ABSTRACT

PERCEPTION OF DISFLUENCIES IN ENGLISH

This study examines the effects of speaker gender and disfluency type (‘uh’ vs. ‘um’) on listeners’ perceptions of English speakers, in terms of both status and solidarity traits. Participants were 33 native speakers of American English who were asked to listen to and rate recordings of six speakers (three males, three females) on several status and solidarity traits. The recordings were spliced together to make three different versions for each speaker (‘um’, ‘uh’, and fluent). Each participant heard and evaluated one version of each of the six speakers’ utterances, using Latin-square design to ensure that each set included one recording from each of the six speakers, two of which contained ‘uh,’ two of which contained ‘um,’ and two of which were without disfluencies. The results showed a small but significant effect of gender on status perceptions, but the study did not find a significant difference in perceptions of ‘uh’ versus ‘um.’

Alyssa Nicole Tobar
December 2016
PERCEPTION OF DISFLUENCIES IN ENGLISH

by

Alyssa Nicole Tobar

A thesis
submitted in partial
fulfillment of the requirements for the degree of
Master of Arts in Linguistics
in the College of Arts and Humanities
California State University, Fresno
December 2016
APPROVED
For the Department of Linguistics:

We, the undersigned, certify that the thesis of the following student meets the required standards of scholarship, format, and style of the university and the student’s graduate degree program for the awarding of the master’s degree.

Alyssa Nicole Tobar
Thesis Author

Michael Shepherd (Chair)  Linguistics

Chris Golston  Linguistics

Jidong Chen  Linguistics

For the University Graduate Committee:

______________________________
Dean, Division of Graduate Studies
AUTHORIZATION FOR REPRODUCTION

OF MASTER’S THESIS

X I grant permission for the reproduction of this thesis in part or in its entirety without further authorization from me, on the condition that the person or agency requesting reproduction absorbs the cost and provides proper acknowledgment of authorship.

Permission to reproduce this thesis in part or in its entirety must be obtained from me.

Signature of thesis author:
ACKNOWLEDGMENTS

I want recognize and acknowledge the following individuals and institutions for making this thesis possible. To my mother, who has always supported me in everything I do. To my father, who motivated me to always do my best. To my sister, who supported me and motivated me to be the best role model I could be. To my uncle, who has always treated me as his own daughter and believed in me every step of the way. To the members of my thesis committee for all the valuable time they invested in this thesis, their inexhaustible patience, and their invaluable insight, and to my committee chair in particular, Michael Shepherd, who helped me through every step of this process and inspired me to do this thesis. To each of my peers and colleagues who helped me to keep going when I didn’t think it was possible. To my employers and work colleagues – Cheryl Chan, Kristi Miller, Katie Gorman, and Sande Wu – whose support and understanding allowed me to have both a career as a student and teacher at the same time. To all my instructors in the Linguistics Department who listened to my endless questions and concerns.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF TABLES</th>
<th>vi</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td>CHAPTER 1: INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>CHAPTER 2: LITERATURE REVIEW</td>
<td>3</td>
</tr>
<tr>
<td>Terminology</td>
<td>3</td>
</tr>
<tr>
<td>Disfluency Production</td>
<td>3</td>
</tr>
<tr>
<td>Status and Solidarity</td>
<td>4</td>
</tr>
<tr>
<td>Rationale for the Current Study</td>
<td>5</td>
</tr>
<tr>
<td>CHAPTER 3: METHOD</td>
<td>7</td>
</tr>
<tr>
<td>Participants</td>
<td>7</td>
</tr>
<tr>
<td>Materials</td>
<td>7</td>
</tr>
<tr>
<td>Procedure</td>
<td>8</td>
</tr>
<tr>
<td>CHAPTER 4: RESULTS AND DISCUSSION</td>
<td>10</td>
</tr>
<tr>
<td>Results</td>
<td>10</td>
</tr>
<tr>
<td>Discussion</td>
<td>12</td>
</tr>
<tr>
<td>CHAPTER 5: CONCLUSION</td>
<td>13</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>14</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>17</td>
</tr>
<tr>
<td>APPENDIX A: SCREENSHOT OF A STUDY ITEM</td>
<td>18</td>
</tr>
<tr>
<td>APPENDIX B: SUMMARY OF RAW RATINGS</td>
<td>20</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1  *Status and Solidarity Traits Used in the Current Study* ........................................... 4
LIST OF FIGURES

Page

Figure 1. Mean solidarity ratings for ‘uh’- and ‘um’- containing utterances by male vs. female speakers. ................................................................. 11

