

Comparing hesitation markers in Sanjuanero Spanish

LAMAR A. GRAHAM
University of Georgia
lamar.a.graham@gmail.com

ABSTRACT: Discourse markers, according to Portolés (2001), are set phrases that convey inferential meaning about what a speaker is communicating and can serve various connective, recapitulative, or reformulative functions. Within the domain of discourse markers, there is a subset known as conversational markers which only carry pragmatic weight with little or no semantic meaning, such as *eh* ‘um’, *este* ‘this’, *hombre* ‘man’, and *bueno* ‘well’. Two particular conversational markers that are in competition with one another in San Juan de Puerto Rico are *eh* and *este*. These are known as hesitation or metadiscursive markers, whose primary discursive purpose is to fill space. This current study aims to compare the frequency of *este* versus the more conventional *eh* in San Juan, using age, gender, educational attainment as sociolinguistic variables and clausal or sentential position as a syntactic variable. Results indicate that *este* is the preferred hesitation marker in San Juan and that age and sentential position are significant factors in the choice of *este* over *eh* in spoken discourse.

Keywords: hesitation markers, San Juan, sociolinguistics, variation.

0. INTRODUCTION

In spoken and written discourse, there are two principal classifications of utterances: lexical items and discourse markers. By contrast, lexical items work together to define the theme of the discourse, while markers, which may or may not be lexicalized, perform a different function altogether. According to Portolés (2001:25-26),

“discourse markers are invariable linguistic units, do not demonstrate a syntactic function in the frame of sentential predication, and possess a coincident task in discourse: to guide, in agreement with its distinct morphosyntactic, semantic, and pragmatic properties, the inferences that are produced in communication” (translation mine).

The most studied markers are those that carry semantic meaning as well as pragmatic weight, such as *es decir* ‘that is to say’, *sin embargo* ‘however’, and *en fin* ‘in the end/finally’.

- (1)
- a. Ana es de Teruel y, *por tanto*, es habladora.
‘Ana is from Teruel and, *therefore*, is talkative.’
 - b. Ana es de Teruel y, *sin embargo*, es habladora.
‘Ana is from Teruel and, *nevertheless*, is talkative.’

(Portolés, 2001: 30)

The italicized phrases *por tanto* and *sin embargo* are the discourse markers in the above examples. Even though both phrases are adverbial adjuncts (*Ana es de Teruel y es habladora* is a well-formed sentence), the inclusion of each changes the implied meaning of the utterance but does not disrupt the relationship between the two elements (Torres & Potowski 2008: 263). In example (1a), the phrase *por tanto* carries the implicature that Ana’s being from Teruel is the cause of her loquaciousness; that is, people from Teruel are characteristically talkative. This contrasts with example

(1b) which contains the phrase *sin embargo*. The implicature here is that Ana is talkative despite being from Teruel, and that people from that area are not known to be talkative.

Unlike other categories of discourse markers which carry semantic as well as pragmatic weight, conversational markers only carry pragmatic weight. Galué (2002: 29) states that the main purpose of conversational markers in discourse is “to organize information or maintain one’s turn within the conversation” (translation mine). Markers such as *eh* ‘um’, *este* ‘this’, *hombre* ‘man’, and *bueno* ‘well’ when used as conversational markers do not maintain what semantic meaning they previously had; in this sense, they serve as space fillers while the speaker searches for a word or reformulates his or her thoughts (Antúñez Pérez 2005: 3). In written discourse, conversational markers are not found in narrative and are only present in transcriptions of oral production (Prada 2001: 86).

- (2) No es una cosa que, *digamos*, nació para eso, mm, cómo lo explico, déjame explicarte.
‘It’s not a thing that, *say*, I was born for, mm, how do I explain ti, let me explain it to you.’
(SANJ_H23_023)
- (3) Por ejemplo recientemente, *tú sabes*, yo me compré una computadora, nunca en mi vida me había comprado una computadora y recientemente, hace como dos semanas y media, fue que finalmente decidí comprarme una computadora porque sí soy profesora.
‘For example, recently, *you know*, I bought myself a computer, never in my life had I bought a computer and it was recently, two and a half weeks ago, that I finally decided to buy myself a computer because I’m a professor.’
(SANJ_M23_058)

Hesitation discourse markers, often classified as metadiscursive conversational markers, are a little-studied phenomenon in Spanish. In her study of the use of conversational discourse markers in the Spanish of Caracas, Galué (2002:44-45) found *este* to be the marker that appeared most often.

