

Folk notions of *um* and *uh*, *you know*, and *like**

JEAN E. FOX TREE

Abstract

The current study measures laypeople's uses of um, uh, you know, and like, including folk notions of meanings, self-assessments of use, history of discussing use, and attitudes toward the words. Unlike the prevalent idea in the popular press that these discourse markers are interchangeable speaker production flaws, respondents in this study demonstrated that people do possess folk notions of meanings and uses that dramatically distinguish markers from each other. Um and uh were thought to indicate production trouble, you know was thought to be used in checking for understanding and connecting with listeners, and like defied definition. The folk notions of um, uh, and you know accord well with researchers' ideas about the meanings of these words. The use of like may be too subtle for laypeople to articulate. Most researchers' views of like involve some kind of discrepancy between what's said and what's meant. Even if they cannot state a meaning, people do treat the different markers differently.

Keywords: discourse markers; fillers; you know; like; spontaneous speech; meaning.

1. Introduction

One popular idea of what *um* and *uh*, *you know*, and *like* mean is that they mean nothing and are the hallmarks of youth. In a *Time* article, Monica Lewinsky was described as 'young, all right, with the lingering baby fat and the uhs and you-knows of a teenager'.¹ In a *Doonesbury* comic, a press reporter was drawn interviewing a 'Mystery Martyr': 'You're a teenaged girl? Unbelievable!' 'Why's that?' says the Mystery Martyr. Answers the reporter, 'You don't use the word "like".'²

Um/Uh, you know, and like are often treated as a trio to avoid at all costs. In the following, a communications professor is quoted describing how ‘filler words are representative of poor communications skills’, highlighting these particular three (Lynch 2002):

‘The use of filler words (“like,” “you know,” “umm” and “you know what I mean?”) has always been a problem, and I find that much of the time, the students who use them the most do not even realize they are doing it,’ he said. ‘It has become a way of speech because it is easy, it is the path of least resistance. I would even go so far as to say that there is a correlation in our culture between communication skills and character development.’

A Web search will pop up many sites advocating the elimination of these words, with *um/uh, you know, and like* almost always represented, as in the following (Berkley 2002):

Verbal viruses are meaningless fillers that speckle our speech, distract from your message, drain our impact and annoy listeners. I call them verbal viruses because they seem to be contagious and we pick them up without being aware of it. The most common verbal viruses are: ‘uh’ ‘um’ ‘like’ ‘you know’ ‘well’ ‘okay’ and ‘sort of’. They also include annoying mouth sounds and lip smacks.

Because these three expressions are frequently lumped together as interchangeable and useless, with no meaning and no function, they formed the most diagnostic set to test whether people do think of discourse markers as interchangeable and useless, or not. The current study explores whether these top irritants might actually have differentiable meanings or functions.

Unlike the general public, most researchers who have studied *um, uh, like, and you know* agree that they are meaningful and functional, although there is ongoing debate about what they mean and how they are used, and how interchangeable they are. There is a steadily increasing amount of research work on each of these markers, with the most work done on *um* and *uh*, followed by *like*, and the least on *you know*. Some of the proposed meanings are quite different from each other, while others overlap. I will highlight here some of the diverse research findings for each of the markers, but refer the reader in search of a more thorough understanding to review articles: for *ums* and *uhs*, see Clark and Fox Tree (2002); for *like*, see Fuller (2003); and for *you know*, see Fox Tree and Schrock (2002).

Um and *uh* have been described as (i) indicators of upcoming delay (Clark and Fox Tree 2002; Smith and Clark 1993), (ii) indicators of speech production trouble (Brotherton 1979; Reynolds and Paivio 1968), (iii) indicators of discourse newness (Arnold et al. 2003), (iv) indicators of

dishonesty (Hosman and Wright 1987; Vrij and Winkel 1991), and (v) reflections of speakers' greater options among words (Schachter et al. 1991), among others. There is evidence that *um* and *uh* should be treated as different markers (Clark and Fox Tree 2002; Fox Tree 2001; Smith and Clark 1993), with *um* predictive of a longer upcoming delay than the *uh*. For the purposes of the current study, however, *um* and *uh* were treated as one group; the hypothesized meanings seemed too similar for the sensitivity of the method. That is, participants in the current study were unlikely to distinguish between the two in answering survey questions about the meanings and uses of the words.

Like has been described as (i) an indicator of loose use of language (Andersen 1998, 2000), (ii) a hedge or approximator (Jucker et al. 2003; Underhill 1988), (iii) an indicator of problems with word choice (Siegel 2002), (iv) a focuser (Underhill 1988), and (v) an indicator that what's said is different from what's meant (Schourup 1985), among others.

You know has been described as (i) an invitation to fill out the speaker's meaning (Jucker and Smith 1998), (ii) a moderator of mood (Östman 1981; Stenström 1990), (iii) a device to decrease social distance (Stubbe and Holmes 1995), (iv) a device to avoid a pause (Schourup 1985), and (v) a device to stall for time when faced with a speech production problem (Erman 1987; Holmes 1986), among others.

