Journal of the Society for American Music (2015), Volume 9, Number 4, pp. 445–454.
 doi:10.1017/S175219631500036X

AmeriGrove II: Perspectives and Assessments

The Grove Dictionary of American Music. 2nd ed. Edited by Charles Hiroshi Garrett. Oxford and New York: Oxford University Press, 2013.

Introduction

CHRISTINA BAADE and EMILY GALLOMAZZEI

Cue: Lalo Schifrin's "Theme from *Mission: Impossible*." If the return of Tom Cruise, explosions, and fast-paced chases to the screen this summer is any indication, Americans (and the global audience for Hollywood action film) love an impossible mission. Certainly, U.S. music scholars do. Reading the eight reviews of the second edition of *The Grove Dictionary of American Music* (hereafter, *AmeriGrove II*), edited by Charles Garrett (with a large and distinguished editorial team and nearly fifteen hundred contributors), as well as dipping frequently into its entries, we were deeply impressed both by the quality and ambition of the eight-volume, 5.4 million-word encyclopedia and by the "gargantuan" task (as Leta Miller puts it) with which the reviewers had been charged.

We will focus our introduction upon the reviews themselves, asking (to rephrase from Richard Crawford), "How does one read *AmeriGrove II*?"[1] John Koegel, Christina Baade's predecessor as book reviews editor for *JSAM*, envisioned this multipart review of *AmeriGrove II* and recruited eight intrepid scholars to each consider a distinct topic, to which they all brought their own perspectives, strategies, and voices. The reviewers and topics are as follows: Berndt Ostendorf and Wolfgang Rathert, overview, European perspective; John Graziano, overview, U.S. perspective; Glenda Goodman, music before 1800; Douglas Shadle, nineteenth-century music; Kip Lornell, folk and traditional music; Leta Miller, twentieth-century art music; Sherrie Tucker, popular music and jazz, 1900–50; and Theo Cateforis, popular music and jazz, 1950–present.[2]

Reading their reviews was an exhilarating experience that helped us better appreciate the achievement of *AmeriGrove II* in encompassing the growth and changes in "American" music studies, allowing us to reflect on the current state of the field, particularly in relation to big questions of belonging: Who are the "we" who study American music? Who and what is included? How do we decide? These are challenging questions indeed, given how many of us (readers and reviewers) have

[1] Richard Crawford, "Amerigrove's Pedigree: On *The New Grove Dictionary of American Music*," *College Music Symposium* 27 (1987): 174–75, quoted by Berndt Ostendorf and Wolfgang Rathert in their review of *AmeriGrove II* in this issue.

[2] Many thanks to John Koegel for envisioning this multipart review and engaging in the persuasive tour de force involved in recruiting this outstanding team of reviewers. Thanks also to Karen Ahlquist for arranging to publish all of the reviews in a single issue. Finally, our deep appreciation to all the reviewers for taking on this tremendous task and carrying it out so well!

contributed to the volume and our temporal and personal closeness to the scholarly conversations that shaped it.

Few reviewers could resist praising *AmeriGrove II*'s scope or remarking upon its size; all of them engaged critically with the dictionary as a document showing the development of the field of U.S. music studies and its central concerns. Glenda Goodman, noting the ideological nature of such projects, linked *AmeriGrove* to Diderot and d'Alembert's *Encyclopédie* (1770–77), a project that "attempted to map the world of knowledge," and Noah Webster's *Dictionary of the American Language*, a lexicographical and nation-building endeavor. Indeed, music studies is one of the few disciplines to have sustained this Enlightenment commitment to cataloguing its field. With a family of *Grove* resources available at the stroke of a key (at least for those with access to good academic libraries), exhaustiveness was not the goal for *Amerigrove II*. It is explicitly a selective and nationally bounded resource: "a repository of historically significant information" related to American music (i.e., "musical life and cultures of the region now covered by the fifty states, the District of Columbia, and US territories," viii, vii). For reviewers, this selectivity was both a strength—Theo Cateforis praised the ways in which scholars' voices and interpretations made their way into the entries—and an invitation to critique and engagement. *AmeriGrove II*, like the first edition (hereafter, *AmeriGrove I*), edited by H. Wiley Hitchcock and Stanley Sadie and published in 1986, is less a map than a self-portrait of our field, inviting us all to greater reflexivity about what it is we study and why.