- (4) Conozco gente que estudia en La Central, ¿no?// *este* /// no sé // *este* // me parece que loh // que-la-las personas que estudian en La Católica, ¿no? *este* // se- son unas personas que son más inclinadas a- a ser menos protestatarias / *este* [...] cabezas caliente. / cualquiera de esos / incluso son más sumisas// ¿no? *Este*. / con todos estos problemas del- de- de- de la Constitución
‘I know people that go to La Central, right? *Uh*, I don’t know, uh, it seems to me that the – that the people that go to La Católica, no? *uh* – are people that are more inclined not to protest, uh... [...] hot heads. Any of them are also more submissive... no? *Uh*... with all these problems with the Constitution...’
(Galué 2002:42; transcription original)

Este carried no semantic value, but its pragmatic uses were, as expected, reformulative and turn-extending. The speaker inserted *este* in order to avoid silence while he or she thought of what to say next. Interestingly, in Galué’s analysis, *este* was the discourse marker (of eight studied; also *claro* ‘clearly’, *okey* ‘okay’, *la verdad* ‘the truth’, *mira/mire* ‘look’, *fíjate* ‘listen up’, *¿entiendes?* ‘understand?’, and *verdad* ‘right’) whose frequency was the highest observed.

Nevertheless, among discourse markers, the treatment of *este* has been relatively scarce in the literature. Another study of hesitation markers by Soler Arechalde (2008) focused on the uses and distribution of *este* by speakers from Mexico City. According to her, *este* is very widely used by speakers in the city despite schoolteachers actively discouraging its use, “[...] por considerarlo un elemento vacío e innecesario, muestra de un manejo pobre e inadecuado de la lengua oral” (156). Also, states Soler Arechalde, though the metadiscursive use of *este* cannot be found in any dictionary or grammar of the Spanish language, its use in Latin America is widespread. This includes Mexico City, in which previous studies of hers showed that *este* was more widely encountered in popular speech (*habla popular*) than formal, educated speech (*habla culta*). In this study, limited to a corpus of *habla culta*, frequencies of *este* were compared based on several variables, including but not limited to gender, communicative context, and topic at hand. Gender, being the only variable applicable to a speaker group at large, was shown to be significant; men produced two thirds of all instances of *este* in the corpus.

In a previous investigation (Graham 2010), the relative frequency of demonstrative pronouns and adjectives in Puerto Rican Spanish was studied. An inordinately high number of instances of the Spanish masculine singular proximal demonstrative adjective *este* were encountered. It was then determined that this particular use of *este* was inconsistent with demonstrative behavior and was instead a more specialized use of the word –as the literature supports, it was a high frequency discourse marker used by this speaker group in order to denote pauses in discourse¹.

Thus, *este* and *eh*, the two conversational markers that I compare, both translate to English *uh* or *um* and are used in the same fashion (Montes, 1999; Torres, 2002), and so it is prudent to study the two as free variants. Based on Galué's (2002) and Soler Arechalde's (2006, 2008) previous work in Caracas and Mexico City, respectively, my working hypothesis is twofold: that *este* is the most frequently used metadiscursive conversational marker in San Juan compared to *eh* and its variants, and that the social factors of gender, age, and educational attainment, along with the syntactic factor of sentential position, significantly affect the speaker's choice of *este* as opposed to a variant of the competing conversational marker *eh*. More specifically, these are the effects that each variable is predicted to have on hesitation marker production:

- Men are predicted to produce *este* significantly more frequently than women –based on Soler Arechalde (2008).
- Speaker age is predicted to be inversely proportional to frequency of *este* production –loosely based on Graham (2010)².
- More educated speakers are predicted to produce *este* less frequently –based on Soler Arechalde (2006).

It is important to note that since predicting where a speaker will insert a hesitation marker is impossible, the focus is on which marker –*este* or *{eh}*³– is more favored in the contexts where hesitations occur.

1. METHODOLOGY

The corpus used in this study was compiled by the *Proyecto para el estudio sociolingüístico del español de España y de América* (PRESEEA – <http://preseea.linguas.net>) (2003), a collection of sociolinguistic interviews. This corpus was chosen because of the timeliness of the interviews and because of its ease of searching and the protocol used to categorize the speakers. In a previous work with these data (Graham, 2010), it was found that all hesitation markers were parsed by commas or ellipses, meaning that *este* was easily distinguishable from its demonstrative homonym. Each interview was labeled according to the gender, generation (age range), and educational attainment of the speaker. The PRESEEA protocol is as follows:

| | |
|-------------|--------------------------------|
| Gender: | H (male, from <i>hombre</i>) |
| | M (female, from <i>mujer</i>) |
| Generation: | 1 (age 20-34) |
| | 2 (age 35-54) |
| | 3 (age 55+) |

¹ In Graham (2010), hesitation marker *este* appeared more frequently than any of the demonstratives under study.

