

UNIVERSIDAD CENTRAL DE VENEZUELA  
FACULTAD DE HUMANIDADES Y EDUCACIÓN  
COMISIÓN DE ESTUDIOS DE POSTGRADO  
ÁREA DE LINGÜÍSTICA



LOS SUJETOS PRONOMINALES DE PRIMERA PERSONA EN EL  
ESPAÑOL DE CARACAS 2004-2010

Vanessa Pérez Brabandere

Caracas, octubre de 2010

UNIVERSIDAD CENTRAL DE VENEZUELA  
FACULTAD DE HUMANIDADES Y EDUCACIÓN  
COMISIÓN DE ESTUDIOS DE POSTGRADO  
ÁREA DE LINGÜÍSTICA

**Maestría en Lingüística**

LOS SUJETOS PRONOMINALES DE PRIMERA PERSONA EN EL  
ESPAÑOL DE CARACAS 2004-2010

AUTORA: VANESSA PÉREZ BRABANDERE

Trabajo de grado para  
optar al grado de Magíster  
Scientiarum en Lingüística

CARACAS, OCTUBRE DE 2010

## **AGRADECIMIENTOS**

Quisiera expresar un agradecimiento especial a mi tutora Paola Bentivoglio, quien no sólo me brindó cuidadosas lecturas y parte de su tiempo en todas las etapas de la tesis, sino que con su ejemplo me transformó en una mejor persona.

Agradecimiento que extiendo a todos los profesores del área, quienes, además de impartir conocimientos, llenan de sabios consejos la vida de aquellos que, como yo, tenemos el honor de conocerlos.

Por último, pero no menos importante, doy gracias a todos esos amigos y familiares que me han acompañado durante todo este largo proceso. En particular a Zayra, mi hermano Leonard y mi Nonna.

## RESUMEN

La variación presencia~ausencia del sujeto pronominal de primera persona en la lengua española ha sido estudiada por diferentes autores. En las gramáticas (Real Academia Española 2009, Gili y Gaya 1976, Alarcos Llorach 1994, entre otros) se afirma que el pronombre está presente sólo cuando se quiere enfatizar el referente del sujeto o bien eliminar la ambigüedad creada por la ausencia del pronombre; en los estudios variacionistas (Bentivoglio 1987, Silva-Corvalán 1982 y 2003, entre otros) se ha podido comprobar que el pronombre está presente por otras razones, sobre todo de carácter semántico-discursivo. De acuerdo con la segunda perspectiva, en la presente investigación se analizará la variación entre presencia y ausencia del sujeto pronominal de primera persona singular (*yo*) y plural (*nosotros/as*) en el Corpus sociolingüístico de Caracas PRESEEA 2004-2010. La muestra que analicé se basa en 36 entrevistas hechas a hablantes caraqueños, 18 hombres y 18 mujeres, cuyo grado de instrucción es de al menos 15 años de escolaridad formal; seleccioné los 100 primeros casos con un verbo finito en 1ª persona singular o plural, por un total de 3.600 cláusulas. Cada caso se codificó de acuerdo con un conjunto de variables lingüísticas y extralingüísticas. Las variables lingüísticas son: presencia y ausencia de sujeto pronominal de primera persona; número gramatical del sujeto; ambigüedad de la forma verbal; clase semántica y tiempo del verbo; polaridad de la cláusula (cláusula afirmativa o negativa); cambio de turno en la conversación; conectividad discursiva; distancia referencial; y efecto repetición o *priming*; mientras que las extralingüísticas son edad y sexo de los hablantes. Analizo cualitativamente la frecuencia de los verbos. Los casos codificados fueron sometidos a un análisis probabilístico mediante el programa GoldVarb X (Sankoff, Tagliamonte y Smith 2005), lo que permitió determinar el efecto de las variables independientes sobre la variable dependiente (presencia~ausencia del sujeto). Los resultados permitirán caracterizar la manifestación del sujeto de primera persona en el español actual de Venezuela y serán comparados con los obtenidos en investigaciones anteriores.

## ÍNDICE

	Páginas
0. INTRODUCCIÓN	1
1. PLANTEAMIENTO DEL PROBLEMA	4
1.1. Variación presencia~ausencia del pronombre 1ª persona	4
1.2. Justificación	5
1.3. Objetivos	7
1.3.1. General	7
1.3.2. Específicos	7
2. MARCO TEÓRICO	9
2.1. Bases teóricas	9
2.2. Antecedentes	11
3. METODOLOGÍA	24
3.1. El corpus	26
3.2. Selección de casos	26
3.3. Exclusiones	26
3.4. Variables de análisis	29
3.5. Análisis de los datos del corpus	30
4. EJEMPLOS Y RESULTADOS	32
4.1. Ejemplos, frecuencias y prueba de $\chi^2$	32
4.1.1. Número gramatical del sujeto	32
4.1.2. Ambigüedad	34
4.1.3. Tiempo/aspecto/modo del verbo	38
4.1.4. Polaridad de la cláusula	43
4.1.5. Cambio de turno	44
4.1.6. Clase semántica del verbo	47
4.1.7. Conectividad discursiva	50
4.1.8. Efecto repetición	54
4.1.9. Distancia referencial	57
4.1.10. Variables extralingüísticas	61

4.2. GoldVarb X	65
4.2.1. Modelo con Conectividad	66
4.2.2. Modelo con Efecto repetición/Distancia referencial	69
4.2.3. Comparación de resultados de los dos modelos con estudios previos	72
4.2.4. Ajuste del modelo	75
4.3. Frecuencia del verbo	77
5. CONCLUSIONES	80
6. REFERENCIAS BIBLIOGRÁFICAS	83
7. ANEXOS	89

## 0. INTRODUCCIÓN

Bajo el concepto de pronombres personales se agrupa un número limitado de formas cuyo significado remite a la persona gramatical. Hay tres personas (1ª, 2ª, y 3ª) y dos números gramaticales (singular y plural). Según Sedano (en preparación:146) la 1ª persona corresponde al emisor o a un grupo que lo incluye; la 2ª persona corresponde al interlocutor o a un grupo que lo incluye; la 3ª persona corresponde a la/s persona/s, objeto/s o evento que no son ni el emisor ni el receptor.

Los verbos conjugados del español ofrecen marcas formales que dependen de la persona y número del sujeto de la oración. Sedano (en preparación:139) explica que:

La estructura de un verbo está generalmente compuesta por un lexema o raíz + morfemas flexivos. Una expresión verbal como *bailábamos*, por ejemplo, tiene la siguiente estructura:

<b>Bailábamos</b>			
Lexema	Morfemas flexivos		
	Vocal temática	Tiempo/aspecto/modo	Persona/número
<i>Bail</i>	<i>-á</i>	<i>-ba</i>	<i>-mos</i>

Los morfemas flexivos de los verbos son los relacionados con tiempo, aspecto, modo, persona y número, y también vocal temática.

En nuestro idioma, un mismo morfema puede contener de forma sincrética la información relacionada con tiempo aspecto y modo, e incluso con persona y número. En *cant-o*, por ejemplo, el morfema *-o* contiene la información sobre el modo (indicativo), el tiempo (simultaneidad con el *origo*)<sup>1</sup>, el aspecto (imperfectivo), la persona (1ª) y el número (singular).

En la conjugación, las tres personas gramaticales (1ª, 2ª, y 3ª) pueden tener varios morfemas: en *bail -á -ba -mos*, por ejemplo, la 1ª persona plural del copretérito

---

<sup>1</sup> Sedano (en preparación) explica que el *origo* es un punto de referencia temporal deíctico “que normalmente coincide con el ‘ahora’ del emisor, es decir, con el momento en que éste emite su enunciado” (140).

de indicativo sólo se expresa a través del morfema *-mos* en las tres conjugaciones, en ambos modos (indicativo y subjuntivo), y en todos los tiempos.

Una característica inherente al sistema de la lengua española y de algunas lenguas románicas (italiano y portugués, por ejemplo) y no románicas (árabe, mandarín, etc.), es que para expresar el sujeto se puede emplear o no un pronombre personal, así como también un sustantivo, cuando se trata de la 3<sup>a</sup> persona, alternancia conocida como LA EXPRESIÓN O MANIFESTACIÓN VARIABLE DEL SUJETO. Todo hablante del español, independientemente de su dialecto o nivel socioeconómico, hace uso en sus enunciados de la expresión variable del sujeto. En la presente investigación se analiza la variación entre la presencia de un sujeto pronominal de primera persona singular y plural<sup>2</sup> vs. el cero fónico en español, según se ilustra en el ejemplo (1):

- (1) a. /sujeto pronominal/  
 YO<sup>3</sup> *tengo* mis proyectos personales de escritura (c31hf.06)<sup>4</sup>  
 b. /cero fónico/  
 Porque Ø no me *veo* haciendo otra cosa (c31hf.06)

La alternancia que me propuse analizar ha sido y sigue siendo objeto de estudio para muchos investigadores, especialmente los variacionistas, quienes han logrado delimitar alguno de los factores que influyen en la variación.

---

<sup>2</sup> En adelante, cuando mencione *sujeto* me estaré refiriendo al sujeto gramatical de primera persona singular y plural.

<sup>3</sup> Resaltaré en versales el sujeto pronominal que acompañe el verbo *y*, en cursiva, el verbo. Cuando el sujeto no está expresado, empleo el símbolo del cero lingüístico (Ø).

<sup>4</sup> El código alfanumérico que identifica al hablante se lee de izquierda a derecha de la siguiente manera: i) c = Caracas; ii) grado de instrucción: 3 = estudios universitarios; iii) grupo etario: 1 = 20-34 años, 2 = 35-54 años, y 3 = 55 años o más; iv) sexo: h = hombre y m = mujer; y v) letra de identificación del hablante en la casilla (a, b, c, d, e, y f).



En Venezuela se cuenta con una investigación pionera (Bentivoglio 1980 y 1987)<sup>5</sup> dedicada al estudio de la alternancia del sujeto explícito de primera persona (*yo como, nosotros comemos*) en contraposición al implícito ( $\emptyset$  *como,  $\emptyset$  comemos*) en una muestra de habla de doce caraqueños pertenecientes a un mismo grupo de edad (30-45 años), y a dos niveles socioeconómicos (alto y bajo).