Despite researchers' outlining all the ways *um*, *uh*, *like*, and *you know* may be useful in communication, popular press accounts persist in propagating the view that these expressions are first and foremost speaker production flaws. This tension is mirrored in other discussions of word meaning. For example, most people considered that the new millennium began January 2000, but some reasoned that people should have taken it as beginning January 2001 (Butters 2000). Linguists and lexicographers might consider a word's meaning to be one thing, but looking at how people use a word may reveal a different meaning (Lawrence 2002). At the same time, intuitions are also not always on the mark; people consider *plural* to mean 'more than one', but evidence such as *.8 pounds* suggests a better definition would be 'not singular' (Lawrence 2002: 325).

The discussion of discourse marker meaning and use can be thought of as a magnified version of the discussion of ordinary word meaning. People can conceive of markers as part of a regional dialect (Watts 1989), with the concomitant much documented strong attitudes people hold toward dialect variation (Preston 1998, 2000)—and an added twist: people harbor negative views of the words, and stigmatize those who use them, while at the same time being unaware of their own uses (Watts 1989). Butters (2002: 328) recommends that researchers 'make use of all three kinds of data—the linguists' intuitions, the informants' intuitions, and

objective empirical data—in arriving at the most accurate conclusions about the meanings that exist “in the human brain”’.

The difference in researchers’ and laypeople’s approaches to discourse markers is a particularly dramatic example of a clash of beliefs about meaning and use. Can these conflicting views be reconciled? On the one hand, most people have little linguistic insight into the workings of their languages, so researchers can be legitimately unconcerned about laypeople’s disparaging attitudes toward these discourse markers. The fact that people think *um*, *uh*, *you know*, and *like* are useless does not mean that they are. But the nagging question remains that if speakers are using these markers for production or comprehension purposes, at some level they should be aware of what they are using these words to do. The current study measures laypeople’s understanding of *um*, *uh*, *you know*, and *like* to compare the popular and scientific views of these words. Folk notions of meanings and uses, self-assessments of use, history of discussing use, and attitudes toward the words were measured.

2. Method

2.1. *Participants*

One hundred and five University of California Santa Cruz undergraduate students who were native English speakers participated in exchange for course credit, such that 35 people responded to each questionnaire centering on one of the three markers (grouping *um* and *uh* together). They were fairly uniform in age, from 17 to 26 with a mean of 19. They were 82% female, reflecting the proportion of women in our major. The majority, 90%, were from California, with 63% of the Californians from Northern California, reflecting the location of Santa Cruz. The majority, 72%, reported that they spoke another language, with varying degrees of proficiency. The proportion of men to women, their ages, and the proportion of monolinguals were similar across the three discourse marker questionnaires.

2.2. *Research design, including coding categories for questionnaire responses*

Three 14-item questionnaires were prepared consisting of 10 questions about the discourse markers of interest (either *um* and *uh*, *like*, or *you know*) and the 4 demographic questions discussed in the participants section (see Appendix). The questionnaires were administered between participants with 35 people per questionnaire. Each participant was asked

to fill out the questionnaire in its entirety. Questionnaires took between 10 and 20 minutes to complete. For the *like* questionnaire, people were given two examples of the kind of *like* to be considered, 'we were like sitting by the train tracks' and 'the bushes were like full of thorns and stuff'. For the *you know* questionnaire, people were given the example 'he just sings his music and you know dances around a bit'. The examples were taken from a corpus of spontaneous stories told by UCSC students.

For open-ended questions, coding categories were allowed to emerge from the data. That is, similar responses were grouped together and idiosyncratic responses were lumped into an *other* category. This was a risky approach, as responses could potentially vary so greatly that no distinct pattern would emerge. As will be seen, however, patterns did emerge for all open-ended questions except the one probing how use varies across situations; in this one case, answers were too disparate to code as anything other than *yes* or *no*.

2.2.1. *Self-assessments of use.* Self-assessment of use was measured with the questions (1) 'How frequently do you use "um" or "uh" [like/you know]?' with responses on a 1 to 5 scale labeled *never*, *rarely*, *sometimes*, *often*, and *all the time* (an option of *I don't know* was also provided; these responses were not included in the analyses); (2) 'Does your use of "um" or "uh" [like/you know] vary depending on whether you are talking to a professor, to a friend, to a family member (or to someone else), or do you think you speak about the same to everyone?' with responses coded as (a) same with everyone, (b) more with friends and/or less with authority, (c) less with friends and/or more with authority, (d) sometimes more, sometimes less with friends/authority, (e) other, (f) no response or ambiguous response; and (3) 'Does your use of "um" or "uh" [like/you know] vary depending on the situation you are in, or do you think you speak about the same no matter what the situation?' with responses coded as *yes* or *no*.

