

THE UNIVERSITY OF CHICAGO

Does Linguistic Discrimination Occur  
In US Music Audiences?

By

Madison Martin

June 2022

A paper submitted in partial fulfillment of the requirements for the  
Master of Arts degree in the  
Master of Arts Program in the Social Sciences

Faculty Advisor: Danielle Bolling

### **Abstract**

The present study experimentally investigated the impact of lyric language on listener experience and the likelihood for the participant to support the song in the future, to determine if linguistic discrimination occurs in the US popular music industry. Online survey participants ( $n = 360$ ) were randomly presented one of two versions of the same experimental song (with lyrics in English or lyrics in a non-English language), and asked questions about their listening experience. In addition to two experimental songs, participants were presented with three control songs. The language of the lyrics caused significant differences in listener experience for the first experimental song. The language identity of the participants (monolingual vs. bilingual) was also influential on listener experience measures, but only for the Spanish version of the first experimental song. The likelihood to support the Spanish version of the song by streaming it was significantly lower than for the English version, and this difference was not explained by differences in the other experience variables, demonstrating that linguistic discrimination does occur in US audiences. Finally, the lyric language manipulation in the second experimental song (with English vs. Korean lyric versions) was not successful, so further research that includes different lyric languages and more varied language backgrounds of the participants is needed, in order for the results to be generalizable to non-English languages other than Spanish.

## Does Linguistic Discrimination Occur in US Music Audiences?

### Introduction

#### Linguistic Discrimination

The United States has a long history of discrimination against the use of languages other than English, which continues to have implications in many contexts (Baron, 2019). While the term “linguistic discrimination” is sometimes used to refer to how individuals make sense of language by sorting speech into “meaningful elements, such as phonemes and morphemes”, in the current study, the phrase “linguistic discrimination” refers to denying an individual the right to equality in social outcomes based on the language they speak or the way in which they speak (Craft et al., 2020, p. 391).

Starting in childhood, humans attach stereotypes to the way others speak and these stereotypes can impact how we connect with others through our social preferences, like believing people with certain accents are smarter or kinder, or influencing who we want to be friends with (Kinzler & DeJesus, 2013). Even bilingual children of various language acquisition backgrounds prefer native-accented English speakers over accents that might be more familiar to them, like those of their parents (DeJesus et al., 2017). While the idea of what is a native accent or “standard American English” is contested by sociolinguists, it is generally agreed upon that it is adjacent to long-standing institutions of power, like wealth and higher education, and should be devoid of strong regional markers (Lippi-Green, 2012, pp. 57-62). By adulthood, the process of using the indexicality of speech to assign individuals or groups to social categories and their related stereotypes (whether positive or negative) is an ingrained cognitive process (Craft et al., 2020).

The real-world implications of this process of linguistic discrimination based on spoken communication can be seen in many different contexts in the US (Craft et al., 2020). For

example, speakers using non-English-languages can face discrimination in pursuing their education (Clements & Petray, 2021; Lippi-Green, 2012), in the workforce, or in the process of finding employment (Kutlu & Wiltshire, 2020; Peng et al., 2020; Roessel et al., 2019). This type of discrimination is also disruptive in essential tasks such as finding a place to live (Squires & Chadwick, 2006) and accessing medical care (Becerra et al., 2015).

Finally, linguistic discrimination has been documented in various forms of media where spoken communication is present, like radio advertisements (Reinares-Lara et al., 2016), or movies and television (Dave, 2017; Dragojevic et al., 2016). In the US, while there are a few languages or accents that sometimes have more positive connotations (e.g., people with British-English accents being seen as more intelligent and of a higher SES), non-standard accented English is often associated with negative stereotypes (e.g., people with New York accents being seen as more arrogant and less friendly than other accents) (Shah, 2019). However, little research has addressed accent bias or non-English-languages in the music sector of media.

### **Lyrics and Music Perception**

While, to our knowledge, most research on linguistic discrimination in media has not expanded to music in the form of lyrics, the importance of lyrics in the perception of music has been well-established. Lyrics can greatly influence the emotional experience of the listener by either increasing the emotional response (as in sad songs) or decreasing the emotional response (as in happy songs) (Ali & Peynircioglu, 2006; Brattico et al., 2011; Mori, 2009; Mori & Iwanaga, 2014). Listeners will use music, even songs which have sad lyrics, as a coping mechanism involved in their own emotional regulation (Lamont, 2011; Mori & Iwanaga, 2014).

The content of the lyrics is important even beyond the general emotional response (like happy or sad) that they impart to the listener. For example, lyrical content of songs seems to be

correlated with their popularity, as songs with simpler or more repetitive lyrics chart higher or for longer (Nunes et al., 2015; Varnum et al., 2021). Trends in lyrics of the most popular songs of the moment seem to reflect shifting cultural norms, like more open expression of sexuality (McAuslan & Waung, 2018; Napier & Shamir, 2018). Lyrics can even reflect larger social changes like periods of socioeconomic struggle, with recent studies looking at how lyrics are changing during the current Covid-19 pandemic (Pettijohn & Sacco, 2009; Putter et al., 2021).

### **Current Study**

Overall, although it is known that lyrics are highly influential in the perception of music by audiences, it is not yet known how the language of the lyrics of a song affects their impact on aspects of the listening experience. The present study examined the influence of language of the lyrics of a song on various aspects of the listening experience to determine if linguistic discrimination occurs in US music audiences. A survey presented participants with one of two versions of a song, which differed only on the language that the lyrics were in, then measured their general experience listening to the song. Specifically, the study was conducted to investigate if any differences in likelihood to support the song (by streaming or purchasing it) between the two language versions would be significant even after considering the participants' own language background and any variance in the other experience measures.