En la presente investigación estudio la EXPRESIÓN del sujeto de primera persona, cuando hay la opción de expresarlo u omitirlo con verbos en forma finita, en un corpus de entrevistas –grabadas entre 2004-2010– a treinta y seis hablantes nacidos en Caracas, todos con un grado de instrucción alto. Tomo en cuenta variables tanto lingüísticas como extralingüísticas; entre las primeras incluyo algunas discursivas y pragmáticas que, hasta donde tengo conocimiento, no han sido consideradas en estudios anteriores sobre el español de Venezuela. Con respecto a las extralingüísticas tomo en cuenta sólo el sexo y la edad del hablante, pues en el corpus que analizo solamente estaba disponible, hasta el momento de hacer el estudio, el nivel 3, cuyos 36 hablantes tienen más de quince años de educación formal.

En el estudio presento, en primer lugar, el planteamiento del problema, la justificación y los objetivos. Luego, describo el marco teórico y los antecedentes sobre los que se basa el análisis, así como la metodología que emplearé. Por último, presentaré los resultados obtenidos, las conclusiones, y señalaré la bibliografía consultada.

---

<sup>5</sup> Bentivoglio 1980 y 1987 son dos versiones de un mismo texto, en inglés y en español, respectivamente. En adelante me referiré sólo a la versión en español (Bentivoglio 1987).

## 1. PLANTEAMIENTO DEL PROBLEMA

### 1.1. Variación presencia~ausencia del pronombre 1ª persona

El hecho de que en el español, como en el portugués, el italiano y el catalán, el verbo conjugado tenga un morfema que remite al sujeto determina que no sea obligatoria la presencia de un elemento (pronombre o sustantivo) que desempeñe la función gramatical de sujeto. Autores de gramáticas como Gili y Gaya 1976 y Alarcos Llorach 1994 consideran que hay sólo dos circunstancias en que nuestro idioma necesita manifestar más claramente la expresión del sujeto contenida en la forma verbal: “a) Cuando se quiere resaltar la participación del sujeto en la acción, como insistiendo en que es aquél y no otro. [...]; b) Cuando pueda haber ambigüedad. [...]. Fuera de estos casos, el empleo del sujeto unido a la forma verbal comunica al estilo español extraordinaria pesadez” (Gili y Gaya, 1976: 23-24).

La expresión del sujeto pronominal en español ha sido objeto de muchos estudios variacionistas a partir de la investigación de Barrenechea y Alonso 1977 así como de las tesis de maestría de Silva-Corvalán 1977<sup>6</sup> y de Bentivoglio 1980. A pesar de que en los diferentes estudios hay similitudes entre las variables propuestas para explicar que la presencia o ausencia del pronombre personal no se debe al azar, los resultados en cuanto a qué variables afectan a la expresión del sujeto difieren de un estudio al otro (cf. Travis 2007); todos han demostrado, sin embargo, que las

---

<sup>6</sup> La tesis de Silva-Corvalán 1977 nunca ha sido publicada por completo; la primera publicación parcial es un artículo que apareció en 1982 (Silva-Corvalán 1982).

razones por las que un hablante utiliza o no un pronombre personal con función de sujeto, cuando este no es obligatorio, son más complejas que las señaladas en las gramáticas mencionadas:

Si de veras la presencia de los pronombres sujeto dependiera solamente del deseo del hablante de enfatizar lo que él/ella hace, piensa, etc., y/o por la necesidad de aclarar las terminaciones verbales ambiguas, un excesivo número de pronombres sujeto quedaría sin justificación alguna. (Bentivoglio, 1987:21)

Partiendo de los estudios mencionados anteriormente, me propuse verificar si los factores sensibles al discurso delimitados en investigaciones posteriores (AMBIGÜEDAD CONTEXTUAL, CAMBIO DE TURNO, CONECTIVIDAD DISCURSIVA, CLASE SEMÁNTICA DEL VERBO, DISTANCIA REFERENCIAL, y EFECTO REPETICIÓN), y los factores sociales (EDAD y SEXO) contribuyen significativamente a la expresión variable del sujeto de primera persona en una muestra del habla actual de Caracas.

Es importante mencionar que el uso del pronombre sujeto constituye también un problema en la enseñanza del español como segunda lengua, sobre todo para hablantes de lenguas como el inglés o el francés, en las que el pronombre es obligatorio.<sup>7</sup>

## 1.2. Justificación

La presente investigación se justifica porque no hay estudios recientes sobre la variación presencia~ausencia del pronombre sujeto de primera persona en el español de Caracas, ya que las únicas muestras de este dialecto, analizadas hasta ahora son las

---

<sup>7</sup> En inglés hay algunos contextos en los que se permite que el sujeto gramatical no sea expreso, pero no se trata de la mayoría de los casos.

del corpus de Caracas 1977 (Bentivoglio 1987), en la que sólo se estudia el habla de doce (12) personas de un mismo grupo generacional (30-45 años) y dos niveles socioeconómicos (bajo y alto), mientras que la segunda investigación extraída del corpus del habla culta de Caracas, grabado entre 1968 y 1977, usa una muestra de doce (12) hablantes de edad comprendida entre los 30 y 45 años (Bentivoglio 1999). Además, los estudios realizados en otros dialectos del español han demostrado la relevancia de ciertas variables que, hasta ahora, no han sido empleadas para analizar la variabilidad del sujeto en ningún dialecto de Venezuela. Vale la pena entonces preguntarse una vez más en qué contextos resulta necesaria la expresión del sujeto, y qué factores lingüísticos o extralingüísticos podrían estar condicionando la presencia o ausencia de los sujetos de primera persona en el español actual de Caracas.

La muestra de habla que analicé procede de treinta y seis (36) hablantes caraqueños, hombres y mujeres en igual número, pertenecientes a tres grupos generacionales (20-34 años, 35-54 años, +55 años). El corpus analizado no está estratificado socioculturalmente, ya que todos los hablantes tienen como mínimo quince (15) años de educación formal. Esto se debe, como ya lo he mencionado en la introducción, al hecho de que las transcripciones de los hablantes de los otros dos niveles socioculturales (1 y 2) no están a la disposición de los investigadores puesto que aún están en fase de corrección. Esta carencia, sin embargo, no debería constituir un problema ya que según Samper Padilla, Hernández Cabrera y Troya Déniz 2006 tiene un interés especial “el análisis de muestras relativamente amplias de hablantes del nivel de instrucción superior, que, al menos teóricamente, están más cerca de la

norma estándar hispánica” (91). Por esta razón, considero que el desarrollo de la investigación propuesta incorpora al estudio de la expresión variable del sujeto de primera persona un corpus relativamente extenso con la posibilidad de comparar los resultados con los de otros corpus de habla caraqueña, grabados hace aproximadamente cuarenta años—el del habla culta fue grabado entre 1968 y 1977— y treinta años —el sociolingüístico se grabó en 1977. Cabe notar que el corpus de 1977 tiene características sociales similares a las del corpus objeto de estudio de la presente investigación.

He preferido adoptar una perspectiva variacionista, pues, como señala Silva-Corvalán 2003, quien ha propuesto explicaciones convincentes de la expresión variable del sujeto en el español, el método variacionista “es el único que nos permite avanzar nuestro conocimiento” de algunos factores semántico-pragmáticos y “el único que nos permite estudiar los diversos dialectos del español” (859).

### 1.3 Objetivos

#### 1.3.1. General

- Determinar los factores sintáctico-semántico-pragmáticos que influyen en la expresión variable del sujeto pronominal de primera persona en el español hablado actualmente en Caracas.

#### 1.3.2. Específicos

- Averiguar si los factores lingüísticos co-varían con los factores sociales inherentes al corpus (edad y sexo).

- Obtener datos y argumentos que expliquen la presencia~ausencia del pronombre sujeto con el fin de poderlos aplicar en la enseñanza del español como segunda lengua.
- Comparar los resultados obtenidos en el análisis del corpus Caracas 2004-2010 con los estudios anteriores sobre otros corpus del español hablado.

## 2. MARCO TEÓRICO

### 2.1. Bases teóricas

La investigación se inscribe dentro del marco de la sociolingüística variacionista, disciplina que estudia la variación presente en las estructuras lingüísticas –en los cinco niveles de la lengua: fonológico, morfológico, sintáctico, léxico y semántico– y las correlaciona con las estructuras sociales (Bright 1974).

La sociolingüística variacionista se basa en una noción teórica básica –la variable dependiente lingüística– que se ha conceptualizado como dos o más maneras de decir la misma cosa.

El éxito logrado en los estudios de variación fonológica tanto por Labov 1966, 1969, 1972, 1975, como por Sankoff 1978, motivó a los sociolingüistas a aplicar las mismas técnicas y métodos al análisis de casos de aparente variación sintáctica. Lavandera, en su artículo de 1978 titulado *Where does the sociolinguistic variable stop?*, cuestionó la posibilidad de extender la noción de variable sociolingüística a otros niveles descriptivos además del fonológico, puesto que la variación sintáctica implica resolver el problema de las posibles diferencias de significado que pueden estar asociadas a cada variante de una misma variable. En fonología, dos o más formas, o alófonos, de un fonema no cambian el significado de la palabra en la que está el segmento que se analiza, y menos aún el del enunciado: un ejemplo es la alternancia del fonema oclusivo dental sonoro /d/ en posición intervocálica, el cual en un término como *dedo* se puede retener ['dedo], debilitar ['de $\delta$ o], o elidir ['de $\emptyset$ o], sin

que haya cambios en el significado, mientras que las variantes de una variable sintáctica no son claramente dos o más formas de decir lo mismo, como también argumentan tanto Romaine 1981 y 1984, como García 1985.