To arrive at their analyses, each reviewer had to read *AmeriGrove II* quite differently from the ways most "users" usually read the encyclopedia. Most users will likely encounter *AmeriGrove II* online, entering search terms into the *Grove Music Online* interface to find specific entries to answer a quick factual question or get an overview of research in a given area. By contrast, the reviewers all focused on the standalone version of *Amerigrove II*, in hard copy or in PDF, although some also referred to the online version. Of course, it was unrealistic to read and comment upon the entire encyclopedia. We were fascinated by the multiple ways in which the reviewers engaged with this "impossible" text, drawing out important themes while also negotiating its sheer size: they sampled, browsed, followed thematic threads, and drew comparisons between the new edition and the old (as well as with other references in the *Grove* family). They attended not only to what was covered but also to how many words were assigned to a given topic, suggesting that assigned word count was an important indicator of perceived significance. They also read a great deal, closely and critically. Many focused their engagement upon the wealth of essay-length entries surveying theoretical concepts, major genres, social categories, important historical events, and cities—offering insights both compelling and expertly informed.

In the rest of this introduction, we would like to bring one more reading strategy to *AmeriGrove II*: that of text analysis using a digital humanities tool, Voyant, developed by Geoffrey Rockwell and Stéfan Sinclair.[3] If our expert reviewers could

[3] Stéfan Sinclair, Geoffrey Rockwell, and the Voyant Tools Team, Voyant Tools (2012) (web application), http://docs.voyant-tools.org. Our ability to carry out this analysis was enabled by

https://doi.org/10.1017/S175219631500036X Published online by Cambridge University Press

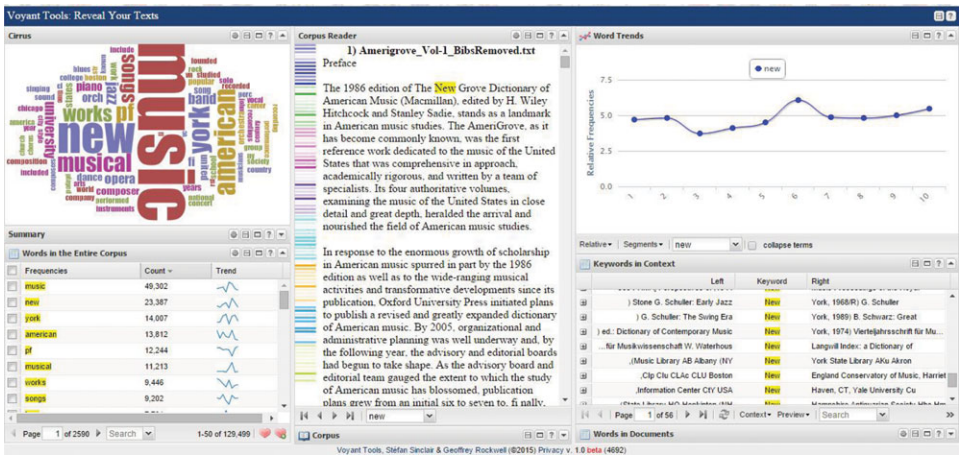

Figure 1. Screenshot of Voyant Tools interface with *AmeriGrove II* corpus. (Stéfan Sinclair, Geoffrey Rockwell, and the Voyant Tools Team, Voyant Tools [2012], http://docs.voyant-tools.org).

not read the entire *AmeriGrove*, the Voyant tool, designed specifically to assist in analyzing large corpuses, could do so with ease. Impossible Mission solved? Not so fast. As Stephen Ramsay asserts in *Reading Machines*, the point of text analysis is less to bring empirical, scientific fact to humanities inquiry and more to open up new interpretive possibilities: "Text analysis, because it allows navigation of the unread . . . [is] capable of presenting the bare, trivial truths of textuality in a way that allows connection with other narratives—in particular, those narratives that seek to install the text into a network of critical activity."[4] Thus inspired, we set about putting our quantitative findings with Voyant into dialogue with the critical observations of the reviewers.