Figure 2. Mean status ratings for ‘uh’- and ‘um’- containing utterances by male vs. female speakers. ................................................................. 11
CHAPTER 1: INTRODUCTION

Consider the following utterance from Clark and Fox Tree (2002): “Well, I mean this uh Mallet said Mallet was uh said something about uh you know he felt it would be a good thing if uh if Oscar went.” Or this one from Liberman (2014): “I guess um you go first.” These utterances contain examples of disfluencies (‘um’, ‘uh’, ‘you know’), which are extremely frequent in everyday speech. In fact, Liberman’s (2014) analysis of a corpus of 26.1 million words of spoken American English shows that if ‘uh’ and ‘um’ are counted as words, they together account for 2.00% of men’s speech and 1.45% of women’s speech. Liberman also found some striking gender differences. Not only are ‘uh’ and ‘um’ about 38% more frequent in men’s speech than in women’s (as shown above), ‘uh’ is about 25% more frequent in men’s speech than in women’s speech, and ‘um’ is about 22% more frequent in women’s speech than in men’s speech.

Other research on the production of disfluencies has focused on the circumstances in which people use them. These studies have shown that native speakers deploy disfluencies in instances of cognitive or planning difficulty (Bortfeld, Leon, Bloom, Schober, & Brennan, 2001; Finlayson & Corley, 2012), and as a signal to new information (Arnold, Fagnano, & Tanenhaus, 2003; Bailey & Ferreira, 2003). Still other studies (Clark & Fox Tree, 2002; Fox Tree, 2007) have examined listeners’ beliefs about the functions of disfluencies. According to these studies, listeners tend to perceive disfluencies as marking uncertainty, delay, unfamiliarity, etc. It remains unclear, however, whether different disfluencies are perceived as signaling different degrees of uncertainty or what affects these perceptions. Moreover, no study has examined listeners’ perceptions of speakers who deploy different disfluencies nor the factors that affect these perceptions. The
current study examines listeners’ perceptions of disfluency use by male and female speakers in terms of status and solidarity traits.
CHAPTER 2: LITERATURE REVIEW

Terminology

The bits of speech at the heart of the present study have been referred to in various ways in the literature, including *disfluencies* (Arnold et al., 2003; Bailey & Ferreira, 2003; Bortfeld et al., 2001; Finlayson & Corley, 2012; Laserna, Seih, & Pennebaker, 2014; Merlo & Mansur, 2004), *(speech/disfluent) fillers* (Clark & Fox Tree, 2002; O’Connell & Kowal, 2005), *delay markers* (Fox Tree, 2001; Kam & Edwards, 2008), *pragmatic markers* (Tottie, 2014), *interjections* (Myers, Bakker, St. Louis, & Raphael, 2012), and *filled pauses* (Fox Tree, 2007; Laserna et al., 2014; Tottie, 2011). The current study follows the majority in adopting the term “disfluencies.”

Disfluency Production

Most research on disfluencies has focused on their production, including why and how speakers use them. Tottie’s (2011) study of 10.4 million words of formal and informal spoken English from the British National Corpus found that men deploy more disfluencies than women in both formal (1.90% vs. 1.35%) and informal (1.00% vs. 0.81%) situations. Moreover, Tottie (2011) found that speakers over 60 years old deploy fillers more than younger speakers (1.24% vs. 0.77% for speakers 0-14 and 0.80% for speakers 15-24). Tottie (2011) also found, based on a 138,296-word subportion of the corpus for which the speakers’ socioeconomic status was recorded, that higher-SES speakers deploy more fillers than lower-SES speakers (0.98% vs. 0.78%).

Tottie (2014) later sought to compare the relative frequency of disfluencies in American and British English. Based on 62,350 words from the Santa Barbara Corpus of Spoken American English, he concluded that American speakers tend to
deploy ‘uh’ and ‘um’ much less frequently than British speakers (0.75% vs. 1.45%). It should be noted, however, that Tottie’s 62,350-word American corpus is tiny by corpus linguistics standards (cf. his 10.4 million words of spoken British English), Liberman’s (2005, 2014) analyses of 23.1 million words of spoken American English put the rate of disfluencies in American English at 1.7%, which is much more similar to Tottie’s estimate for British English.

