The Eurovision Song Contest on *YouTube*:
A Corpus-Based Analysis of Language Attitudes

Dejan Ivković
York University

Abstract

This article examines language attitudes expressed in the comments sections of *YouTube* web pages. A corpus-based analysis of language attitudes was carried out on data taken from a range of *YouTube* pages featuring video uploads of songs performed in the Eurovision Song Contest (ESC) between 2003 and 2010. Comments of *YouTube* users were analyzed using the WordSmith concordancer to reveal patterns of overtly-expressed language attitudes. The analysis shows that language is a statistically frequent topic in ESC *YouTube* discussion boards. Generally, singing in a non-native language elicits mostly negative comments, but compared to the comments in English, the French language comments tend to be more about the status of French relative to that of English and less about overall linguistic diversity in the ESC. Finally, the performance of the German 2010 winning song elicits mixed positive/negative responses, appearing controversial to commenters because it was delivered in what appears to be a mélange of standard British English, a “Cockney-like” accent, and a German accent. The study sheds light on multilingualism and language choice in the ESC and offers an input parameter for language policy and language planning related to the Contest.

Introduction

According to Garrett, “a combination of a more quantitatively focused work searching for the broader patterns in the landscapes with more qualititative approaches, giving us a deeper understanding of the competing ideological stances involved, would link with a view of language attitudes research” (p. 157). Such an approach may also include a media-based treatment from a language attitudes perspective that “gives a picture of the social and cultural meanings of specific linguistic landscapes and how language attitudes can influence them” (p. 156).

The Internet is a domain-specific instantiation and extension of the broader linguistic landscape, what we call the virtual linguistic landscape (Ivković, 2012; Ivković & Lotherington, 2009). One site on this landscape, the Eurovision Song Contest (ESC), represents a particularly rich opportunity for language attitude researchers. Multilingual in character, the ESC, which originated in 1956, is present on *YouTube* through uploads of songs performed in the Contest. As a “song contest on a political stage” (Raykoff, 2009) and a locus of intensive language contact and language contestation, the ESC involves multilingual use and reflects linguistic capital and power relations among the nations and languages of Europe.

This article examines folklinguistic attitudes – that is, non-linguists’ views of languages and language variation (e.g., Preston, 1996) – of commenters on Eurovision Song Contest *YouTube* pages, by triangulating the results of corpus-based analyses with the societal treatment approach (Garrett, 2010; McKenzie, 2010), as described further below.
The *YouTube* online forums associated with the ESC have a large number of users from varied linguistic backgrounds who, because of their interests in song performance, are particularly attentive to language-related issues, such as the accent of the performers and the choice of language of the songs. For these reasons, *YouTube* provides “an ideal space for dominant ideologies of language to be contested” (Chun & Walters, 2011, p. 352). Commentaries are made by forum participants from disparate locations on a variety of topics, one of the most prominent being language, including language features and perceptions of language use.

The corpus-based analysis of language attitudes in this study was carried out on data taken from *YouTube* pages featuring video uploads of songs performed in the ESC between 2003 and 2010. I first discuss language choices in the ESC and the corpus-based method as an approach to language attitudes research. In the central part of the article, I present three case studies of language attitudes expressed in the comments sections of *YouTube* web pages. I conclude with a discussion of the potential of the research to inform ESC language policy.

**Traditional Approaches to Language Attitudes**

An attitude is a mental state and a concept in social psychology closely related, but not identical, to the concepts of belief, opinion, stereotype, and value (Allport, 1935). According to social psychologists Eagly and Chaiken, an “attitude is a psychological tendency that is expressed by evaluating a particular entity with some degree of favour or disfavour” (p. 1, emphasis in original). As attitudes are subjective and not directly observable (Allport, 1935; Eagly & Chaiken, 1993), scholars have to rely on both overt and covert verbal and non-verbal messages – as well as emotional reactions – in order to explore other people’s attitudes.

It is no surprise that attitudes emerge from language contestation. Attitudes to language have been the object of psycho-sociological research from the 1920s to the present (C. Baker, 2006; Gardner, 2002). The study of language attitudes has positioned itself as an interdisciplinary field drawing from both social psychology and linguistics. Its domain of investigation is linguistic phenomena, but its methods of analysis and theoretical underpinnings primarily stem from social psychology. The units of analysis for attitude range from specific or “whole” languages (Garrett, 2010, p. 10) to language features to language contact phenomena such as code-switching.

Three main approaches have been identified in research on language attitudes (e.g., Garrett, 2010; Ryan, Giles, & Hewstone, 1988): direct measures, indirect measures (or “the matched guised technique”), and societal treatment of language varieties. In the direct approach (measures), respondents are typically asked to evaluate language use or language phenomena. Attitudes to various language phenomena are often elicited explicitly in the form of questionnaires or surveys (e.g., MacKinnon, 1981). The indirect approach to studying language attitudes involves covert and more subtle techniques. Typically, the indirect approach uses the matched guise technique, whereby respondents are exposed to different readings of the same text by a single speaker. Here, respondents are not consciously aware of what is being rated. Compared to the direct approach, the indirect approach is more likely to elicit people’s private attitudes, as opposed to the attitudes they would be willing to share explicitly with other people (Garrett, 2010, p. 41). In this approach, the societal desirability bias, or expressing what is believed to be socially acceptable or desirable, is minimized. Finally, the societal treatment
The studies approach relies on the analysis of existing texts and is the least obtrusive. In this type of research, attitudes are inferred from various kinds of observed behaviour (Garrett, p. 52), verbal or non-verbal. Garrett cites several studies as examples of the societial treatment approach (pp. 46-51), including an analysis of letters to the editors in African newspapers (Schmied, 1991). In Garrett’s framework, the present study represents the societial treatment approach, since the written comments that are taken as evidence of attitudes were collected from the media, rather than elicited by researchers.

A Corpus-based Approach to Language Attitudes

Some of the limitations inherent in the traditional methodologies for studying attitudes can be addressed using the methods and tools of computational analysis. The direct approach is subject to the “societal desirability bias,” whereby the respondents are likely to give answers that they perceive to be socially acceptable or desirable. C. Baker (1995) states that “people may be affected in their response to an attitude test by the research and the perceived purpose of the research” (p. 19). Further, both the direct and indirect approaches draw inferences from pre-conceived sets of questions and non-spontaneously occurring language. Last, a non-computational societal treatment approach is constrained by the number of media discourses analyzed; typically, generalizations are based on smaller and more limited datasets.

A corpus-based approach to analyzing language-related attitudes can complement the traditional methods and approaches. As Kennedy (1998) notes, “in the case of corpus-based research, the evidence is derived directly from texts. In this sense, corpus linguistics differs from approaches to language which depend on introspection for evidence” (pp. 7-8). The main advantage of the corpus approach in research on attitudes, as in discourse analysis more broadly, “is the use of computational tools for investigating distributional patterns across [a range of texts]” (Upton & Ann Cohen, 2009, p. 587).