Before we discuss our findings, a few more words about Voyant, the methods we used, and the strengths and limitations of text analysis are in order. Voyant, like other popular digital humanities tools, produces word frequency tabulations, tracks word trends, allows users to investigate how keywords are used in context, and provides a range of visualization options, among other possibilities (Figure 1). The key to sharing results from these projects is visualization. One type of visualization Voyant generates is the word cloud. Word clouds offer a visually striking way into texts, but they are also imprecise and methodologically opaque (Figures 2 and 3). For example, the two word clouds pictured here are the same size, although in a more accurate representation, *Amerigrove II* would be twice as big as *Amerigrove I*; further, they do not reveal how the text was manipulated to produce the results. More useful, we found, were the word frequency lists generated for each edition,

Anna-Lise Santella of Oxford University Press, who facilitated our access to text files for both editions of *AmeriGrove*. We also received significant support in facilities and expertise from the Lewis & Ruth Sherman Centre for Digital Scholarship at McMaster University and its administrative director, Dale Askey. A special thank you to Dr. Paige Morgan, a postdoctoral fellow at the centre, whose expertise and helpfulness were invaluable throughout our research process.

[4] Stephen Ramsay, *Reading Machines: Toward an Algorithmic Criticism* (Champaign: University of Illinois Press, 2011), 16–17, 78–79. Many thanks to Dr. Jennifer Askey for recommending this book.

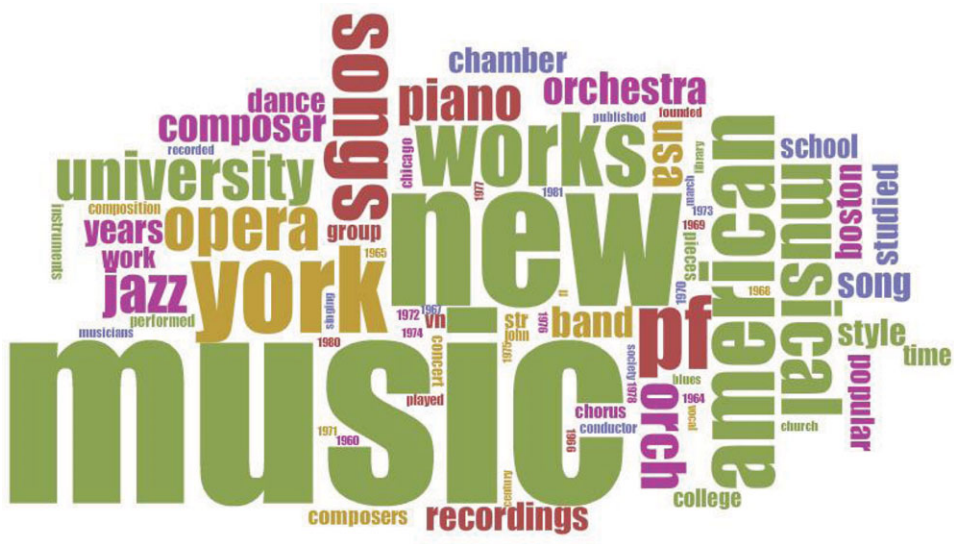

Figure 2. Word cloud: most frequent terms in *Amerigrove I.* Generated by Voyant Tools.

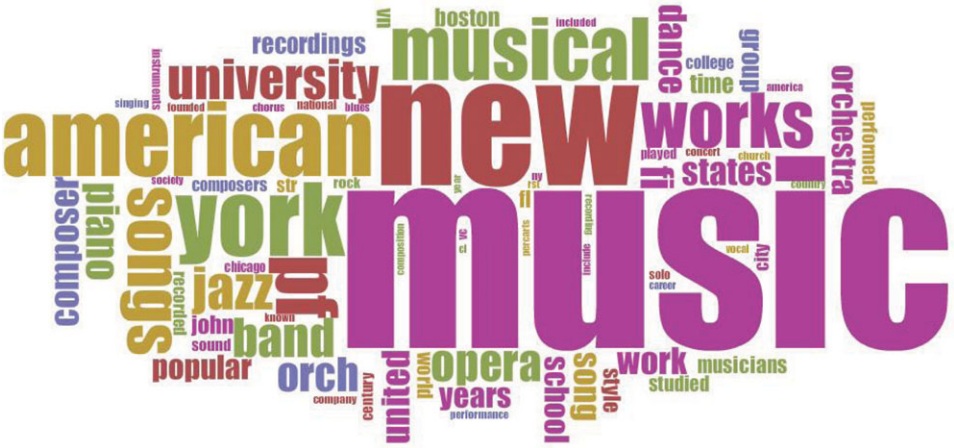

Figure 3. Word cloud: most frequent terms in *Amerigrove II.* Generated by Voyant Tools.