² The small sample in this study indicated that there may be an age preference for this particular marker, as it was more frequently produced by younger speakers.

³ Where it appears, the morpheme notation *{eh}* represents all encountered allomorphs of *eh* [e], including *ehh* [e:] and *ehm* [em].

- Education: 1 (less than high school education)
 2 (high school education)
 3 (postsecondary education)

Eighteen interviews, nine each of men and women, all of various ages and education levels, were selected for data extraction. There are three men and women from each generation. From each generation, there was one interviewee chosen for each education level. The exception to this rule is the educational distribution of the first-generation speakers; there were no interviewees with less than a high school education, so another high school graduate was selected to fill that space. Therefore, the distribution of speakers is as follows:

| Gender | Categories | | | | | | | Total | |
|--------|------------|---|---|-------|-------------------|---|---|-------|-------|
| | Generation | | | Total | Educational level | | | | Total |
| | 1 | 2 | 3 | | 1 | 2 | 3 | | |
| Male | 3 | 3 | 3 | 9 | 2 | 4 | 3 | 9 | |
| Female | 3 | 3 | 3 | 9 | 2 | 4 | 3 | 9 | |
| Total | 6 | 6 | 6 | 18 | 4 | 8 | 6 | 18 | |

Table 1 - Distribution of speakers.

Wordsmith® Concord (2005) was the program used in searching for tokens for this study. In order to find the desired tokens, certain search strings had to be used. Beginning with *este*, care had to be taken not to return any demonstrative adjectives. Fortunately, the transcribers of the PRE-SEEA San Juan corpus included partial stops in their representations of the participants' speech; these were shown as either commas or ellipses. The necessary search string, therefore, had to require that there be punctuation adjacent to the end of the word in order to return as many of the correct number of hesitation marker *este* as possible. Also, since *este* as a discourse marker is invariable in its production, there was no need to search for any variants of the word. *Eh*, on the other hand, can be produced in a number of ways, and the transcriptions reflected this. In order to return *eh*, *ehh*, and *ehm* (the three most evident variants) as well as any other possible representations, wildcards were used.

A total of 1,148 instances of *este* and *{eh}* were collected from the 18 transcriptions. 113 of these instances were excluded, either because they were produced by the interviewer or because they were repetitions. This left 1,035 target tokens of *este* and *{eh}* for analysis.

The data were coded for use in GoldVARB (Sankoff *et al.*, 2006). The occurrence of *este* or *{eh}* was the dependent variable of study. Four factors were included in the coding string: gender of the interviewee, age (generation), educational attainment, and the position of the coda within the sentence. The original coding scheme for use in GoldVARB was as follows:

| Variable and code | Meaning |
|-----------------------------------|--|
| Choice of discourse marker | |
| E | <i>este</i> |
| e | <i>eh, eh, or eh</i> |
| Generation | |
| 1 | younger generation (20-34) |
| 2 | middle generation (35-54) |
| 3 | elder generation (55+) |
| Gender | |
| m | male |
| f | female |
| Educational attainment | |
| P | primary education (no high school diploma) |
| S | secondary education (high school diploma) |
| H | higher education (postsecondary degree) |
| Position within sentence | |
| i | sentence-initial |
| n | sentence-internal |
| f | sentence-final |

Table 2 - GoldVARB coding scheme.

There were only three instances of sentence-final discourse markers, which inhibited analysis in GoldVARB. Therefore, the three (f) instances were recoded to be included with the sentence-internal instances (n). A sample token of (e3mHn) would indicate use of {*eh*}, by an elder speaker, male, college- or university-educated, and that the token occurred in sentence-internal position, as follows:

- (5) Si tú me conociste a m[i] amigo, *eh*, de comer carne de cerdo, a menos que no sea que tenga que por cuestión de vida o muerte, voy a buscar la manera por lo menos una vez en la semana de comer carne cerdo.
 ‘If you knew me, friend, *uh*, about eating pork, unless it’s not a matter of life or death, I’m going to find a way to eat pork at least once a week.’
 (SANJ_H33_034)

Once the tokens were coded, they were entered into GoldVARB. A binomial one-step analysis was used to find the factor weights of each group, and a two-step analysis was performed in order to determine which factors were and were not significant.