Based on the previous research, we hypothesized that the manipulation of the language of the lyrics would cause the songs with non-English lyrics to be significantly less likely to be supported by listeners in the future (by streaming or purchasing the song) when compared to the English-lyric counterparts, showing explicit consumer/audience bias based on language in music (1A). Additionally, we hypothesized that the language identity of the participants would not influence their listener experiences or the potential demonstration of implicit and explicit biases,

as previous research shows that even bilingual individuals demonstrate linguistic bias due to the overwhelming assumed dominance of English in the US (1B). Finally, we hypothesized that the differences in support behaviors between the two language versions of the songs would not be explained by differences in enjoyment or emotional reaction, therefore demonstrating that linguistic discrimination found in other areas of life and media extends into the music industry and song (2).

## **Methods**

### **Participants**

Participants of this online survey study were 18 years of age or older and residents of the United States ( $n = 360$ ). While all participants were fluent in English, 155 (43%) were bilingual or multi-lingual (fluent in multiple languages other than English). Specifically, 108 participants (30%) indicated they were equally fluent in another language as they were in English, while 47 participants (13%) indicated that they were more fluent in a language other than English (but still were fluent in English). All participants were recruited by Qualtrics using their own panel sourcing procedures. This study was approved by the Institutional Review Board (IRB) of University of Chicago, and all participants gave their informed consent to complete the study before the survey began.

### **Materials and Procedures**

Participants were given a survey through the Qualtrics platform which instructed them to listen to snippets of five songs in total and answer questions about their listening experience for each song. In order to place the study in the context of the US popular music industry, the songs were sourced from the Billboard charts from the last 10 years (2011-2021), as Billboard has long

been an industry standard for charting and success-based measures for the American Music Industry (Lafrance et al., 2017). In order to experimentally manipulate which language the lyrics were in, while ensuring all other sonic variables stayed consistent, two experimental songs were chosen that had been officially released by the artist in a version with English lyrics, and in a version with non-English lyrics. To reflect the two most popular non-English languages present in the Billboard charts, one song with a Spanish version and one song with a Korean version was chosen (Zellner, 2020). Three additional control songs were also included (one non-English and two English). To minimize any potential harm to the participants, all songs were free of explicit language (the clean version of “Candy” by Doja Cat was used). Finally, all participants were monetarily compensated for their participation in the study by Qualtrics.

In order to create the song snippet stimuli for the survey, the five songs were purchased on iTunes and then edited to be 90 seconds long, each starting from the beginning of the song. The 90 second length was chosen based on the preview length for songs on iTunes, and on the general length of the first verse and chorus of a standard radio-friendly song.

In the survey, participants were asked to listen to the snippet of the song and then answer questions about their experience, including: how much they enjoyed it, if they have heard the song before (to control for familiarity effects; Ali & Peynircioglu, 2010), what language they perceived the song to primarily be sung in, if they would categorize the song as foreign language (to see if certain languages are considered more or less foreign by the US audience), and to describe the main emotional response they had to the song. Specifically, the participants were asked about the emotional response valence (how positive or negative their feelings were), and their response’s intensity (how strong their emotional response was; Ali & Peynircioglu, 2010; Mori & Iwanaga, 2014). Then, participants who indicated that they had not heard the song before were asked two additional questions to measure audience support behavior (e.g., how likely they

would be to support the song in the future by streaming, or how likely they would be to support the song by buying it on a platform like iTunes), as these are both important factors in charting/popularity (Billboard, 2021).

In a final section, which did not include any music samples, the participants were asked about their general music preferences and behaviors, to determine the relative importance of lyrics and genre in the enjoyment of music, as well as how certain factors influence how participants seek out new music to consume.

## **Design**

Participants were randomly assigned to one of two groups and were given one of two versions of the survey in order to test the impact of lyric language on listener experience. Specifically, the two groups had surveys which were the same except for which language version was provided as the stimulus for the two experimental songs (“Who Says”/ “Dices” by Selena Gomez and “Regular” by NCT). Version 1 of the survey included these five songs: “Who Says (English version)” by Selena Gomez & The Scene, “Regular (Korean version)” by NCT, “Zitti e Buoni” by Måneskin, “The A-Team” by Ed Sheeran, and “Candy (clean version)” by Doja Cat. Version 2 of the survey included these five songs: “Dices (Spanish version of Who Says)” by Selena Gomez & The Scene, “Regular (English version)” by NCT, “Zitti e Buoni” by Måneskin, “The A-Team” by Ed Sheeran, and “Candy (clean version)” by Doja Cat. The order of appearance of the five songs and their associated questions was also randomized for every participant. This design provided for a between-subjects analysis based on the manipulation of the language of the lyrics for the two experimental songs.

## **Statistical Analyses**



In order to identify the impact of lyric language on the eight listener experience variables of the two experimental songs, between-groups comparisons were conducted with the chi-square test of independence for categorical variables (familiarity, lyric language identification, and classification of song as foreign or not), independent-samples t-tests for the interval variables (enjoyment and emotional response intensity), and Wilcoxon rank sum test for the ordinal variables (likelihood to stream and likelihood to purchase the song). The emotional response valence was split into two different categorical variables (presence of positive response yes/no, and presence of negative response yes/no), and analyzed separately with chi-square tests of independence. The same process was completed for the analysis of the influence of language identity but was done separately for each language version of the song, for a total of four analyses. Finally, multiple linear regression analyses were performed on the support behavior measures (streaming and purchasing the song in the future) with all previously mentioned listener experience measures as predictor variables, as well as language identity as a predictor variable, to see if differences in support behaviors were still significant between the two language versions after controlling for variances in experience.

