces this to about 10,000 usable songs. After all processing, over 2 million distinct chord annotations are obtained. The chord annotation data is used as is (e.g., without normalization into key of C). We are working on transposing into the key of C, but we have run into key detection issues with the standard approach using key profiles [17]: presumably, that relies upon support from notes in the melody (omitted from chord annotations).

After the chord-annotated lyrics are downloaded, postprocessing is needed to ensure that user commentary and other additional material are not included. This is based on a series of regular expressions. [3] The lyrics are all converted into a tabular format that more directly reflects the

[2] See www.chordie.com; this was crawled in September of 2011.

[3] The Perl code for reproducing the experiments is available at www.cs.txstate.edu/%7Eto17/chord-meaning-from-lyrics.

```
C  They're gonna put me in the
F  movies <l>
C  They're gonna make a big star out of
G  me <l> We'll
C  make a film about a man that's sad and
F  lonely <l> And
G7 all I have to do is act
C  naturally <l> <v>
```

Figure 3. **Sample chord annotations extracted from lyrics.** Each chord instance in figure 2 has a separate line.

Contingency Table Cells

X\Y	+	-
+	XY	$X\neg Y$
-	$\neg XY$	$\neg X\neg Y$

G versus 'film'

	+	-
+	14	231,223
-	85	1,557,047

Table 1. **Contingency tables.** The left shows the general case, and the right shows the data for chord G and 'film'.

line-level alignment of chords and the corresponding text. Specifically, this uses a tab-separated format with the current chord name along with words from the lyrics for which the chord applies. There will be a separate line for each chord change in the song. Figure 3 illustrates this format. This shows that special tokens are also included to indicate the end of the line and verse.

3.2 Chord Sequence Token Co-occurrence

As mentioned above, the simple approach to deriving word associations is based on co-occurrence statistics. Several metrics have been proposed to measure this [4]; for example, *Chi Square* analysis determines the extent to which co-occurrence counts differs from that due to chance (e.g., difference of joint probability from the product of the marginals).

Given the tabular representation of the chord annotations with lyrics words, the next stage is to compute the co-occurrence statistics. This first tabulates the contingency table entry for each pair of chord and target token, as illustrated in table 1. (Alternatively, chord sequences can be of length four, as discussed later. These are tabulated using a sliding window over the chord annotations, as in n-gram analysis.) This table shows that the chord G co-occurred with the word 'film' 14 times, out of the 231,237 total instances for G. The word itself had 99 occurrences, and there were 1,557,047 instances where neither the word 'film' nor the chord G occurred. Next, a variety of co-occurrence metrics are derived using these tabulations, including Dice, Jaccard, mutual information, Chi square, and G^2 log likelihood [4, 10]. These are defined as shown in figure 4.

3.3 Alignment via GIZA

Using the IBM models [2] for word alignment has been shown to outperform simple co-occurrence metrics [14]. For this, we use the GIZA toolkit (specifically *GIZA++* version 2). Given its development for machine translation,

$$Dice(X,Y) = \frac{2 \times P(X=1,Y=1)}{P(X=1)+P(Y=1)} \qquad Jaccard(X,Y) = \frac{f(X=1,Y=1)}{f(X=1,Y=1)+f(X=1,Y=0)+f(X=0,Y=1)}$$

$$MI(X,Y) = log_2 \frac{P(X=1,Y=1)}{P(X=1) \times P(Y=1)} \qquad AvgMI(X,Y) = \sum_x \sum_y \frac{P(X=x,Y=y) \times log_2(P(X=x,Y=y))}{P(X=x) \times P(Y=y)}$$

$$\chi^2(X,Y) = \sum_{i,j} \frac{P(obs[ij]-exp[ij])^2}{exp[ij]} \qquad G^2(X,Y) = 2 * \sum_{i,j} exp[ij] \times log(\frac{obs[ij]}{exp[ij]})$$

$Dice(G, film) = 0.000121$; $Jaccard(G, film) = 0.000061$; $MI(G, film) = 0.129199$
$AvgMI(G, film) = 0.0000001$; $X^2(G, film) = 0.129045$; $G^2(G, film) = 0.125770$

Figure 4. **Common co-occurrence metrics.** Using the counts shown in table 1, these statistics can be directly computed, resulting in the values shown for the chord G and word 'film'.

```
C F  They're gonna put me in the movies<l>
C G  They're gonna make a big star ... me<l>
C F  We'll make a film about a man that's \\
        sad and lonely<l>
F G7 C And all I have ... act naturally<l>
```

Figure 5. **Alternative chord annotations extracted from lyrics.** Chords for same verse line in figure 3 are together.