2. ANALYSIS OF DATA

Of the 1,035 desirable tokens, 563 were instances of *este* (54.2%), while 472 were some variant of *eh* (45.8%). The analysis as performed by GoldVARB shows observed frequencies and factor weights as follows. In each table, capital ‘E’ represents application of the dependent variable (occurrences of *este*), and lowercase ‘e’ represents non-application (occurrences of a variant of *eh*).

| Age group | | <i>este</i> | <i>eh</i> | Total N | Total % |
|---------------------|---|-------------|-----------|---------|---------|
| Younger speakers | N | 193 | 66 | 259 | 25 |
| | % | 74.5 | 25.5 | | |
| Middle-age speakers | N | 314 | 190 | 504 | 48.7 |
| | % | 62.3 | 37.7 | | |
| Elder speakers | N | 56 | 216 | 272 | 26.3 |
| | % | 20.6 | 79.4 | | |
| Total | N | 563 | 472 | 1,035 | 100 |
| | % | 54.4 | 45.6 | | |

Table 3 - Generation (age) frequency table.

The youngest generation heavily favors the use of *este* over a variant of *eh*, with a frequency of 74.5%. At the other end of the spectrum, the elder generation tends not to use *este*, with a frequency of only 20.6%. The middle generation uses *este* more frequently than the elder generation but not as frequently as the younger generation. Interestingly, the data also show that the middle generation was found generally to use hesitation markers more often than either of the other two generations.

- (6) Trece años, *este*, yo me mudé de de Hato Rey a Carolina cuando tenía seis... y a los trece años fue que yo me mudé de allí.
 ‘At age 13, *uh*, I moved from – from Hato Rey to Carolina when I was six... and it was at age 13 that I moved away from there.’
 (SANJ_H13_011 – younger speaker)
- (7) Sí, porque Bayamón está muy hacinado y como que hay tanto revolú en donde, a todas horas hay tapones, *este*, como que hay mucho mucho hacinamiento.
 ‘Yes, because Bayamón is very congested and since there’s so much confusion where, at every hour there are traffic jams, *uh*, since there’s lots and lots of congestion.’
 (SANJ_M22_053 – middle-age speaker)
- (8) Una carrera que, *este*, que, tú sabes que yo ganara mucho dinero, una secretaria.
 ‘A career in which, *uh*, you know, I could earn a lot of money, a secretary.’
 (SANJ_M31_063 – elder speaker)

| Gender | | <i>este</i> | <i>eh</i> | Total N | Total % |
|--------|---|-------------|-----------|---------|---------|
| Male | N | 305 | 287 | 592 | 57.2 |
| | % | 51.5 | 48.5 | | |
| Female | N | 258 | 185 | 443 | 42.8 |
| | % | 58.2 | 41.8 | | |
| Total | N | 563 | 472 | 1,035 | 100 |
| | % | 54.4 | 45.6 | | |

Table 4 - Gender frequency table.

Both men and women were shown to use *este* more frequently than a variant of *eh*, women proportionally more often than men. Overall, by raw count, men used hesitation markers more often

than women, both *este* and *eh*; yet when given a choice, men did not show as clear a tendency to use one marker over the other.

- (9) Aquí, *este*, hasta se me olvida el nombre de este barrio aquí. *Este*, eso es, *este*, Caimito Bajo.
‘Here, *uh*, I almost forgot the name of this neighborhood here. *Uh*, that is, *uh*, Caimito Bajo.’
(SANJ_H21_013 – man)
- (10) *Este*, mis prioridades son, en ese orden, *eh*, la vida primero porque, *este*, con la vida viene la salud y... si yo tengo vida y tengo salud.
‘*Uh*, my priorities are, in that order, *um*, life first because, *uh*, with life comes health and... yes, I have life and I have health.’
(SANJ_M23_058 – woman)

As will be shown, however, the difference based on gender was not found to be statistically significant.

| Educational level | | <i>este</i> | <i>eh</i> | Total N | Total % |
|---------------------|---|-------------|-----------|---------|---------|
| Higher education | N | 274 | 291 | 565 | 54.6 |
| | % | 48.5 | 51.5 | | |
| Secondary education | N | 207 | 100 | 307 | 29.7 |
| | % | 67.4 | 32.6 | | |
| Primary education | N | 82 | 81 | 163 | 15.7 |
| | % | 50.3 | 49.7 | | |
| Total | N | 563 | 472 | 1,035 | 100 |
| | % | 54.4 | 45.6 | | |

Table 5 - Educational attainment frequency table.