## **Results**

### **Participant Demographics**

In total, 360 participants and their responses to the survey were recorded. While language identity was collected on two different dimensions, the more specific dimension (monolingual vs. equally fluent in English and another language vs. more fluent in another language vs. English) was not found to be significant for any of the subsequent analyses that included language identity as a factor, so it has not been included in the rest of the results. For bilingual participants ( $n = 155$ , 43%), the most common additional languages spoken other than English

were Spanish ( $n = 110$ ), Other (not included in the top 10 most common languages other than English spoken in the US) ( $n = 22$ ), and Tagalog ( $n = 8$ ). See Table 1 for more information on languages other than English spoken by the bilingual participants.

**Table 1**

*Frequency of Non-English Language Spoken by Bilingual Participants*

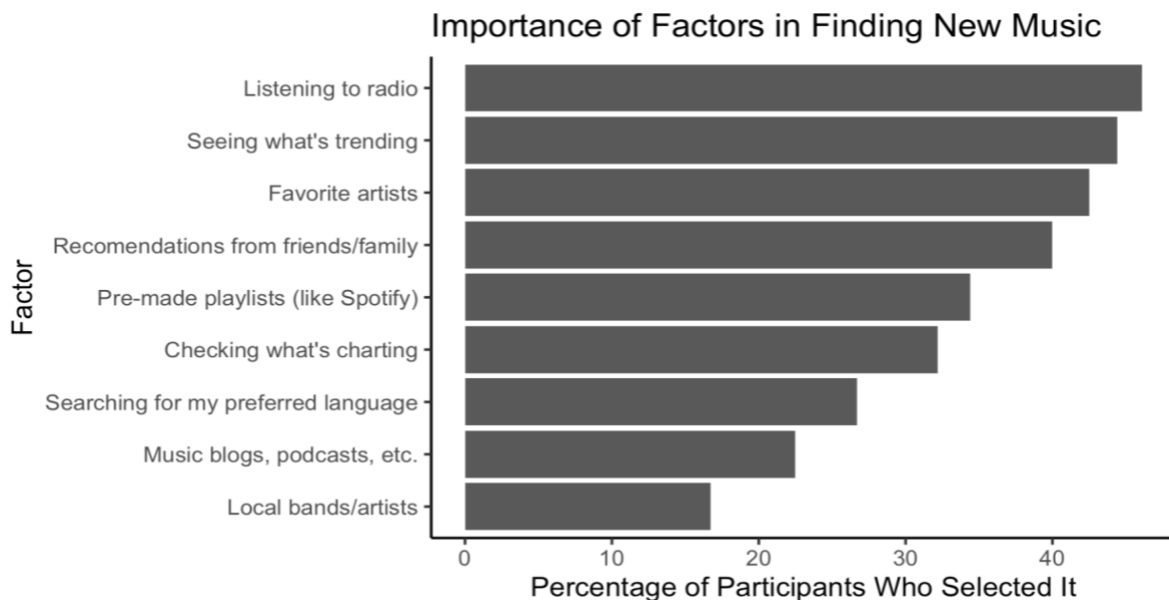
Language	Equally fluent Bilinguals ( $n = 108$ )		Most Fluent in Language Other Than English (ESL) Bilinguals ( $n = 47$ )	
	$n$	%	$n$	%
Arabic	2	1.9	0	0
Chinese	1	.9	2	4.3
French	2	1.9	1	2.1
German	3	2.8	1	2.1
Haitian Creole	0	0	0	0
Italian	0	0	0	0
Korean	1	.9	0	0
Other	9	8.3	13	27.7
Russian	0	0	0	0
Spanish	84	77.8	26	55.3
Tagalog	4	3.7	4	8.5
Vietnamese	2	1.9	0	0

### General Music-related Behaviors

In the final section of the survey, participants were asked a few questions about their general behaviors as music consumers in order to get a baseline reading about listener attitudes and behaviors when it comes to music, lyric language, and other preferences. When asked to select which tactics they find most important to use when finding new music to listen to, the most frequently selected responses were “Listening to what's being played on the radio”, “Seeing what's trending (on apps like TikTok, Twitter, YouTube, etc.)”, and “Searching for new music from my favorite artists”. Only a small percentage of participants (8%) responded that they do not listen to or look for new music. Out of the nine possible factors to choose from (they were instructed to select all that apply), the participants selected only 3.16 options on average, with 71% of participants selecting only 1-4 options. See Figure 1 for the importance of each factor in finding new music.

**Figure 1**

*Percent of Participants Who Selected Each Factor as Important in the Process of Finding New Music*



When asked to rank eight factors by importance in enjoying music the most important factor on average was relatability of lyrics ( $M = 4.29$ ), while the least important factor on average was freshness ( $M = 4.68$ ). See Table 2 for information on each factor and its average rank.

**Table 2**

*Rank of Importance of Factors in Enjoying Music*

Factor	Rank Mean	Rank Standard Deviation	Overall Average Rank
Fits into my favorite genre	4.61	2.27	6
Is by one of my favorite artists	4.41	2.30	2
Has relatable lyrics	4.29	2.21	1
Has lyrics that are in a language that I understand/speak fluently	4.63	2.29	7
Fits my mood or for different situations (working out, driving to work, studying, etc.)	4.46	2.31	3
Is popular or trending	4.48	2.39	5
Is nostalgic or has a lot of memories attached to it	4.46	2.24	4
Is fresh or something I haven't heard before	4.68	2.31	8

*Note.* When asked to rank the importance of these attributes of a song or album for the enjoyment of music, on a 1-8 scale where 1 is most important and 8 is least important.

