

Juri Matsubara
Professor Ann Bradlow
Independent Study
Spring 2002

Dialect Differentiation between Speech and Song

Introduction

Through personal listening to music from different English-speaking countries, an observation grew: singers tended to lose much of their regional speech dialects when singing. For example, singers from Britain who had a noticeable regional dialect during interviews, such as Elton John and George Michael, did not sound too different when singing from other singers with various regional dialects, such as Australian, Canadian, Scottish, Irish and Welsh English. This observation grew to include foreign music as well: Enrique Iglesias, who has a strong Spanish accent while speaking English, does not sound as Spanish while singing in English, although some of his songs retain a Spanish musical style.

Some past research focuses on perception of vowels produced by singers. Hollien et al. (2000) examined perceptual differences in vowels that are sung at extremely high versus low fundamental frequencies. They found a profound effect of pitch on vowel perception such that the subjects tended to mistakenly identify high-pitch vowels as the vowel with the next highest formant, just above the sung frequency. Although the typical singer on a modern radio station does not sing at extremely high pitches, if his or her regularly spoken vowels are affected by singing in a higher range, this could lead to the singers sounding different when singing versus speaking.

Other research concerns application of song as an aid in speech production. Glover et al. (1996) found that being asked to sing significantly reduces stuttering in subjects who stutter ranging from 18 to 47 years old. Their data suggest that singing enables these subjects to impose a structure on speech production that overcomes stuttering. This also suggests that the subjects are generating fluent speech in their mind, which is obstructed by stuttering in speech but can be overcome when singing. These findings suggest that singing may be used as a way to improve

unwanted foreign accents. However, fluent speech in the mind would not be present in this case, and therefore this may mean that students of English as a second language (ESL) should devote considerable concentration on mentalizing fluent speech before attempting singing exercises.

Schunk (1999) offers a glimpse of applications of song in education. She found that ESL students retained vocabulary significantly better when singing words accompanied by signs, as opposed to speaking them while seeing the signs. This offers evidence that song taps into parts of the brain slightly different from speech. The melody, which imposes pitch and note structure, renders speech into production that differs by both frequency and length to simple speaking. These differences somehow lead to results that are beneficial.

Some questions that the current study pursues are: do regional dialects diminish through singing as a universal phenomenon? Does the phenomenon carry across to foreign accents as well? The implications of the study lead to practical uses of song such as follows: Can singing be an effective tool in foreign-language learning or foreign accent reduction? Would it be an effective tool in general education also? Finally, can singing be an effective means of cross-cultural communication?

The main hypothesis of this study is that different speech dialects do in fact become standardized through singing; that is, the dialects are perceived to be less different from each other. It is predicted that native listeners will have a harder time distinguishing dialects when judging sung language versus spoken language.

Methods

The chosen regional English dialects for this study are: Australian, British, Canadian, Irish, Scottish, United States and Welsh English. Canadian and Welsh English were only included for the song portion of the experiment because there were no adequate stimuli for them in speech. The chosen foreign English accents are: Cantonese, French, Japanese, Korean, Spanish, Tagalog and Zulu.

Song samples are all English-language commercial songs sung by professional artists or bands. First the experimenter ran through a background check with most of the songs, especially the native English samples. This was done by finding an interview with the singer, such as Shirley Manson of the band Garbage. After confirming that the singer has a prominent regional

dialect when speaking, the experimenter searched for one of the more obscure songs by the singer or band. Obscurity was valued so that subjects of the experiment would not identify dialects by simple familiarity with the background of the singer.

Speech samples are all taken from an online database of people all over the world reading out a standard text, which is a small paragraph. These people are not professional speakers and have not practiced saying the paragraph extensively. Therefore the English of the foreign dialects is sometimes “choppy,” the speaker often pausing or stuttering. The website is maintained by a team of people led by Steven Weinberger at George Mason University.

All song and speech samples were cut into smaller stimuli files with Cool Edit 2000. All music samples have between 27 to 34 syllables and end at a coherent phrase of the song. The standard speech paragraph was cut into a smaller file of 37 syllables. The decision to have similar number of syllables among the stimuli, instead of similar duration of the sample, was made in order to present subjects with roughly the same amount of linguistic information for each music and speech sample. Appendix A lists the lyrics in each song stimulus and the sentence of the truncated speech stimulus.