2.2.2. *History of discussing use.* History of discussing use was measured by (1) 'Have you discussed the use of "um" or "uh" [like/you know] in a formal setting such as a class, lab, or section?' and (2) 'Have you discussed the use of "um" or "uh" [like/you know] in an informal setting such as with friends or family?' Responses were coded as *yes*, *no*, or *other*.

2.2.3. *Attitudes.* Attitudes were measured by (1) asking participants to check off which one of the following statements they agreed with most: (a) Communication would be better if people didn't use 'um' or 'uh'

[like/you know], (b) Communication would be worse if people didn't use 'um' or 'uh' [like/you know], (c) Saying 'um' or 'uh' [like/you know] doesn't affect communication, (d) I don't know whether communication is affected by saying 'um' or 'uh' [like/you know] or not; and by (2) asking participants to check off which one of the following statements they agreed with most: (a) There are times when I try to avoid using 'um' or 'uh' [like/you know], and I succeed, (b) There are times when I try to avoid using 'um' or 'uh' [like/you know] but it doesn't work, (c) I don't try to avoid using 'um' or 'uh' [like/you know].

2.2.4. *Folk notions.* Folk notions of what the words mean were assessed with the questions (1) 'What, if anything, do you think "um" or "uh" [like/you know] means?' and (2) 'Why do you think people use "um" or "uh" [like/you know]?' People were also asked to describe when they themselves used the discourse marker most frequently.

As can be expected for open-ended questions, many of the responses were idiosyncratic. For example, one student wrote that *like* is used to make fun of people, another wrote that it is used to break ideas into sub-units, and a third wrote that it is used 'to "add meaning" to what they say'. Similarly, *you know* proposals included 'fills in the blank for details unremembered or unnecessary to describe', 'make[s] life easier, simpler', and '[indicates] we have conveyed a story sufficiently'. When describing personal uses, people reported that *like* is used when a person is 'hyper, happy, just fooling around', *um* is used when a person is 'engaged in rapid dialogue', and *you know* is used when 'speaking off the top of my head'.

Nonetheless, many responses were not idiosyncratic and the following coding categories emerged for markers' meanings: (a) 'I don't know', (b) 'nothing', or 'it's a filler', (c) 'it means the same as X' where X could be *um*, *you know*, *like*, *hella*, *dude*, *I mean*, or *and* (that is, a specific word and not merely the catch-all label *filler*), (d) the speaker needs time to think, remember, or find words, (e) the speaker is checking for the listener's understanding or connecting with listeners, (f) a combination of meanings listed above, (g) other meanings (the idiosyncratic responses), and (h) no response, or only the propositional meaning of the words (e.g., *like* means 'to compare two things').

In answer to the question about use, as opposed to meaning, people's answers fell into two additional categories, (i) because they are nervous, insecure, or embarrassed, and (j) because others use it and/or 'it's a habit'. The use of *you know* to leave information off-record was suggested by two respondents, but because it was an infrequent response (despite being the hypothesized use of several researchers'), this response was coded as idiosyncratic.

Responses to the question about personal use were coded as (a) 'I don't know', (b) 'I never say it', (c) when trying to remember, planning what to say, or having difficulty speaking, (d) when telling a story with friends and feeling comfortable, (e) when nervous, (f) when explaining or describing, (g) a combination of circumstances listed above, and (h) other responses.

Each respondent's answers were assigned to only one category. This simplifies comparison across markers. The only exceptions to this rule were combinations involving the meaning or use *time to think*. This was the only meaning or use to frequently co-occur with others. Specifically, among the meaning responses there was one case of *holding the floor* and *time to think* (*um*) and five cases of *nervousness* and *time to think* (four *ums* and one *you know*). Among the use responses, there were three cases of *holding the floor* and *time to think* (two *ums* and one *like*), and two cases of *nervousness* and *time to think* (both *ums*). Because of their prevalence, these combinations, and only these combinations, were grouped together under the category *time to think*.

3. Results

3.1. *Self-assessments of use*

On average, people used a discourse marker somewhere between 'sometimes' and 'often'; 3.6 (0.82 SD); only three people said 'I don't know'. There were no differences across discourse markers, $F(2, 99) = 0.19$, $P = 0.83$.

Most respondents, 66%, claimed to use these discourse markers more with friends and less with authority figures. Only 12% claimed the reverse, and 10% claimed their uses were the same across addressees (the remaining 10% either did not answer, or gave ambiguous or unusual answers). The people who described the words as more likely with friends generally focused on the less-guarded nature of casual conversation. The people who described the words as more likely with authority figures generally focused on increased nervousness leading to increased markers. One person, who completed the *um/uh* survey, described both these scenarios. People viewed *like* and *you know* from the standpoint of being more likely with friends (26 and 28 people) versus more likely with authority figures (2 people and 1 person). *Ums* and *uhs*, on the other hand, were split across these viewpoints (15 people versus 10 people).

In addition to indicating that they used these discourse markers more often with some addressees than others, almost everyone, 91%, indicated that they used the words more often in some situations. This percentage

was similar across discourse markers. The situations people described varied greatly, precluding a categorization scheme.