The table shows that use of *este* versus a variant of *eh* is virtually even between the lowest- and the highest-educated groups. Interviewees who had completed only high school favored the use of *este* 67.4 percent of the time.

- (11) No, nosotros vivíamos *este*, o sea, como, *este*, papi no se llevaba mucho con, con, con abuela Lola que era la mamá de mami.
‘No, we lived *uh*, rather, since, *uh*, Daddy didn’t get along very well with, with, with, Grandma Lola who was Mommy’s mom.’
(SANJ_M21_050 – primary education speaker)
- (12) Pues fíjate Bayamón, *este*, actualmente han hecho mejoras y están haciendo mejoras en ah, *eh*, en el casco en y los alrededores.
‘Well, let’s see, Bayamón, *uh*, recently has made improvements and is currently making improvements in ah, *um*, in downtown and in the surrounding areas.’
(SANJ_H22_018 – secondary education speaker)
- (13) O sea, ella, cosió el traje y yo se lo bordé en las, en perlas, *este*, y fue tan bonito porque, entre las dos, hacer todo eso, en realidad fue una experiencia fabulosa.
‘Or rather, she sewed the suit and I embroidered it with pearls, *uh*, and it was so beautiful because, between the two [tasks], to do all of that, in reality, was a fabulous experience.’
(SANJ_M33_070 – higher education speaker)

Within the education factor group, speakers who had not completed high school tended not to use hesitation markers very often at all compared to those who had had more education. However, it must be noted that there were two fewer speakers in the group with less than a high-school education (none of the younger speakers fell into that group) and, consequently, two more speakers in the group of those that held a high-school diploma.

| Context | | <i>este</i> | <i>eh</i> | Total N | Total % |
|-------------------|---|-------------|-----------|---------|---------|
| Sentence internal | N | 469 | 417 | 886 | 85.6 |
| | % | 52.9 | 47.1 | | |
| Sentence initial | N | 94 | 55 | 149 | 14.4 |
| | % | 63.1 | 36.9 | | |
| Total | N | 563 | 472 | 1,035 | 100 |
| | % | 54.4 | 45.6 | | |

Table 6 - Sentential position frequency table.

Este and variants of *eh* occur much more frequently in sentence-internal position. Within that context, *este* occurs very slightly more often than $\{eh\}$. In sentence-initial position, *este* is the marker that is seen more often with an observed frequency of 63.1%.

- (14) *Este*, eran \$135 millones de dólares en ventas tributables y la diferencia 15 millones a las Fuerzas Armadas e instituciones que no pagan arbitrio por concepto de la compra de, *eh*, producto de tabaco.
‘*Uh*, it was \$135 million in taxable sales and the other \$15 million went to the Armed Forces and institutions that don’t pay into the discretionary fund in the way of buying, *um*, tobacco products.’
(SANJ_H33_034)
- (15) Sí quedan y entonces, *este*, va a la escuela que están vendiendo los famosos gallitos, cuando entonces Tito dice que pagó un peso por un gallito.
‘Yes they stay and then, *uh*, he goes to the school where they’re selling the famous birdies, when Tito says that he paid a peso for a birdie.’
(SANJ_M22_053)

Below are the factor weights of each independent variable. Group 1 represents the age factor group, group 2 gender, group 3 education, and group 4 position of the marker within the sentence.

| <i>Group and factor</i> | Weight |
|----------------------------|---------|
| Age | |
| Younger speakers | 0.701 |
| Middle-age speakers | 0.595 |
| Elder speakers | 0.179 |
| Sentential position | |
| Sentence-initial | 0.661 |
| Sentence-internal | 0.472 |
| Gender | |
| Female | [0.536] |
| Male | [0.473] |
| Education | |
| Primary education | [0.562] |
| Secondary education | [0.540] |
| Higher education | [0.460] |

Table 7 - Factor weights for all groups.

Age and sentential position were shown to be significant in this study, while gender and educational attainment, whose factor weights are in square brackets, were shown to be insignificant. The ranges in factor weights for age and sentential position are .522 and .149, respectively.