Limitations of the Corpus-based Approach

At the same time, computational approaches to language attitudes have limitations that are inherent in the corpus-based methodology, in general, and in the “web-for-corpus” approach, in particular. Since the data comprise exclusively verbal text, visual and other non-linguistic markers, which otherwise may provide valuable secondary clues, are overlooked (P. Baker, 2006, p. 17). Further, the logged data do not provide the means to validate the researcher’s interpretations or inferences, in contrast to ethnographic approaches. With regard to the virtual ethnography approach, for instance, Androutsopoulos (2009) states that systematic observation and interviews with Internet actors can alert “researchers to emic categories and views, which may act as correctives to their assumptions and interpretations” (p. 16). To address the latter limitation and strengthen the findings, the present study triangulates the results of the quantitative corpus measures and discourse analysis with media commentaries.

Research Questions

Initial observations based on my reading of online comments posted by YouTube users led to a sense that language-related discussions constitute one of the most frequent topics on YouTube websites devoted to the Eurovision Song Contest. Moreover, it seemed that positive attitudes
towards singing in one’s native language in the contest prevailed, as opposed to singing in English by non-native speakers. These observations gave rise to two research questions:

RQ1: How prominent is the topic of language in the ESC YouTube discussion forums, as measured by the presence of language-related lexical items?

RQ2: What is the predominant attitude towards singing in English by non-native speakers versus singing in one’s native language? More specifically, how did the commenters react to the choice of English to represent France in 2008?

The first question is addressed through analyses of the frequencies, keyness (salience of lexical items relative to their occurrence in another comparable corpus), and dispersion of language-related lemmas (the set of different forms of words) across the corpus. The second question is addressed in two separate case studies that analyse attitudes about languages: English versus native language, and français versus anglais (French versus English). The choice of the third case study was driven by the results of the frequencies and distribution analyses, which pointed at the word ‘accent’ in the German sub-corpus that was collected from the comments on the ESC 2010 German entry.

Data and Methodology

Corpus Creation

To address the research questions, a corpus was compiled from the ESC users’ comments to illustrate language attitudes on YouTube. The comments are mostly in English, while comments in other languages, including Dutch, Russian, Serbian/Croatian, Scandinavian, and French, are found in the respective sub-corpora. The corpus creation consisted of a controlled collection of four sub-corpora, based on the following predefined parameters: the year of competition, the winning song, and the number of comments posted (see Table 1). In addition, two smaller sub-corpora, Belgium 2003-2008 (SC_B) and France 2008 (SC_F), were collected from the comments on recent non-winning songs. In these two sub-corpora, the language choices of the entries were more controversial: SC_B was collected from the comments on two Belgian entries (2003, 2008), both songs being performed in “imaginary” or made-up languages, while SC_F was collected from the comments on the French entry performed entirely in English (2008).

The general macro corpus (GENERAL_ESC) comprises six sub-corpora. Four balanced sub-corpora (SC), each having on average 150,000 tokens (strings of characters separated by spaces, such as ‘singing,’ but also non-words, such as ‘ouio’) were created from the comments on the video uploads of the ESC winning songs between 2007 and 2010: Serbia 2007 (SC_S), Russia 2008 (SC_R), Norway 2009 (SC_N), and Germany, 2010 (SC_G) (Table 1). At the time of collection, these were also the uploads with the largest number of viewings and posted comments. Two smaller sub-corpora, associated with two non-winning songs, were collected from a number of web pages with fewer – and predominantly bilingual – comments, those being English and Dutch, and English and French, respectively. These two sub-corpora contain around 45,000 and 93,000 tokens, respectively. GENERAL_ESC also served as a reference corpus used to determine relative frequencies, key word inventory, collocations, and dispersion. The total corpus size (GENERAL_ESC) is 769,846 tokens.
In the initial phase of the study, language-related lemmas and collocations were identified through analyses of lexical frequencies and distribution. Subsequently, the selected lexical items and patterns describing language attitudes were analyzed using the Wordsmith concordancer.

The analysis of word frequencies and distribution served two purposes: first, to address the first research question concerning the frequency of language attitudes as a discussion topic in the ESC YouTube comments and, second, to identify specific language-related lexical items relevant to this study. The analysis of collocations and concordances served to flag potentially new and specific topics; this was the principled motivation for dividing the corpus into six sub-corpora of varied sizes, as my hypothesis was that different uploads of songs would elicit different types of reactions from YouTube users, and I wanted to zero in on particular language-related topics (e.g., accent or the use of a particular language). In addition, dividing the corpus facilitated making potential generalizations through the analysis of keyness and dispersion across the sub-corpora.

The British National Corpus (BNC) serves in this study as an external control corpus. The standard BNC corpus is used alongside the specialized ESC corpus “to discover how far certain features are distinctive of the discourse under examination and how far they occur elsewhere in the language as a whole” (McEnery & Wilson, 2003, p. 115).

**Conceptual Lemma**

To describe occurrences of individual lexical items, that is, lemmas, representing concepts in multiple languages, the term *conceptual lemma* is used. A conceptual lemma is “a representation of a language independent unit of meaning, either as a bilingual […] or as a multilingual conceptual lemma” (Gibbon & Lüngen, 2000, p. 303), unless individual languages are compared, in which case it is referred to as just a *lemma*. In this study, conceptual lemmas are represented in English, whether the word comes from English or another language, and whether it was encoded using the standard spelling or an orthographic variant. They are capitalized and in brackets, to indicate that they are abstractions and conceptual representations. In this study, two language-related conceptual lemmas are chosen `/LANGUAGE/` and `/ACCENT/`. The search term `langu*/sprache/taal` in English, French, German or Dutch, for example, produces instances (tokens) of the conceptual lemma `/LANGUAGE/`, with the wild card asterisk (*) retrieving all the tokens that start with the character string before the wild card. The search term `langu*` may...
yield the following results: languages, LANGUAGE, language (misspelled in English), or langue (in French).

**Coding of Attitude Valence and Attitude Types**

Three raters were asked to evaluate independently argument types that YouTubers had used to explain or comment on language choices in songs in the ESC. The raters were also instructed to determine the attitude valence of attributes collocating with the lemma ACCENT. To measure the range of attitudes to language and the ratio between positive and negative attitudes, a semantic differential scale was used (Snider & Osgood, 1969). Data from the corpus were classified according to a seven-point attitude scale with the following levels: very positive (score 3), positive (2), positive to ambivalent (1), ambivalent (0), ambivalent to negative (-1), negative (-2), and very negative (-3). Numerical scores from +3 to -3 points were allocated to each level accordingly. The raters were advised to consider the examples within the context, allowing for the possibility that the same adjective or proposition would be rated as both positive and negative in different contexts (as occurred, e.g., in the case of Cockney). In cases of disagreements, the raters were asked to negotiate a final rating.

