

THE ELEMENT OF VOCAL DISGUISE IN SOCIOPHONETIC DECEPTION: A REPORT ON ONGOING RESEARCH

Brittany Sobol

Introduction

THE PURPOSE OF THIS PAPER is to report my current findings within my M.A. (Linguistics) research, which focuses on vocal disguise within forensic linguistics. Specifically, I aim to answer the question of what individuals are most likely to do with their voice when they do not want to be recognized by a listener. Beyond this, I am interested in whether specific sociolinguistic characteristics – gender and place of origin – have an effect on the disguise choices that individuals make.

In the following sections, I provide some background on the study of voice disguise, summarize the data collection process and methodology of my project, and finally, offer a brief explanation of my results to date.

Background

Over the last 45 years, there has been an increasing amount of research into what it means to disguise one's voice. A disguised voice, such as that used by the perpetrator of a crime (Dumas 1990; French 2013; Hollien and Majewski 1977), might be achieved through simple changes to a speaker's inherent vocal properties like pitch range, phonation type¹, and degrees of nasality. Another way to disguise one's voice is to alter the socially learned properties of one's accent by imitating features of a foreign accent. The range of accent-related options that speakers might choose from, as well as the range of variation that exists between speakers in the production of a disguised voice has not been subjected to a systematic study. In this project, I will investigate vocal disguise using the variationist sociolinguistic framework (Labov 1966; Tagliamonte 2007) by viewing disguise as a form of style-shifting (Eckert and Rickford 2001): a situational change in the grammatical, phonological, phonetic and lexical properties used by a speaker.

First, I will examine a set of vowel properties speakers alter when they disguise their own voices. Second, to determine sociolinguistic factors that

¹Phonation is the type of sound produced by one's vocal folds, such as whispering, yelling, or speaking normally.

might govern or influence phonetic shift between normal speaking styles and disguised voices, I will consider the role that a speaker's gender plays in the selection of a voice disguise: do males alter the same phonetic features that females do? Third, I plan to study how regional dialect variation is involved in disguising one's voice: do speakers of the same variety of English alter similar phonetic properties when asked to disguise their voices, and do speakers of different varieties select different properties? For the purposes of this paper, I will focus on results relating only to the first question: what vowel properties do speakers tend to alter when they are disguising their voice?

This study will contribute to our understanding of forensic linguistic theories regarding vocal deception with respect to the use of accent as a disguise. As well, it will enhance connections between the realms of forensic linguistics and sociolinguistics, in that it will consider vocal disguise as a form of situational style-shifting.

Before I turn to my study, several terms require definition. First and foremost, I use the term "modal" interchangeably with the words "normal," or "typical." That is, modal articulation is the form of articulation that a speaker uses most often when speaking normally, in a day-to-day setting. Similarly, modal voice and modal phonation are the speaking voice and phonation that a speaker would use in a typical setting. On the opposite end of the voice spectrum lies voice disguise. In this study, voice disguise assumes that an individual will change the modal qualities of their voice in order to become less identifiable by a listener.

Style-shifting and pitch

Much of the current research in vocal disguise touches on pitch change. For example, De Decker (2015: 4) found that in his style-shifting study, 42% of the shifts observed in vowels between non-quoted and quoted speech involved a shift in F0.² If we are to think of vocal disguise as a form of style-shifting, which my research aims to support, then it is plausible that a shift in pitch would also be observed in data between modal and disguised voices.

Masthoff (1996) also presents some interesting findings related to pitch in his article on vocal disguise: of the individuals who chose to disguise their voice, 31% raised their pitch and 23% lowered their pitch; the only form of disguise which was more common was whispering (1996: 165). Interestingly, Masthoff found that only males raised their pitch during disguise, while only females lowered their pitch (1996: 166).

Finally, Zhang (2007: 156) found that while using an Automatic Speaker Recognition program, lowering of pitch was the voice modification

²Within phonetics, a type of acoustic measurement used to determine the pitch of one's voice.

third most likely to result in incorrect recognition (55% correct recognition), where whispering was lowest (0% correct recognition), followed by chewing gum (45% correct recognition).