3.2. *History of discussing use*

On average, 14% of respondents had discussed the use of a discourse marker in a formal setting, and 66% had discussed it in an informal setting. In a formal setting, most had discussed *um/uh* (26%), fewer had discussed *like* (13%), and the least had discussed *you know* (3%; χ^2 analysis could not be performed because of low cell values). There was a similar pattern for the informal setting: most had discussed *um/uh* (91%), fewer had discussed *like* (56%), and even fewer had discussed *you know* (47%), $\chi^2(2) = 16.6$, $P < 0.001$. So, people are more likely to have talked about the use of *ums* and *uhs*, then *likes*, and lastly *you knows*.

3.3. *Attitudes*

With respect to whether discourse markers helped or hindered communication, most respondents, 60%, said either that they did not know (35%) or that discourse markers had no effect (25%). About a third, 36%, thought they hindered communication. A small minority, 4%, wrote that they helped. Responses were similar across discourse markers, $\chi^2(4) = 7.2$, $P = 0.13$ (to meet the assumptions of χ^2 , the *help* category was not included in the analysis).

About two-thirds of respondents tried to avoid these discourse markers, with more claiming success (73% of those who try) than failure at this endeavor. There was a difference however in which markers respondents tried to avoid, with 77% of respondents trying to avoid *like* or *um/uh*, and only 43% trying to avoid *you know*, $\chi^2(4) = 13.8$, $P = 0.008$. So, the markers people discussed more were the ones they were most concerned about avoiding in speech.

3.4. *Folk notions*

Although there was great variation in the meanings people assigned to *um* and *uh*, *you know*, and *like*, patterns emerged (see Table 1, which summarizes answers to the question ‘What, if anything, do you think “um” or “uh” [like/you know] means?’). Each participant’s data is represented once in the table; so, there are 35 *um/uh* responses, 35 *you know* responses, and 35 *like* responses. *Um* and *uh* were described as indicating a need for time to think by 89% of respondents; only 4% of the other two discourse markers were assigned this meaning. *You know* was described

Table 1. *The number of times respondents identified a particular meaning*

| | <i>um</i> or <i>uh</i> | <i>you know</i> | <i>like</i> |
|---|------------------------|-----------------|-------------|
| 'It means nothing' | 4 | 1 | 7 |
| 'The same as <i>um</i> / <i>you know</i> / <i>like</i> /etc.' | 0 | 2 | 11 |
| Speaker needs time to think | 31 | 1 | 2 |
| Speaker is checking for understanding | 0 | 26 | 3 |
| Other | 0 | 5 | 12 |

Table 2. *The number of times respondents identified a particular use for others*

| | <i>um</i> or <i>uh</i> | <i>you know</i> | <i>like</i> |
|---|------------------------|-----------------|-------------|
| 'It means nothing' | 7 | 0 | 1 |
| 'The same as <i>um</i> / <i>you know</i> / <i>like</i> /etc.' | 0 | 2 | 3 |
| Speaker needs time to think | 24 | 0 | 3 |
| Speaker is checking for understanding or connecting with listener | 0 | 17 | 1 |
| Because others use it | 0 | 1 | 12 |
| Out of nervousness, insecurity, or embarrassment | 1 | 4 | 1 |
| Other | 3 | 11 | 14 |

Table 3. *The number of times respondents identified a particular use for themselves*

| | <i>um</i> or <i>uh</i> | <i>you know</i> | <i>like</i> |
|--------------------------------------|------------------------|-----------------|-------------|
| I don't know | 5 | 7 | 6 |
| When I am having difficulty speaking | 14 | 0 | 4 |
| When I am comfortable speaking | 0 | 9 | 13 |
| When I am nervous | 13 | 2 | 3 |
| When I am explaining or describing | 0 | 11 | 2 |
| Other | 3 | 6 | 7 |

as being used to check for listener understanding by 77% of respondents; only 4% of the other two discourse markers were assigned this meaning. *Like* was described as meaning nothing or meaning the same as some other marker by 51% of respondents; only 10% of the other two discourse markers were assigned one of these meanings. In this table and in Tables 2 and 3, the *other* responses included no responses, combination responses, and idiosyncratic responses.

Similar patterns emerged when people ascribed uses to others; 69% of respondents described *um* and *uh* as indicating a need for time to think, 49% described *you know* as being used to check for listener understanding, and 34% described *like* as being used out of habit, which makes sense given they could not come up with a meaning for *like* (see Table 2, which

summarizes answers to the question ‘Why do you think people use “um” or “uh” [like/you know]?’).

Though diffused, the pattern was still discernable when people described the reasons why they used the discourse markers. People were more likely to describe their use of *um* and *uh* as a result of speech production trouble or nervousness and their use of *like* and *you know* as a result of comfort with speaking and explaining/describing (see Table 3). As with the questions about meaning and other’s use, responses for *like* varied more than for *um* or *uh* or *you know*.