La raíz del problema de la existencia o no de variación sintáctica radica en la cuestión del significado: algunos investigadores se inclinan por la posible sinonimia o variación estricta, mientras que otros sostienen que “en todo contexto diferentes formas tienen significados diferentes” (Silva-Corvalán, 2003:858). Sankoff 1988, al proponer la noción de distribución complementaria débil en la comunidad,<sup>8</sup> afirma que hay dos posibilidades cuando nos encontramos ante una posible variación sintáctica: i) las dos formas en variación tienen la misma función, es decir, su significado se neutraliza en el discurso; o ii) cada una cumple funciones diferentes y, por lo tanto, estas funciones alternan. De acuerdo con lo señalado, para afirmar que estamos frente a un caso de variación lingüística, pese a que las formas sean diferentes, lo realmente importante es que las variantes puedan considerarse como “dos o más realizaciones equivalentes de un elemento común que dicen lo mismo” (Silva-Corvalán 2001:135).

Los estudios previos en los que se han aplicado el método variacionista y las técnicas de análisis estadístico muestran que la variación entre la presencia y la ausencia de la forma pronominal del sujeto no es arbitraria y es funcionalmente

---

<sup>8</sup> En fonología se usa el término de distribución complementaria para la relación entre dos sonidos que presenten similitud fonética, de manera que, dado un contexto, sólo puede aparecer un sonido, pero no el otro. Un ejemplo es la consonante nasal alveolar [ŋ], que se velariza ante consonante velar [ˈaʊŋke], mientras que el alófono [n] no puede aparecer en esa posición.



equivalente. En investigaciones se han propuesto algunas de las variables que condicionan tal alternancia, aun cuando en futuras investigaciones podrían añadirse otras.

## 2.2. Antecedentes

Reseñaré, brevemente, diversos estudios relacionados con la presencia~ausencia de los pronombres personales. He revisado varias investigaciones sobre la presencia~ausen Pelly 1984, Morales 1986 y 1997, D’Introno 1989, Fernández Soriano 1999, Luján 1999, Enríquez 2000, Hurtado 2001, Lamoglia Duarte 2003, Otheguy y Zentella 2007, Meneses de Santana 2008 , pero en este apartado sólo me ocuparé de aquellos en los que se analiza la expresión variable del sujeto de primera persona, o en los que se proponen variables que puedan servir de apoyo a mi investigación.

En las páginas 13 y 14 se pueden apreciar los cuadros 1 y 2, respectivamente. En el cuadro 1 presento, en orden cronológico, las variables o conjunto de factores utilizados en algunas de las investigaciones en las que se ha tratado de explicar la manifestación variable del sujeto tanto en el español como en el portugués de Brasil; este cuadro 1 refleja los factores empleados para el estudio del sujeto de primera persona, factores que he agrupado en sintácticos, semánticos, pragmáticos y de modalidad textual, con el fin de facilitar su identificación y de poder apreciar cómo el estudio de la expresión variable del sujeto se ha hecho más complejo a lo largo del último trentenio. En el cuadro 2 se ofrecen los detalles de los corpus empleados en las

investigaciones mencionadas en el cuadro 1. Las especificaciones de los corpus se encuentran en Esgueva y Cantarero 1981; de Pineda 1983; Morales y Vaquero 1990; Bentivoglio y Sedano 1993; de Paiva y Lamoglia Duarte 2003; Blanco Canales 2004, así como en las publicaciones de la Universidad Nacional Autónoma de México 1971, de la Universidad Central de Venezuela 1979, y de la Universidad Nacional de Buenos Aires 1987.

Desde las primeras investigaciones, el estudio de la *presencia y ausencia* del pronombre se ha hecho más complejo, haciendo que el punto de vista se haya extendido desde la gramática a la semántica y a la pragmática, lo que abarca una gran parte del espectro del análisis lingüístico: en el cuadro 1 se mencionan los autores en la columna (1) de acuerdo con la fecha de publicación de sus investigaciones. Cantero 1976 toma en cuenta la AMBIGÜEDAD DE LA FORMA VERBAL, TIEMPO, ASPECTO Y MODO ( de ahora en adelante, TAM) DEL VERBO y el ÉNFASIS; Barrenechea y Alonso 1977, sólo la PERSONA y NÚMERO GRAMATICAL así como la AMBIGÜEDAD DE LA FORMA VERBAL; Silva-Corvalán 1977 estudia la CORREFERENCIALIDAD entre el sujeto de la cláusula bajo estudio y el de la cláusula anterior y también la AMBIGÜEDAD; Bentivoglio 1987 toma en cuenta todas las variables anteriores –excepto la PERSONA GRAMATICAL, pues sólo estudia la primera persona– y añade el NÚMERO GRAMATICAL, EL CAMBIO DE TURNO, la SEMÁNTICA DEL VERBO y el ÉNFASIS. Cifuentes 1980-1981 estudia el NÚMERO, la AMBIGÜEDAD, TAM DEL VERBO y el ÉNFASIS; Morales 1982, en un corpus bilingüe, quiere determinar si la variación se debe a la influencia del inglés (lengua en que el sujeto es obligatorio) o a la COMPENSACIÓN DE LA ELISIÓN DE /-S/ en la flexión verbal de la segunda persona, junto con variables como la AMBIGÜEDAD, el ÉNFASIS, el CAMBIO DE TURNO, y la CORREFERENCIALIDAD; Enríquez 1984 estudia variables como NÚMERO, TIEMPO, ASPECTO Y MODO VERBAL, SEMÁNTICA DEL VERBO, AMBIGÜEDAD, POLARIDAD, ORACIONES COMPUESTAS, y ESTILO DEL ENUNCIADO (*informal/formal*). Hochberg 1986, junto al NÚMERO y al

ÉNFASIS, estudia la AMBIGÜEDAD y la ELISIÓN DEL MORFEMA DE PERSONA VERBAL /-s/, y asevera que la ambigüedad se ha extendido a más personas gramaticales, lo que, por lo tanto, produce una mayor frecuencia de expresión del pronombre; Miró Vera y de Pineda 1990 se ocupan de la ambigüedad causada por formas verbales homónimas y estudian las variables PERSONA, AMBIGÜEDAD, POLARIDAD, DEPENDENCIA DE LA CLÁUSULA y MODALIDAD DISCURSIVA en la lengua *hablada* vs. la *escrita*. Cameron 1992, en su disertación doctoral, estudia las variables NÚMERO, TAM DEL VERBO y CAMBIO DE REFERENCIA; Paredes Silva 1993 estudia las variables AMBIGÜEDAD, ÉNFASIS y amplía la variable CORREFERENCIALIDAD proponiendo una mucho más compleja que llama CONECTIVIDAD DISCURSIVA; Martín Rojo y Meewis 1993 sostienen que la variación del pronombre se puede explicar con la TEORÍA PRAGMÁTICA DE TÓPICO/COMENTARIO; Bayley y Pease-Álvarez 1996 tratan de determinar si la expresión del sujeto en el habla de los niños bilingües español/inglés está sujeto a las limitaciones reportadas para las variedades que no están en contacto con otra lengua y, para ello, proponen las variables NÚMERO, CAMBIO DE REFERENCIA, y AMBIGÜEDAD. Cameron 1997 continúa sus análisis de la expresión del sujeto y trata de caracterizar la variable CAMBIO DE REFERENCIA basándose en la teoría de la accesibilidad, propuesta por Ariel 1990, y la de continuidad de tópico, de Givón 1983; Blanco Canales 1999 emplea las variables NÚMERO GRAMATICAL, SEMÁNTICA DEL VERBO, CORREFERENCIALIDAD y AMBIGÜEDAD; Bentivoglio 1999, en una muestra del habla culta de Caracas, analiza la variación mediante las variables AMBIGÜEDAD, ÉNFASIS, CAMBIO DE TURNO, CORREFERENCIALIDAD, y SEMÁNTICA DEL VERBO (ésta

última variable tiene variantes diferentes a la de su investigación de 1987). Stewart 2001, basándose en la teoría de la cortesía desarrollada por Brown y Levinson (1978, 1987), sostiene que el uso variable de *yo* se debe a procesos relacionados con la protección de la imagen (*face*) por parte del hablante; Cameron y Flores Ferrán 2003 comienzan analizando la CORREFERENCIALIDAD y, luego, deciden incorporar a su investigación la variable *PRIMING* (la cual se llamaré de ahora en adelante, EFECTO REPETICIÓN), basada en un concepto bien conocido de la Psicolingüística; Paredes Silva 2003 desarrolla la variable correferencialidad adaptándola a un corpus oral (su estudio anterior era sobre un corpus escrito) y, además, incluye las variable NÚMERO y ÉNFASIS. Quijada Bringtown 2006, en su estudio sobre la 3<sup>a</sup> persona, utiliza las variables NÚMERO, AMBIGÜEDAD, CONECTIVIDAD DISCURSIVA, SEMÁNTICA DEL VERBO y CAMBIO DE TURNO; Samper Padilla, Hernández Cabrera, y Troya Déniz 2006, adoptan las variables AMBIGÜEDAD, CAMBIO DE REFERENCIA, y SEMÁNTICA DEL VERBO; Travis 2007 utiliza las variables TIEMPO, ASPECTO Y MODO (TAM) DEL VERBO, TIPOS DE TEXTOS, DISTANCIA REFERENCIAL, y EFECTO REPETICIÓN; finalmente, Torre-Cacoullós y Travis 2008 usan las variables AMBIGÜEDAD, CAMBIO DE TURNO, TIPOS ORACIONALES, SEMÁNTICA DEL VERBO, CORREFERENCIALIDAD, y EFECTO REPETICIÓN.

Como se aprecia en el cuadro 1, las variables más utilizadas en los estudios sobre el sujeto variable son la AMBIGÜEDAD en quince (15), el CAMBIO DE REFERENCIA en diez (10), el NÚMERO GRAMATICAL en diez (10), el TIPO DE VERBO en

ocho (8), el ÉNFASIS en (8), el CAMBIO DE TURNO en cinco (5), y TIEMPO/ASPECTO/MODO (TAM) DEL VERBO en cinco (5). La CONECTIVIDAD DISCURSIVA y el EFECTO REPETICIÓN han sido empleados en tres (3) investigaciones; la ELISIÓN FUNCIONAL DE LA –S, la TIPOLOGÍA TEXTUAL, y la POLARIDAD solamente se han tomado en cuenta en dos (2). Las variables usadas en una sola investigación son la TEORÍA DE LA ACCESIBILIDAD, la TEORÍA DE LA CORTESÍA, la TEORÍA PRAGMÁTICA TÓPICO/COMENTARIO, y la DISTANCIA REFERENCIAL.