**Status and Solidarity**

The present study is concerned with listeners’ perceptions of speakers’ use of disfluencies, and the perceptions of this type are conventionally measured on a mix of status and solidarity scales (Brown, 1965). According to Georgieva (2014), solidarity is “the dimension that accounts for the level of co-operation and social harmony amongst speakers and status refers to “social factors as ‘social status’ and ‘power’” (p. 13). Status traits used in previous language-perception studies include “educated” and “wealthy” (from Ryan & Carranza, 1975) and “confident,” “successful,” and “ambitious” (from Stewart, Ryan, & Giles, 1985), as well as “intelligent” and “successful” (used in both studies). For solidarity traits, both Ryan and Carranza and Stewart et al. used “friendly,” “kind,” and “trustworthy,” with Stewart et al. also adding “sincere.” The status and solidarity traits used in the current study are summarized in Table 1.

<table>
<thead>
<tr>
<th>Status</th>
<th>Solidarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>wealthy</td>
<td>honest</td>
</tr>
<tr>
<td>intelligent</td>
<td>trustworthy</td>
</tr>
<tr>
<td>educated</td>
<td>kind</td>
</tr>
<tr>
<td>confident</td>
<td></td>
</tr>
</tbody>
</table>
Rationale for the Current Study

Although there is no definitive work on perceptions of disfluencies, Fox Tree (2007) did look at folk notions about ‘uh,’ ‘um,’ ‘you know,’ and ‘like.’ In this study, 105 undergraduates were given a 14-item questionnaire about their attitudes toward, and use of, these disfluencies. That study revealed very specific folk notions about how different disfluencies are deployed and what they signal. For example, ‘uh’ and ‘um’ were thought to indicate problems with production and nervousness. If people possess folk notions about the deployment of different disfluencies, these notions may also have an impact on their perception, i.e., if people believe that someone will deploy a certain disfluency to indicate nervousness then they may also believe that the person using it lacks confidence. Overall, Fox Tree (2007) found that disfluencies are generally perceived negatively, as indicating, for example, that someone needs more time to think or is nervous. Because disfluencies are generally seen as something to avoid, speakers’ use of disfluencies will also likely tend to be negatively interpreted on certain scales (confidence, intelligence, etc.).

The current study seeks to further examine the perceptions of disfluencies (specifically ‘uh’ and ‘um’) by testing listeners’ perceptions of utterances containing ‘uh’ or ‘um.’ Given Liberman’s (2005, 2014) findings that there are gender differences in the production of ‘uh’ and ‘um’, the current study also hopes to find a similar pattern in listeners’ perceptions of these disfluencies.

This thesis examines the effects of speaker gender and disfluency type (‘uh’ vs. ‘um’) on listeners’ perceptions of English speakers, in terms of both status (intelligence, education, wealth, and confidence) and solidarity (honesty, trustworthiness, and kindness). Three other scales – funny, professional, and
interesting – were treated as distractors. The hypotheses behind this study are as follows:

1). There will be a significant difference in listeners’ perceptions of male versus female speakers’ use of ‘uh’- versus ‘um’-containing utterances.

2). Participants will rate ‘um’ higher than ‘uh’ on the status scales.

The first hypothesis was inspired by Liberman (2005, 2014), who found widely different rates of ‘uh’ and ‘um’ usage by male versus female speakers. The second hypothesis was based on a pilot study that exploring general differences between ‘uh’ and ‘um,’ which showed that participants tend to rate speakers using ‘um’ higher on the status trait of wealth.
CHAPTER 3: METHOD

Participants

A total of 33 participants were recruited from a linguistics course at a large public university in Central California. There were 13 males and 20 females, with an average age of 23 (range 18-45). Data from 8 additional participants were excluded due to those participants’ being non-native English speakers \((n = 5)\), guessing the purpose of the study \((n = 2)\), or being noticeably distracted while completing the study \((n = 1)\).

Materials

To develop the stimuli, six speakers (three self-identified males, three self-identified females) were asked to watch a short animated film from the series “Simon’s Cat,” after which they were audio-recorded (using a Samsung Galaxy 3 smartphone) answering eight open-ended questions about the film. This yielded 3-5 minutes of audio from each speaker. The next step was to develop a set of three natural-sounding and otherwise-identical stimulus recordings for each speaker that differed only in their containing an ‘uh,’ an ‘um,’ or no disfluencies. This was done by using Praat (Boersma & Weenink, 2016) to splice spontaneous instances of ‘uh’ and ‘um’ from each speaker’s original recording into otherwise-fluent 2-3 second utterances selected from elsewhere in the same recordings. The final stimulus set thus consisted of 18 audio recordings – one containing ‘uh,’ one containing ‘um,’ and one without disfluencies for each of the six speakers. To verify that the spliced recordings were as natural as possible, a volunteer was asked to listen to the recordings and identify which ones had been altered through splicing. The volunteer was unable to determine which recordings had been altered or in what way, suggesting that all 18 recordings were sufficiently natural.
Note that the version without disfluencies was only included as a distractor. Without it, all stimulus items would have contained disfluencies, and participants might have noticed that.

