# Vocal Fry Use in Adult Female Speakers Exposed to Two Languages

***Todd A. Gibson, †Connie Summers, and *Sydney Walls,** **Baton Rouge, Louisiana, and †El Paso, Texas*

**Summary: Objective.** Several studies have identified the widespread use of vocal fry among American women. Popular explanations for this phenomenon appeal to sociolinguistic purposes that likely take significant time for second language users to learn. The objective of this study was to determine if mere exposure to this vocal register, as opposed to nuanced sociolinguistic motivations, might explain its widespread use.

**Study Design.** This study used multigroup within- and between-subjects design.

**Methods.** Fifty-eight women from one of three language background groups (functionally monolingual in English, functionally monolingual in Spanish, and Spanish-English bilinguals) living in El Paso, Texas, repeated a list of nonwords conforming to the sound rules of English and another list of nonwords conforming to the sound rules of Spanish. Perceptual analysis identified each episode of vocal fry.

**Results.** There were no statistically significant differences between groups in their frequency of vocal fry use despite large differences in their amount of English-language exposure. All groups produced more vocal fry when repeating English than when repeating Spanish nonwords.

**Conclusions.** Because the human perceptual system encodes for vocal qualities even after minimal language experience, the widespread use of vocal fry among female residents in the United States likely is owing to mere exposure to English rather than nuanced sociolinguistic motivations.

**Key Words:** Vocal fry–Perceptual judgmentss–Adult women–Bilingual–Sociolinguistic factors.

## INTRODUCTION

Vocal fry is a voice register characterized by irregular vibratory patterns, low fundamental frequency, and low energy, which results in a voice perceived as a rapid series of clicks rather than the steady buzz associated with modal register phonation.[1,2] Previously associated with a male voice quality,[3,4] a number of recent studies have identified widespread use of this voice register among American women.[5–7] Its use appears influenced by both linguistic and sociolinguistic motivations.

Linguistic motivations for the use of vocal fry appear related to syntax because it is most likely to appear at the end of either sentences[4] or paragraphs.[8,9] This has been interpreted to mean that vocal fry serves as a cue to syntactic boundaries[10] and may even serve pragmatic functions such as marking the end of a conversational exchange.[11] We recently showed that lexical stress also influences its use.[12] Lexical stress in English is characterized by greater energy and higher fundamental frequency than surrounding syllables. These linguistic cues compete with the lower energy and lower fundamental frequency associated with vocal fry, with the result being that vocal fry is most likely to appear in unstressed syllables.

Sociolinguistic motivations for the use of vocal fry perhaps have received more attention. Vocal fry use by women has been associated with toughness,[13] commiseration,[14] and insecurity.[15] Yuasa[7] compared vocal fry use by young American women speaking English and young Japanese women speaking Japanese. She found that American women used vocal fry almost 80% more frequently than did Japanese women. Yuasa proposed that vocal fry use among young upwardly mobile American women was a strategy used to appear more masculine to advance in a male-dominant society. Success for Japanese women, she reported, likely would not be enhanced by vocal fry use because of different cultural standards.

Another as yet underexplored possibility is that vocal fry use in female American English speakers might merely be the result of being exposed to vocal fry. That is, vocal fry might be sufficiently pervasive in the speech of American English-speaking women as to be in a virtuous cycle in which frequent use begets more frequent use. Although some languages use vocal fry to distinguish one speech sound from another,[16] English does not. However, a significant body of literature has shown that lexical representations can include information like vocal fry even when it typically is not part of a language's sound system. For example, several studies have found that memory for words is enhanced when listeners hear the same talker at time one and time two.[17–19] This has been interpreted to mean that irrelevant speaker qualities like intonation, fundamental frequency, and speaking rate[17,20,21] are encoded in the mental lexicon, even after a single exposure. Co-occurring, nonhuman irrelevant information (eg, barking dogs and ringing telephones heard in the background) appears also to be encoded into lexical representations[22] after a single presentation.

To our knowledge, Yuasa's[7] is the only study to compare directly the use of vocal fry cross-linguistically. Her Japanese-speaking participants, however, were studying or working in California and thus likely exposed, however obliquely, to the sounds of English being produced in the surrounding majority language. Although it was not reported, many, if not all, likely were bilingual speakers proficient in both Japanese and English because they were either students at the University of California, Berkeley, or working at the Japan Pacific Resource Network in Oakland, California. Because the Yuasa study was a between-subjects and not within-subjects design, we cannot know if these Japanese participants would have changed the frequency of their

Accepted for publication November 8, 2016.

From the *Louisiana State University, Baton Rouge, Louisiana; and the †University of Texas at El Paso, El Paso, Texas.

Address correspondence and reprint requests to Todd A. Gibson, Louisiana State University, 84 Hatcher Hall, Baton Rouge, LA 70803. E-mail: toddandrewgibson@lsu.edu

Journal of Voice, Vol. 31, No. 4, pp. 510.e1–510.e5

0892-1997



http://dx.doi.org/10.1016/j.jvoice.2016.11.006

vocal fry use across their two languages. Sociolinguistic and virtuous cycle hypotheses, however, would make different predictions regarding this use.