De todas las variables mencionadas es necesario describir con más detalle la CONECTIVIDAD DISCURSIVA, el EFECTO REPETICIÓN, y la DISTANCIA REFERENCIAL. Estas tres variables han surgido de la reflexión teórica en torno a la variable CAMBIO DE REFERENTE, que en las primeras investigaciones sólo tenía dos variantes: *referencia igual o diferente*.

Como se aprecia en el Cuadro 2, Paredes Silva 1993 propone una variable escalonada, la CONECTIVIDAD DISCURSIVA, con varios grados. La autora afirma que cuanto más estrecha es la conexión entre un referente/sujeto y su mención previa, menor es la necesidad de explicitarlo a través de un pronombre o de un sustantivo (en caso de que el sujeto sea de tercera persona). En 1993, la lingüista brasileña analiza un corpus escrito y propone seis grados de conectividad discursiva. En su investigación del 2003, en la que analiza un corpus de habla, aumenta a siete grados la escala de conectividad discursiva originalmente propuesta en 1993, pues considera necesario tomar en cuenta la intervención del entrevistador, quien muchas veces

estimula y mantiene la referencia a través de las preguntas dirigidas al hablante, aunque esto produzca un corte en la secuencia. Paredes Silva 2003 ejemplifica los siete niveles de la escala de la siguiente manera:

- 1°. La conexión óptima, grado 1, corresponde a la permanencia, en la función de sujeto, del mismo referente/tópico, en el mismo plano discursivo (manifestado por el mantenimiento del sistema de tiempo-aspecto-modo verbal), como en el ejemplo siguiente:
  - (1) *Eu já desfilei* na S. Clemente, *ø desfilei* no Arrastão da Cascadura, *ø desfilei* no Folião do Botafogo, *ø desfilei* em uma porção de escola mas só que a escola de coração é o Salgueiro.<sup>9</sup>  
 ‘Yo ya *desfilé* en San Clemente, *ø desfilé* en el Arrastão de Cascadura, *ø desfilé* en el Folião de Botafogo, *ø desfilé* en una cantidad de escuelas pero sé que la escuela del corazón es la de Salgueiro’.
- 2°. El grado 2 indica un cambio de plano: a pesar de que se mantiene el mismo tópico como sujeto, hay un cambio de plano discursivo, del tipo figura-fondo, hecho-opinión, realidad-irrealidad, reflejada en modificaciones en el sistema de tiempo-aspecto-modo de las oraciones adyacentes.
  - (2) *Eu sempre morei em casa, nunca morei em apartamento. Estou satisfeita. Eu não queria* morar em apartamento não.  
 ‘Yo siempre viví en casa, nunca viví en apartamento. *Estoy* satisfecha. *Yo no quería* vivir en apartamento, ¡no!’
- 3°. En el grado 3, se mantiene el mismo referente del sujeto, a pesar del cambio de turno. En el ejemplo (3) el referente de la hablante, denominada como *a senhora* ‘usted’ por el entrevistador, es retomado en el turno de habla del entrevistado como *yo*.
  - (3) – *A sra. tem* lembrança de seus avós?  
 – Não, *eu* só *vi* meu avô, porque minha avó eu não conheci  
 ‘–¿*Usted tiene* recuerdos de sus abuelos?  
 – No *yo* solo *vi* a mi abuelo, porque a mi abuela yo no la conocí.’
- 4°. El grado 4 supone la interferencia de cláusulas impersonales entre el sujeto en estudio y su última mención, y representa un quiebre en la continuidad del referente como sujeto, como se ve en ejemplo de abajo:
  - (4) – *Eu pinto, eu desenho, eu pinto (hes) eu compro bonequinha de papel e tem* aqueles desenho pra pintar, *eu pinto*.  
 ‘– Yo pinto, yo dibujo, yo pinto (hes) *yo* compro muñequitas de papel y *hay estos* dibujos para pintar, *yo pinto*’.

---

<sup>9</sup> Paredes Silva 2003 especifica que las palabras subrayadas ilustran el caso en estudio, mientras que las palabras en cursivas remiten a la cláusula que precede el caso.

- 5°. El grado 5 se refiere a la interferencia de cláusulas cuyos sujetos remiten a otras personas gramaticales diferentes a la primera persona, pues desvían la atención del referente/tópico en estudio, como es el caso del sintagma *mis nietos* en el ejemplo:
- (5) E eu acredito que teve uma melhora muito grande. *Meus netos são* felizes, se relacionam muito bem com a mãe, e com o pai e são crianças verdadeiras que eu espero que continuem assim.  
 ‘Y yo creo que tuvo una mejoría muy grande. *Mis nietos son* felices, se relacionan muy bien con la mamá, con el papá y son niños de verdad que yo espero que continúen así’.
- 6°. En el grado 6, el referente que en las cláusulas previas tenía otra función desempeña el rol de sujeto gramatical. Se trata, por lo tanto, de un cambio de función. En el ejemplo (6) el referente/tópico predominante en las oraciones anteriores es *Dios*; el referente en primera persona cumple un papel secundario (como indican las menciones a través del pronombre oblicuo) incluso cuando asume la posición de tópico/sujeto en la última oración.
- (6) *Deus [...]* é o que me dá vontade, o que me dá força, sabe, tudo que me dá inspiração para mim fazer alguma coisa, isso é meu Deus, sabe? Tudo que faz as coisas andar certinha. Não sei explicar bem o que é, sabe? É aquilo que *me* faz ser o que eu sou.  
 ‘*Dios [...]* es el que me da voluntad, el que me da fuerza, sabe, todo lo que me da inspiración para hacer algunas cosas, eso es mi Dios, ¿sabe? Todo lo que hace que las cosas salgan bien. No sé explicar bien lo que es, ¿sabe? Es aquello que *me* hace ser lo que yo soy’.
- 7°. El grado 7 concierne al cambio de tópico/subtópico discursivo y representa un desvío en el rumbo del discurso. En el ejemplo (7) el tema del hablante sobre el tiempo de permanencia en el barrio representa una ruptura en la secuencia, retomada luego como respuesta a la pregunta de la entrevista.
- (7) – E você vai ao cinema?  
 – Olha, a última vez que *eu fui* ao cinema (risos) (falando rindo) eu tenho 20 anos de Campo Grande  
 – Sei  
 – Tá, eu já fui umas três vezes foi muito  
 ‘– ¿Y usted va al cine?’  
 – Oiga, la última vez que *yo fui* al cine (risas/habla riendo) yo tengo 20 años de Campo Grande  
 – Yo sé  
 – *Yo si fui* unas tres veces fue mucho’. (Paredes Silva, 2003:105-106; mi traducción del original al español).

Paredes Silva (1993, 2003) propone e implementa la CONECTIVIDAD DISCURSIVA a partir del CAMBIO DE REFERENCIA, pero no es la única que trata de optimizar esta última variable. Cameron 1992 se da cuenta de que los sujetos no



expresos se ven ligeramente favorecidos en ambientes correferenciales y ligeramente desfavorecidos en contextos de cambio de referencia (ver también Cameron y Flores-Ferrán 2003:47). Cameron 1997 trata de explicar el resultado obtenido en su estudio anterior y para ello emplea la TEORÍA DE LA ACCESIBILIDAD O CADENA DE TÓPICOS; en un estudio más reciente realizado con Flores-Ferrán en el 2003, desarrolla y describe la variable EFECTO REPETICIÓN. El EFECTO REPETICIÓN –término que viene de la investigación de divulgación-activación en Psicolingüística, donde se lo conoce también como *persistence*, *perseverance*, *perseveration*, o, en términos más coloridos, efecto “lobos de la misma camada”– consiste en el hecho de que los hablantes, sistemática, aunque inconscientemente, repiten formas y estructuras similares. Es decir, si hay una secuencia de referentes, hay que ver si todos mantienen la misma forma o no. En el caso de la primera persona, la forma variable es pronombre *vs.* Ø

Cameron y Flores-Ferrán (2003:48) codifican el EFECTO REPETICIÓN. Para ello toman en cuenta tanto el referente de primera persona de la cláusula bajo estudio (*target* ‘blanco’) como el referente de primera persona previo (*trigger* ‘gatillo’), que debe estar presente en una de las diez cláusulas anteriores, pues en caso contrario se considerará que el gatillo no tiene referente anterior de primera persona. Los autores sostienen que se puede demostrar estadísticamente la presencia del EFECTO REPETICIÓN si se comparan las probabilidades inherentes a la repetición de sujetos expresos en el segundo referente (la cláusula en estudio) en aquellas cláusulas con gatillo expreso y en aquellas con gatillo nulo.

Travis (2007:3) señala que el EFECTO REPETICIÓN puede ser léxico y estructural; ambas nociones son necesarias al analizar la presencia~ausencia del pronombre, pues una de las variantes, la del sujeto expreso, implica repetición léxica (*yo, nosotros/as*), mientras que la del sujeto no-expreso es únicamente un fenómeno estructural.

Otra variable relacionada tanto con la TEORÍA DE LA ACCESIBILIDAD como con el CAMBIO DE REFERENCIA es la DISTANCIA REFERENCIAL; en ella se asigna a cada expresión referencial un índice numérico que da cuenta de la distancia textual, medida en número de cláusulas interpuestas entre esa expresión y la anterior mención del mismo referente en el texto. Esta medida es inversamente proporcional a la accesibilidad referencial de las entidades del discurso: a menor distancia referencial, mayor accesibilidad, y a mayor distancia referencial, menor accesibilidad (Vázquez Rozas 2004:34).

Travis 2007 también sostiene que el EFECTO REPETICIÓN y la DISTANCIA REFERENCIAL actúan como tendencias contrarias, pues cuando el sujeto es expreso, el EFECTO REPETICIÓN favorece un sujeto expreso, mientras que la DISTANCIA REFERENCIAL favorece un próximo sujeto no-expreso; en cambio, cuando hay una gran distancia después de un sujeto no-expreso, el EFECTO REPETICIÓN favorece un sujeto no-expreso, en tanto que la DISTANCIA REFERENCIAL favorece un sujeto expreso, es decir, la DISTANCIA REFERENCIAL es inversamente proporcional al EFECTO REPETICIÓN.