Methodology

My study consisted of 21 participants (13 women, 8 men) who took part in a vocal disguise experiment. One of the participants, Speaker 019, did not speak clearly enough for her recording to be analyzed and was excluded from the data analysis. Recruitment was completed through several means: posters were placed around the Memorial University campus, announcements were made via social media, and in-person recruitment was carried out through classroom visits.

Once enough participants were gathered, meetings were set according to participants' convenience. They were instructed to meet at Memorial's Speech Sciences and Language Acquisition Lab (SSLAL) at the designated time. Upon their arrival, participants were told to read and fill out two forms: a consent form and a brief survey indicating their gender, age, and place of origin.

The recording process was as follows: participants were brought into a sound attenuated room in the SSLAL. They were seated in front of a recorder (Audio Technica AT831b condenser microphone and Marantz PMD 670 solid state recorder, sampled at a frequency of 22kHz and 16-bit depth to uncompressed WAV format). Participants were, shown a printed version of "Comma Gets a Cure" (Honorof, McCullough and Somerville 2000) and asked to read the text twice straight through: the first time in their normal speaking voice, and the second time in a disguise of their own choosing. They were told that the disguise could be as simple or as elaborate as they wanted, but that it was to remain the same from the beginning to the end of their reading. Should a participant make a mistake while reading, they were told to go back a couple of words and continue reading. Participants were given two to three minutes to briefly familiarize themselves with the passage and were asked if they had any questions. The recording then took place. Once both recordings were completed, the researcher (author) returned to the room and turned off the recorder, and if the participant had no questions regarding their involvement, they were free to go.

Recordings were transferred from the recorder to a MacBook Pro. Each sound file was renamed to ensure both anonymity and organization (e.g. Speaker_001). Each recording of the text was broken down into individual sentences, and timestamps were added for each version of the participants' readings.

Vowel analysis focused on 15 vowels (represented by Wells' (1982) keywords THOUGHT, TRAP, STRUT, SCHWA (comma in Wells 1982), LOT, PRICE, DRESS, FACE, KIT (no stress), KIT, FLEECE (no stress),

FLEECE, GOAT, FOOT, GOOSE) and 5 properties: pitch, duration, intensity (loudness), F1³ frequency, and F2⁴ frequency. Values were obtained for each participant's vowel productions using automatic scripts that were run through Praat, a computer program for phonetic research (Boersma and Weenink, 2010). Both the disguise and non-disguise condition were statistically compared through the use of within-subjects t-tests for each acoustic and vowel normal/disguise pair via R, a statistical analysis program widely used in linguistic research (R Core Team, 2015). The condition (normal vs. disguise) served as the Independent Variable.

Results

Table 1 Breakdown of Shifts in Disguise

Vowel	Lexical Set	Pitch	Duration	Intensity	F1	F2	TOTAL	%
EH1	DRESS	13	6	7	8	8	42	11.7%
AH0	SCHWA	10	9	7	6	4	36	10.0%
IH1	KIT str.	12	3	6	4	5	30	8.3%
AA1	LOT	8	3	3	8	5	27	7.5%
AY1	PRICE	8	5	5	3	5	26	7.2%
IH0	KIT	9	6	5	3	3	26	7.2%
IY0	FLEECE	8	4	3	2	8	25	7.0%
AO1	THOUGHT	9	6	3	5	1	24	6.7%
EY1	FACE	9	3	2	6	3	23	6.4%
UW1	GOOSE	7	3	4	1	8	23	6.4%
AE1	TRAP	7	5	4	3	4	23	6.4%
IY1	FLEECE str.	5	4	5	2	3	19	5.3%
AH1	STRUT	5	4	1	2	2	14	3.9%
OW1	GOAT	4	3	2	1	4	14	3.9%
UH1	FOOT	5	0	1	1	0	7	1.9%
TOTAL		119	64	58	55	63	359	
%		33%	17.8%	16.1%	15.3%	17.5%		

Table 1 shows the rate at which each vowel category and acoustic property (pitch, duration, intensity and F1 and F2) exhibited a statistically significant difference ($p < .05$) across the two conditions. As is depicted in the table, the type of disguise most used by participants was alteration of pitch (33%). This result supports findings put forth by De Decker (2015), Masthoff (1996), and Zhang (2007), suggesting that pitch alteration is one of the most commonly used forms of vocal disguise.