4. Discussion

Almost every survey respondent admitted to using the marker they were evaluating at least sometimes, usually with friends, and almost everyone believed they could control their use of each marker to some extent under different circumstances. Although people claimed to use *um*, *uh*, *you know*, and *like* equally frequently—rating their uses, on average, between ‘sometimes’ and ‘often’—they differed by marker in their circumstances of use, their history of discussing uses, and their attitudes. They described themselves as using *like* and *you know* mostly with friends, but *um* and *uh* with both friends and authority figures. Although not frequently discussed in school, many people reported discussing their use of these words with friends and family. *Um* and *uh* were the most popular topic, then *like*, then *you know*. Two-thirds of respondents tried to avoid using these words (usually *um*, *uh*, and *like*), and most were successful. But only one-third thought the words hindered communication.

Unlike the prevalent idea that the markers are interchangeable speaker production flaws, respondents in this study demonstrated that people do possess folk notions of meanings and uses that dramatically distinguish markers from each other. *Um* and *uh* were thought to indicate production trouble, *you know* was thought to be used in checking for understanding and connecting with listeners, and *like* defied definition. People’s hypothesized meanings and the uses they ascribed to others were well aligned. There was more slippage when people described their own uses.

Overall, though, folk notions of the uses for *um*, *uh*, and *you know* accord with what researchers have argued. Like the folk notion, most researchers’ views of *um* and *uh* involve some kind of speaker production difficulty; a review of work in the area concluded that *um* and *uh* indicate upcoming delay (Clark and Fox Tree 2002). Like the folk notion, most researchers’ views of *you know* involve some kind of speaker–listener interaction; a review of work in the area concluded that *you know* indicates

that the speaker is leaving some information off-record and ‘inviting addressee inferences’ (Fox Tree and Schrock 2002: 744; Jucker and Smith 1998). There was no uniform folk notion for *like*, unless one wants to describe *lack of meaning* as the folk notion. The *lack of meaning* meaning accords with the observation that there is a wide range of intuitions about *like*’s use (Siegel 2002: 39). But it may be that *like*’s use is too subtle for laypeople to articulate. Most researchers’ views of *like* involve some kind of discrepancy between what’s said and what’s meant; a review of work in the area concluded that *like* is a ‘loose talk marker’ (Andersen 1998: 149).

Even if laypeople cannot articulate precisely what *like* means, they do have a sense for how it can be used. They recognize that discourse markers cannot substitute for each other without changing meaning. Written-out transcriptions of three spontaneously produced utterances were presented to 58 students in an undergraduate psychology class. The transcription appeared in either its original form or in one of two alternative forms with the other discourse markers replacing the original, as in the following:

- (1) Original: We had to go pick up um my buddy Mark’s girlfriend.
 Alternative 1: We had to go pick up you know my buddy Mark’s girlfriend.
 Alternative 2: We had to go pick up like my buddy Mark’s girlfriend.
- (2) Original: We were dressed for concert attire, you know, ten-tennis shoes, jeans, sweatshirt.
 Alternative 1: We were dressed for concert attire, like, ten-tennis shoes, jeans, sweatshirt.
 Alternative 2: We were dressed for concert attire, um, ten-tennis shoes, jeans, sweatshirt.
- (3) Original: There was a girl in a red shirt that had issues being in a crowd and didn’t really like being touched by anybody else. So when anyone like came near her she full-on pushed people away from her.
 Alternative 1: ... when anyone you know came near her she full-on pushed people away from her
 Alternative 2: ... when anyone um came near her she full-on pushed people away from her

Students read one of the above nine items, followed by a forced choice question probing either the *um/uh* folk notion meaning, the *you know* folk notion meaning, or the *like* meaning supported by a number of researchers, as in the following:

- (4) ‘We had to go pick up um my buddy Mark’s girlfriend.’
Which of the following is most likely to be true? (circle letter)
- Speaker couldn’t think of whom to pick up.
 - Speaker is checking that listener knows what he means when he says ‘my buddy Mark’s girlfriend.’
 - Speaker is indicating that ‘my buddy Mark’s girlfriend’ only loosely reflects what’s on his mind.

Each item was assessed by between 4 and 9 students.

One possible outcome is that students ignore the discourse markers entirely. If the context biased toward one or another inference, they would unanimously select the meaning that corresponded with that example regardless of which discourse marker they read. For example, students would select ‘Speaker couldn’t think of whom to pick up’ for the utterance ‘We had to go pick up um my buddy Mark’s girlfriend’ regardless of whether they read *um*, *you know*, or *like* between the words *up* and *my*. Another possible outcome is that students do not take discourse markers into account when making their selection; they may notice them, but not associate them with particular meanings. In this case, students would select a particular meaning, on average, a third of the time. A third possible outcome is that students prefer one of the meanings over the others, and select that meaning regardless of which context or discourse marker they read.