**Frequency and Distribution Results**

**Frequency and Keyness**

While simple frequency measurements reveal the total number of occurrences of particular lemmas in the corpus, representations of keyness provide clues about “the statistically significantly higher frequency of particular words or clusters in the corpus, or a comparable specialized corpus” (Baker et al., 2008). Frequency lists of lexical items in the ESC corpus reveal the prominence of words and collocations pertaining to language attitudes. Language-related and potentially language-related lemmas, such as LANGUAGE, ACCENT, PRONUNCIATION, COCKNEY, TONGUE, FRENCH, RUSSIAN, are far more frequent in the ESC corpus than they are in the British National Corpus.

Table 2 shows the relative frequency of the semantic fields identified in the first 50 non-grammatical lemmas. The semantic fields referring to the contest and location are inherently associated with the nature of the event (e.g., song, win, points, country, Europe). In addition, evaluative/affective (e.g., beautiful, love, sucks) and cognitive/declarative (e.g., think, say) fields identified in the frequency list of GENERAL_ESC are expected and frequent in a general corpus. Furthermore, the lexical items from the semantic fields of profanity (e.g., fuck, shit) and the one denoting one’s sexual orientation (e.g., gay) frequently occur in social networking sites (Thelwall, 2008). Language-related lexical items (e.g., language, accent, English language, Russian language), on the other hand, are not inherently linked to the ESC. Hence, their presence in the corpus may be interpreted as salient.

In the BNC, the frequency ranking for the lemma LANGUAGE is 537, while the ranking for the lemma ACCENT is 5,737. By way of comparison, in GENERAL_ESC, including grammatical words, these two language-related lemmas belonging to the same semantic field rank 147 /LANGUAGE/ and 125 /ACCENT/. Cumulatively, these two lemmas, indexing the semantic field of ‘language,’ account for one of the most discussed topics (Table 2), after discussions about the ESC itself, songs, and countries.
Semantic field | Lemmas | Frequency | Rank |
---|---|---|---|
Eurovision | Song-songs-singer; Eurovision; win-won-winner; points; voice; music; video; Dima (name of the winner in 2008) | 29,800 | 1 |
Affective/ Evaluative | Like (note: can also be a conjunction); love; sucks; lol Good-better-best; bad; nice; cute; great; stupid; beautiful | 23,488 | 2 |
Location | Russia; Serbia; Greece; country-countries; Germany; France; Europe; Turkey | 9,014 | 3 |
Cognitive or declarative | Think; know; say | 4,800 | 4 |
Profanity | Fuck; shit | 2,552 | 5 |
Potentially language | English; Russian; French | 2,333 | 6 |
Language | Language; accent | 1,603 | 7 |
Other | People, go, have | 1,579 | 8 |
Sexual orientation | Gay | 1,459 | 9 |

Table 2. Frequencies of lemmas according to semantic field in the GENERAL_ESC corpus

We may thus posit that the keyness of the language-related lemmas is indicative of the significantly greater than expected presence of language-related discussions in GENERAL_ESC than in a generic corpus such as the BNC. Moreover, internal keyness, or salience of certain keywords relative to GENERAL_ESC, may be indicative of the presence or absence of specific attitudes, such as attitudes towards accent or language choice in individual sub-corpora. Keyness of the lemmas from Table 2 in a particular sub-corpus, focusing on language-related lemmas (Table 3), was measured against GENERAL_ESC minus the sub-corpus in question. For example, keyness of the Serbian sub-corpus (SC_S) was measured relative to GENERAL_ESC, excluding SC_S.

<table>
<thead>
<tr>
<th>Sub-corpus</th>
<th>‘LANGUAGE’ or first language</th>
<th>‘ENGLISH’</th>
<th>‘ACCENT’</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SC_B (Dutch)</td>
<td>26</td>
<td>null</td>
</tr>
<tr>
<td>2</td>
<td>SC_F (French)</td>
<td>25</td>
<td>36</td>
</tr>
<tr>
<td>3</td>
<td>SC_S (Serbian)</td>
<td>3</td>
<td>null</td>
</tr>
<tr>
<td>4</td>
<td>SC_R (Russian)</td>
<td>6</td>
<td>22</td>
</tr>
<tr>
<td>5</td>
<td>SC_N (Norwegian)</td>
<td>9</td>
<td>null</td>
</tr>
<tr>
<td>6</td>
<td>SC_G (German)</td>
<td>7</td>
<td>19</td>
</tr>
</tbody>
</table>

Table 3. Keyness ranking for language-related lemmas by sub-corpus

Table 3 shows ranking of the lemma LANGUAGE, including the lemmas referring to the performer’s first language, e.g., SERBIAN (3), RUSSIAN (6), and NORWEGIAN (9), and the lemma ENGLISH. The relatively high ranking of the lemma LANGUAGE and the lemmas referring to the performer’s first language (e.g., FRENCH, RUSSIAN, SERBIAN) suggests that the topic of language is typical of the genre and is likely to occur in other ESC YouTube pages, as well. In contrast, the lemma ACCENT is salient only in SC_R & SC_G (rows 4 & 6). It is important to note that ACCENT in SC_G is the only language-related lemma with the highest keyness ranking of 1 in its sub-corpus, indicating that the topic of accent is the most salient in
this sub-corpus. Other lemmas from Table 2 that are salient in all six sub-corpora include /EUROVISION/, /SING/, /SONG/, and the lemmas referring to the performer’s country, while the lemma /GAY/ ranks high in SC_S (5), SC_R (7), and SC_N (9).

Further, the keyness pattern for the lemma ENGLISH may be interpreted as indicative of the saliency of the following discussion topics: In SC_F, it is singing in English instead of singing in French, while in SC_G and SC_R, it is singing in English with a mixed/heavy accent. Singing in the English language appears not to be an issue in SC_B, SC_S, or SC_N. In 2003 and 2008, Belgium was represented by a song performed in an imaginary language; Serbia was represented by a song in Serbian; while Norway’s 2009 winning song, performed in English, did not give rise to many posted comments. At the same time, the very absence of language-related comments in the case of the 2009 Norwegian winning entry may indicate covert “compliance” with English as a language choice, which in itself may be construed as an implicit language attitude.

Dispersion and Dispersion Plots

A dispersion analysis tells us “how evenly morphemes or words or patterns/constructions are distributed across (parts of) a corpus” (Gries, 2010, p. 5); this can be represented visually by a dispersion plot (P. Baker, 2006, p. 60). The genre of the Internet forum is heteroglossic (Androutsopoulos, 2009) and multiparticpant (Herring, 2010; Marococcia, 2004), typically lacking the cohesion that a single-authored text affords. Therefore, claims about representativeness need to be confirmed at the sub-corpus level to determine if a word or collocation indexing a particular language attitude is typical. Another reason for using dispersion and the graphical grid in this study is to illustrate corpus structure and cross-linguistic frequencies of specific lemmas in the respective languages visually. Figures 1.1–1.3 display dispersion plots of the conceptual lemmas /ACCENT/ and /LANGUAGE/, and the collocations sing* in and chante* en.