If we are to think of vocal disguise as a form of style-shifting, then this result correlates well with De Decker's 2015 findings. It seems that individuals are more likely to shift their pitch when they are taking part in a form of

³Refers to the height of the tongue in the mouth.

⁴Refers to how far back or forward the tongue is in the mouth.

speech that requires them to deviate from their modal voice, be that in a quotative form⁵ or in disguise mode.

Unlike Masthoff (1996), my results do not show a clear divide across gender. Of my participants, four lowered their pitch in their disguise mode (two females and two males), while nine tended to raise their pitch (five females and four males) supporting Masthoff's finding that it is more likely an individual will raise their pitch than lower it when partaking in a disguise condition.

Further analysis will aim to discover whether one's gender or place of origin has an overall effect on how one chooses to disguise their voice. Potential patterns found in duration, intensity, F1, and F2 will be considered, as well. I will also analyze the disguises of each participant impressionistically to see if any patterns emerge in chosen disguise forms (e.g. taking on a foreign accent, whispering, etc.).

This study will add to the growing field of forensic linguistics, and help to strengthen connections between forensic linguistics and sociolinguistics. My data will be added to an online corpus (in progress) so that it may be shared and help other researchers who are interested in the parameters of vocal disguise, particularly when voice identification is not necessary. Most importantly, my research will add knowledge and understanding to what is considered a small but emerging specialty within the field of linguistics.

Brittany Sobol is an M.A. candidate
in the Department of Linguistics,
Memorial University.

⁵The form of speech used when an individual quotes another individual in conversation, such as 'And Mike said, "I love it!"'

References

- Boersma, Paul and David Weenink. 2010. Praat: Doing phonetics by computer. Available at: <http://www.praat.org/>. Accessed October 27, 2013.
- De Decker, Paul. 2015. Phonetic shifting across narrative and quoted speech styles. *Proceedings of Meetings on Acoustics* 20.1: 1-14. Acoustical Society of America.
- Dumas, Bethany K. 1990. Voice identification in a criminal law context. *American Speech* 65.4: 341-348.
- Eckert, Penelope and John R. Rickford (eds.). 2001. *Style and Sociolinguistic Variation*. Cambridge: Cambridge University Press.
- French, Peter. 2013. An overview of forensic phonetics with particular reference to speaker identification. *International Journal of Speech Language and the Law* 1.2: 169-181.
- Hollien, Harry, and Wojciech Majewski. 1977. Speaker identification by long-term spectra under normal and distorted speech conditions. *The Journal of the Acoustical Society of America* 62.4: 975-980.
- Honorof, Douglas N., Jill McCullough and Barbara Somerville. 2000. Comma gets a cure: A diagnostic passage for accent study. Available at: <http://www.dialectsarchive.com/comma-gets-a-cure>. Accessed 30 November 2013.
- Labov, William. 1966. *The Social Stratification of English in New York City*. Washington D.C.: Centre for Applied Linguistics.
- Masthoff, Herbert R. 1996. A report on a voice disguise experiment. *International Journal of Speech Language and the Law* 3.1: 160-167.
- R Core Team. 2015. R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Available at: <https://www.r-project.org/>. Accessed 01 March 2015.
- Tagliamonte, Sali A. 2007. Quantitative analysis. In Robert Bayley and Ceil Lucas (eds.), *Sociolinguistic Variation: Theories, Methods, and Applications*, 190-214. Cambridge: Cambridge University Press.
- Wells, J.C. 1982. *Accents of English*, vol. 1. Cambridge: Cambridge University Press.
- Zhang, Cuiling. 2007. An acoustic study on disguised voices. *International Journal of Speech Language and the Law* 14.1: 154-156.