To determine if participants generally prioritize lyric language or the sound/genre of a song, they were asked two mirror questions about their usual listening behaviors. First, they were asked if they avoid listening to a song if the lyrics are in a language they don't understand or speak, even if they like the sound or genre. The majority of participants either responded that

they don't avoid it, but prefer songs with lyrics they can understand (31%), or that yes, they do avoid it (24%, for 55% total of the participants). However, a large portion also responded that no, they enjoy music they can't understand (without translation) (39%), with the rest (6%) stating that they don't pay attention to lyrics at all, only sound or genre. Second, they were asked if they avoid listening to a song if they like the lyrics but don't like the overall sound or genre. Like the previous question, the majority of participants either responded that they don't avoid it, but prefer songs with a specific sound or genre (36%), or that yes, they do avoid it (23%, for 59% of the total number of participants). Finally, a large portion stated that they enjoy music outside of their preferred genre (35%), with the rest responding that they don't pay attention to sound or genre, only to lyrics (6%).

## **Lyric Language and Listener Experience**

### ***Control Songs***

Between-group comparison analyses on each of the eight listener experience measures were conducted for each of the three control songs and none of these comparisons were significant, so no baseline difference between the two groups existed on these specific measures. Additionally, a between-group comparison of the number of bilingual participants was conducted to ensure random assignment, and there was no significant difference in this variable between the group who was given survey version 1 ( $n = 84$ ) and the group who was given survey version 2 ( $n = 71$ ),  $X^2(1, 360) = 3.27, p = .098$ . Additionally, the participants' accuracy in identifying whether the lyrics were in English or a non-English language was high for all control songs, including the Italian control song ("Zitti e Buoni" by Måneskin, 75%), the first English control song ("The A-Team" by Ed Sheeran, 90%), and the second English control song ("Candy" by

Doja Cat, 86%). When analyzing the accuracy of identifying which specific language the lyrics were in on the control songs between each group, none of the comparisons were significant. See Table 3 for a full list of language identification for the control songs in the survey.

**Table 3**

*Identification of Lyric Language for Survey Control Songs*

Language	Song 3 (Italian)		Song 4 (English)		Song 5 (English)	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Arabic	6	1.6	3	0.8	1	0.3
Chinese	5	1.4	2	0.6	5	1.4
English	93	25.8	324	90.0	309	85.8
French	26	7.2	5	1.4	2	0.6
German	6	1.6	3	0.8	1	0.3
Haitian Creole	1	0.3	0	0	0	0
Italian	44	12.2	0	0	3	0.8
Korean	5	1.4	2	0.6	4	1.1
Other	40	11.1	9	2.5	16	4.4
Russian	13	3.6	0	0	1	0.3
Spanish	113	31.4	12	3.3	9	2.5
Tagalog	5	1.4	0	0	4	1.1
Vietnamese	3	0.8	0	0	5	1.4

### ***Experimental Song 1***

For experimental song 1 (“Who Says” by Selena Gomez), the majority of participants were not familiar with the song (having not heard it before or not being sure if they had heard it before) (62%), but the results of the chi-squared tests showed there was a significant difference in familiarity between the two language versions,  $X^2(2, 360) = 22.1, p = .011$ , with participants being more familiar with the English version than with the Spanish version. The manipulation of the language of the lyrics was successful, as there was a significant difference in whether the participants identified the lyrics as in English or in a non-English language,  $X^2(1, 360) = 166.4, p = .002$ , with many participants correctly identifying if the song was in English (91% for the English version) or in a non-English language (76% for the Spanish version). See Table 4 for a full list of lyric language identification for experimental song 1.

This manipulation of lyric language also produced a significant difference in the categorization of the song as foreign or not,  $X^2(1, 360) = 75.0, p = .002$ , with more participants categorizing the Spanish version of the song as foreign (69%) than the English version (24%). However, there was a lack of consensus overall with the identification of which specific non-English language each song was in, with only around half (53%) of participants correctly identifying the language in the Spanish version as Spanish. However, this confusion was not due to a general inability of participants to identify English vs. non-English languages. In the three control songs, the majority of participants on average (83%) were able to accurately identify if the song was in English or not.

There were also significant differences in the emotional responses to experimental song 1 based on the language of the lyrics. First, the enjoyment of the English version of the song ( $M = 4.16, SD = 1.01$ ) was significantly higher than the enjoyment of the Spanish version ( $M = 3.68,$

$SD = 1.06$ );  $t(357) = 4.41, p = .009$ . While there was no significant difference in the number of participants who had a positive emotional reaction to the song ( $p = .42$ ), the language of the lyrics produced a significant difference in the number of negative emotional reactions to the song,  $X^2(1, 360) = 8.40, p = .004$ , with the number of participants reporting negative emotions being higher in the Spanish version condition (36%) than in the English version condition (22%). Finally, the intensity of the reported emotional reaction to the English version of “Who Says” ( $M = 2.53, SD = 1.10$ ) was significantly higher than reported with the Spanish version of the same song ( $M = 2.23, SD = .96$ );  $t(351) = 2.76, p = .006$ .

Lastly, for experimental song 1 the reported likelihood of the participants supporting the song by streaming it in the future was higher for the English version ( $Mdn = 4$ ) than the Spanish version ( $Mdn = 3$ ), and this difference was significant  $U(n = 188) = 5331, p = .005$ . However, there was no significant difference in the likelihood of purchasing the song in the future ( $p = .25$ ).

**Table 4**

*Identification of Lyric Language for Survey Experimental Song 1*



Language	Survey Version 1 (English, $n = 180$ )		Survey Version 2 (Spanish, $n = 180$ )	
	$n$	%	$n$	%
Arabic	0	0	4	2.2
Chinese	1	.6	4	2.2
English	164	91.1	43	23.9
French	2	1.1	5	2.8
German	1	.6	1	.6
Haitian Creole	1	.6	0	0
Italian	2	1.1	4	2.2
Korean	0	0	5	2.8
Other	3	1.7	14	7.8
Russian	0	0	2	1.1
Spanish	4	2.2	96	53.3
Tagalog	0	0	2	1.1
Vietnamese	2	1.1	0	0

### *Experimental Song 2*

For experimental song 2 (“Regular” by NCT), the song was generally not familiar to the participants (83% had not heard it before), but there was no significant difference in familiarity between the two lyric language conditions of the song (English vs. Korean) ( $p = .59$ ). Contrary to song 1, the success of the language manipulation of this song was unclear. Participants were not generally able to accurately identify if the song’s lyrics were in English (77% for the English

version) or non-English language (only 64% for the Korean version), and the difference was not significant between the two lyric language conditions ( $p = .09$ ). Additionally, the difference between the identification of the song as foreign between the two language conditions was only marginally significant,  $X^2(1, 360) = 42.9, p = .052$ , with the Korean version being more likely to be classified as foreign (71%) than the English version (36%). See Table 5 for a full list of lyric language identification for experimental song 2.