which we were able to analyze—not only for a word's frequency (i.e., its number of occurrences) relative to that of other words within each edition but also for how a word's frequency differed between the editions.

To determine whether the trends (i.e., the word frequency differences between editions) we discovered were significant, we employed a keyword significance test often used by digital humanities scholars and corpus linguists: *the log likelihood test* (LL). Although generating the word frequency lists was a useful starting point, their values need to be normalized before they can add to our discussion: in this case, the LL test compares the frequency of a word in two corpuses (*Amerigrove I* and *II*) against their total word count to give a relative representation of the word in both.

https://doi.org/10.1017/S175219631500036X Published online by Cambridge University Press

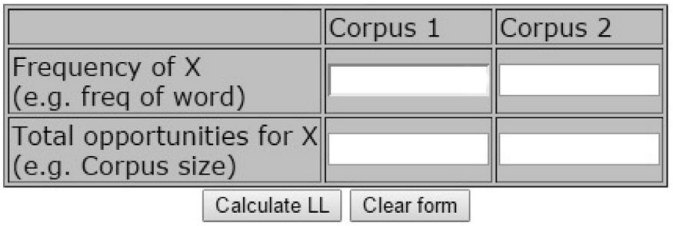

Figure 4. Screenshot of log likelihood calculator (from Tony McEnery and Andrew Hardie, "Statistics in Corpus Linguistics," Support Website for *Corpus Linguistics: Method, Theory and Practice* [Cambridge: Cambridge University Press, 2012]. http://corpora.lancs.ac.uk/clmtp/2-stat.php).

To form a basis of comparison, the text from both editions of *Amerigrove* was analyzed using Voyant. Before entering each corpus into the Voyant system, the text had to be reworked to suit our purposes. For example, the bibliographies accompanying entries were removed. This information would not tell us much about *how* this new edition talks about U.S. music and could even skew the final results gathered, particularly with regard to publication details like cities, presses, and dates. Once the text was entered, we applied a standard English "stop words" list in order to exclude common words like "the," "and," and "in" from our results. Applying this stop words list ensured that the "top terms" explored in this introduction were representative of musical genres, practices, and populations. The LL calculator (Figure 4) then compared the frequency of the top fifty words within the corpus of *Amerigrove I* vs. *II* to generate the log likelihood, which shows whether the difference in frequency can be considered significant (Table 1).

We found performing word frequency analysis a useful point of entry into these two large corpuses, especially when considered in relation to the reviewers' in-depth explorations. Given the importance the reviewers ascribed to the size of entries, we found the ability to tabulate the number of words throughout both editions a particularly compelling application of the tool. However, the *context* of terms—especially terms used in multiple ways—was harder to decipher. Although many of our findings correlated with themes identified by the reviewers, we also found many intriguing results we were unable to fully explain given our constraints of time and expertise.

One of our most striking findings is also obvious from a glance at the word clouds: that both editions had a similar "most used" vocabulary after the stop words were removed. They included obvious choices like "music" (and "musical"), "american" (and "usa" in *Amerigrove 1*, "united" and "states" in *Amerigrove II*), "song" (and "songs"); the intriguingly prominent "university"; and words linked to chronological accounts of genres and lives: "became," "early," and "later." Given that we retained works and recordings lists (even though we stripped the bibliographies), the prominence of "works" is unsurprising, while the high ranking of "pf" (and "piano") leaves no doubt as to the centrality of the piano in a range of musical genres. Investigating the ranks of "new" and "york" shows "New York" to be second only to "music" on both word frequency lists, confirming the city's preeminence in U.S. musical life. Finally, the almost geometric increase in the references (by

https://doi.org/10.1017/S175219631500036X Published online by Cambridge University Press

Table 1. Top fifty words in *AmeriGrove I* and *II* (after English stop words list applied) with log likelihood. The log likelihood number expresses the difference in word frequency between the two editions. The higher the number, the more statistically significant the difference. Generated with Voyant Tools. Log likelihoods calculated with McEnery and Hardie's tool.