Nuanced pragmatic or sociolinguistic motivations for vocal fry (eg, indexing toughness, insecurity, commiseration, masculinity, or providing emphasis) that are culture-dependent likely require a significant amount of acculturation before being adopted by a second language learner. Indeed, a large body of research on interlinguistic pragmatics has identified that the appropriate use and understanding of language in the second language context is positively correlated with second language proficiency.[23–25] It is unlikely, therefore, that individuals with very limited American English exposure would learn the sociolinguistic motivations for vocal fry. On the other hand, if vocal fry is a pervasive quality of the vowel system for American English-speaking women, it might be adopted without regard to its sociolinguistic uses and with minimal exposure owing to the sensitivity of the human perceptual system to encode such information into lexical representations.

In the current study, we compared predictions made by the sociolinguistic versus the virtuous cycle hypothesis. We were able to make these comparisons by reanalyzing data drawn from a previous investigation of working memory in Spanish-English bilingual speakers. Participants included functionally monolingual American English speakers, Spanish-English bilingual speakers, and functionally monolingual American Spanish speakers. All participants were living in El Paso, Texas, a border city, which has a large Spanish-speaking community and a large bilingual community. Therefore, functionally monolingual American English speakers likely were exposed to Spanish in the community (eg, supermarket, restaurants, radio), and functionally monolingual Spanish speakers likely were exposed to English in the community.

We reasoned that the sociolinguistic hypothesis would predict that, because the functionally monolingual groups lacked the type of experience necessary to learn the nuanced sociolinguistic uses of vocal fry, their vocal fry usage would not change across languages. However, the bilingual group likely would have experienced sufficient experience in English and Spanish to know the sociolinguistic motivations related to vocal fry use. Therefore, the sociolinguistic hypothesis would predict that the bilingual group's vocal fry use would be greater in English than in Spanish. On the other hand, the virtuous cycle hypothesis would predict that the frequency of vocal fry use would be greater in English than in Spanish for all three groups because they likely had been exposed to both English and Spanish in the community. According to this hypothesis, the vocal fry register likely would be encoded in the English lexical representations of all three groups owing to the sensitivity of the human perceptual system.

Although two groups of our participants were functionally monolingual, we eliminated the need for individuals to be proficient in both languages by asking participants to repeat a set of nonwords (syllables that sound like words but have no meaning, like *woogalamic*[26]) based on the rules of the English sound system and another set based on the rules of the Spanish sound system. Despite a thorough literature review, we found no reports of widespread use of vocal fry among female Spanish speakers in other countries.

## METHOD

### Participants

Sixty female participants repeated Spanish and English nonwords as part of a larger study on bilingualism and working memory. The data from two participants (both from the functionally monolingual English group) were not included because their use of vocal fry was more than 3.5 standard deviations (SDs) above the mean. The final sample included 58 participants.

Language history questionnaires provided language background and demographic information including the age at which individuals were regularly exposed to English (which indexed cumulative language experience), the current percentage of daily experience in English versus Spanish, and a language self-rating measure based on a scale of one to five, with one signifying no ability and five signifying native proficiency. We created three language experience groups based on information gathered from the questionnaire. To determine whether to use cumulative or current English experience to create language experience categories, we created a percent lifetime English measure by subtracting participants' age of first regular English experience from their current age and dividing this difference by their current age. We then calculated the percentage of current daily exposure to English and calculated a measure of internal consistency to determine if the two measures tapped into the same or different constructs. A Cronbach alpha of .94 indicated that the two measures tapped into the same underlying construct. Therefore, we used current English experience as the variable to create language experience categories because two individuals had not provided information regarding age of first regular English experience.

Individuals with 70% daily English experience or more were treated as functionally monolingual in English (n = 25). Individuals with 30% daily English experience or less were categorized as functionally monolingual in Spanish (n = 23), and those whose daily English experience was between 69% and 31% were treated as bilingual (n = 10). These categories corresponded to self-rating measures of English and Spanish ability. That is, functionally monolingual speakers' strong language was statistically significantly better than their weak language, and there was no statistically significant difference in the self-ratings between English and Spanish for bilingual speakers (see Table 1 for participant details).

### Materials

Participants repeated both English and Spanish nonwords. English nonwords were taken from Dollaghan and Campbell,[27] a frequently used measure of short-term memory for sounds. These included four items each at one-, two-, three-, and four-syllable lengths. No individual syllable was an English word, but the rules for English sounds applied to all stimuli such that all syllables potentially could become an English word (eg, /nɑɪb/ is not an English word but could become one). Primary stress landed on the second syllable for four-syllable English nonwords but on the first syllable for all other English syllable lengths. Spanish nonwords were taken from Gibson et al[28] and included four items each at two-, three-, four-, and five-syllable