The difference in lyric language conditions did not produce a significant difference in the emotional responses to experimental song 2, including enjoyment ( $p = .31$ ), positive ( $p = .59$ ) or negative ( $p = .74$ ) emotional reaction, or intensity ( $p = .35$ ). Finally, there were no significant differences between the English and Korean versions of the song on support behaviors, so the participants who were not previously familiar with the song were not more likely to stream the song in the future ( $p = .25$ ) or purchase the song in the future ( $p = .31$ ).

**Table 5**

*Identification of Lyric Language for Survey Experimental Song 2*

Language	Survey Version 1 (Korean, $n = 180$ )		Survey Version 2 (English, $n = 180$ )	
	$n$	%	$n$	%
Arabic	7	3.9	6	3.3
Chinese	6	3.3	0	0
English	65	36.1	138	76.7
French	5	2.8	4	2.2
German	3	1.7	2	1.1
Haitian Creole	1	.6	3	1.7
Italian	4	2.2	1	.6
Korean	30	16.7	1	.6
Other	23	12.8	9	5.0
Russian	5	2.8	0	0
Spanish	26	14.4	14	7.8
Tagalog	1	.6	0	0
Vietnamese	4	2.2	2	1.1

### Language Identity of Participants and Listener Experience

For experimental song 1, there were no significant differences between monolingual and bilingual participants on any of the eight listener experience measures for the group who heard the English version of the song. As for the group who heard the Spanish version of the song, the results of the chi-squared analysis showed there was a significant difference in how the monolingual and bilingual participants identified the language of the lyrics,  $X^2(11, 180) = 19.4$ ,

$p = .036$ , with more bilingual participants (71%) accurately identifying the song as Spanish than monolingual participants (42%) who were able to do so.

There was also a significant difference with the Spanish version in the emotional response to the song between the two language identity groups,  $X^2(1, 180) = 4.9, p = .027$ , with more bilingual participants having a positive emotional response (78%) than monolingual participants who experienced the same (61%). Similarly, results of the independent-samples t-test indicated there was a significant difference in enjoyment between the two groups, with bilingual participants ( $M = 3.90, SD = 1.03$ ) enjoying the Spanish version more than the monolingual participants ( $M = 3.54, SD = 1.06$ );  $t(141.7) = -2.23, p = .028$ . Finally, while the difference between the two language identity groups in the likelihood to purchase the Spanish version of song 1 in the future was not significant, there was a significant difference in the likelihood to support the song by streaming it,  $U(n = 107) = 1341, p = .004$ . Specifically, the bilingual participants were more likely to stream the song ( $Mdn = 3$ ) than the monolingual participants ( $Mdn = 2$ ).

As for experimental song 2, the only listener experience measure that showed a significant difference between the two language identity groups for the English version of the song was emotional response intensity, with the bilingual participant having a more intense response ( $M = 2.54, SD = 1.05$ ) than the monolingual participants ( $M = 2.18, SD = 1.02$ );  $t(140.99) = -2.23, p = .028$ . This difference in intensity was regardless of the valence of the emotional response, as neither the positive nor the negative emotional response frequencies between the two groups were significantly different ( $p = .92$  and  $p = .74$ , respectively). For the Korean version of experimental song 2, the results of the individual analyses on the eight listener

experience measures indicated no significant differences between monolingual and bilingual participants.

### **Lyric Language and Support Behaviors**

For experimental song 1 (“Who Says” by Selena Gomez), results of the multiple linear regression indicated that there was a collective significant effect between the listener experience variables and the lyric language condition on the likelihood for the participant to stream the song in the future ( $F(8, 351) = 95.16, p = .002, R^2 = .68$ ). The individual predictors were examined further and indicated that enjoyment ( $t = 3.05, p = .002$ ), familiarity ( $t = -27.19, p < .001$ ), positive emotional response ( $t = 3.17, p = .002$ ), negative emotional response ( $t = -2.35, p = .019$ ), and language variation ( $t = -2.02, p = .044$ ) were significant predictors in the model. Specifically, when controlling for the other significant variables, the participants who were in the Spanish version group were still less likely to stream the song in the future (estimate =  $-.25$ ) than in the English version group. The multiple linear regression conducted on the likelihood to purchase the song in the future was also statistically significant ( $F(8, 351) = 79.42, p < .001, R^2 = .63$ ), but only enjoyment ( $t = 2.85, p = .005$ ) and emotional intensity ( $t = 3.14, p = .001$ ) were significant predictors in the model, not lyric language ( $p = .33$ ). When the language identity variable (monolingual vs. bilingual) was added to the models for both support measures it was not significant, either alone or as an interaction with language variation.