AmeriGrove I		*AmeriGrove II*		More or Less Frequent in *Amerigrove I**	Log Likelihood*
Term	Frequency	Term	Frequency		
music	22,490	music	49,302	Less	110.36
new	12,298	new	23,387	More	27.00
york	7,786	york	14,007	More	63.82
pf	6,965	american	13,812	Less	12.54
american	6,499	pf	12,244	More	81.83
works	6,165	musical	11,213	Less	26.14
songs	5,521	works	9,446	More	274.15
musical	5,106	songs	9,202	More	122.37
opera	4,421	jazz	7,701	More	2.44**
university	4,348	university	7,479	More	67.98
jazz	3,940	early	6,784	More	0.75**
orch	3,932	band	6,721	Less	66.47
became	3,586	became	6,479	More	27.26
piano	3,578	opera	6,404	More	278.38
early	3,428	orch	6,291	More	126.4
usa	3,109	including	6,182	Less	122.57
orchestra	2,986	piano	6,098	More	62.39
composer	2,850	united	5,359	Less	3105.09
band	2,781	states	5,341	Less	2997.72
recordings	2,723	composer	5,215	More	17.02
later	2,690	dance	5,133	Less	9.69
song	2,690	later	5,063	More	8.16
style	2,647	orchestra	5,004	More	20.26
years	2,628	song	4,999	More	11.41
boston	2,602	work	4,952	Less	14.51
studied	2,581	began	4,890	Less	3.24**
chamber	2,474	popular	4,883	More	0.16**
popular	2,447	school	4,807	Less	0.41**
dance	2,358	years	4,789	More	17.03
including	2,358	group	4,347	Less	0.78**
school	2,347	fl	4,277	Less	52.31
began	2,319	recordings	4,229	More	110.09
work	2,231	time	4,128	More	9.93
time	2,226	style	4,014	More	126.18
str	2,132	boston	3,952	More	122.75
group	2,107	john	3,915	Less	4.56
college	2,085	musicians	3,810	Less	23.20
vn	2,007	vn	3,794	More	5.34
composers	2,004	str	3,744	More	25.42
pieces	1,916	studied	3,733	More	163.91
chorus	1,893	world	3,726	Less	151.26
concert	1,875	performed	3,702	More	0.07**
performed	1,851	city	3,683	Less	50.08
instruments	1,848	recorded	3,548	Less	3.55**
composition	1,841	composers	3,464	More	29.49
john	1,829	college	3,441	More	50.92
chicago	1,818	century	3,359	Less	6.99
conductor	1,810	rock	3,341	Less	139.89
played	1,795	sound	3,337	Less	141.23
1970	1,743	played	3,256	More	12.7

*Log likelihoods are given for words in the *AmeriGrove II* column.
**This log likelihood suggests that the difference in frequency between the two editions is *not* significant.

https://doi.org/10.1017/S175219631500036X Published online by Cambridge University Press

Table 2. Occurrences of century references in *AmeriGrove I* and *II*. Generated with Voyant Tools. Century references include dates and ordinals (e.g., 1886 and 19th).

Century	Aggregated Word Frequency *Amerigrove I*	Aggregated Word Frequency *Amerigrove II*
16th (1500s)	295	312
17th (1600s)	462	636
18th (1700s)	2690	3559
19th (1800s)	17862	25698
20th (1900s)	97071	164020

date and title) to each ensuing century is remarkable within both editions, as is the striking growth for all periods between editions. Goodman's and Douglas Shadle's observations that their centuries have been "repopulated" seems to be borne out by these numbers—as are their observations that coverage of the eighteenth and nineteenth centuries is dwarfed by coverage of the twentieth. In *AmeriGrove II*, there are 30,205 appearances of pre-twentieth century dates and titles, whereas twentieth century dates and titles *increased* from the first edition by more than *twice* that number (Table 2).