Travis (2007:16) codifica la DISTANCIA REFERENCIAL tomando en cuenta hasta un máximo de diez (10) cláusulas, tanto principales como subordinadas, producidas tanto por el hablante como por el interlocutor. La investigadora analiza esta variable según las cláusulas que pueda haber entre dos sujetos. Las variantes son las siguientes: i) no hay ninguna cláusula; ii) hay una cláusula; iii) hay de dos a cuatro cláusulas; y iv) hay cinco o más cláusulas.

En mi investigación, al igual que en Travis 2007, tomaré en cuenta hasta un máximo de diez (10) cláusulas que precedan al referente en estudio.

### 3. METODOLOGÍA

#### 3.1. El corpus

El corpus está constituido por treinta y seis transcripciones del Corpus Sociolingüístico de Caracas 2004-2009, correspondientes a muestras de habla que fueron grabadas según los criterios establecidos en el Proyecto para el Estudio Sociolingüístico del Español de España y de América (PRESEEA).<sup>10</sup> La muestra de Caracas forma parte del corpus del habla de varias ciudades hispanohablantes, que – según el *desideratum* del proyecto– permitiría estudiar las distintas variedades del español hablado en sus distintos matices así como en su unidad. La recolección del Corpus Sociolingüístico de Caracas 2004-2010 (CSC 2004-10) ha sido coordinado por Paola Bentivoglio, Irania Malaver y María José Gallucci, de la Universidad Central de Venezuela, y financiado por el Consejo de Desarrollo Científico y Humanístico (CDCH) de esa universidad (Bentivoglio y Malaver 2006).

El corpus está compuesto por ciento ocho (108) grabaciones de habla semiformal del mismo número de hablantes caraqueños –nacidos en Caracas, de padres caraqueños–, con una duración mínima de cuarenta y cinco minutos cada una. En la grabación participaron el hablante y dos entrevistadores, quienes trataron de que la muestra fuera lo más espontánea posible (Bentivoglio y Sedano 1993:6; López Morales 2004:45-48; Silva-Corvalán 2001:52-53). La transliteración de las

---

<sup>10</sup> Para más información sobre el proyecto PRESEEA, se puede visitar la página [www.linguas.net/preseea/contenido/metodologia2.asp](http://www.linguas.net/preseea/contenido/metodologia2.asp)

grabaciones se hizo de la manera más fiel posible al texto grabado, según las pautas establecidas por el PRESEEA.

Los hablantes seleccionados para la presente investigación han cursado o están cursando estudios superiores, y están equitativamente distribuidos según los siguientes parámetros sociales: i) edad: tres grupos generacionales (12 hablantes por grupo) identificados con los números 1 (de 20 a 34 años), 2 (de 35 a 54 años), y 3 (55 años o más); y ii) sexo: 18 hombres/18 mujeres. A partir de estos parámetros, cada una de las transcripciones se identifica con un código que resume las características sociales de cada hablante. El código alfanumérico se lee de izquierda a derecha de la siguiente manera: i) c = Caracas; ii) grado de instrucción: 3 = alto, iii) grupo generacional: 1, 2, o 3; iv) sexo: h = hombre y m = mujer; v) letra que identifica al hablante de los otros cinco hablantes con igual caracterización social (*a, b, c, d, e, y f*); y vi) año en el que se realizó la grabación (04, 05, 06, 07, 08, 09). A modo de ilustración, el código *c32ma.05* corresponde a una transcripción del Corpus de Caracas (c), grabado en el año 2005, del habla de una mujer (m), cuya edad está comprendida entre los 35 y 54 años (grupo generacional 2), perteneciente al nivel sociocultural alto (3), y cuyo código identificativo es “a”. La distribución de hablantes en el corpus analizado se puede apreciar en el cuadro 3:

Cuadro 3: Distribución de los hablantes de nivel sociocultural 3 del CSC 2004-2010

	1: 20-34 años		2: 35-54 años		3: 55 años o más	
	Hombres	Mujeres	Hombres	Mujeres	Hombres	Mujeres
	6	6	6	6	6	6
	12		12		12	
<b>TOTAL</b>	<b>36</b>					

### 3.2. Selección de casos

En la presente investigación, los textos del corpus analizado se dividieron en cláusulas finitas, entendidas estas como unidades lingüísticas que poseen un verbo conjugado en primera persona del singular o plural. Cada una de estas unidades lingüísticas es un *caso*, y el conjunto de los mismos constituye el *universo de casos* de la investigación. En el estudio se han tomado en cuenta los primeros cien (100) casos de cada hablante, lo que supone un universo de tres mil seiscientos (3.600) casos. He excluido del análisis las intervenciones de los dos entrevistadores.

### 3.3. Exclusiones

No se tomaron en cuenta las cláusulas incompletas, en las que el hablante se interrumpía antes de pronunciar la forma verbal conjugada, lo que, en la mayoría de los casos, se debía a vacilaciones en el discurso.

He excluido también expresiones como: *vamos a decir*, *digamos*, *no sé cuanto*, etc., que son marcadores discursivos, pues como señalan María Antonia Martín Zorraquino y José Portoles Lázaro (1999: 4057-4058):

Los ‘marcadores del discurso’ son unidades lingüísticas invariables, no ejercen una función sintáctica en el marco de la predicación oracional –son,

pues, elementos marginales– y poseen un cometido coincidente en el discurso: el de guiar, de acuerdo con sus distintas propiedades morfosintácticas, semánticas y pragmáticas, las inferencias que se realizan en la comunicación.

Esto es, el significado de los marcadores contribuye al procesamiento de lo que se comunica y no a la representación de la realidad comunicada.

Algunos ejemplos de marcadores discursivos excluidos se presentan en (2):

- (2) a. una parroquia muy sencilla de gente trabajadora, humilde, Ø **digamos** lo que se denominaba en la sociedad clase media media (c33hb.05)  
 b. reconozco que es el héroe cultural del, del país y Ø **no sé cuánto** (c33md.05)

También se han excluido del análisis las cláusulas en las que es obligatoria la presencia del pronombre cuando éste: i) ocurre a la derecha de un verbo conjugado, porque no se puede determinar cuándo es posible o no la variación<sup>11</sup> (Bayley y Pease Álvarez 1996:90; Cameron 1997:33), como en (3a); ii) está modificado por un refuerzo adjetival o adverbial (*mismo, sólo, también, hasta*, entre otros), pues este exige la realización fonética del sujeto (Cameron 1997:33; Silva-Corvalán 2003:851), como en (3b); iii) aparece en un contexto contrastivo, dado que el sujeto pronominal es requerido por la situación de contraste (Bayley y Pease Álvarez 1996:90; Cameron 1997:33; Silva Corvalán 2001:155, 2003:850) como en (3c):

- (3) a. si esta vaina es así en la universidad yo agarro y estudio Administración y soy tremendo administrador, ¿que sé yo? ¿no?, *me imaginaba* YO varias estupideces (c31ha.04)  
 b. mi abuela no me compraba juguetes, YO mismo me los *compraba* (c33hd.08)  
 c. Y YO lo *veía* a él y ÉL me *veía* (c32mb.05)

---

<sup>11</sup> Si bien los autores dan por descontado por qué se excluye este caso, no explican sus razones, dejando al lector la interpretación. Para mayor precisión quisiera aclarar que esta exclusión se debe a la falta de variación y a la imposibilidad de determinarla, pues, cuando un sujeto está ausente, la posición que se le asigna por defecto es antepuesto al verbo y por lo tanto no pueden determinarse los contextos variables.

Debo señalar que no he encontrado ningún caso de pronombre usado como respuesta a preguntas formuladas con el pronombre interrogativo *quién/es*, pero es necesario mencionar que, según Cameron (1997:33), Bayley y Pease-Álvarez (1996:90), cuando esto se da, la presencia del pronombre es obligatoria.

No analicé tampoco las cláusulas cuyo verbo conjugado en 1ª persona del plural tenía como sujeto un cuantificador (*los dos, todos*, etc.), ya que, en general, se puede considerar que las formas antes mencionadas suplen al pronombre<sup>12</sup> o a un sujeto plural compuesto por un pronombre de primera persona singular y un nombre o pronombre de otra persona (*ella y yo*), puesto que, de acuerdo con Cantero (1976:236), en las cláusulas de ese tipo el sujeto siempre está expreso, ya que la composición contiene al *yo+otros* que forma el *nosotros*. Estos casos se ilustran en (4):

- (4) a. LOS DOS *estábamos* mal (c33he.08)  
 b. ELLA Y YO nos la *llevamos* muy bien (c31hc.05)

Mientras seleccionaba los casos, encontré otros que debían excluirse: i) *vamos a + infinitivo*, pues no hay ningún caso en el que se observa la presencia del pronombre, por lo tanto, la única forma parece ser la que se ilustra en (5a) ; y ii) las interrogativas de tipo fático en primera persona, según se ilustra en (5b):

---

<sup>12</sup> Sigo a Hugo Cifuentes 1980-1981, quien decide no integrarlos a su análisis: “En la muestra analizada se han encontrado los siguientes sujetos contruidos con *-mos*, que no son ni *nosotros* ni *nosotras* (21 en total): *todos* (10), *casi todos* (1), *pocos* (1), *dos* (1), *las dos* (2), *los padres* (1), *los tres* (1), *las tres* (1), *tres profesores* (1), *un grupo de lo mas heterogéneo* (1), *un grupo bastante pequeño*(1). [...] estos casos no fueron computados a pesar de que por lo menos en algunos de ellos se podría catalizar *nosotros*” (751).