For example, Figure 1.1 shows dispersion (density and distribution evenness of the spread) of search terms within five sub-corpora and the GENERAL_ESC corpus (‘File’) of ‘lang*,’ ‘sprach*,’ and ‘taal’ in English/French, German, and Dutch, each retrieving instances of ‘language’ in respective languages. The horizontal axis displays the file name followed by the search term (‘File’), number of tokens (including word and non-words) (‘Words’), the total number of occurrences of the search term (‘Hits’), the number of occurrences per 1,000 tokens in the (sub)-corpus. Each vertical black line represents one occurrence. The green bar illustrates how even the spread of the search term is across the corpus (‘Plot’); the more even the distribution, the higher the dispersion value (‘Dispersion’), ranging from >0 to +/-1.
Language appears to be a topic readily discussed in a wide range of ESC YouTube threads. As Figures 1.1 and 1.2 show, the lemma LANGUAGE and collocation sing* in are relatively evenly distributed in GENERAL_ESC (rows 14 & 10). Such distribution indicates that discussing language, in general, and discussing which language to sing in, in particular, is a common theme in a wide range of ESC YouTube forums. In addition, the dispersion rate of the English/French lemma LANG* (Figure 1.1) is relatively even across GENERAL_ESC (.782; row 14), peaking at the beginning in SC_B, where the “Imaginary language” is discussed (.798; row 1). In contrast, the lemma ACCENT (Figure 1.3) has unstable distribution across GENERAL_ESC (.322, row
10) and occurs significantly only in SC_G and SC_R, where having either a strong foreign accent (SC_R) or a mixed accent (SC_G) is perceived as salient.

The analysis of frequency, keyness, and dispersion of lemmas has shown (1) a high frequency and salience of language-related conceptual lemmas, such as /LANGUAGE/, /ACCENT/, /ENGLISH/, and /FRENCH/ relative to BNC; and (2) collocational preferences, coupling these lemmas with phrases, such as sing in/chanter en, native language, and words, such as Cockney. Motivated by these findings, the next section presents the concordancing analysis of attitudes to language(s) and accent.

Concordancing Results

A concordance analysis of the words and collocations related to attitudes towards accents and specific languages was carried out on the GENERAL_ESC corpus. For the purpose of this analysis, words and collocates arguably not referring to language attitudes were ignored, e.g., “My native language is [language name]” (see Figure 2.1, row 11). Where the number of occurrences was too large, as was the case for singing in and accent, samples were retrieved based on a particular collocation criterion, such as singing in a native language or has a/an * accent. In the case of chanter en the sample was exhaustive and consisted of all 72 instances extracted from the corpus. To illustrate the attitudes towards singing in English, French, or one’s native language, a number of multi-party discussion threads or polylogues (Marococcia, 2004) were selected and analyzed as representative examples.

Specifically, attitudes towards singing in a native versus non-native language, the use of French versus English in the ESC, accent mixing, and attitudes towards accent were analyzed. These language-related attitudes and their saliency were identified and verified in the quantitative analysis and the preliminary concordance analysis, based on the following search terms: language (687 occurrences), langue (60), English (1,553), French (316), francais/français (181), and accent (862). The initial search revealed language-related collocates, such as singing in (356)/chanter en (72), honte (‘shame’ in French) (56), made up (32), imaginary (20), mother/native tongue (10), and native language (37).

<table>
<thead>
<tr>
<th>Collocate with /LANGUAGE/</th>
<th>sing* in</th>
<th>%</th>
<th>chante* en</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>/ENGLISH/</td>
<td>135</td>
<td>38.0</td>
<td>46</td>
<td>63.8</td>
</tr>
<tr>
<td>/FRENCH/</td>
<td>10</td>
<td>2.8</td>
<td>19</td>
<td>26.4</td>
</tr>
<tr>
<td>/RUSSIAN/</td>
<td>15</td>
<td>5.8</td>
<td>1</td>
<td>1.4</td>
</tr>
<tr>
<td>/NATIVE/</td>
<td>39</td>
<td>11.0</td>
<td>null</td>
<td>null</td>
</tr>
<tr>
<td>/MADE UP- IMAGINARY/</td>
<td>52</td>
<td>14.6</td>
<td>null</td>
<td>null</td>
</tr>
<tr>
<td>/OWN- YOUR-HER-HIS-ONE’S-THEIR/</td>
<td>31</td>
<td>8.7</td>
<td>1</td>
<td>1.4</td>
</tr>
<tr>
<td>OTHER</td>
<td>74</td>
<td>20.8</td>
<td>5</td>
<td>7.0</td>
</tr>
<tr>
<td>Total</td>
<td>356</td>
<td>100</td>
<td>72</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 4. Distribution of collocates sing* in/chant* en + attribute + /LANGUAGE/

As shown in Table 4, the most frequent collocate with both sing* in and chanter en is the conceptual lemma /ENGLISH/ (38% and 63.8%, respectively). The French language discussions,
henceforth indexed by the collocation *chant* en, tend to favour the distinction between singing in English versus singing in French. In contrast, in the English language comments, the debates around “singing in English versus singing in a native/one’s own language” dominate. Furthermore, the high frequency of the collocates imaginary and made-up account for the discussions on the Belgium 2003 and 2008 language choice, with most of occurrences coming from SC_B. Related to the lemma ACCENT, the collocates fake (27), Cockney (18), implAssible and its variants (32) (mocking the non-native pronunciation of the Russian singer Dima Bilan), were also identified.

In the following section, I explore these collocates as they occur in the comments, focussing on two case studies, both concerning attitudes about specific languages: the attitudes towards singing in a foreign versus native language and the attitudes towards singing in English versus singing in French. The first case study is addressed through the analysis of the English collocate *sing* in in GENERAL_ESC, while the second case study uses the predominantly French-language sub-corpus (ESC_F) to analyze concordances of the French collocate *chante* en, which occurs almost exclusively in this sub-corpus.

**Attitudes about Specific Languages**

**Foreign versus Native Language**

Out of 39 postings with the collocation “native language” (Figure 2.1), 28 (72%) express mostly positive attitudes towards the proposition “one should sing in one’s native language,” four (10%) are mostly negative, and seven (18%) are ambivalent or purely factual. Four types of arguments, each explaining a function of the choice of language both in the ESC and in general, can be identified in the concordances of *native language/tongue* (Figures 2.1 & 2.2) and Threads I & II (see further below). These functions are: evaluative/affective, instrumental, pragmatic, and symbolic.