3. DISCUSSION AND RESULTS

Age appears to be the most prevalent factor in choosing to use *este* over a variant of *eh*. This is so because the largest discrepancy was seen in this factor group as opposed to sentential position. The use of *este* by younger speakers is fifty percentage points higher than that of elder speakers, which is a very telling contrast. This is not to say that *este* was not used at all by older speakers; it was observed 21% of the time. However, the tendency was still toward a variant of *eh*. The middle generation can be viewed as the intermediate stage in the shift toward *este*, as the frequency was not nearly as high as that of the younger generation but much higher than that of the elders.

Sentential position was shown as the other significant factor in choosing a hesitation marker. However, its influence on such a choice cannot be easily compared to age. Even though there is a clear favoring toward using *este* in sentence-initial position, there are not nearly as many instances of conversational markers in that position as there are in sentence internal position; the ratio of sentence-initial markers to sentence-internal markers is nearly 1 to 6. Such an uneven distribution of markers can certainly skew results one way or the other, which is what appears to have happened here.

There appears not to have been any interaction between the two significant factors of age and sentence position, or between either significant factor and any of the insignificant factors. This is not unexpected, as these two factors belong to two different categories: age is a sociocultural variable, and sentential position is a syntactic variable. The GoldVARB analysis bore this out; in a variable rules analysis, factors that interact return different factor weights when calculated together. There was no change in the factor weights for age and sentential position worth mentioning⁴.

4. CONCLUSION

As mentioned at the beginning of this study, two hypotheses were proposed to be tested: that *este* is the hesitation conversational marker most often used in San Juan, and that gender, age, education level, and sentential position would significantly influence the choice of *este* or *{eh}* as a marker. Just as was shown in Galué's (2002) study of conversational markers, *este* was shown to be the most favored, occurring in 54.2% of all instances of hesitation markers. One observed similarity between this study and that of Galué is that in both, the speech of a major Caribbean city (Caracas being the capital and largest city of Venezuela) was studied. This raises the question of whether *este* as a hesitation marker is characteristic of North American or Caribbean Spanish. To test this, a further study of this marker would include other Hispanophone countries as a factor group. Furthermore, only two of the four tested factors—age and sentential position—were shown to significantly affect the choice between *este* and *eh*. Educational attainment, though significant in Soler Arechalde's (2006) study, was not determined to be a significant factor here; this indicates a potentially salient difference in the stigma attached to *este* in Mexico City versus in San Juan. Nor was gender determined to be significant, unlike in Soler Arechalde's (2006, 2008) studies; this finding possibly indicates a lesser degree of gender marking in San Juan. It was noted that, though the relative frequencies of the *este-eh* pair were minimally different between men and women, men showed a higher overall frequency of use of these two hesitation markers. This study in no way concludes that men use more hesitation markers overall than women do; though it was not examined, there exists the possibility that women in San Juan preferred other hesitation markers such as *pues* 'well' and *bueno* 'okay' instead. A future study might prove beneficial in testing this and would once again apply Galué's (2002) findings to Puerto Rican data, revealing either a consistency or a contrast be-

⁴ Though there were no interacting factors in the analysis, the cross-tabulations of said factors can be found in the appendix at the end of this article.

tween the two areas. This study could also be improved by including more syntactic factors, such as the part of speech preceding or following the marker, or whether the marker follows a turn in the conversation or is within the same speaker's turn.

At this point in time, it does not appear that *eh* is disappearing from use in San Juan, since younger speakers still use it. There are a few explanations as to the difference in frequency between the two markers. The more frequent use of *este* as a conversational marker as opposed to *eh* by younger speakers can be viewed in two ways: as an indication of a shift in preference over the generations, or as a marker of age. However, further study would be necessary to test which of the two situations is actually occurring. If one were to investigate the choice between *este* and *eh* as a marker of age or a shift in preference, a series of follow-up studies would need to take place with the same speakers, but at different times in their lives: once when they entered middle age, and again at age 55 or older (in accordance with the PRESEEA protocol). The observed frequencies would then be compared to those of the current study and analyzed: if the same speakers in the future produce *este* as frequently as they do now, the logical conclusion would be a shift in preference. On the other hand, if there is a much stronger tendency toward *eh* or a variant thereof, it can be said that *este* is more of a marker of youth.