- (5) a. sin embargo es muy difícil explicarle, explicarle a la ca... a la mente que Ø *vamos a estar* sumergidos sesenta metros (c31hf.06)
- b. y ¿Ø qué *estaba diciendo?* (c31hc.05)

### 3.4. Variables de análisis

La variable dependiente es la alternancia entre la presencia de un sujeto pronominal *versus* el cero fónico, según se ilustra en el ejemplo (6):

- (6) a. /sujeto pronominal/<sup>13</sup>  
NOSOTROS *nos metíamos* en la parte de atrás del Volkswagen (c31hd.05)
- b. /cero fónico/  
todavía ahorita Ø nos ... nos *vemos* pues (c31hd.05)

Para el presente análisis, utilicé diez (10) variables lingüísticas, así como las dos variables sociales inherentes a los hablantes del corpus, SEXO y EDAD (ver sección 3.1.). Las variables lingüísticas incluyen la variable dependiente, PRESENCIA~AUSENCIA DEL PRONOMBRE, y nueve variables independientes: i) el NÚMERO GRAMATICAL DEL SUJETO; ii) la CONECTIVIDAD DISCURSIVA; iii) la AMBIGÜEDAD VERBAL; iv) el CAMBIO DE TURNO; v) la CLASE SEMÁNTICA DEL VERBO; vi) el TIEMPO, ASPECTO Y MODO DEL VERBO; vii) la POLARIDAD DE LA CLÁUSULA; viii) el EFECTO REPETICIÓN; y ix) la DISTANCIA REFERENCIAL.

Todos los ejemplos de las variables lingüísticas independientes se encuentran en la sección dedicada al análisis de los resultados, con el fin de que los ejemplos que ilustran cada variable no estén separados de los resultados arrojados por el análisis.

---

<sup>13</sup> A partir de ahora, en la parte (a) de todos los ejemplos se presentan los casos en los que se expresa el sujeto, que se señala en versales; y en la parte (b), los casos donde no se expresa el sujeto, lo que se indica con el signo Ø.

### 3.5. Análisis de los datos del corpus

El análisis se hizo en varias etapas. Primero, extraje manualmente todas las cláusulas que cumplían los criterios propuestos para la presente investigación. Los sujetos de 1ª persona, tanto en singular como en plural, se codificaron<sup>14</sup> en hojas de Excel. El análisis cuantitativo incluye las frecuencias absolutas y relativas de los casos analizados, con y sin pronombre expreso, y el cálculo de estimaciones de fiabilidad y significación de los mismos con la prueba del  $\chi^2$ , que estima la probabilidad de que la distribución de los valores obtenidos no sea aleatoria en cada una de las variables propuestas.<sup>15</sup> El análisis de regresión múltiple utilizado es el programa multivariante GoldVarb X (Sankoff, Tagliamonte, Smith 2005), diseñado especialmente para el estudio de la variación lingüística. Este programa, en cualquiera de sus versiones, permite calcular la probabilidad general de que, en un contexto dado, aparezca una de las variantes (presencia~ausencia) de la variable dependiente, cuando un grupo de variables independientes –lingüísticas y extralingüísticas– actúa simultáneamente. Es necesario aclarar que en una investigación variacionista es importante determinar tanto la influencia como la no influencia de una variable independiente sobre la dependiente.

---

<sup>14</sup> Codificar un caso es rotularlo de tal manera que todas sus categorizaciones lingüísticas y sociales sean factorizadas en una secuencia alfanumérica, ordenada y biunívoca, de tantos dígitos como variables se usen.

<sup>15</sup> Para más detalles sobre la prueba del  $\chi^2$  véase el anexo 2.

El programa GoldVarb X realiza dos procedimientos: el análisis binomial de primer nivel y el análisis binomial de subida y bajada.<sup>16</sup>

El análisis binomial de primer nivel, o método directo, arroja las probabilidades de que la variante de la variable dependiente aparezca con cada uno de los factores implicados. En el análisis de subida y bajada, *Binomial up & down*, el programa ofrece varios modelos posibles para dar una explicación a la combinación de los datos y selecciona el más adecuado, es decir, aquel

que explique la variación lingüística en relación con los factores contemplados. Para tomar esa decisión, el programa parte de un análisis de probabilidades de los grupos de factores tomados de uno en uno [...] para luego ir combinándolos [...] determinando en cada momento la probabilidad y el grado de significación de cada uno de los niveles contrastados (Hernández Campoy y Almeida, 2005:278).

Debido al hecho de que CONECTIVIDAD DISCURSIVA y EFECTO REPETICIÓN/DISTANCIA REFERENCIAL no pueden analizarse conjuntamente, pues hay interacción entre ellas, hice análisis estadísticos separados de cada una de estas variables conjuntamente con las otras variables independientes.

---

<sup>16</sup> Para una descripción más detallada del análisis de regresión realizado por GoldVarb X, véase el anexo 3.

#### 4. EJEMPLOS Y RESULTADOS

##### 4.1. Ejemplos, frecuencias y prueba de $\chi^2$

En este apartado presentaré, en primer lugar, los textos que ejemplifican las variantes de cada variable propuesta. Luego, los cuadros con las frecuencias absolutas y relativas obtenidas al cruzar cada una de las variables (lingüísticas y extralingüísticas) con la variable dependiente (presencia~ausencia del sujeto de 1<sup>era</sup> persona), y además, los resultados de la prueba del  $\chi^2$ . Cuando es posible, he añadido los resultados obtenidos en otras investigaciones en las que se ha utilizado las mismas variables con el fin de propiciar la comparación.

##### 4.1.1. Número gramatical del sujeto

La primera variable independiente es el NÚMERO GRAMATICAL DEL SUJETO DE PRIMERA PERSONA y tiene dos variantes: *singular* y *plural*, como se ilustra en los ejemplos (7) y (8):

###### (7) /Singular/

- a. YO *quiero seguir* estudiando (c32mb.05)
- b. Ø *saqué* muy buenas notas (c32ha.04)

###### (8) /Plural/

- a. NOSOTRAS *queremos tener* hijos (c33ma.04)
- b. Ø *vivíamos* alrededor de esa cuadra (c33ha.05)

El cuadro 4 muestra los resultados obtenidos al analizar el NÚMERO GRAMATICAL DEL SUJETO (*singular* o *plural*) según la presencia o ausencia del pronombre personal de primera persona:

Cuadro 4: Presencia o ausencia de la 1ª persona según NÚMERO GRAMATICAL DEL SUJETO

Número gramatical		Presencia	Ausencia	Total	%
Singular	N	947	1.838	2.785	77
	%	34	66		
Plural	N	125	690	815	23
	%	15	85		
Total	N	1.072	2.528	3.600	100
	%	30	70		

$\chi^2$  corregido por Yates = 104,17      p < 0,0000

El cuadro 4 muestra que, de un total de 2.785 cláusulas en primera persona del singular, 946 (34%) tienen el pronombre expreso, mientras que tan solo 125 de las 815 cláusulas plurales analizadas lo tienen, lo que representa el 15%. La prueba del  $\chi^2$  permite confirmar que hay una relación estadísticamente significativa entre las dos variables. Los resultados obtenidos en otros estudios en relación con la presencia del pronombre se muestran en el cuadro 5:

Cuadro 5: NÚMERO GRAMATICAL DEL SUJETO en este y en otros estudios

Número gramatical	Estudios		Presencia %	Ausencia %
Singular	<b>Pérez Brabandere 2010</b>	<b>Caracas 2004-2010</b>	<b>34</b>	<b>66</b>
	Blanco Canales 1999	Alcalá de Henares 1997	54	46
	Bentivoglio 1987	Caracas 1977	46	54
	Cifuentes 1980-1981	Santiago de Chile 1971	34	66
	Barrenechea y Alonso 1977	Buenos Aires 1967-1975	24	76
Plural	<b>Pérez Brabandere 2010</b>	<b>Caracas 2004-2010</b>	<b>15</b>	<b>85</b>
	Blanco Canales 1999	Alcalá de Henares 1997	14	86
	Bentivoglio 1987	Caracas 1977	16	84
	Cifuentes 1980-1981	Santiago de Chile 1971	17	83
	Barrenechea y Alonso 1977	Buenos Aires 1967-1975	20	80

Como podemos apreciar en el cuadro 5, en el corpus de habla actual de Caracas la presencia del pronombre singular solo es de 34%, porcentaje muy inferior al obtenido

en el corpus de Caracas 1977 (46%) por Bentivoglio 1987, pero idéntico al obtenido en Santiago de Chile 1971 (Cifuentes 1980-1981). A su vez, el resultado obtenido por Blanco Canales 1999 en el corpus de Alcalá de Henares 1997 es muy superior a la media con un 54% mientras que Barrenechea y Alonso 1977 obtuvieron el porcentaje más bajo de todos (24%) en su estudio del corpus de Buenos Aires.

#### 4.1.2. Ambigüedad

La segunda variable independiente es la ambigüedad de la forma verbal. En el cuadro 6 se pueden apreciar las formas verbales consideradas ambiguas porque no revelan si el sujeto es de primera o tercera persona singular. Para que la forma verbal se considere morfológicamente ambigua, debe haber dos (o más) participantes, generalmente con las mismas características semánticas, por ejemplo, [+ humano]. La primera persona del plural no permite la ambigüedad gracias a la terminación *-mos*.

Cuadro 6: Formas verbales ambiguas de primera y tercera persona en español

EJEMPLO: YO/ÉL/ELLA AMABA	CONJUGACIONES VERBALES (NOMENCLATURA DE BELLO)	
	Modo	Tiempo
1. <i>amaba</i>	Indicativo	Copretérito
2. <i>había amado</i>	Indicativo	Antecopretérito
3. <i>amaría</i>	Indicativo	Pospretérito
4. <i>habría amado</i>	Indicativo	Antepospretérito
5. <i>ame</i>	Subjuntivo	Presente
6. <i>amara</i> (1 <sup>ra</sup> forma)	Subjuntivo	Pretérito
7. <i>amase</i> (2 <sup>da</sup> forma)		
8. <i>haya amado</i>	Subjuntivo	Antepresente
9. <i>hubiera amado</i> (1 <sup>ra</sup> forma)	Subjuntivo	Antepretérito
10. <i>hubiese amado</i> (2 <sup>da</sup> forma)		
11. <i>amare</i>	Subjuntivo	Futuro
12. <i>hubiera amado</i>	Subjuntivo	Antefuturo

Fuente: Seco (2001:276–278)

Las variantes de la AMBIGÜEDAD son las propuestas por Paredes Silva 1993:

i) morfológicamente no-ambigua, como en (9); ii) morfológicamente ambigua, pero contextualmente no-ambigua, como en (10); y iii) morfológica y contextualmente ambigua,<sup>17</sup> como en (11):

- (9) /no ambigua/  
 a. bueno S. mira YO *conocí* a D. en Fuente Real (c31mf.05)  
 b. sí, si Ø *recuerdo* una anécdota particular (c33mf. 09)
- (10) /morfológicamente ambigua/  
 a. también recuerdo en esa época, que YO *vivía* en La Pastora (c32me. 06)  
 b. yo era de los, de los niños que Ø *agarraba* la, la, los animalitos (c32hc.05)
- (11) /morfológica y contextualmente ambigua/  
 a. mi mamá, el año antes de que YO *entrara*, ella hizo un semestre (c31hd.05)  
 b. mi hermana es sumamente permisiva, bueno, eso que Ø *te estaba contando*, de la reunión que ella invitó unas amigas (c32mb.05)

<sup>17</sup> Si bien una forma es morfológica y contextualmente ambigua solamente cuando no está presente el pronombre sujeto, en la variante “presencia” se incluyen los casos en que la cláusula sería morfológica y contextualmente ambigua si el pronombre no estuviera presente.