All the truncated stimuli were pasted into Cool Edit with ten seconds of silence between each stimulus and ten seconds after the last stimulus, and five seconds before the first stimulus. Two separate files were created in this way, one with all song samples and the other with all speech samples. The first stimulus in both files is a sample from the U.S. that the subjects were asked to consider as standard U.S. English. The order of the stimuli was chosen randomly. The order of the song stimuli are as follows: 1) sample U.S. 2) Cantonese 3) Welsh 4) U.S. 5) Tagalog 6) Scottish 7) Spanish 8) Korean 9) Irish 10) Zulu 11) Canadian 12) Japanese 13) French 14) Australian 15) British. The order of the speech stimuli are as follows: 1) sample U.S. 2) Korean 3) British 4) Scottish 5) Zulu 6) Japanese 7) Irish 8) Spanish 9) U.S. 10) Cantonese 11) Australian 12) French 13) Tagalog. A CD-R with all the labeled individual stimulus files will be included.

Finally, a paper survey was prepared with four parts: song, speech, feedback and background information. For song and speech, subjects were asked the same basic question of whether the stimulus was produced by a native or non-native English speaker, and then asked to identify the native dialect or language of the singer or speaker. For the song section subjects were given a chance to check a line if they recognized the singer or song, so the experimenter

could check this effect on their performances. The song and speech section of the test were done while playing the stimulus files on Cool Edit 2000. After the listening and answering sections, the feedback questions asked whether song or speech was easier in terms of discerning native dialects or languages, any strategies used for the task, and any additional comments. Lastly the background questions asked for native language of the subject, exposure to other languages and music tastes. A simplified survey is shown in Appendix B.

Upon completion of the survey, subjects were debriefed on the purpose of the experiment – to find a perceptual difference between song and speech. The experimenter offered to go through the stimuli briefly with the right answers. All but one subject asked to know the answers.

Results

There were three pilot subjects to make sure the test ran smoothly. Nine test subjects were run. All test subjects but one were Northwestern University students; the other was a University of Illinois – Urbana Champaign student. Seven were monolingual, one able to understand two languages but only able to speak one, and one bilingual. There were no significant differences between the groups, but the one subject who can understand two but only speak one language got one of the highest scores in the speech section, getting 8 out of 12 correct. However, a monolingual subject also performed just as well, so no conclusions can be made about differences between the groups.

Subjects were generally more accurate in identifying dialects in speech than in song. On average subjects answered the specific dialect or language perceived 26.2% correctly in song, and 42.6% correctly in speech. On average subjects answered whether the singer or speaker was a native or non-native English speaker 81.7% correctly in song, and 99.1% correctly in speech. All subjects gave feedback that the speech section was easier in making dialect distinctions, and this is directly seen in their results.

Two subjects said they used no strategies for the task, but the others wrote a brief description. Two subjects commented that they eliminated choices that were already chosen and tried not to circle them again. For the song section, two subjects wrote down that the style of the music sounding more foreign or domestic influenced their answers. For the speech section, two

subjects tried to match how the stimuli sounded to people they know. Also, two subjects listened for non-native pronunciations in the vowels or the letter 'r.' Finally four subjects listened for stuttering, mistakes, or stumbling, all which led them to conclude that the stimuli that sounded uncomfortable were not native.

A table has been constructed for each kind of stimulus: song and speech (Table Song and table Speech). The rows show the actual stimulus, or what the subjects were presented with, and the columns present what the subjects answered. The yellow cells show the correct answers. It is apparent that there are more responses in the yellow cells for speech than song. Another significant finding is shown in the light blue cells. More dialects and languages were identified as U.S. English in song than speech. A non-U.S. native was identified as U.S. 25 times in song, but only twice in speech. Recognizing the singer or song did not have a significant effect on performance; three instances of recognition led to the correct answer, but in 10 instances recognition led to an incorrect response, six of those being U.S. English. It seems that for these subjects there is a tendency to assume that familiar singers or songs are from the U.S. It would be interesting to perform this experiment on subjects at a different university, for example in Britain, and see whether those subjects would answer British English more frequently than these subjects.

In both song and speech, some dialects or languages were particularly confusable with other choices, as indicated by more than half of the subjects choosing the same incorrect choice. For example, in song, seven subjects answered 'U.S.' for the Australian English sample. Some subjects explained that the singer, Natalie Imbruglia, sounded familiar, and therefore they assumed she was from the States. Others said her voice had no signs of a regional dialect. In speech, two cases are most noteworthy – Japanese and Zulu. Five subjects answered 'Korean' for the Japanese stimulus, which could be explained by the similarity of the two languages or the close association of their foreign accents by the subjects. Also, five subjects answered 'Cantonese' for Zulu. This was a result of people not being familiar with a Zulu English accent added to their relative familiarity with a Cantonese English accent; yet this confusion may indicate some sort of link between Zulu and Cantonese, which could be pursued in a later study.

Discussion

The results show that accurate perception of dialect variation is in fact greatly reduced from song to speech. Subjects are remarkably accurate in identifying whether a given speech stimulus is produced by a native or non-native English speaker, and are moderately accurate in identifying even the exact regional dialect or language of the speech stimulus. This accuracy has been shown to be diminished for perception of song stimuli.