REFERENCES

- ANTÚNEZ PÉREZ, I. (2005), "Verbal fillers: a paralinguistic element present in communicative competence", in *Investigación y educación*, 16, 1-21.
- GALUÉ, D. (2002), "Marcadores conversacionales: un análisis pragmático", in *Boletín de Lingüística*, 18, 27-48.
- GRAHAM, L. A. (2010), "Demonstratives as referential expressions in Puerto Rican Spanish", unpublished manuscript.
- MONTES, R. G. (1999), "The development of discourse markers in Spanish: interjections", *Journal of Pragmatics*, 31, 1289-1319.
- PORTOLÉS, J. (2001), *Marcadores del discurso*, Barcelona, Ariel.
- PRADA, J. J. (2001), *Marcadores del discurso en español: Análisis y representación*, MA thesis, Montevideo, Universidad de la República.
- PRESEEA (2003), *Metodología Proyecto para el Estudio Sociolingüístico del Español de España y de América*. Available at: < <http://preseea.linguas.net> > [Accessed 15 March 2011].
- SANKOFF, D, TAGLIAMONTE, S. A. & SMITH, E. (2006), "GoldVARB X", Department of Linguistics, University of Toronto, Toronto, Canada. Available at: < http://individual.utoronto.ca/tagliamonte/Goldvarb/GV_index.htm >.
- SOLER ARECHALDE, M.^a A. (2006), "El uso de *este*... en el habla de la ciudad de México", Alba Valencia Espinoza (Comp.): *Actas del XIV Congreso Internacional de la ALFAL*, Santiago de Chile. Available at: < http://www.mundoalfal.org/cdcongreso/cd/dialectologia_sociolingüistica/soler.html >.
- (2008), "Algunos factores determinantes y contextos de uso para el marcador discursivo *este*... en el habla de la ciudad de México", en *Anuario de Letras*, 46, 155-168.
- TORRES, L. (2002), "Bilingual discourse markers in Puerto Rican Spanish", in *Language in Society*, 31:1, 65-83.
- TORRES, L. & POTOWSKI, K. (2008), "A comparative study of bilingual discourse markers in Chicago Mexican, Puerto Rican, and MexiRican Spanish", in *International Journal of Bilingualism*, 12:4, 263-279.

APPENDIX

These are the cross-tabulations of the independent variables as calculated by GoldVARB. Since it was determined that none of the variables were interacting, they are included here and not the main body of the report.

Age (horizontal) and gender (vertical)

| | 1 | | 3 | | 2 | | Σ | |
|----|-----|-----|-----|-----|-----|----|------|----|
| | + | % | + | % | + | % | + | % |
| m | 117 | 80: | 23 | 13: | 165 | 60 | 305 | 52 |
| e: | 29 | 20: | 149 | 87: | 109 | 40 | 287 | 48 |
| Σ: | 146 | : | 172 | : | 274 | | 592 | |
| f | 76 | 67: | 33 | 33: | 149 | 65 | 258 | 58 |
| e: | 37 | 33: | 67 | 67: | 81 | 35 | 185 | 42 |
| Σ: | 113 | : | 100 | : | 230 | | 443 | |
| Σ | 193 | 75: | 56 | 21: | 314 | 62 | 563 | 54 |
| e: | 66 | 25: | 216 | 79: | 190 | 38 | 472 | 46 |
| Σ: | 259 | : | 272 | : | 504 | | 1035 | |

Age (horizontal) and education (vertical)

| | 1 | | 3 | | 2 | | Σ | |
|----|-----|-----|-----|-----|-----|----|------|----|
| | + | % | + | % | + | % | + | % |
| H | 57 | 66: | 32 | 18: | 185 | 62 | 274 | 48 |
| e: | 29 | 34: | 150 | 82: | 112 | 38 | 291 | 52 |
| Σ: | 86 | : | 182 | : | 297 | | 565 | |
| S | 136 | 79: | 18 | 64: | 53 | 50 | 207 | 67 |
| e: | 37 | 21: | 10 | 36: | 53 | 50 | 100 | 33 |
| Σ: | 173 | : | 28 | : | 106 | | 307 | |
| P | 0 | --: | 6 | 10: | 76 | 75 | 82 | 50 |
| e: | 0 | --: | 56 | 90: | 25 | 25 | 81 | 50 |
| Σ: | 0 | : | 62 | : | 101 | | 163 | |
| Σ | 193 | 75: | 56 | 21: | 314 | 62 | 563 | 54 |
| e: | 66 | 25: | 216 | 79: | 190 | 38 | 472 | 46 |
| Σ: | 259 | : | 272 | : | 504 | | 1035 | |