En el cuadro 7 se presentan las frecuencias obtenidas al cruzar la AMBIGÜEDAD DE LA FORMA VERBAL con la variable dependiente:

Cuadro 7: Presencia o ausencia del pronombre de 1ª persona según la AMBIGÜEDAD DE LA FORMA VERBAL

Ambigüedad		Presencia	Ausencia	Total	%
No ambigua	N	809	2.257	3.065	85
	%	26	74		
Morfológicamente ambigua	N	181	257	439	12
	%	41	59		
Morfológica y contextualmente ambigua	N	82	14	96	3
	%	85	15		
Total	N	1.072	2.528	3.600	100
	%	30	70		

$$\chi^2 = 186,91 \quad 2 \text{ g. d. l.} \quad p < 0,0000$$

Los resultados del cuadro 7 muestran claramente que la presencia de un pronombre explícito es mayor (85%) cuando la forma verbal es *morfológica y contextualmente ambigua* que cuando la forma es sólo *morfológicamente ambigua* (41%) o no ambigua del todo (26%). La prueba del  $\chi^2$  señala que la relación entre las dos variables es estadísticamente significativa. Los resultados obtenidos en otros estudios en relación con la presencia del pronombre pueden apreciarse en los cuadros 8 y 9:



Cuadro 8: AMBIGÜEDAD DE LA FORMA VERBAL en este y en otros estudios

Ambigüedad	Estudios		Presencia %	Ausencia %
No ambigua	<b>Pérez Brabandere 2010</b>	<b>Caracas 2004-2010</b>	26	74
	Paredes Silva 1993	Río de Janeiro 1992-1993	81	19
	Bentivoglio 1987	Caracas 1977	44	56
	Barrenechea y Alonso 1977	Buenos Aires 1967-1975	20	80
Morfológicamente ambigua	<b>Pérez Brabandere 2010</b>	<b>Caracas 2004-2010</b>	41	59
	Paredes Silva 1993	Río de Janeiro 1992-1993	57	43
	Bentivoglio 1987	Caracas 1977	75	25
	Barrenechea y Alonso 1977	Buenos Aires 1967-1975	43	57
Morfológica y contextualmente ambigua	<b>Pérez Brabandere 2010</b>	<b>Caracas 2004-2010</b>	85	15
	Paredes Silva 1993	Río de Janeiro 1992-1993	10	90
	Bentivoglio 1987	Caracas 1977	50	50
	Barrenechea y Alonso 1977	Buenos Aires 1967-1975	100	0

En el cuadro 8 se pueden apreciar los resultados de los estudios en los que, al igual que en la presente investigación, se emplearon tres variantes para la variable AMBIGÜEDAD. Lo extraño es que los resultados de las diferentes investigaciones son contradictorios, aunque debe tomarse en cuenta que el corpus de Río de Janeiro 1992-1993 es del portugués escrito. Si sólo tomáramos en cuenta los corpus del español, en los que se han analizado las tres variantes, se destacaría que los valores encontrados por Bentivoglio 1987 para la variante *morfológicamente ambigua* es mucho más alta (75%) en el corpus de Caracas 1977 que la del corpus caraqueño actual (41%) y del bonaerense de 1967-1975. En este último corpus, Barrenechea y Alonso 1977 reportan que la variante morfológica y contextualmente ambigua favorece la presencia del pronombre de manera categórica (100%); también en el habla actual de Caracas la presencia del pronombre alcanza un porcentaje muy alto (85%).

Cuadro 9: AMBIGÜEDAD DE LA FORMA VERBAL en otros estudios que sólo toman en cuenta la Ambigüedad morfológica

Ambigüedad	Estudios		Presencia %	Ausencia %
No ambigua	Torre-Caucoulllos y Travis 2008	Nuevo México 1992-1995	34	66
	Samper, Hernández y Troya 2006	Las Palmas de Gran Canarias 1985	26	74
	Samper, Hernández y Troya 2006	Madrid 1985	26	74
	Samper, Hernández y Troya 2006	Salamanca 1985	33	67
	Samper, Hernández y Troya 2006	Sevilla 1972-1973	28	72
	Blanco Canales 1999	Alcalá de Henares 1997	21	79
	Bentivoglio 1999	Caracas 1968-1977	41	59
	Miró Vera y de Pineda 1990	Sevilla 1972-1973	27	73
	Ambigua	Torre-Caucoulllos y Travis 2008	Nuevo México 1992-1995	36
Samper, Hernández y Troya 2006		Las Palmas de Gran Canarias 1985	36	64
Samper, Hernández y Troya 2006		Madrid 1985	43	57
Samper, Hernández y Troya 2006		Salamanca 1985	29	71
Samper, Hernández y Troya 2006		Sevilla 1972-1973	43	57
Blanco Canales 1999		Alcalá de Henares 1997	21	79
Bentivoglio 1999		Caracas 1968-1977	66	34
Miró Vera y de Pineda 1990		Sevilla 1972-1973	28	72

En el cuadro 9, la variable AMBIGÜEDAD tiene solamente dos variantes: *ambiguo* y *no ambiguo*. La presencia pronominal para las formas *no ambiguas* varía entre un mínimo de 21% en Alcalá de Henares a un máximo de 41% en el corpus del habla culta de Caracas, mientras que cuando hay *ambigüedad* los porcentajes varían entre un mínimo de 21% y un máximo de 66% en el habla culta de Caracas, con valores intermedios en las otras ciudades.

#### 4.1.3. Tiempo, aspecto y modo del verbo

En vista de que varios autores, entre ellos Cantero 1976 y Travis 2007, sostienen que el TIEMPO-ASPECTO-MODO (TAM) DEL VERBO influye en la presencia~ausencia del pronombre, decidí comprobar si esto es relevante para la alternancia objeto del presente estudio. El tiempo, el aspecto y el modo son tres

categorías lingüísticas que afectan a toda la cláusula, pero presentan una morfología particular en todas las formas verbales conjugadas: i) *presente*; ii) *copretérito*; iii) *futuro*; iv) *pretérito*; v) *pospretérito*; vi) *antepresente*; vii) *antecopretérito*; viii) *presente de subjuntivo*; ix) *pretérito de subjuntivo 1<sup>era</sup> forma y 2<sup>da</sup> forma*; y x) *antepresente y la 1<sup>era</sup> forma del antepretérito de subjuntivo*, según puede observarse en los ejemplos respectivos de (12) a (21):

- (12) /Presente/  
 a. YO *tengo* una sobrina en la laguna de Tacarigua (c33hc.06)  
 b. si usted no se casa con mi hija Ø lo *mato* (c33ha.04)
- (13) /Copretérito/  
 a. YO no *sabía* que te gustaban los viejos (c33hd.08)  
 b. bueno, Ø *íbamos* a escondidas (c33ha.04)
- (14) /Futuro/  
 a. supongo que en caso de presentársele alguna emergencia pues YO le *podré brindar* la mano (c31ma.04)  
 b. de La Candelaria Ø *iremos* a otro sitio y así sucesivamente (c31ma.04)
- (15) /Pretérito/  
 a. NOSOTROS *estuvimos* arreglando nuestro cuarto (c31ha.04)  
 b. total en pleno baile Ø *me desnudé* (c33ha.04)
- (16) /Pospretérito/  
 a. YO *tendría* como, qué sé yo, doce años, trece años (c33me.07)  
 b. me gusta muchísimo la naturaleza, si por mi fuera, si tuviera tiempo, este, Ø me *iría* siempre a la playa (c33mf.09)
- (17) /Antepresente/  
 a. primero descubrí que, por lo menos YO *he descubierto* que, que yo me transformo (c31hd.05)  
 b. más nunca Ø los *he visto* (c31hc.05)
- (18) /Antecopretérito/  
 a. pero mi sorpresa fue que no era nada de lo que YO *había leído* (c33mc.05)  
 b. cuando me gradué, Ø le *había dicho* a mi mamá que quería trabajar en algo (c31mc.04)

- (19) /Presente de subjuntivo/  
 a. de verdad, no es que YO *menosprecie* eso (c32hc.05)  
 b. yo bailo bien, vamos no es que Ø *sea* pedante, sino que bailo bien (c33hd.08)
- (20) /Pretérito de subjuntivo 1<sup>era</sup> forma y 2<sup>da</sup> forma/  
 a. sí no fuera por mí mamá, NOSOTROS *viviéramos* debajo de un puente, seguramente (c32mb.05)  
 b. me trataba, no sé, como si Ø *fuese*, no sé, un niño de doce años (c31he.05)
- (21) /Antepresente y antepretérito de subjuntivo/<sup>18</sup>  
 a. no recuerdo que YO me *haya*, ya en la carrera, *planteado*: coye ¿será esto lo que me gusta? (c31hd.05)  
 b. cualquier otro gobierno que Ø *hubiéramos tenido* aquí en este país hubiera tratado mejor el problema de La Guaira (c33hc.06)

Los resultados se presentan ordenados de acuerdo con la frecuencia de aparición en el cuadro 10:

---

<sup>18</sup> Uní estas dos categorías verbales porque, en total, sólo encontré 3 casos: 1 de antepresente (con ø) y 2 de la 1<sup>era</sup> forma de antepretérito (con pronombre), y se podían fusionar en una sola variante puesto que ambas formas son del modo subjuntivo y de pasado.