Results demonstrate that students both pay attention to the discourse marker, and that the discourse marker affects their choices of meanings. When *um* was read, 50% of respondents selected the *um* folk notion meaning of needing time to think. When *you know* was read, 41% of respondents selected the *you know* folk notion meaning of checking for listener understanding. When *like* was read, 61% of respondents selected the meaning of *like*’s indicating a loose use of language. At the same time, there was evidence that students generally favored the *like* meaning; 45% of all responses were the *like* meaning, compared to 29% *um* meaning and 26% *you know* meaning. Another way to think of this is that students thought the ‘loose use of language’ meaning could describe examples containing *like*, *um*, and *you know*.

5. Folk notions versus actual use

Some comparisons can be made between folk notions of the use of the markers and actual use of the markers. At the same time that the data presented in this study were collected, another group of students partici-

pated in a study involving telling and retelling personal stories to members of their subject pool class (Fox Tree 2006). The use of *likes*, *you knows*, and *ums* and *uhs*, among other discourse markers, was explored in this story-telling study. *Like* was the most prevalent discourse marker in this corpus, comprising almost 3% of all words used, on average (Fox Tree 2006). *Ums* and *uhs* accounted for almost 2% of all words used, and *you knows* only 0.2%. Although these percentages are not reflected in the subjective judgments of frequency of use illustrated by the current folk notions study, they do mirror people's history of discussing the markers—they are more likely to talk about the words they used more often, *um*, *uh*, and *like*. The percentages also mirror people's desire to eliminate the use of markers—they are more likely to want to restrict their uses of the more frequent markers, *um*, *uh*, and *like*.

The way people used the markers in the story-telling study modeled some of the intuitive concepts people have about their uses. In the story-telling study, close examination was given to the markers that were recycled in similar locations across two tellings of the same personal story to different addressees. *Ums* and *uhs* and *likes* were the most frequently recycled discourse markers: 9% of *ums* and *uhs* and 22% of *likes* were re-used in similar locations on successive tellings of a story. But they were not re-used in the same way.

Recycled *ums* and *uhs* occurred either at the beginning of the narrative, before elements of a list, or in remembering a name or a time (not including special cases where the *um* or *uh* was immediately preceded by a conjunction, often pronounced together with the conjunction as in *andum*; see Fox Tree 2006 for discussion of this exclusion). These are all locations where speakers might be more likely to have production trouble. *Ums* and *uhs* were also three times as likely to occur before actual evidence of production trouble (repetitions, restarts, and discourse markers) than *likes* were (and this study did not include the most common evidence of production trouble following *ums* and *uhs*, silent pauses [Clark and Fox Tree 2002], so the amount of production trouble is likely much greater than that reported by looking at repetitions, restarts, and discourse markers alone).

In contrast to *ums* and *uhs*, recycled *likes* occurred before elements of a story that could best be thought of as highlighting that the upcoming information is a loose rendering, such as the first telling 'drive over like bumpy roads' and second telling 'drive over like unpaved roads' (Fox Tree 2006: 731). The *likes* are not just meaningless speech tics, because addressees who retold the speaker's story to the research assistants sometimes also recycled the speaker's *likes*, as in the speaker's 'he was *like* pretending to be Jamaican' and the addressees' 'he was *like* some fake

Jamaican guy' (Fox Tree 2006: 736). And people retelling a story with no *likes* at all often spontaneously inserted *likes* into similar locations—locations that can be thought of as needing a loose rendering, such as exactly what kind of lion's cage the pet ant should live in (referring to a Monty Python sketch; Fox Tree 2006).

You know was the only other marker that recycled, not including markers that recycled as part of quotations (for example, when the quotee is reported as having begun their quote with *oh* in both stories; Fox Tree 2006). In the two cases of recycled *you knows*, *you know*'s use can be thought of as coordinating speaker and listener understanding, as in the first telling 'but they're *you know* they don't they don't run very well' and the second telling 'it's my my favorite car but *you know* they're not they're not great cars' (Fox Tree 2006: 734).

So, for *um*, *uh*, and *you know*, there is some evidence that people from the same speech community are in fact using the markers in the way that they intuitively verbalize that the markers are used. The notable exception is *like*, which people do use systematically, even though they cannot articulate what it means.

6. Conclusion

Taken together, the evidence presented here suggests that people do have a sense for a prototypical meaning or function for *um/uh*, *you know*, and *like*. This view of a prototypical meaning flies in the face not only of popular press accounts of universal absence of meaning across all markers, but of scientific accounts of these words' multifunctionality (Erman 1987; Holmes 1986). Multifunctional approaches have come under fire because of their theoretical pitfalls, expressed succinctly as 'If [discourse markers] can do so much, how do we know what they are doing at all?' (Fox Tree and Schrock 2002: 736). In an alternative approach to discourse markers, researchers consider what the primary function of a particular marker could be without enumerating a list of potential meanings and uses (Andersen 2000; Blakemore 2002; Clark and Fox Tree 2002; Fox Tree and Schrock 2002; Jucker 1993; Smith and Jucker 2000).