As the reviewers observe, this explosion of twentieth-century coverage responds to significant changes in the field, including the growth of research in non-classical genres, especially popular music; the musics of African Americans and other racialized people in the Americas; and broader questions of cultural hybridity, transnational movement, and power. The word frequency lists register how seriously the *AmeriGrove II* editorial team responded to these developments. A scan of the word frequency charts confirms that with words like "rock," "country," "guitar," "broadway," "gospel," "traditional," and even "band," "album," and "studio" making striking leaps in the rankings with large, statistically significant increases in representation.

Above all, however, "jazz" stands out as the most frequently occurring genre term in *AmeriGrove II* (with 7701 appearances), reflecting both the increased coverage of jazz in the new edition and the generous word counts for many entries, which, as Cateforis point out, often align with word counts assigned to canonical classical topics and people. A quick look also shows that "jazz" has surpassed "opera" as the top-ranked genre. However, this "switch" in position requires a closer investigation. The occurrences of both words has grown, but although the frequency with which "opera" appears in the second edition is significantly lower than in the first edition, the differences for "jazz" are not statistically significant. As Sherrie Tucker notes, the coverage of jazz is greater in the new edition, but it was also well represented in the first edition. Meanwhile, occurrences of "opera" have not increased at the same rate as have the popular terms listed above.

Where goes opera, so goes art music? Finding the answer using text analysis is more challenging than in the case of jazz or rock, for which single-word genre designations are widely used. One reason, as Leta Miller observes, is that it has been difficult to arrive at a satisfactory overarching term to describe "art" or "classical" music (both of these words are quite low in the frequency lists for both editions). A

https://doi.org/10.1017/S175219631500036X Published online by Cambridge University Press

Table 3. Occurrences of references to racial and ethnic groups in *AmeriGrove I* and *II* (preliminary). Generated with Voyant Tools.

Categorization	Frequency *Amerigrove I*	Frequency *Amerigrove II*	More or Less Frequent in *Amerigrove I*	Log Likelihood
African American[a]	2120	4984	Less	35.99
Asian[b]	235	825	Less	61.99
Native American[c]	1109	1725	More	44.34
Latino[d]	636	2873	Less	397.63
White[e]	1828	3062	More	38.50

[a]Represents primarily words incorporating "black," "African," and "Afro-."
[b]Represents words incorporating "Asian" and "Oriental."
[c]Represents words incorporating "Indian," "native," "First Nations," and "aboriginal." This count is incomplete: it does not include names of nations and peoples.
[d]Represents primarily words incorporating "Latin," "Latino/a," "Spanish-," "Mexican," and "Puerto Rican." This count is incomplete: it does not include a full list of nationalities.
[e]Represents words incorporating "white," "European," and "Caucasian."

second reason is that art music operates as an unmarked category in musicology: Van Cliburn is a "pianist," not a classical pianist, whereas Teddy Wilson is a "jazz pianist." Our workaround was to turn our attention to specifically classical genres and terminologies, such as "opera," "chamber," and "symphony." Two trends are observable when we compare *AmeriGrove I* and *II*. First, with the exception of "chamber," the count for each of these terms grew.[5] Second, all of these terms have a lower overall *rate* of representation in *AmeriGrove II*. At least at the quantitative level, *AmeriGrove II* has not stinted art music, but it has increased its representation of a number of other non-classical genres to represent a much wider swath of U.S. musical life.

What about other sorts of diversity, such as ethnicity, sexuality, and gender? The reviewers all praised the quality, scope, and theoretical sophistication of the entries on race, gender, sexuality, and ethnic groups and their musics, as well as the expansion of biographical coverage of women, queer, and racialized people. Our analysis confirmed that the occurrence of words used to describe African Americans, Asian Americans, Latina/os, and Native Americans increased in the second edition; the growth in the rate of occurrence was statistically significant for all of these groups (with the probable exception of Native Americans; Table 3). Because we had to identify and search for each individual term, using text analysis to investigate representation for racialized groups in publications dating from 1986 and 2013 demonstrated to us the degree to which preferred language has changed ("the bare, trivial truths of textuality," indeed). It also presented significant methodological challenges, particularly for Latina/os and Native Americans, groups often discussed in *AmeriGrove* with reference to their nations (e.g., "Mexican" and "Choctaw"). The politics of naming and representation are important, and we acknowledge that our preliminary analysis only scratched the surface. Several of the reviews attend to these questions with far more nuance.