There are two possible explanations: the act of singing affects vowels, or a song's melody overrides some of the characteristic intonations of a dialect. As seen in Hollien et al. (2000) earlier, high-pitched vowels tended to be mistaken for vowels with the next higher formant. If any higher pitch compared to regular speaking could affect the vowels produced, perceptual differences found in this experiment could be due to difficulties that arise from transformed vowels in song. Another explanation is that dialect differentiation is difficult for song because characteristics of a regional dialect diminish in song. Specifically, if subjects categorize dialects through various intonation patterns or vowel pronunciations, these indicators would greatly decrease in song because song alters both pitch and length of regular speech. The melody of the song would diminish various intonation patterns, while words drawn out for a long duration would diminish vowel differences. These two explanations could explain the results of the current study.

However, the current study is a pilot experiment and has limitations. Some subjects gave feedback that the style of the music influenced their answers. This can account for all the subjects getting the Zulu song correct – the music, they say, was a dead give-away. Also, the singers in this experiment are practiced professionals, whereas the speakers from the online archive are presented with unfamiliar text, and therefore may give away clues by stuttering, speaking slowly, or completely pausing.

Future directions for this study would be a full-fledged study with the same idea. 10-15 people with various dialects would be asked to speak and sing the same set of words. The song portion would be sung to a simple tune such as 'Mary had a Little Lamb.' Two different sets of subjects would be asked to evaluate the stimuli; they would not judge both song and speech produced by a certain speaker. Otherwise the evaluators may recognize a voice and simply write down what they had written before for that stimulus.

The implications for this study could extend to the study of English as a Second Language and international communication. The idea is that, perhaps singing what one needs to convey would be more effective in communication, if a strong dialect could be diminished. It is still not clear whether the act of singing in higher pitches than regular speech, the act of imposing a tune over the natural pitches of a dialect or the elongated vowels are what leads to dialect diminution, but the first step has been taken to show the “universality of sung speech.”

References

Glover, Helen; Kalinowski, Joseph; Rastatter, Michael. “Effect of instruction to sing on stuttering frequency at normal and fast rates.” *Perceptual & Motor Skills*. Vol 83(2) Oct 1996, 511-522.

Hollien, Harry; Mendes-Schwartz, Ana P; Nielsen, Kenneth. “Perceptual confusions of high-pitched sung vowels.” *Journal of Voice*. Vol 14(2) Jun 2000, 287-298.

Schunk, Heather A. “The effect of singing paired with signing on receptive vocabulary skills of elementary ESL students.” *Journal of Music Therapy*. Vol 36(2) Sum 1999, 110-124.

Weinberger, Steven H. Speech accent archive. 2002. <http://classweb.gmu.edu/accent>

Tables:

Table 1: Speech

	U.S.	British	Australian	Irish	Scottish	Spanish	French	Japanese	Cantonese	Korean	Tagalog	Zulu
U.S.	9	0	0	1	1	0	0	0	0	0	0	0
British	0	5	4	4	2	1	0	0	0	0	0	0
Australian	0	2	2	2	4	0	0	0	0	0	0	0
Irish	0	1	0	2	2	0	0	0	0	0	0	0
Scottish	0	1	1	0	6	0	0	0	0	0	0	0
Spanish	0	0	0	0	0	5	0	0	0	1	4	0
French	0	0	0	0	0	0	3	8	0	0	1	1
Japanese	0	0	0	0	0	0	0	4	1	1	0	0
Cantonese	0	0	0	0	0	0	0	0	4	3	0	5
Korean	0	0	1	0	0	1	0	5	2	1	0	1
Tagalog	0	0	0	0	0	0	1	0	2	3	3	1
Zulu	0	0	0	0	0	0	0	0	0	0	1	1

Table 2: Song

	U.S.	British	Australian	Irish	Scottish	Canadian	Welsh	Spanish	French	Japanese	Cantonese	Korean	Tagalog	Zulu
U.S.	5(1)	3(3)	7(2)	1	1	2	3(1)	0	1	2	3(1)	0	2	0
British	0	2(1)	0	1	2(1)	0	4(1)	0	1	0	0	0	0	0
Australian	0	1	0	1	1	2	0	0	1	2	0	0	0	0
Irish	1	0	0	2	3	1	0	0	0	0	0	1	1	0
Scottish	1	2	0	1	0	0	0	0	0	0	0	0	0	0
Canadian	1	0	2	2	0	2	0	0	0	2	1	0	1	0
Welsh	0	0	0	1	1	1	1(1)	0	0	0	0	0	0	0
Spanish	0	0	0	0	0	0	0	5	0	0	1	1	0	0
French	0	1	0	0	0	0	1	3	2	1	0	0	0	0
Japanese	0	0	0	0	0	0	0	0	0	0	1	4	2(1)	0
Cantonese	0	0	0	0	0	0	0	0	1	1	2	1	2	0
Korean	0	0	0	0	0	0	0	0	1	1	1	2	1	0
Tagalog	1	0	0	0	1	1	0	0	0	0	0	0	0	0
Zulu	0	0	0	0	0	0	0	0	0	2	0	0	0	9

Number in parantheses indicates how many out of the total in specific cell recognized the song or speaker.