General		Lyrics		General		Lyrics	
Word	Freq	Word	Freq	Word	Freq	Word	Freq
the	.057	*i*	.033	with	.007	is	.005
and	.028	the	.028	as	.006	all	.005
of	.027	a	.028	at	.005	for	.005
to	.026	*you*	.024	this	.005	we	.005
a	.023	and	.018	they	.005	can	.004
in	.019	to	.017	be	.005	but	.004
that	.013	in	.011	are	.005	so	.004
i	.011	it	.010	have	.005	don	.004
it	.010	me	.010	we	.005	re	.004
is	.010	my	.009	but	.005	ll	.004
for	.009	of	.008	his	.005	d	.004
you	.008	on	.007	from	.004	*love*	.004
was	.008	that	.007	not	.004	no	.004
he	.007	your	.006	n't	.004	she	.004
on	.007	be	.005				

Table 2. **Top words in corpus.** General word frequencies based on Corpus of Contemporary American English [3], and word frequencies for lyrics based on Chordie.

GIZA requires the specification of the source and target languages. Most work in statistical MT treats English as the source language and another language like French as the target. For the experiments discussed here, the chords are treated as the source and the target the words (mostly English). In our case, running the tool with the reverse direction produces negligible differences. In addition, GIZA normally includes a preprocessing stage that groups tokens in classes based on similar usages. However, that stage is omitted here because there is no context with which to determine the classes.

IBM Model 1, the simplest one in GIZA, follows: [2]

$$P_\theta(t, a|s) = P_\theta(l|s) P_\theta(a|l, s) P_\theta(t|a, l, s)$$

where s is the source language, t is the target language, l is target sentence length, a is the alignment, and θ are the overall parameters. The alignments are hidden and estimated via an HMM.

Prior to using GIZA, each column is put into separate files. Then, the toolkit preprocessing utilities convert them into a combined sentence file. (To avoid problems, lines that are too long or that contain garbage are discarded using the toolkit's utility to clean the input files.) GIZA only relies upon line correspondence in the two files when establishing alignments. Figure 5 shows how the input might be formatted. In addition, an optional step is used to group chords on the same line into sequences of four chords (e.g., $C\_D\_C\_D$), which are treated as individual tokens in the alignment.

4. OVERVIEW OF DATA

Before discussing the experiments, we present characterizations of the data involved. Naturally, lyrics are different from general English. Table 2 illustrates some differences in relative frequency for the top words. Comparing the two word frequency listings, we can see some peculiarities with respect to lyrics, such as the most common word being 'I' rather than 'the' and that 'you' moves up to the top 5. The word 'love' moves up in rank dramatically (271 to 27), and the word 'your' moves up a bit as well (from 69 to 14).

Frequency information for common chords and for chord sequences is shown in table 3. This illustrates that the major chords dominate the others, accounting for 64% of total occurrences. The B chord is an oddball, occurring less frequently than both of the minor chords Am and Bm, as well as being just a little more frequent than its minor. Note that the top of the sequence listing is skewed towards major chords; minor chords do occur in about half of the sequence types.

5. EXPERIMENTS

Two separate groups of experiments are performed. We first present an evaluation of the meanings attached to individual chords, using the common happy-versus-sad attribution regarding major versus minor chords. We also evaluate arbitrary chord sequences, using external annotations for songs meanings. External song-level annotations

Single		Sequence		Single		Sequence	
Ch.	Freq	Seq.	Freq	Ch.	Freq	Seq.	Freq
G	.154	CGCG	.005	G7	.010	CFCF	.003
C	.124	GCGC	.005	D7	.008	DCGD	.003
D	.124	EEEE	.004	A7	.008	GCGD	.002
A	.094	DGDG	.004	E7	.007	DAGD	.002
E	.068	GDGD	.003	Gm	.006	GCDG	.002
F	.061	GDCG	.003	Eb	.006	AGDA	.002
Am	.053	EAEA	.003	Em7	.006	AEDA	.002
Em	.047	DADA	.003	Am7	.005	DGCG	.002
B	.026	ADAD	.003	Cm	.005	GDAG	.002
Bm	.022	AEAE	.003	B7	.005	CGDG	.002
Dm	.019	CGDC	.003	C7	.005	DAED	.002
Bb	.015	FCFC	.003	Cadd9	.004	GDEmC	.002

Table 3. **Chord frequency.** This shows the frequency of chords and sequences (i.e., 4-grams) in Chordie.

$happy$: happy, blessed, blissful, bright, golden, halcyon, prosperous, laughing, riant, cheerful, contented, content, glad, elated, euphoric, felicitous, joyful, joyous, felicitous, fortunate, glad, willing, well, chosen, felicitous

sad:, sad, bittersweet, doleful, mournful, heavyhearted, melancholy, melancholic, pensive, wistful, tragic, tragical, tragicomic, tragicomical, sorrowful, deplorable, distressing, lamentable, pitiful, sorry, bad

Figure 6. **Synonyms for happy & sad.** Via WordNet 2.1.

are used in order to keep the evaluation objective, as there is no available resource with segment-level annotations.