For experimental song 2 (“Regular” by NCT), results of the multiple linear regression indicated that there was a collective significant effect between the listener experience variables, the lyric language condition, and participant language identity on the likelihood for the participant to stream the song in the future ( $F(10, 349) = 52.41, p < .001, R^2 = .59$ ). The

individual predictors were examined further and indicated that enjoyment ( $t = 3.46, p < .001$ ), familiarity ( $t = -21.25, p < .001$ ), positive emotional response ( $t = 4.06, p < .001$ ), emotional response intensity ( $t = 2.14, p = .033$ ), and the interaction between language variation and language identity ( $t = -2.10, p = .037$ ) were significant predictors in the model. The multiple linear regression conducted on the likelihood to purchase the song in the future was also statistically significant ( $F(10, 349) = 41.65, p < .001, R^2 = .53$ ), but only enjoyment ( $t = 2.97, p = .003$ ) and emotional intensity ( $t = 4.85, p < .001$ ) were significant predictors in the model, not lyric language ( $p = .32$ ), participant language identity ( $p = .41$ ), or the interaction between the two ( $p = .32$ ).

### Discussion

This study aimed to investigate if linguistic discrimination seen in other contexts in the United States extended into the popular music industry, specifically to address how linguistic discrimination might play a role in consumer perception and experience of music. Hypothesis 1A suggested that due to the manipulation of the lyric language, the experimental songs with non-English lyrics would be less likely to be supported by the participants than the English lyric versions of the same song, demonstrating a general consumer bias or preference against non-English songs. Hypothesis 1B suggested that the language identity of the participants (being monolingual or bilingual) would not influence their experience of the song or their likelihood to support the song in the future based on the language of the lyrics. Finally, Hypothesis 2 suggested that differences in support behaviors would not be explained solely by differences in audience preference or bias, demonstrating that the principles of linguistic discrimination seen in other contexts do apply to music consumption in the US popular music industry.

### **Differences in Listener Experience Based on Lyric Language**

This study found that the Spanish version of experimental song 1 was less likely to be supported in the future by the participants (in the form of streaming) when compared to the English version of the song, which confirms Hypothesis 1A. The language of the lyrics also influenced many other listener experience dimensions, including the level of enjoyment of the song, the frequency of negative emotional responses, and the intensity of those emotional responses. These findings are in line with research that demonstrates the importance of lyrics in the perception of music, especially in the formation of an emotional response (Ali & Peynircioglu, 2006; Brattico et al., 2011; Mori, 2009; Mori & Iwanaga, 2014). Specifically, the non-English version of experimental song 1 was enjoyed less and was more likely to produce a negative emotional response in the participants when compared to the English version, and emotional responses were less intense on average in the Spanish condition, regardless of the participants' own language abilities. This suggests the presence of a general preference for songs with English lyrics, but these results alone are not sufficient to make claims about any potential linguistic discrimination.

As for experimental song 2, there was no significant difference between the two language version conditions in the ability to identify if the lyrics were in English or a non-English language. Therefore, it is difficult to interpret differences in listening experiences to the two versions of song 2 (English vs. Korean lyrics), since these differences cannot be attributed to successful language manipulation. However, in contrast to experimental song 1, there were no demonstrated differences on the experience measures between the two groups including level of enjoyability, emotional response, and support behaviors. This, interpreted alongside the lack of

difference in language identification, further reflects the failure of the lyric language manipulation for this song.

It is important to note that this failure of the manipulation of the lyric language was not due to an overall inability to identify languages. As shown in Table 3, participants in both groups were clearly able to accurately identify whether the three control songs were in English or a non-English language, so the issue was not with a baseline lack of accuracy in our sample. One potential limitation of this study that might explain the failure of the manipulation was the presence of three Spanish words in song 2, which may have caused unnecessary confusion during the identification. However, this would not explain the participants who identified a language other than Korean or Spanish, which was a large portion of the participants (69%).

### **Impact of Listener Identity and Interaction with Lyric Language**

The participant's language identity (monolingual vs. bilingual) was influential in some dimensions of experience with experimental song 1, but only in the Spanish-language condition, which partially contradicts Hypothesis 1B. While prior research suggests that linguistic bias is experienced equally by monolingual and bilingual individuals (DeJesus et al., 2017; Souza et al., 2013), our results found that biases toward a non-English language were mediated by language background. Specifically, monolingual participants were less likely to enjoy the Spanish version of the song, have a positive emotional response to the song, and were less likely to stream the song in the future.

The discrepancy in the importance of language identity between the two language versions of experimental song 1 could be due to a few different factors. First, the prior research on linguistic biases in monolingual vs. bilingual individuals was conducted on children and



adolescents, but the current study was completed with adult participants. It is possible that differences in biases between these two groups become salient in adulthood, and therefore were measured in this study. Second, the majority of bilingual participants were fluent in Spanish (71%), so the differences in the Spanish lyric condition could be due to familiarity with the language instead of a true difference between attitudes towards Spanish lyrics between the monolingual and bilingual groups. The intelligibility of lyrics has been positively correlated with the perceived importance of lyrics (Condit-Schultz & Huron, 2015), so the increased intelligibility of the song for the Spanish-speaking participants could have made the lyrics more important in their listening experience when compared to the monolingual participants, outside of any linguistic biases. In order to assess the relationship between language identity, lyric language, and listener experience whilst controlling for intelligibility, future research should replicate this study with a more diverse range of bilingual participants with varying lyric languages.

### **Evidence of Linguistic Discrimination and Implications for Music Industry**

Finally, participants were less likely to support the non-English version of experimental song 1 by streaming it in the future even when controlling for all other experience measures, which confirms Hypothesis 2 and suggests that linguistic discrimination does occur in US popular music audiences. To our knowledge, this is the first study to expand upon previous research on the presence of linguistic discrimination in many other contexts (Baron, 2019; Craft et al., 2020; Lippi-Green, 2012), and demonstrate its effects in the music industry. Specifically, while enjoyment, emotional response (positive and negative), and familiarity were significant

predictors of the likelihood to stream, the Spanish version of the song was still less likely to be streamed when controlling for these influences.