Appendix A: Stimuli

Lyrics of the truncated section of each song follow. Number inside parentheses indicates the number of syllables in the whole section.

Sample U.S.

Billy Joel – Honesty (33)

honesty is such a lonely word everyone is so untrue; honesty is hardly ever heard and mostly what I need from you

Cantonese

Faye Wong – If You Were Mine (27)

we'll be so happy just you and me, dancing through the night holding you tight; I will treat you so right with my heart

Welsh

David Gray – Please Forgive Me (32)

for I know not what I do; feels like lightning running through my veins, every time I look at you every time I look at you

U.S.

Semisonic – Follow (29)

take me wherever you go, help me forget tomorrow; love me your best and I know, all of the rest will follow

Tagalog

Regina Velasquez – I Will Always Love You (29)

I'll be there, I'll be there waiting; I will always love you I will always take you there's no one who loves you like I do

Scottish

Garbage – Cherry Lips (29)

such a delicate boy in a hysterical realm I'm in an emotional landslide and on physical terms (29)

Spanish

Julio Iglesias – Hey (34)

it's true I've had so many other loves to share but I would close my eyes and you'd be there no matter where I go you are everywhere

Korean

BoA – Amazing Kiss (30)

amazing kiss shining stars and I'll miss the flame that used to be ours; ours alone cuz we made it happen... our miracle

Irish

The Corrs – No More Cry (31)

I'm telling you smiling for you only I am trying for you solely I'm praying for you only no more cry no more cry

Zulu

Miriam Makeba – I Shall Sing (27)

I shall sing now sing my song yes be it right and be it wrong; in the night in the day anyhow and anyway

Canadian

Chantal Kreviazuk – This Year (29)

this year I'll reach that pinnacle, this year I'll get to the top; faith only lies where she's got that energy this year I'm never gonna stop

Japanese

Shiina Ringo – Rinne Highlight (28)

sometimes I believe me believe me you and I find you're gonna left me far behind I'm trying to believe you aren't believe you

French

Vanessa Paradis – Cut Cut Brother (31)

and it's harvest time again. cut cut again my friend. no one wants to share the pain. cut cut brother. and it's harvest time again

Australian

Natalie Imbruglia – Satellite (29)

but with all this space can we keep it alive keep it alive so I'm asking you baby are you coming down tonight

British Coldplay – See You Soon (33)

in a bulletproof vest where the windows are closed I'll be telling my best I'll see you soon in a telescope lens when only one's friend

Truncated standard text spoken by all speech stimuli follows. It has 37 syllables.

Please call Stella. Ask her to bring these things with her from the store: Six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for her brother Bob.

Appendix B: Survey

Music Survey:

Please consider the first sample to be native U.S. English. For each following music sample, please answer the following questions: (you will have 10 seconds between each sample; circle your first intuition)

0. Listen to Sample. Don't write anything.

1. Is this a native speaker of English? Yes No

If yes, what do you think is the native dialect of the singer?

(Australian English, British Eng, Canadian Eng, Irish Eng, Scottish Eng, U.S. Eng, Welsh Eng)

If no, what do you think this person's native language is?

(Cantonese (Hong Kong), French, Japanese, Korean, Spanish, Tagalog (The Philippines), Zulu (Africa))

__ Check here if you think you recognize the singer/song.

(1-14 identical)

Speech Survey:

Please consider the first sample to be native U.S. English. For each following speech sample, please answer the following questions: (you will have 10 seconds between each sample; circle your first intuition)

0. Listen to Sample. Don't write anything.

1. Is this a native speaker of English? Yes No

If yes, what do you think is the native dialect of the speaker?

(Australian English, British Eng, Irish Eng, Scottish Eng, U.S. Eng)

If no, what do you think this person's native language is?

(Cantonese (Hong Kong), French, Japanese, Korean, Spanish, Tagalog (The Philippines), Zulu (Africa))

(1-12 identical)

Feedback Questions

Which section was easier for you to make native dialect/language judgments? (Circle one)

Music/speech

Did you use any strategies when making the judgments?

Any additional comments:

General Information

Your age:

Your native language(s):

Your exposure to any other languages through study or people:

What kind of music do you listen to generally? (country/genre)