5.1 Results for individual chords

The first evaluation covers the meaning attached to individual chords, such as that $Cmaj$ is 'bright' whereas $Cmin$ is 'somber'). To confirm the typical associations attributed to major versus minor chords (i.e., happy and sad, respectively), we compare the inferred word associations with synonyms reflecting this dichotomy. Figure 6 shows the synonyms for 'happy' and 'sad' from WordNet. [4] The idea is to check the most common chord associated with each of these synonym sets, seeing how often a major chord is chosen for a $happy$ word versus a minor chord for a sad word.

Specifically, we tabulate the average metric assigned to true and false positives for major versus minor chords. Figure 7 summarizes the result. For major chords, synonyms for 'happy' are assigned an average score of 81.1 (using X^2), whereas synonyms for 'sad' are assigned an average score of 39.2. Likewise, for minor chords, synonyms for 'sad' have an average score of 77.7, compared to 62.1 for 'happy'. As a baseline, a random value was used in place of the co-occurrence metric. As shown in the figure, there are much fewer true positives for the major chords (e.g., average scores for good versus bad nearly the same).

[4] See http://wordnet.princeton.edu.

Total 186 cases with score 12541.236 (avg 67.426)
Major good: 81 with 6573.275 (avg **81.152**) (A,contented) bad: 35 with 1326.313 (avg 37.895) (C,wistful) baseline: average scores *52.4* and 51.4, respectively
Minor good: 19 with 1476.133 (avg **77.691**) (Bm,tragic) bad: 51 with 3165.515 (avg 62.069) (Am,bright) baseline: average scores *32.9* and 43.7, respectively

Figure 7. **Evaluation of individual chord meanings.** This tests how well the metric decides whether synonyms for 'happy' ('sad') should go with a major (minor) chord.

5.2 Results for chord sequences

To evaluate the performance in learning chord sequence meaning, we compare the output against the Mood Tag Dataset (MTD) prepared by Hu et al. [6]. [5] Table 4 lists the meaning categories used in the MTD, along with the words used to define the categories. For example, category $G11$ is for sincerity and is defined in terms of 'earnest' and 'heartfelt'. This data set only provides song-level annotations, so we count how often the inferred chord sequence meanings match the song-level meanings for all the songs incorporating the chord sequence. For example, if a particular song contains 10 distinct chord sequences, and if six of the sequences were labeled with the meaning category corresponding to the song annotation, then the score for the song would be 0.6. As the MTD categories are defined in terms of words, we check for word overlap from the top words associated with a chord sequence with those from the meaning category. Although a lenient measure, the word-chord alignment process being evaluated has the handicap of dealing with over 10,000 meaning categories (i.e., all lyric words).

To test against the MTD, we just need the chord annotations for each of the songs covered. The annotations are for specific combinations of artist and album, so the songs are downloaded individually via the web interface to ensure the right version is used (if available). Out of 3,470 songs that are annotated, only 2,160 chord annotation files were obtained. Songs can be labeled with more than one category. If so, when verifying whether a chord is a match, we check the associated word for membership in any of the lists. The results are promising when using GIZA for the alignment using special tokens for chord sequences. The resulting alignment shows high precision, specifically at 89.5% (1,779 chord sequences out of 1,987). However, this comes at the expense of recall, with no suggestions for many of the chord sequences. In comparison, using average mutual information yields about 70,000 more taggings, but the precision drops to 20%. The baseline for this is 25.9%, which is the relative frequency for the most common category ($G12$).

[5] This dataset was used in MIREX-2011. See www.music-ir.org/mirex/wiki/2011:Audio_Tag_Classification.

Label	Freq	Examples
G12	.259	calm, comfort, quiet, ... tranquility
G15	.182	sad, sadness, unhappy, ..., sad song
G5	.115	happy, happiness, ..., mood: happy
G32	.095	romantic, romantic music
G2	.084	upbeat, gleeful, ...
G16	.073	depressed, blue, dark, ... gloomy
G28	.039	anger, angry, choleric, ...
G17	.028	grief, heartbreak, ... sorrowful
G14	.022	dreamy
G6	.022	cheerful, cheer up, ... sunny
G8	.018	brooding, contemplative, ... wistful
G29	.018	aggression, aggressive
G25	.012	angst, anxiety, ... nervous
G9	.009	confident, encouraging, ... optimistic
G7	.007	desire, hope, hopeful, ...
G11	.006	earnest, heartfelt
G31	.006	pessimism, cynical, pessimistic, ...
G1	.005	excitement, exciting, exhilarating, ...