*Evaluative/affective function arguments* are based on subjective premises or emphasize the importance of feelings: what something feels like; how something sounds. An example of a generalized claim is, “People always sound better when they sing in their native tongue” (Figure 2.2, row 3). Examples of more specific claims are, “Lena’s song would sound great in German. Immer noch fantastisch Lena. (‘Still fantastic Lena.’)” (Figure 2.2, row 2) and “It [singing in Serbian, a Slavic language] reminds me of my native language polish [another Slavic language]. It lets me feel the real emotion of the song” (Figure 2.1, row 24).

*Instrumental function arguments* claim that language choice is a means to win the contest. Since the rules changed, most recently in 1999, countries are allowed to send entries in the language of their choice. One commenter complained, “The worst thing happened when the rules changed...you don’t have to sing in native language...now nearly everyone sings in English” (Figure 2.1, row 29). However, as another commenter observed, “Generally, it’s the songs in English that win. So from a competitive standpoint, it’s better to sing in English” (Thread II, Commenter A). The UK, “having their native language as the most popular in the world should be at an enormous advantage” (Figure 2.1, row 30), whereas “a past requirement in EuroVision was that you had to sing in your native language, which unfortunately made it harder for the nations that spoke other languages to win” (Figure 2.1, row 9), “in this competition. but it seems
artists are reluctant to sing in their mother tongue, as most winners are English singing” (Figure 2.2, row 8).

Pragmatic function arguments emphasize the role of English as a lingua franca. According to this view, singing in English facilitates broader comprehension in the contest, as “English is a universal language, and it means all countries can understand the lyrics” (Thread I, Commenter B).

Symbolic function arguments hinge on the symbolic value of the presence versus absence of specific languages. “Native language is definitely a great national value” (Figure 2.1, row 3) and a symbol of identity (Figure 2.2, row 11). Language also represents a country (Figure 2.1, rows 4, 23, 31); for example, “Eurovision is not about being able to understand every single word that’s being said, it’s about receiving an impression of the different countries’ cultures...and therefore...every singer should sing in his mother tongue!” (Thread I, Commenter C). Further, singing in one’s native language is perceived as a tool in the fight against Anglicisation/Americanization (Figure 2.1, rows 15, 21) and an affirmation of cultural distinctiveness and diversity, which needs to be preserved (Figure 2.1, row 16) and celebrated (Figure 2.1, row 25).

Figure 2.1. Concordance of the collocation native language
In this sample, attitudes towards singing in one’s native language were categorized by commenters as mostly positive. Conversely, attitudes towards using English, as opposed to singing in one’s first language or mother tongue, were categorized as mostly negative, with a few statements rated as ambiguous or factual.

The arguments are exemplified in two discussion threads from the corpus (Threads I & II). In Thread I, Commenter A puts forward the argument that every country should be represented by a song performed in a domestic language. In response, Commenter B argues that one should instead sing in English as a *lingua franca* to facilitate wider understanding. According to Commenter C, in contrast, the pragmatic principle of information delivery is less important than the symbolic function of language, since language choices index cultural differences. Last, Commenter D implicitly agrees with Commenter C by offering a counterargument to Commenter B’s claim. S/he maintains that if the main criterion for the choice of English is its “universality,” then the preferred language choice in the ESC should be Chinese, since Chinese “is the language the most spoken in the world.”

**Thread I**

*Commenter A*

It’s a good song but every country should to sing in his languages in Eurovision

*Commenter B*

I disagree, English is the universal language, and it means all countries can understand the lyrics.

*Commenter C*

I disagree, Eurovision is not about being able to understand every single word that's being said, it's about receiving an impression of the different countries' cultures.... and therefore, I also think that every singer should sing in his mother tongue! (Figure 2.2, row 4)
Commenter D
Why should it be English the language the most spoken in the world is Chinese

In Thread II, according to Commenter A, singing in English is instrumental in winning the competition and achieving international fame and recognition. She or he notes that for the past 10 years, except in 2007 when a song in Serbian won (see Figure 2.1, row 23), all other winning entries were in English. Consequently, singing in English should significantly improve one’s chances to win the Contest (Commenter A). Commenter C supports this view, pointing out that the ESC statistics show that the entries in English are most likely to win. The BBC commentator, Terry Wogan, would likely agree. Wogan states, “ABBA had this little-known song called Waterloo; they sung it in Swedish and lost the Swedish Song Contest. And the following year they sung it in English and got into final.”

Commenter B, however, gives precedence to the symbolic function of language selection over the denotative value of the message. Equating English only with globalization is contested by Commenter C, who questions the interpretation of globalization with English as the sole language of international communication.

Thread II

Commenter A
Generally it's the songs in English that win. So from a competitive standpoint, it's better to sing in English. Also, if one wishes to be an international singer...yeah. If you think this song sucks, you should look at the winners from the past ten years...

Commenter B
what is this international singer bullshit that you have to sing in english to be popular overseas? ...its just a bunch of BS the globalists want you to think so they can make English the only language on the planet

Commenter C
Most bands that sing in English do better internationally. I am not saying this is a good thing, just a fact. It isn't made up by globalists. By the by, do you even know what that term means? It does not at all mean making everything English speaking.

Commenter A
I never said "overseas," and I never said "you have to sing in English." Many popular bands that span over Europe are English singing, and yes, most do...not all.

While some commenters do not substantiate the claim that one should sing in one’s native/first language, the majority of discussants propose some rationale regarding the language choice in the ESC, but also in general. These rationales can be categorized as affective/evaluative, instrumental, pragmatic, or symbolic.
**Français versus Anglais**

Compared to the postings in English, the French language postings tend to be more about the position of French relative to that of English and less about the overall linguistic diversity in the ESC and the global linguistic ecology (cf. Mühlhäusler, 1995). The most frequent collocate with *chanter en* (to sing in) is *anglais* (63.8%), followed by *français* (26.4%) (Table 4). The two lemmas overwhelmingly co-occur in the expression *chanter en*. This is in contrast to the English equivalent of the same phrase, *sing in*, which frequently co-occurs with the expressions *native language/own language/foreign language*, as shown in Table 4.

Of 72 comments, 46 (64%) were classified by commenters as mostly negative towards the attitude ‘one should sing in English as opposed to French,’ eight as mostly positive (11%), and 18 as ambivalent or factual (25%). The following types of arguments for or against the attitude that ‘one should sing in French’ were identified in the concordances of *chanter en* (Figure 3), almost all collocating either with *anglais* or *français*.

*Affective/evaluative function arguments.* It is a shame (*honte; dommage*) for France to be represented by a song in English (Figure 3, rows 29, 54 (*honte*); row 39 (*dommage*)). In SC_F, the word *honte* occurs 62 times, 37 times with reference to language choice.