Age (horizontal) and sentential position (vertical)

| | 1 | | 3 | | 2 | | Σ | |
|----|-----|-----|-----|-----|-----|----|------|----|
| | + | % | + | % | + | % | + | % |
| n | 153 | 72: | 45 | 20: | 271 | 60 | 469 | 53 |
| e: | 60 | 28: | 177 | 80: | 180 | 40 | 417 | 47 |
| Σ: | 213 | : | 222 | : | 451 | | 886 | |
| i | 40 | 87: | 11 | 22: | 43 | 81 | 94 | 63 |
| e: | 6 | 13: | 39 | 78: | 10 | 19 | 55 | 37 |
| Σ: | 46 | : | 50 | : | 53 | | 149 | |
| Σ | 193 | 75: | 56 | 21: | 314 | 62 | 563 | 54 |
| e: | 66 | 25: | 216 | 79: | 190 | 38 | 472 | 46 |
| Σ: | 259 | : | 272 | : | 504 | | 1035 | |

Gender (horizontal) and education (vertical)

| | m | | % | | f | | % | | Σ | | % | |
|------|-----|-----|-----|----|------|----|---|---|---|---|---|---|
| | + | - | + | - | + | - | + | - | + | - | + | - |
| H E: | 149 | 48: | 125 | 49 | 274 | 48 | | | | | | |
| e: | 159 | 52: | 132 | 51 | 291 | 52 | | | | | | |
| Σ: | 308 | : | 257 | | 565 | | | | | | | |
| S E: | 112 | 65: | 95 | 70 | 207 | 67 | | | | | | |
| e: | 60 | 35: | 40 | 30 | 100 | 33 | | | | | | |
| Σ: | 172 | : | 135 | | 307 | | | | | | | |
| P E: | 44 | 39: | 38 | 75 | 82 | 50 | | | | | | |
| e: | 68 | 61: | 13 | 25 | 81 | 50 | | | | | | |
| Σ: | 112 | : | 51 | | 163 | | | | | | | |
| Σ E: | 305 | 52: | 258 | 58 | 563 | 54 | | | | | | |
| e: | 287 | 48: | 185 | 42 | 472 | 46 | | | | | | |
| Σ: | 592 | : | 443 | | 1035 | | | | | | | |

Gender (horizontal) and sentential position (vertical)

| | m | | % | | f | | % | | Σ | | % | |
|------|-----|-----|-----|----|------|----|---|---|---|---|---|---|
| | + | - | + | - | + | - | + | - | + | - | + | - |
| n E: | 234 | 47: | 235 | 60 | 469 | 53 | | | | | | |
| e: | 261 | 53: | 156 | 40 | 417 | 47 | | | | | | |
| Σ: | 495 | : | 391 | | 886 | | | | | | | |
| i E: | 71 | 73: | 23 | 44 | 94 | 63 | | | | | | |
| e: | 26 | 27: | 29 | 56 | 55 | 37 | | | | | | |
| Σ: | 97 | : | 52 | | 149 | | | | | | | |
| Σ E: | 305 | 52: | 258 | 58 | 563 | 54 | | | | | | |
| e: | 287 | 48: | 185 | 42 | 472 | 46 | | | | | | |
| Σ: | 592 | : | 443 | | 1035 | | | | | | | |

Education (horizontal) and sentential position (vertical)

| | H | | % | | S | | % | | P | | % | | Σ | | % | |
|------|-----|-----|-----|-----|-----|----|------|----|---|---|---|---|---|---|---|---|
| | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - |
| n E: | 216 | 47: | 179 | 65: | 74 | 48 | 469 | 53 | | | | | | | | |
| e: | 242 | 53: | 96 | 35: | 79 | 52 | 417 | 47 | | | | | | | | |
| Σ: | 458 | : | 275 | : | 153 | | 886 | | | | | | | | | |
| i E: | 58 | 54: | 28 | 88: | 8 | 80 | 94 | 63 | | | | | | | | |
| e: | 49 | 46: | 4 | 12: | 2 | 20 | 55 | 37 | | | | | | | | |
| Σ: | 107 | : | 32 | : | 10 | | 149 | | | | | | | | | |
| Σ E: | 274 | 48: | 207 | 67: | 82 | 50 | 563 | 54 | | | | | | | | |
| e: | 291 | 52: | 100 | 33: | 81 | 50 | 472 | 46 | | | | | | | | |
| Σ: | 565 | : | 307 | : | 163 | | 1035 | | | | | | | | | |