Table 4. **Mood Tag Dataset.** Categories for MTD along with sample words used to define them. $Freq$ gives relative frequency, out of 6,490 total assignments.

6. CONCLUSION

This paper has demonstrated how to learn the meaning of chord sequences from lyrics annotated with chords. Two separate approaches have been illustrated. The simple approach uses co-occurrence statistics derived from contingency tables. The preferred approach uses word alignment tools designed for statistical machine translation.

For future work, we will look into additional aspects of music as features for modeling meaning (e.g., tempo and note sequences). In addition, as this approach could be used to suggest chord sequences that convey moods suitable for a particular set of lyrics, future work will investigate its use as a songwriting aid; in fact, this was the original motivation for the research.

By using resources intended for guitarists, the current work is more suitable for popular music than other types (e.g., classical). A long-term research goal is to develop a framework for learning similar associations from scores that include lyrics (e.g., operas). Other long-term aspects to be addressed include getting access to more data and integrating audio analysis into the process. In principle, voice recognition over lyrics could ameliorate sparse data problem, provided that the natural noise in songs can be sufficiently filtered.

Acknowledgments Hupahu Ballard helped validate our experiments. Scott Heftler provided useful tips on the processing. Hu Xiao provided access to the MTD. Douglas Eck suggested the extension using speech recognition over audio.

7. REFERENCES

[1] M. Barthet, A. Anglade, G. Fazekas, S. Kolozali, and R. Macrae. Music recommendation for music learning: Hotttabs, a multimedia guitar tutor. In *Proc. Workshop on Music Recommendation and Discovery (WOMRAD-2011)*, pages 7–13, 2011.

[2] P. F. Brown, V. J. Pietra, S. A. D. Pietra, and R. L. Mercer. The mathematics of statistical machine translation: Parameter estimation. *Computational Linguistics*, 19:263–311, 1993.

[3] M. Davies. The 385+ million word corpus of contemporary American English (1990-2008+): Design, architecture, and linguistic insights. *International Journal of Corpus Linguistics*, 14:159–90, 2009.

[4] T. Dunning. Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, 19(1):61–74, 1993.

[5] W. A. Gale and K. W. Church. Identifying word correspondences in parallel texts. In *Fourth DARPA Workshop on Speech and Natural Language*, pages 152–157, 1991.

[6] X. Hu, J. S. Downie, and A. F. Ehman. Lyric text mining in music mood classification. In *Proc. ISMIR*, pages 411–6, 2009.

[7] D. Jurafsky and J. Martin. *Speech and Language Processing*. Prentice Hall, Upper Saddle River, New Jersey, 2000.

[8] G. Leech. *Semantics*. Middlesex, Penguin Books, 1974.

[9] R. Macrae and S. Dixon. Guitar tab mining, analysis and ranking. In *Proc. ISMIR, Miami, Florida.*, 2011.

[10] C. D. Manning and H. Schütze. *Foundations of Statistical Natural Language Processing*. MIT Press, Cambridge, Massachusetts, 1999.

[11] C. McKay, J. A. Burgoyne, et al. Evaluating the genre classification performance of lyrical features relative to audio, symbolic and cultural features. In *Proc. ISMIR*, 2010.

[12] M. McVicar, T. Freeman, and T. D. Bie. Mining the correlation between lyrical and audio features and the emergence of mood. In *Proceedings ISMIR*, 2011.

[13] M. McVicar, Y. Ni, R. Santos-Rodriguez, and T. D. Bie. Leveraging noisy online databases for use in chord recognition. In *Proc. 12th International Society on Music Information Retrieval (ISMIR)*, 2011.

[14] F. J. Och and H. Ney. A systematic comparison of various statistical alignment models. *Computational Linguistics*, 29(1):19–51, 2003.

[15] T. O'Hara. Inferring meaning of chord sequences from lyrics. In *Proc. Workshop on Music Recommendation and Discovery (WOMRAD-2011)*, pages 40–43, 2011.

[16] C. Schmidt-Jones and R. Jones, editors. *Understanding Basic Music Theory*. Connexions, 2007. http://cnx.org/content/col10363/latest.

[17] D. Temperley. A Bayesian approach to key-finding. In C. Anagnostopoulou, M. Ferrand, and A. Smaill, editors, *Music and Artificial Intelligence*, pages 195–206. Berlin, Springer-Verlag, 2002.

[18] D. Torres, D. Turnbull, L. Barrington, and G. Lanckriet. Identifying words that are musically meaningful. In *Proc. ISMIR*, September 2007.

[19] D. Turnbull, L. Barrington, et al. Semantic annotation and retrieval of music and sound effects. *IEEE TASLP*, 16 (2), 2008.

[20] B. Whitman and D. Ellis. Automatic record reviews. In *Proc. ISMIR*, 2004.