*Instrumental function arguments.* Singing in English puts everyone on equal footing (Figure 3, row 45), but may also be advantageous compared to singing in other languages, including French (Figure 3, row 58). One commenter poses the dilemma of whether to sing in English and win, or sing in French and lose, “You gotta know whether people prefer a guy who sings in French without any chance of winning or a guy who sings in English ([the] international language) and who for once has a chance to win” (French original in Figure 3, row 69).

*Pragmatic function arguments.* According to one commenter, ”it is normal for Sebastian to sing in English since everyone does it” (Figure 3, row 30). Moreover, English is “an easy language to learn” and facilitates communication among different nations (Thread III, Commenter C).

*Symbolic function arguments.* Performers who represent France should sing in French, not in English (Figure 3: 35, 50, 51, 62), because not being represented by the country’s national language may be perceived as abandonment of French on account of English (Figure 3, rows 29, 32, 46) and loss of the ethnolinguistic identity of the French (Thread III, Commenters A & D).

In Thread III, Commenter A equates cultural diversity with the “return” of French in the ESC. The argument is only partially supported by Commenter B, who advocates the return of national languages and stresses the importance of the performers’ artistic qualities.

**Thread III**

*Commenter A*

La langue française regagnera le jour où la diversité culturelle sera respectée. quand les abrutis cesseront de chanter en anglais et que la langue sera réintroduite comme LA langue du grand prix européen de la chanson.

‘The French language will again see the light of day when cultural diversity is respected.'
When the morons stop singing in English and when the [French] language is reintroduced as THE language of the ESC.’

Commenter B
… je suis tout à fait d'accord avec vous. En tant que fan du Concours Eurovision depuis plus de 15 ans, je remarque que la culture anglaise touche tous les pays. Je suis pour le retour des langues nationales, de l'orchestre et surtout des ARTISTES à l'Eurovision.

‘I completely agree with you. As a fan of the Eurovision Contest for more than 15 years, I think that Anglophone culture influences all countries. I am for the return to Eurovision of national languages, the orchestra, and especially ARTISTS.’

Commenter C
... le Français c'est pas la langue des pays de l'est non plus. Donc a par chauvinisme (orgueil) a la con aucun intérêt d'utiliser le Français plus que l'Anglais. De plus l'Anglais s'est imposé en tant que langue internationale, c'est une langue facile a apprendre et très pratique (on peut communiquer avec n'importe quelle nationalité grâce a l'Anglais) personne ne peut remettre ça en cause.

‘... French isn’t the language of the eastern countries, either. So chauvinism (pride) stupidly is not a good reason to use French more so than English. Besides, English has imposed itself as an international language, it is an easy language to learn and very practical (you can communicate with someone of any nationality thanks to English) no one can deny that.’

Commenter D
Une chanson en anglais pour représenter la France ! Que veut-on donner de nous comme image? Celle d'un pays conquis, colonisé, un pays qui a renoncé à son identité ? Notre langue n'est pas qu'à nous, elle est présente sur les cinq continents ! Le pays principal de la francophonie donnant l'image de l'abandon de sa langue? Quelle honte!

‘A song in English to represent France! What kind of image are we trying to create of ourselves? An image of a conquered, colonized country, which has renounced its identity? Our language is not just ours, but it is found on all five continents! Is the principal country of the francophone world sending the signal that it is abandoning its language? How shameful!’
In Thread III, Commenter C adopts a pragmatic stance in relation to the use of French as opposed to English. Commenter C emphasizes what she or he perceives as relevant factors. First, French has lost its privileged position in some parts of the world. Second, English is an easy language to learn and has already established itself as a lingua franca. Commenter D, in contrast, represents a voice of desperation and blame. S/he deplores the neglect of French, claiming that the loss of linguistic prestige is directly associated with the loss of identity as well as France’s cultural and political advantage as the leader of the francophone world.

Besides Andorra, Spain, Portugal, and Serbia, France is one of the few countries that either does not permit or typically does not send songs in English to the ESC. However, in 2008 France sent an entry in English with only a few words in French, which provoked a debate that reached as far as the French Parliament. According to the ESC portal ESC Today, Mr. Gonnot from the ruling party stated that “[my] fellow [French] citizens don’t understand why France is giving up defending its language in front of hundreds of millions of television viewers around the world.”

Marc Teissier du Cros, the song producer, replied that the singer may sing “in English, French and even in Italian according to his inspiration.” He went on to say that in 2008, “half the Eurovision candidates were singing in English,” expressing doubt that “singing in French is the best way to make oneself understood by the whole world.” Gonnot’s stance supports the
symbolic function argument and is reinforced by the majority of opinions expressed in the SC_F sub-corpus, which advocate the use of French in the ESC as a symbol of the country, its culture, and Francophonie. The stance of Teissier du Cros, in contrast, represents a language-choice-as-instrument position, arguing that what matters is having a chance to win, whereby French is treated equally with any other language. Ironically, despite sending an entry in English that year, France finished 18th out of 25 finalists.

Attitudes towards Lena Meyer’s Accent

In the German sub-corpus (SC_G), collected from the comments on Lena Meyer’s 2010 English-language performance of the song Satellite in Moscow, the lemma ACCENT is the fifth most frequent non-grammatical word after SONG, LENA, LOVE, and LIKE. ACCENT in the singular form occurs as many as 737 times (see distribution, Figure 1.1). If we add the German form AKZENT (14), Slavic AKCENT (8), and creatively spelled variants, the conceptual lemma /ACCENT/ occurs more than 790 times (Table 5) in the sub-corpus, in different languages and spelling/misspelled/mistyped variants. In almost all the instances, the lemma refers to Meyer’s accent in her English-language performance of the song Satellite.

<table>
<thead>
<tr>
<th>N</th>
<th>Word</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>75</td>
<td>Accent</td>
<td>1</td>
</tr>
<tr>
<td>76</td>
<td>Accenct</td>
<td>2</td>
</tr>
<tr>
<td>77</td>
<td>Accent</td>
<td>737</td>
</tr>
<tr>
<td>78</td>
<td>Accenti</td>
<td>1</td>
</tr>
<tr>
<td>79</td>
<td>Accento</td>
<td>2</td>
</tr>
<tr>
<td>80</td>
<td>Accents</td>
<td>18</td>
</tr>
<tr>
<td>81</td>
<td>Accent’s</td>
<td>2</td>
</tr>
<tr>
<td>82</td>
<td>Accentt</td>
<td>1</td>
</tr>
<tr>
<td>83</td>
<td>Accentu</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 5. Word frequency of the conceptual lemma /ACCENT/ in SC_G

In response to this performance, to illustrate the range of attitudes towards language and accent and the ratio between positive and negative attitudes, a semantic differential scale was used. A total of 248 collocates with the English word accent were retrieved based on the following
collocational parameters: 1) Her accent is + attribute; 2) Attribute + accent; 3) I + verb + accent; and 4) Her accent + verb (excluding to be) (Table 6).