Laypeople's understandings of the uses of *um* and *uh*, *you know*, and *like* are not as far off from scientific views as some popular press accounts and Internet sites would have us believe. The stigmatization of these words may come about because they are products of spontaneous talk. Instead of contributing to the propositional content of what's conveyed, discourse markers comment on the way something is expressed

or the way people coordinate their conversation (see Clark's [1996] discussion of *collateral signals*). When speaking spontaneously, people can have trouble getting their ideas out. They can also choose to verbalize only part of their ideas, or to express things loosely. They can indicate these production and coordination issues with discourse markers. Discourse markers can therefore co-occur with occasions of planning trouble, or conversations where information is left off-record, or times when a loose representation of a concept stands in for a more fleshed out version. If the expectation is that the talk be well-prepared and that ideas are fully verbalized, then the presence of *ums*, *uhs*, *you knows*, and *likes* can send a negative message. But if the expectation is merely to exchange ideas with someone in a conversational setting, then *ums*, *uhs*, *you knows*, and *likes* do just what they are supposed to do, and, for the most part, laypeople are aware of these uses.

Appendix: Sample survey questions (*um/uh* survey)

1. How frequently do you use *um* or *uh* when you talk? Circle one:
 1. never
 2. rarely
 3. sometimes
 4. often
 5. all the time
 6. I don't know

2. Select one, and fill in the blank if necessary:
 - a. I say *um* or *uh* most frequently when I am [space for open-ended response follows]
 - b. I don't know when I say *um* or *uh* more than at other times.
 - c. I never say *um* or *uh*.

3. Does your use of *um* or *uh* vary depending on whether you are talking to a professor, to a friend, to a family member (or to someone else), or do you think you speak about the same to everyone? [space for open-ended response follows]

4. Does your use of *um* or *uh* vary depending on the situation you are in, or do you think you speak about the same no matter what the situation? [space for open-ended response follows]

5. Have you discussed the use of *um* or *uh* in a formal setting such as a class, lab, or section? Circle *yes* or *no*.

If so, please describe anything you can remember. [space for open-ended response follows]

6. Have you discussed the use of *um* or *uh* in an informal setting such as with friends or family? Circle *yes* or *no*.
If so, please describe anything you can remember. [space for open-ended response follows]
7. What, if anything, do you think *um* or *uh* means? [space for open-ended response follows]
8. Why do you think people use *um* or *uh*? [space for open-ended response follows]
9. Which of the following statements do you agree with most? Circle one:
 - a. There are times when I try to avoid using *um* or *uh*, and I succeed.
 - b. There are times when I try to avoid using *um* or *uh* but it doesn't work.
 - c. I don't try to avoid using *um* or *uh*.
10. Which of the following statements do you agree with most? Circle one:
 - a. Communication would be better if people didn't use *um* or *uh*.
 - b. Communication would be worse if people didn't use *um* or *uh*.
 - c. Saying *um* or *uh* doesn't affect communication.
 - d. I don't know whether communication is affected by saying *um* or *uh* or not.
11. How old are you?
12. What's your gender?
13. Where were you raised (where have you lived for a significant amount of time in your life)? [space for open-ended response follows]
14. What other languages do you speak, if any, and how well do you speak them? [space for open-ended response follows]

Notes

* This research was supported by faculty research funds granted by the University of California Santa Cruz.

1. M. Carlson, 'Monica, we hardly knew you', *Time*, 15 February 1999, p. 28.
2. G. Trudeau, Doonesbury comic strip, *San Jose Mercury News*, 4 June 2003, p. 8F.