The 18 stimulus recordings were arranged into three sets of six using Latin-square design to ensure that each set included one recording from each of the six speakers, two of which contained ‘uh,’ two of which contained ‘um,’ and two of which were without disfluencies. The presentation order of each stimulus set was determined using restricted randomization (Bailey, 1983; Youden, 1972) to avoid two consecutive stimuli representing the same disfluency type or three consecutive stimuli representing the same speaker gender. Each of the three restricted randomizations was presented both forward and backward, yielding a total of six versions. At the end of the study, participants were asked what they believed its purpose was. Only two participants guessed the purpose of the study, and their data were excluded from the analysis.

Procedure
Each participant was randomly assigned one of the six versions by Qualtrics online survey software (Qualtrics, 2016). The software also presented the instructions and auditory stimuli and recorded the participants’ responses. Despite the study being Internet-based, participants completed it in a quiet office, in the presence of the researcher or an assistant, and using a researcher-provided laptop and Sony MDR 7506 professional headphones.

After listening to each stimulus recording, participants were asked to rate the speaker on a set of 10, 5-point semantic differential scales, four with endpoint labels representing status traits (educated, intelligent, confident, professional, interesting, and wealthy) three with endpoint labels representing solidarity traits
(honest, trustworthy, and kind), and three that were treated as distractors (funny, professional, and interesting). A screenshot of a study item, including the 10 scales, can be found in Appendix A. Each scale had five levels – ranging from “not at all” to “extremely” – with only the first and last points labeled.
CHAPTER 4: RESULTS AND DISCUSSION

Results

The aim of this study was to examine (a) whether there was a significant difference in listeners’ perceptions of male versus female speakers’ use of ‘uh’-versus ‘um’-containing utterances and (b) whether participants rated ‘um’-containing utterances higher than ‘uh’-containing utterances on status scales.

In regards to the first hypothesis, there was indeed a very nearly significant difference in perceptions of ‘uh’ and ‘um’ based on speaker gender. There was a very nearly significant difference in the ratings between genders, as shown in Figures 1 and 2, which show female speakers being rated higher when using ‘uh’ than when using ‘um’; however, male speakers were rated higher when using ‘um’ than when using ‘uh.’ For a summary table of raw ratings by gender, disfluency, and trait, see Appendix B.

A two-way ANOVA comparing ‘uh’ versus ‘um’ and ‘male’ versus ‘female’ speaker for the status ratings showed a very nearly significant interaction of gender and disfluency $F(1, 524) = 3.72, p = .054$, with ‘uh’-using female speakers and ‘um’-using male speakers receiving the most favorable ratings. Another two-way ANOVA, comparing ‘uh’ versus ‘um’ and ‘male’ versus ‘female’ speaker for solidarity ratings, showed a significant effect of speaker gender $F(1, 392) = 5.09, p = .02$, with female speakers being rated as more honest, trustworthy, and kind than their male counterparts.

The second hypothesis – that speakers’ using ‘um’ would be rated higher on status scales than speakers’ using ‘uh’ – was not supported. A $t$-test comparing average ‘um’ and ‘uh’ ratings showed no difference, $t(394) = .85, p = .39$. 
**Figure 1.** Mean solidarity ratings for ‘uh’- and ‘um’- containing utterances by male vs. female speakers.

**Figure 2.** Mean status ratings for ‘uh’- and ‘um’- containing utterances by male vs. female speakers.
Discussion

The results of the current study show that while there was no significant difference in perceptions of speakers’ using ‘uh’-containing versus ‘um’-containing utterances, there was indeed a very nearly significant difference between the perceptions of the different genders based on disfluency type. This difference aligns with Liberman’s (2014) finding that there are gender differences in the deployment of ‘uh’ and ‘um.’ The results are also consistent with Fox Tree’s (2007) findings, insofar as both Fox Tree’s study and the current one show that listeners have different perceptions of different disfluency types. Ultimately, given the exploratory nature of the present study, it is difficult to interpret these results further.
CHAPTER 5: CONCLUSION

The general trend that emerged from this study was that ‘uh’-containing utterances were evaluated more favorably than ‘um’-containing utterances when spoken by females, but that the opposite was true for utterances spoken by males. However, while this trend was very nearly significant for status ratings, it was not found to be significant for solidarity ratings.