<table>
<thead>
<tr>
<th>Very Positive</th>
<th>Positive</th>
<th>Positive to ambivalent</th>
<th>Ambivalent to negative</th>
<th>Negative</th>
<th>Very Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Great (6)</td>
<td>Cute (22)</td>
<td>Natural (3)</td>
<td>Funny (7)</td>
<td>Strong (2)</td>
<td>Weird (20)</td>
</tr>
<tr>
<td>Adorable (4)</td>
<td>Sexy (8)</td>
<td>Nice (2)</td>
<td>Different (5)</td>
<td>Cockney (1)</td>
<td>Fake (13)</td>
</tr>
<tr>
<td>Amazing (2)</td>
<td>Cool (6)</td>
<td>Unique (2)</td>
<td>Interesting (2)</td>
<td>Forced (1)</td>
<td>Cockney (8)</td>
</tr>
<tr>
<td>Awesome (1)</td>
<td>Funny (5)</td>
<td>Cockney (1)</td>
<td>Mix/mixed/</td>
<td>Funny (1)</td>
<td>Bad (5)</td>
</tr>
<tr>
<td>Beautiful (1)</td>
<td>Nice (4)</td>
<td>Decent (1)</td>
<td>Mixture (4)</td>
<td>Heavy (1)</td>
<td>Strong (2)</td>
</tr>
<tr>
<td>Brilliant (1)</td>
<td>Cockney (2)</td>
<td>Funny (1)</td>
<td>Cockney (2)</td>
<td>Stereotypical (1)</td>
<td>Wannabe (2)</td>
</tr>
<tr>
<td>Enchanting (1)</td>
<td>Good (3)</td>
<td>Retarded (1)</td>
<td>Funny (1)</td>
<td>Weird (1)</td>
<td>Broken (1)</td>
</tr>
<tr>
<td>Fabulous (1)</td>
<td>Best fake (1)</td>
<td></td>
<td>Germanesque (1)</td>
<td>Mix (1)</td>
<td>Mix (1)</td>
</tr>
<tr>
<td>Gorgeous (1)</td>
<td>Best thing</td>
<td></td>
<td></td>
<td>Funny (1)</td>
<td>(fkn)</td>
</tr>
<tr>
<td>Fucking (1)</td>
<td>song (1)</td>
<td></td>
<td></td>
<td>Hilarious (1)</td>
<td></td>
</tr>
<tr>
<td>Hot (1)</td>
<td>Mix (1)</td>
<td></td>
<td></td>
<td>(a little)</td>
<td>Irritating (1)</td>
</tr>
<tr>
<td>Lovable (1)</td>
<td>Sexy weird</td>
<td></td>
<td></td>
<td>(a bit)</td>
<td>Ridiculous (1)</td>
</tr>
<tr>
<td>Lovely (1)</td>
<td>(1)</td>
<td></td>
<td></td>
<td>Odd (1)</td>
<td>Don’t like (2)</td>
</tr>
<tr>
<td>Orgasmic (1)</td>
<td>Sweet (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Love (32)</td>
<td>Like (10)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

55*3=165 65*2=130 11*1=11 22*0=0 8*-1=-8 58*-2=-116 29*-3=-87

Total Positive Score = 306 Total Negative Score = -211

Parameters: 1) Accent is + attribute e.g., Her accent is funny; 2) Attribute + accent e.g., She has a funny accent; 3) I + [verb] + accent. e.g., I hate her accent. 4) Her accent + [verb]. e.g., Her accent sucks.

Table 6. Semantic differential scale of attitude collocations and number of occurrences in SC_G

As shown in Table 6, there are 131 predominantly positive and 95 predominantly negative ‘attitudinal indices’ – attributes indexing a particular attitude – in the sub-corpus, while 23 were rated as ambivalent. Positive attitudes also score better overall (306 points), with a considerable number of negative comments (-211 points), accounting for the positive versus negative ratio ≅1.45:1 (i.e., 306:211). Some attributes may be classified in more than one category (e.g., Cockney, fucking, mix), depending on the context. For instance, the adjective-adverb fucking occurs once in a very positive sense as an intensifier of the adjective adorable (Figure 4, row 2), and three times in a very negative sense (Figure 4, rows 1, 3, 4), in her fucking accent and a fucking bad accent. Usage of some attributes may arguably be interpreted both as positive and negative. One commenter, for example, refers to Lena’s accent as “stupidly retarded, yet stupidly good” – a contradiction in terms.

Figure 4. Concordance of fucking accent
Lena’s accent arouses controversy. On the one hand, some YouTube commenters are unfavourably disposed towards her accent “because she sounds like an English villager” and “because she is faking [the accent].” Some, on the other hand, approve of the accent “because it makes the song funny” and “adds something special to the song.” One commenter argues that “[Lena’s song] is popular only because of her funny accent.” According to another commenter, the accent is not only “the best thing about the song, but it makes the song tolerable.” The commenter goes on to say, “I love the accent. The song is so bad that it would not work without the mispronunciation,” possibly referring to the singer’s mixing of different phonological systems. Lena pronounces inconsistently ‘day’ as /dai/ but ‘way’ as /wei/ in two lines that otherwise rhyme, provoking one commenter to write: “I love you Lena but your accent's awful. It's not ‘die’, it's ‘day’. I cannot understand the meaning of lyrics. It's really confusing.” The pronunciation of ‘day’ as /dai/ prompted another German-speaking commenter to discuss the origin of the variation in German. S/he notes (English translation follows in italics): “‘die’ für ‘day’ ist die Aussprache in einigen britischen Regionen, z.B. Manchester, der Dialekt dort heißt Mancunian” (“in some parts of Britain, e.g., in Manchester, whose dialect is called Mancunian, ‘day’ is pronounced ‘die’”).

A number of commenters are nonetheless intrigued by the unconventional mélange of different accents that Lena puts together in her performance of the song Satellite. Lena’s accent is imaginatively described as “a mixture between American and German, stupidly retarded yet stupidly good,” “a fucked up mess of accents,” “sorta Björk-but-not-really-Icelandic-maybe-not-even-human accent,” “Irish-Scottish-Austro-like,” “Germanesque,” “a weird mix of German and Cockney (London),” and “a Cockney/German mix,” depending on the apparently unanswered and unanswerable question of whether Lena was doing what she did consciously, accidentally, or because she could do no better.4

The word Cockney occurs 17 times in the sub-corpus (Figure 5), in all instances referring to Lena’s accent. According to the raters for this study, comments on Lena’s accent range from positive (Figure 5, rows 6, 7) to positive to ambivalent (Figure 5: 13) to ambivalent (Figure 5, rows 12, 15) to ambivalent to negative (Figure 5, row 4) to negative (Figure 5, rows 1, 3, 5, 8, 9, 10, 11, 14) to very negative (Figure 5, row 2). Two propositions were rated as factual (Figure 5, rows 16, 17). According to Coolbrands, a British opinion canvassing enterprise, in a 2008 survey of “cool” accents, Cockney accent ranked fourth/fifth out of 12 accents used in the UK,5 confirming that attitudes towards Cockney accent are mixed.