References

- Andersen, G. (1998). The pragmatic marker *like* from a relevance-theoretic perspective. In *Discourse Markers: Descriptions and Theory*, A. H. Jucker and Y. Ziv (eds.): 147–170. Amsterdam: John Benjamins.
- (2000). The role of the pragmatic marker *like* in utterance interpretation. In *Pragmatic Markers and Propositional Attitude*, G. Andersen and T. Fretheim (eds.), 17–38. Philadelphia: John Benjamins.
- Arnold, J. E., Fagnano, M., and Tanenhaus, M. K. (2003). Disfluencies signal thee, um, new information. *Journal of Psycholinguistic Research* 32 (1): 25–36.
- Berkley, S. (2002). How to cure the ‘verbal virus’: A five-step treatment plan. URL: <<http://www.school-for-champions.com/speaking/verbalvirus.htm#Author>> [accessed on 13 October 2005].
- Blakemore, D. (2002). *Relevance and Linguistic Meaning: The Semantics and Pragmatics of Discourse Markers*. Cambridge: Cambridge University Press.
- Brotherton, P. (1979). Speaking and not speaking: Processes for translating ideas into speech. In *Of Speech and Time*, A. W. Siegman and S. Feldstein (eds.), 179–209. New York: Wiley.
- Butters, R. R. (2000). The ‘real’ meaning of *millennium*. *American Speech* 75 (1): 111–112.
- (2002). Intuition, data, and meaning. *American Speech* 77 (3): 327–328.
- Clark, H. H. (1996). *Using Language*. New York: Cambridge University Press.
- Clark, H. H. and Fox Tree, J. E. (2002). Using *uh* and *um* in spontaneous speaking. *Cognition* 84: 73–111.
- Erman, B. (1987). *Pragmatic Expressions in English: A Study of You Know, You See, and I Mean in Face-to-Face Conversation*. Stockholm: Almqvist and Wiksell International.
- Fox Tree, J. E. (2001). Listeners’ uses of *um* and *uh* in speech comprehension. *Memory and Cognition* 29 (2): 320–326.
- (2006). Placing *like* in telling stories. *Discourse Studies* 8 (6): 723–743.
- Fox Tree, J. E. and Schrock, J. C. (2002). Basic meanings of *you know* and *I mean*. *Journal of Pragmatics* 34: 727–747.
- Fuller, J. M. (2003). Use of the discourse marker *like* in interviews. *Journal of Sociolinguistics* 7 (3): 365–377.
- Holmes, J. (1986). Functions of *you know* in women’s and men’s speech. *Language in Society* 15: 1–22.
- Hosman, L. A. and Wright, J. W. II (1987). The effects of hedges and hesitations on impression formation in a simulated courtroom context. *Western Journal of Speech Communication* 51: 173–188.
- Jucker, A. H. (1993). The discourse marker *well*: A relevance-theoretical account. *Journal of Pragmatics* 19: 435–452.
- Jucker, A. H. and Smith, S. W. (1998). *And people just you know like ‘wow’*: Discourse markers as negotiating strategies. In *Discourse Markers: Descriptions and Theory*, A. H. Jucker and Y. Ziv (eds.): 171–120. Amsterdam: John Benjamins.
- Jucker, A. H., Smith, S. W., and Lüdge, T. (2003). Interactive aspects of vagueness in conversation. *Journal of Pragmatics* 35: 1737–1769.
- Lawrence, W. P. (2002). Do native speakers know what words mean? *American Speech* 77 (3): 325–327.
- Lynch, S. (2002). Filler words become regular practice. *Elon University: The Pendulum Online*. 31 October.
- Östman, J.-O. (1981). *You Know: A Discourse Functional Approach*. Amsterdam: John Benjamins.

- Preston, D. R. (1998). Why we need to know what real people think about language. *The Centennial Review* 42 (2): 255–284.
- (2000). Some plain facts about Americans and their language. *American Speech* 75 (4): 398–401.
- Reynolds, A. and Paivio, A. (1968). Cognitive and emotional determinants of speech. *Canadian Journal of Psychology* 22 (3): 164–175.
- Schachter, S., Christenfeld, N., Ravina, B., and Bilous, F. (1991). Speech disfluency and the structure of knowledge. *Journal of Personality and Social Psychology* 60 (3): 362–367.
- Schourup, L. C. (1985). *Common Discourse Particles in English Conversation*. New York: Garland.
- Siegel, M. E. A. (2002). *Like*: The discourse particle and semantics. *Journal of Semantics* 19: 35–71.
- Smith, S. W., and Jucker, A. H. (2000). *Actually* and other markers of an apparent discrepancy between propositional attitudes of conversational partners. In *Pragmatic Markers and Propositional Attitude*, A. Anderson and T. Fretheim (eds.), 207–237. Philadelphia: John Benjamins.
- Smith, V. L., and Clark, H. H. (1993). On the course of answering questions. *Journal of Memory and Language* 32: 25–38.
- Stenström, A.-B. (1990). Lexical items peculiar to spoken discourse. In *The London-Lund Corpus of Spoken English: Description and Research*, J. Svartvik (ed.), 137–175. Lund University Press.
- Stubbe, M. and Holmes, J. (1995). *You know, eh* and other ‘exasperating expressions’: An analysis of social and stylistic variation in the use of pragmatic devices in a sample of New Zealand English. *Language and Communication* 15 (1): 63–88.
- Underhill, R. (1988). *Like* is, like, focus. *American Speech* 63 (3): 234–246.
- Vrij, A. and Winkel, F. (1991). Cultural patterns in Dutch and Surinam nonverbal behavior: An analysis of simulated police/citizen encounters. *Journal of Nonverbal Behavior* 14 (3): 169–184.
- Watts, R. J. (1989). Taking the pitcher to the ‘well’: Native speakers’ perception of their use of discourse markers in conversation. *Journal of Pragmatics* 13: 203–237.

Jean E. Fox Tree is an Associate Professor of Psychology at the University of California Santa Cruz. She received her A.B. from Harvard University in linguistics, her M.Sc. from the University of Edinburgh in cognitive science, and her Ph.D. from Stanford University in psychology. She studies the comprehension and production of spontaneous speech. Address for correspondence: Psychology Department, Social Sciences II Room 277, University of California, Santa Cruz, CA 95064, USA <foxtree@ucsc.edu>.