This pattern suggests that, for status scales, at least, ‘uh’ and ‘um’ are perceived differently; however, it is difficult to account for this difference in perception due to the exploratory nature of the present study. Most studies have treated ‘uh’ and ‘um’ as generally the same, but the present findings suggest that perceptions of ‘uh’ and ‘um’ may not be so similar. Based on these observations, I suggest that future research examine both production and perception of these two disfluencies simultaneously and do so for a specific group (in terms of age, region, etc.). Otherwise, we risk comparing production data from one demographic with perception data from another. Such research would allow more definitive conclusions about how these disfluencies differ, as well as the circumstances in which these disfluencies are found and how factors such as gender, age, etc. influence their deployment.
REFERENCES
REFERENCES


APPENDICES
APPENDIX A: SCREENSHOT OF A STUDY ITEM
Please listen to the following recording and rate the speaker on the scales below.

- Not at all educated
- Not at all honest
- Not at all trustworthy
- Not at all kind
- Not at all intelligent
- Not at all wealthy
- Not at all funny/humorous
- Not at all confident
- Not at all professional
- Not at all interesting

- Extremely educated
- Extremely honest
- Extremely trustworthy
- Extremely kind
- Extremely intelligent
- Extremely wealthy
- Extremely funny/humorous
- Extremely confident
- Extremely professional
- Extremely interesting
APPENDIX B: SUMMARY OF RAW RATINGS
<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th>Male</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>‘uh’</td>
<td>‘um’</td>
<td>‘uh’</td>
<td>‘um’</td>
</tr>
<tr>
<td><strong>Status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>wealthy</td>
<td>3.09</td>
<td>2.88</td>
<td>2.91</td>
<td>2.91</td>
</tr>
<tr>
<td>intelligent</td>
<td>3.30</td>
<td>3.06</td>
<td>3.03</td>
<td>3.30</td>
</tr>
<tr>
<td>educated</td>
<td>3.18</td>
<td>3.15</td>
<td>3.12</td>
<td>3.33</td>
</tr>
<tr>
<td>confident</td>
<td>3.12</td>
<td>3.00</td>
<td>2.64</td>
<td>2.97</td>
</tr>
<tr>
<td><strong>Solidarity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>honest</td>
<td>3.48</td>
<td>3.58</td>
<td>3.33</td>
<td>3.45</td>
</tr>
<tr>
<td>trustworthy</td>
<td>3.42</td>
<td>3.30</td>
<td>2.91</td>
<td>3.21</td>
</tr>
<tr>
<td>kind</td>
<td>3.48</td>
<td>3.36</td>
<td>3.12</td>
<td>3.33</td>
</tr>
</tbody>
</table>
Non-Exclusive Distribution License
(to archive your thesis/dissertation electronically via the library’s eCollections database)

By submitting this license, you (the author or copyright holder) grant to Fresno State Digital Scholar the non-exclusive right to reproduce, translate (as defined in the next paragraph), and/or distribute your submission (including the abstract) worldwide in print and electronic format and in any medium, including but not limited to audio or video.

You agree that Fresno State may, without changing the content, translate the submission to any medium or format for the purpose of preservation.

You also agree that the submission is your original work, and that you have the right to grant the rights contained in this license. You also represent that your submission does not, to the best of your knowledge, infringe upon anyone’s copyright.

If the submission reproduces material for which you do not hold copyright and that would not be considered fair use outside the copyright law, you represent that you have obtained the unrestricted permission of the copyright owner to grant Fresno State the rights required by this license, and that such third-party material is clearly identified and acknowledged within the text or content of the submission.

If the submission is based upon work that has been sponsored or supported by an agency or organization other than Fresno State, you represent that you have fulfilled any right of review or other obligations required by such contract or agreement.

Fresno State will clearly identify your name as the author or owner of the submission and will not make any alteration, other than as allowed by this license, to your submission. By typing your name and date in the fields below, you indicate your agreement to the terms of this distribution license.

Embargo options (fill box with an X).

- [X] Make my thesis or dissertation available to eCollections immediately upon submission.
- [ ] Embargo my thesis or dissertation for a period of 2 years from date of graduation.
- [ ] Embargo my thesis or dissertation for a period of 5 years from date of graduation.

Alyssa Nicole Tobar
Type full name as it appears on submission

November 1, 2016
Date