Interestingly, there were no significant differences in the likelihood to purchase the song between the two different language versions, even when controlling for other experience variables. Therefore, this study did not produce evidence that linguistic discrimination occurs in the act of purchasing music. However, many US music consumers no longer purchase music in traditional ways, including buying physical albums and CDs, so purchasing now makes up less than 10% of all music consumption (Steele, 2019). While the difference in the likelihood to support the song by purchasing was not significant, the more important support behavior measure was streaming, which did demonstrate linguistic discrimination.

The discriminatory behavior in streaming alone is likely to have many real-world implications for any artist in the US music industry who chooses to perform their songs in non-English languages. In 2015, streaming became the largest source of revenue in the music industry and comprises the majority (82%) of all music consumption in the US (Datta et al., 2018; Steele, 2019). So, not only can diminished streaming behaviors lead to a loss of revenue for an artist singing non-English lyrics, but it can also prevent success in the charts. For example, even though the Billboard Hot 100 chart equates 100 paid subscription streams with one purchase of the song (Billboard, 2018), songs that dominate in streaming regularly top songs which are purchased more due to streaming being a much more popular avenue for music consumption (Molanphy, 2020). Therefore, a significant limiting factor to audience members supporting a song by streaming it (like what was experienced in the non-English version of song 1) could mean the difference between being able to place on the important US charts. For this reason, it would be interesting for future research to address whether the amount of non-English

lyrics in a song (a single word or phrase vs. the entire song being in non-English) would impact the potential linguistic discrimination.

Further research is needed to explore if the results found with experimental song 1 are dependent upon the non-English language being Spanish, or if these results are found with other languages. A limitation of this study was the failure of the manipulation of lyric language for the English/Korean experimental song, which makes the results tied to linguistic discrimination difficult to generalize to non-English languages other than Spanish. This is particularly of note since the majority of the bilingual participants were Spanish speaking, so comparison with other languages would be needed to see if results are due to familiarity with the language or increased intelligibility of the lyrics. However, the makeup of the non-English languages spoken by the participants generally matched that of the United States, which was our target population, as Spanish is the most common non-English language spoken (Thompson, 2021; United States Census Bureau, 2019). Additionally, the between-subjects study design allowed for potential differences between the two randomly assigned groups in general listening preferences or level of biases, although this was addressed with the inclusion and analysis of different types of control songs. Finally, future research controlling for other attributes of music that could impact the listener's experience and likelihood to support that were not able to be accounted for due to limitations of this study, like genre or rhythmic structure, would be useful in better understanding the connection between lyric language and audience support.

### **Conclusion**

This study found that lyric language was influential in the audience's listening experience with popular music, with the Spanish version of the song being generally less liked and less

likely to be supported by the participants than the English version. Contrary to some prior research, language identity of the participant was a mediating factor in the listening experience, but only became significant when listening to non-English songs. However, the lack of success of the lyric language manipulation of the second experimental song prevents these findings from being generalizable to other non-English languages. Most importantly, the Spanish version of the first experimental song was less likely to be supported by the participants in the future through streaming, even when controlling for all other differences in the listening experience, which confirms that linguistic discrimination does occur in the US popular music industry. To our knowledge, this study is the first to explore linguistic discrimination within the popular music industry, and helps to expand upon our understanding of the mechanisms of linguistic discrimination in general. Finally, the findings of this study help to contextualize and highlight the many real-world implications, both monetary and otherwise, that linguistic discrimination could have on musical artists who perform in non-English languages.

## References

- Ali, S. O. & Peynircioglu, Z. F. (2006). Songs and emotions: Are lyrics and melodies equal partners? *Psychology of Music, 34*(4), 511-534.
- Ali, S. O. & Peynircioglu, Z. F. (2010). Intensity of emotions conveyed and elicited by familiar and unfamiliar music. *Music Perception, 27*(3), 177-182.
- Baron, D. (2019). Douglas A. Kibbee: Language and the law: Linguistic inequality in America [Book Review]. *Language Policy, 18*, 315-317.
- Becerra, D., Androff, D., Messing, J. T., Castillo, J., & Camino, A. (2015). Linguistic acculturation and perceptions of quality, access, and discrimination in health care among Latinos in the United States. *Social Work in Health Care, 54*(2), 134-157.
- Billboard. (2018). *Billboard finalizes changes to how streams are weighted for billboard hot 100 & billboard 200*.  
  
<https://www.billboard.com/pro/billboard-changes-streaming-weighting-hot-100-billboard-200/>
- Billboard. (2021) *Billboard charts legend*.  
  
<https://www.billboard.com/p/billboard-charts-legend>
- Brattico, E., Alluri, V., Bogert, B., Jacobsen, T., Vartiainen, N., Nieminen, S., & Tervaniemi, M. (2011). A functional MRI study of happy and sad emotions in music with and without lyrics. *Frontiers in Psychology, 2*(308), 1-16.
- Clements, G. & Petray, M. J. (2021). *Linguistic discrimination in US higher education*. Routledge.
- Condit-Schultz, N. & Huron, D. (2015). Catching the lyrics: Intelligibility in twelve song genres. *Music Perception, 32*(5), 470-483.

Craft, J. T., Wright, K. E., Weissler, R. E., & Queen, R. M. (2020). Language and discrimination:

Generating meaning, perceiving identities, and discriminating outcomes. *Annual Review of Linguistics*, 6, 389-407.

Datta, H., Knox, G., Bronnenberg, B. J. (2018). Changing their tune: How consumers' adoption of online streaming affects music consumption and discovery. *Marketing Science*, 37(1), 5-21.

Davé, S. (2017). Racial accents, hollywood casting, and Asian American studies. *Cinema Journal*, 56(3), 142-147.

Dejesus, J. M., Hwang, H. G., Dautel, J. B., & Kinzler, K. D. (2017). Bilingual children's social preferences hinge on accent. *Journal of Experimental Child Psychology*, 164, 178-191.