Lena’s accent is judged by some commenters as pretentious, that she is trying to put on a Cockney accent (Figure 5, rows 4, 10-12), “exaggerated, but still great” (Figure 5, row 15), as a fake Cockney accent (Figure 5, row 6), and even as Mockney (Mock Cockney) (a portmanteau of ‘mock’ and ‘Cockney’) or “an exaggerated version of lower class English accent” (Figure 5, row 8). Three commenters argue that Lena’s accent is just an imitation – a fake voice, apparently sounding like that of Lily Allen (Figure 5, rows 6, 13) or Kate Nash (Figure 5, row 13), both singer-songwriters who are known for using Cockney accent in their performances.
One can find evidence of similar attitudes in contexts other than these YouTube comments. For example, a Spiegel International journalist describes, not without malice, Lena’s accent as a “Swedish therapist imitating Ali G.” He goes on to say,

Lena’s accent isn’t a mockney accent, the affection of London’s working-class Cockney tone that the like of Blur’s Damon Albarn were accused of using. Nor is it the full-on “jafakean,” the fake Jamaican accent you often hear on the top decks of North London buses, as the preferred slang of the school kids who like to sound like they’re from the ghetto. Instead, it is a mixture that borrows from the two, then adds a shot of mixed-up European, presumably made up of the native German and what sounds like Scandinavian.

It is interesting that the German singer’s accent is compared primarily to the native British speech of British singers rather than to a stereotypical German accent. What seems to provoke some YouTubers and journalists is the accent mixing, which is interpreted as a presumptuous mix of a “double” fake (fake Jamaican from North London) accent seasoned with a whole new mélange of German and Nordic accents. The attitude in the Spiegel column towards Lena’s accent is negative and derogatory. Apparently, what is under attack is not her mockney accent, which is a native British accent, but the fact that she mixes different semiotic resources into a single prosody, prompting another commenter to wonder: “What is up with her accent? It sounds pretty unnatural like its wierd mix of english and german.” While accent mixing and resemblance to other languages and accents are features that generate a considerable number of negative comments, it is likely that along side other factors, such as her appearance, song quality, and voice, Lena’s “intriguing” accent may have also contributed to her winning the Contest in Moscow.

Although the insights from the sub-corpus are specific to this case study, commenters generally seem to be sensitive to foreign accents and pronunciations, as the Russian sub-corpus also indicates: In ESC_R, there are 32 instances of the lemma ‘impassible’ [sic], including variations such as ‘impAssible,’ ‘impaaasible,’ etc, with reference to the Russian singer’s non-native pronunciation of ‘o’ as ‘a,’ apparently in an attempt to sound “more American.”
Conclusion

In this article, it was demonstrated that CBA can provide a useful method for studying language attitudes expressed in web-based social media. Spontaneously produced, naturalistic data from commentaries on ESC YouTube video uploads were compiled in a predominantly bilingual, English and French, corpus. Three case studies were conducted: One examined language choice in general (GENERAL_ESC), while the other two studies focused on the use of French (SC_F) and accent-mixing (SC_G) in different sub-corpora. In particular, the first case study examined language attitudes towards singing in English versus singing in one’s native language. The second case study analyzed comments in French with a focus on attitudes towards singing in English versus singing in French. The third case study examined attitudes towards the accent of the 2010 winner, then the 18-year old German singer Lena Meyer, who had performed the song Satellite in English. The frequency, distribution, and concordancing results of the corpus-based analysis were triangulated with online media views, that is, the opinions of a BBC commentator, a British opinion canvassing agency, a French politician, and a German columnist, to shed additional light on the findings.

The analysis confirmed the observation that language is a frequent topic in all six corpora, as well as that positive attitudes towards singing in one’s native language prevail, addressing RQ1 & RQ2, respectively. Compared to the postings in English, however, the French language postings tend to be more about the status of French relative to that of English and less about overall linguistic diversity in the ESC and the global linguistic ecology. In both case studies, mainly negative attitudes towards the use of English as a non-native language prevail. In the comments, language attitudes were explained by arguments that support affective/evaluative, instrumental, pragmatic, or symbolic functions of language choice.

The third case study showed that the performance of the German 2010 winning song elicited mixed positive/negative responses and provoked much discussion and debate among YouTubers, as the results of the frequency, keyness, and distribution of the lemma ACCENT confirm. Lena Myer’s performance appeared controversial to commenters because it was delivered in what appeared to be a mélange of standard British English, a “Cockney-like” accent, and a Germanic accent. The insights from the online media reviews support the corpus analysis findings, showing that language choice is potentially a political issue, too. Illustrating this point is the debate on the use of French versus English in the ESC among French politicians, as described in the portal ESC Today.

As these case studies demonstrate, the issue of language choice is not uncontroversial for ESC language policy decision makers and those in charge of the national selections. The linguistic conundrum for the language planners is whether to give precedence to the language-choice-as-an-instrument or language-choice-as-a-symbol argument. Hence, it is no surprise that since 1956, the language choice rule has changed several times. Between 1966 and 1972 and between 1977 and 1998, language choice was restricted to the official languages of participating countries. As of 2012, there is no such restriction, and the majority of countries choose to send songs in English. For example, between 1999 and 2011, Germany sent nine songs to the contest in English and only four in German, its national language. Nonetheless, in 2007 in Helsinki, an entry performed entirely in Serbian won, praised by many for the quality of performance despite (or because of?) the language choice. Moreover, as social psychology research into attitudes
attests, attitudes do not always reflect behaviour (Walters, 2007); that is, an overtly expressed attitude supporting the use of national languages does not necessarily result in a vote, even if the song in question is liked.

Acknowledgments

I am grateful to two Language@Internet reviewers for their helpful feedback, and especially to Susan Herring and Keith Walters for their insightful, generous, and detailed critique, feedback, and suggestions. I also thank Ljiljana Vuletić for her comments on an earlier version of the article.

Notes

1. Ivković (2012) describes the virtual linguistic landscape as “visibility and salience of linguistic items and other semiotic markers delineating ethnolinguistic presence and indexing power relations in cyberspace-as-the-public-sphere” (p. 10).
4. I am indebted to Keith Walters for this observation.
6. Alias Sacha Baron Cohen, who portrays the fictional Austrian gay character in the movie Brüno.
7. A British band and its frontman from the mid-1990s, often cited as a prime example of a Britpop band that uses the Mockney accent.

References


Biographical Note

Dejan Ivković [dejan.lj.ivkovic@gmail.com] received a Ph.D. from York University in 2012. His main research interests are in multilingualism and computer-mediated communication.