[5] "Chamber" fell from 2474 occurrences in *AmeriGrove I* to 2195 in *AmeriGrove II*, but "symphony" grew from 1289 to 2395.

https://doi.org/10.1017/S175219631500036X Published online by Cambridge University Press

If tracking references to racialized groups was difficult, using text analysis to investigate the representation of white people in *AmeriGrove* was still more challenging because whiteness tends to be an unmarked category. Critical scholars, especially feminist, critical race, and queer theorists, have long argued that unmarked categories (such as maleness, whiteness, or heterosexuality) reinforce dominance by remaining "relatively invisible as an unstated, privileged norm," as Loren Kajikawa writes in his "Race and ethnicity" entry. Text analysis is blind to what is not named, thus references to African American and black identities outnumber references to European American and white identities within both editions. However, there was a numerical increase in the rate of occurrence for terms relating to white and European identities in *AmeriGrove II*—reflecting, perhaps, the impact on our field of whiteness studies, which seeks to make the dynamics of whiteness (and white privilege) visible and open to critical investigation.

When considering questions of visibility, the changes in our scholarly field and broader culture are perhaps most striking when comparing the representation of sexual and gender minorities documented between the two editions. "Homosexual," "homosexuals," "lesbianism," "transsexual," and "gay" in reference to gay men (in the Bette Midler entry) each appeared once in the 1986 *AmeriGrove I*. In *AmeriGrove II*, there are nearly five hundred occurrences of words relating to transgender, lesbian, gay, bisexual, and other queer identities. The musical and scholarly work that transformed these identities from that which was almost literally unspeakable into significant areas of research is charted in a number of *AmeriGrove II* subject entries, including "Lesbian, gay, bisexual, transgender and queer music," Nadine Hubbs's updating of Philip Brett and Elizabeth Wood's classic entry; "Transgender," by Stephan Pennington; and "Sex, sexuality" by Fred Maus.

Turning to a final set of marked and unmarked categories, "women" made significant gains, from 425 to 1494, although occurrences of the word still fell behind references to "john" (3915), "george" (2326), "william" (2052), "charles" (1886), "james" (1766), "paul" (1726), "david" (1700), "robert" (1538), and "thomas" (1501). A fairer comparison would perhaps be occurrences of "he" and "she." In *AmeriGrove I*, there were roughly 5.7 occurrences of "he" for every "she" (28,135 to 4912); in *AmeriGrove II*, there were roughly 4.5 (45,606 to 10,229). This change in ratio is due to a decreasing rate of occurrence for "he"; there was no statistically significant change in the rate of occurrence for "she," even though there was significant numerical growth (the number doubled). Of course, these raw numbers lack the nuanced analyses to be found in the reviews. It is also worth remembering that the numbers reflect our field and its histories, contexts, and concerns as much as they reflect *AmeriGrove II*, whose editorial team made attending to difference and power, as well as the inclusion of underrepresented groups, a priority—at which they succeeded in many ways.

If our rather basic text analysis offers insights into the shifting representations of genres, centuries, race, sexuality, and gender, it also complements our reviewers' sense that *AmeriGrove II* has become more attentive to border crossing, hybridity, non-compositional musicking, and sound. Consider the surprising entrance of "sound" to the second edition's top fifty words, the decreasing shares of "composer"

and "composition," and the increased rate of representation for "dance," "culture," "international," and even "radio."

Ultimately, our digital humanities adventure with Voyant gave us greater insight into these two monumental corpuses, the tool's colorful screen inviting us to explore these texts in new ways, to dig below the relatively crude analysis we have presented here, and to further interpret the lists it has produced. As our reviewers show, *AmeriGrove II* (and *I*) is a capacious, engaging, impossible text that rewards many sorts of reading—whether by human or machine, online or on paper. We hope you are engaged and challenged by the ways into *AmeriGrove* charted by these reviews, and hope they inspire you in your own impossible missions in American music.

https://doi.org/10.1017/S175219631500036X Published online by Cambridge University Press