Dragojevic, M., Mastro, D., Giles, H., & Sink, A. (2016). Silencing nonstandard speakers: A content analysis of accent portrayals on American primetime television. *Language in Society*, 45(1), 59-85.

Kinzler, K.D. & DeJesus, J. D. (2013) Northern = smart and southern = nice: The development of accent attitudes in the United States. *The Quarterly Journal of Experimental Psychology*, 66(6), 1146-1158.

Kutlu, E. & Wiltshire, C. (2020). Where do negative stereotypes come from? The case of Indian English in the USA. *Proceedings of the Linguistic Society of America*, 5(1), 74-82.

Lafrance, M., Scheibling, C., Burns, L., & Durr, J. (2017). Race, gender, and the billboard top 40 charts between 1997 and 2007. *Popular Music and Society* 41(5), 522-538.

Lamont, A. (2011). University students' strong experiences of music: Pleasure, engagement,

- and meaning. *Musicae Scientiae*, 15(2), 229-249.
- Lippi-Green, R. (2012). *English with an accent: Language, ideology, and discrimination in the United States*. Routledge.
- McAuslan, P. & Waung, M. (2018). Billboard hot 100 songs: Self-promoting over the past 20 years. *Psychology of Popular Media Culture*, 7(2), 171-184.
- Molanphy, C. (2020, Jan 18). *2020 might have just gotten its "Old Town Road"*. Slate.  
<https://slate.com/culture/2020/01/roddy-ricch-the-box-hot-100-tiktok-justin-bieber.html>
- Mori, K. (2009). The influence of the meaning of lyrics on the expressed emotion of music valence. *Proceedings of the second international conference of students of systematic musicology* (pp. 53–58). Ghent, Belgium.
- Mori, K. & Iwanaga, M. (2014). Pleasure generated by sadness: Effect of sad lyrics on the emotions induced by happy music. *Psychology of Music*, 42(5), 643-652.
- Napier, K. & Shamir, L. (2018). Quantitative sentiment analysis of lyrics in popular music. *Journal of Popular Music Studies*, 30(4), 161-176.
- Nunes, J. C., Ordanini, A., & Valsesia, F. (2015). The power of repetition: The repetitive lyrics in a song increase processing fluency and drive market success. *Journal of Consumer Psychology*, 25(2), 197-199.
- Pantos, A. J. & Perkins, A. W. (2012). Measuring implicit and explicit attitudes toward foreign accented speech. *Journal of Language and Social Psychology*, 32(1), 3-20.
- Papinczak, Z. E., Dingle, G. A., Stoyanov, S. R., Hides, L., & Zelenko, O. (2015). Young people's uses of music for well-being. *Journal of Youth Studies*, 18(9), 1119-1134.
- Peng, Y., Genc, E., Nicholson, B., & Markham, M. S. (2020). Not professional enough to be a therapist: International therapists' experience of language discrimination. *Current Psychology*. <https://doi.org/10.1007/s12144-020-00848-4>

- Pettijohn, T. F. & Sacco, D. F. J. (2009). The language of lyrics: An analysis of popular billboard songs across conditions of social and economic threat. *Journal of Language and Social Psychology, 28*(3), 297-311.
- Putter, K.C., Krause, A. E., & North, A. C. (2021). Popular music lyrics and the covid-19 pandemic. *Psychology of Music, 1-16*. <https://doi.org/10.1177%2F03057356211045114>
- Reinares-Lara, E., Martín-Santana, J. D., & Muela-Molina, C. (2016). The effects of accent, differentiation, and stigmatization on spokesperson credibility in radio advertising. *Journal of Global Marketing, 29*(1), 15-28.
- Roessel, J., Schoel, C., Zimmermann, R., & Stahlberg, D. (2019). Shedding new light on the evaluation of accented speakers: Basic mechanisms behind nonnative listeners' evaluations of nonnative accented job candidates. *Journal of Language and Social Psychology, 38*(1), 3-32.
- Ruth, N. & Schramm, H. (2021). Effects of prosocial lyrics and musical production elements on emotions, thoughts and behavior. *Psychology of Music, 49*(4), 759-776.
- Shah, A. P. (2019). Why are certain accents judged the way they are? Decoding qualitative patterns of accent bias. *Advances in Language and Literary Studies, 10*(3), 128-139.
- Souza, A. L., Byers-Heinlein, K., & Poulin-Dubois, D. (2013). Bilingual and monolingual children prefer native-accented speakers. *Frontiers in Psychology, 4*, 1-6.
- Steele, A. (2020, Jan 9). *U.S. music streams topped a trillion in 2019*. The Wall Street Journal. <https://www.wsj.com/articles/u-s-music-streams-topped-a-trillion-in-2019-11578607152>
- Squires, G. D. & Chadwick, J. (2006). Linguistic profiling: A continuing tradition of discrimination in the home insurance industry? *Urban Affairs Review, 41*(3), 400-415.
- Thompson, S. (2021, May 27). *The US has the second-largest population of spanish speakers:*



*How to equip your brand to serve them.* Forbes.

United States Census Bureau. (2019). *Language spoken at home.*

<https://data.census.gov/cedsci/table?q=Language%20Spoken%20at%20Home&tid=ACST1Y2019.S1601>

Varnum, M. E. W., Krems, J. A., Morris, C., Wormley, A., & Grossman, I. (2021). Why are song lyrics becoming simpler? A time series analysis of lyrical complexity in six decades of American popular music. *PLoS* 16(1), 1-18.

Zellner, X. (2020, March 3). *BTS's 'On' becomes 20th non-english-language top 10 in hot 100 history.* <https://www.billboard.com/articles/business/chart-beat/9327176/all-20-non-english-language-songs-hit-hot-100s-top-10/>