Cuadro 10: Presencia o ausencia del pronombre de 1ª persona según el TAM DEL VERBO

Tiempo verbal		Presencia	Ausencia	Total	%
Presente	N	476	1.209	1.685	47
	%	28	72		
Pretérito	N	233	709	942	26
	%	25	75		
Copretérito	N	279	433	712	20
	%	39	61		
Antepresente	N	43	94	137	4
	%	31	69		
Antecopretérito	N	10	21	31	1
	%	32	68		
Presente subjuntivo	N	13	22	35	1
	%	37	63		
Pretérito subjuntivo 1ª forma y 2ª forma	N	11	18	29	1
	%	38	62		
Pospretérito	N	5	12	17	0
	%	29	71		
Futuro	N	1	8	9	0
	%	11	89		
Antepresente y Antepretérito Subjuntivo	N	1	2	3	0
	%	33	67		
Total	N	1.072	2.528	3.600	100
	%	30	70		

$$\chi^2 = 47,0849 \quad 9 \text{ g. d. l.} \quad p < 0,0000378$$

Tal y como lo muestra el cuadro 10, el TAM (TIEMPO, ASPECTO Y MODO) VERBAL es estadísticamente significativo para la presencia explícita del pronombre en el español de Caracas. Los tiempos en los cuales el porcentaje de *presencia* supera la media general de 30% son: *copretérito de indicativo* (39%), *pretérito de subjuntivo en su 1ª y 2ª forma* (38%), *presente de subjuntivo* (37%), *antepresente y antepretérito de subjuntivo* (33%), *antecopretérito de indicativo* (32%) y *antepresente de indicativo* (31%); los tiempos anteriormente mencionados se alejan considerablemente del

mínimo de presencia obtenido por el *futuro* (11% ). La prueba del  $\chi^2$  indica que puede rechazarse la hipótesis nula.

Cantero 1976, en su investigación sobre el habla mexicana, toma en cuenta la variable TAM del verbo y, aunque no señala las frecuencias relativas, da una lista (por orden de aparición) de los tiempos verbales con los que se manifiesta la presencia del pronombre. Esa lista me permitió hacer un cuadro comparativo, ordenado jerárquicamente, con los datos del autor y los obtenidos en la presente investigación. En el cuadro 11 se puede apreciar cuáles son los tiempos verbales que más favorecen la presencia del pronombre de primera persona en ambos estudios:

Cuadro 11: Comparación entre los resultados de Cantero 1976 y los de Caracas 2004-2010

<b>Estudios</b>	
CANTERO 1976: MÉXICO 1968-1970	CARACAS 2004-2010
1. Copretérito de indicativo	1. Copretérito de indicativo
2. Pretérito de subjuntivo	2. Pretérito de subjuntivo 1ª y 2ª forma
3. Pospretérito de indicativo	3. Presente de subjuntivo
4. Presente de subjuntivo	4. Antepresente y antepretérito de subjuntivo
5. Antecopretérito de indicativo	5. Antecopretérito de indicativo
6. Antepretérito de subjuntivo	6. Antepresente de indicativo
7. Antepresente de subjuntivo	

En ambas investigaciones el *Copretérito de indicativo* y el *Pretérito de subjuntivo* encabezan la lista de los TAM que están relacionados con la presencia del pronombre; los otros TAM de la lista suministrada por Cantero 1976 son relevantes, si bien no en el mismo orden, en esta investigación. Sólo el *Pospretérito de indicativo* se escapa a esta red de coincidencias, pues en el presente estudio tiene un valor ligeramente inferior (29%) a la media de presencia del pronombre (30%).

La variable TAM, como se dijo al principio de este apartado, también fue estudiada por Travis 2007, pero al revisar los datos de la autora se nota que cuando va a analizarlos estadísticamente los separa en TAM ambiguos y TAM no ambiguos, de manera que esta variable puede solaparse con la variable AMBIGÜEDAD. Por esta razón decidí no tomar en cuenta esta variable en el segundo nivel de análisis.

#### 4.1.4. Polaridad de la cláusula

Decidí comprobar si la POLARIDAD –categoría destinada a determinar si una cláusula es afirmativa o negativa– tiene algún efecto en el corpus analizado, pese a que esta variable no ha resultado significativa en la investigación de Miró Vera y de Pineda 1990. Las variantes son dos: i) *positiva*; y ii) *negativa*, según se muestra en (22-23):

- (22) /polaridad positiva/  
 a. me involucré con el grupo, NOSOTROS lo *llamábamos* el Grupo Guaica (c33md.05)  
 b. porque recuerda  $\emptyset$  *somos* siete, pero seis somos contemporáneos (c33ma.04)
- (23) /polaridad negativa/  
 a. YO **no** le *doy* pescado a C. porque a él no le gusta (c32ma.04)  
 b. los tres primeros años  $\emptyset$  **no** *tuve* vacaciones de verdad (c32hg.08)

El cuadro 12 muestra la correlación entre la variable presencia~ausencia de la forma pronominal y la POLARIDAD de la cláusula:

Cuadro 12: Presencia~ausencia del pronombre de 1ª persona según la POLARIDAD DE LA CLÁUSULA

<b>Polaridad</b>		<b>Presencia</b>	<b>Ausencia</b>	<b>Total</b>	<b>%</b>
Positiva	N	950	2.229	3.179	88
	%	30	70		
Negativa	N	122	299	421	12
	%	29	71		
Total	N	1.072	2.528	3.600	100
	%	30	70		

$\chi^2$  corregido por Yates = 0,11      p < 0,7452725

En el cuadro 12 se observa claramente que la POLARIDAD de la cláusula no influye en la presencia de un pronombre explícito, lo que es corroborado por el resultado de la prueba del  $\chi^2$ , la cual indica que debe aceptarse la hipótesis nula. En estudios previos, como en el de Enríquez 1984, se señala una ligera tendencia a utilizar más el pronombre *yo* en las oraciones afirmativas que en las negativas. Miró Vera y de Pineda 1990, quienes usan en su investigación otro tipo de análisis estadístico (comparación de medias), afirman que la media da un valor de 1,68, lo que es menor a 1,96 (media mínimo requerida) y, por lo tanto esta variable, al igual que en el presente estudio, resulta estadísticamente no significativa.

#### 4.1.5. Cambio de turno

La quinta variable independiente es el CAMBIO DE TURNO en la conversación; está conformada por dos variantes: *igual turno de habla y diferente turno*. Según Bentivoglio 1987, “el hablante, al referirse a sí mismo tiende a repetir el pronombre sujeto de primera persona independientemente de otros factores, si ocurre una



interrupción” (38). Los ejemplos (24) y (25) ilustran las dos variantes de esta variable:

(24) /igual turno de habla/<sup>19</sup>

- a. **H.:** Es que ... bueno Ø<sub>1</sub> no sé ... es que ... YO<sub>2</sub> a veces pienso en eso pero YO<sub>3</sub> no sé en qué medida ... sea verdad o no (c31hf.05)
- b. **H.:** Ø<sub>1</sub> Tenía ... ahí Ø<sub>2</sub> tenía Ø<sub>3</sub> no sé serán trece, catorce por ahí, este ... ahí también Ø<sub>4</sub> fuimos a un salto que se llama el Salto El Guapo (c32ha.05)

(25) /diferente turno de habla/

- a. **E.:** ¿a qué cree que se deba eso?  
**H.:** YO *creo* que se debe principalmente a, a que se, ehhh el gobernante actual ha intensificado el mensaje de lucha de clases (c32hf.08)
- b. **E.:** ¿amistades? ¿novio?  
**H.:** no, Ø no tengo novio (c31mb.04)

Los resultados que se presentan en el cuadro 13 muestran si el CAMBIO DE TURNO en la conversación influye en la presencia o ausencia del pronombre:

Cuadro 13: Presencia o ausencia del pronombre de 1ª persona según CAMBIO DE TURNO EN LA CONVERSACIÓN

Cambio de turno		Presencia	Ausencia	Total	%
Igual turno de habla	N	1.010	2.389	3.399	94
	%	30	70		
Diferente turno de habla	N	62	139	201	6
	%	31	69		
Total	N	1.072	2.528	3.600	100
	%	30	70		

$\chi^2$  corregido por Yates = 0,07

p < 0,7937867

Los resultados del cuadro 13 indican que el cambio de turno en la conversación no influye en la presencia o ausencia del elemento pronominal, ya que los porcentajes de aparición del pronombre son casi iguales: 30% para *igual turno de habla* y 31% para

<sup>19</sup> En todos los ejemplos, a partir de ahora, cuando la cláusula se inicia con **H.**, se entenderá que se trata del hablante, y cuando se inicia con **E.**, del entrevistador.

*diferente turno de habla*; la prueba del  $\chi^2$  confirma que no se puede rechazar la hipótesis nula.

Los resultados obtenidos en otros estudios en relación con la presencia del pronombre son los reportados en el cuadro 14:

Cuadro 14: CAMBIO DE TURNO en este y otros estudios

<b>Cambio de turno</b>	<b>Estudios</b>		<b>Presencia %</b>	<b>Ausencia %</b>
Igual turno de habla	<b>Pérez Brabandere 2010</b>	<b>Caracas 2004-2010</b>	<b>30</b>	<b>70</b>
	Bentivoglio 1999	Caracas 1968-1977	41	59
	Bentivoglio 1987	Caracas 1977	37	63
	Morales 1982	San Juan, PR 1968	26	74
Diferente turno de habla	<b>Pérez Brabandere 2010</b>	<b>Caracas 2004-2010</b>	<b>31</b>	<b>69</b>
	Bentivoglio 1999	Caracas 1968-1977	66	34
	Bentivoglio 1987	Caracas 1977	49	51
	Morales 1982	San Juan, PR 1968	58	42

Como se puede apreciar en el cuadro 14, la presencia pronominal cuando hay *igual turno de habla*, oscila entre 20% en Puerto Rico y 41% en el Corpus del habla culta de Caracas (1968-1977), mientras que en el Corpus actual de esta misma ciudad el porcentaje baja a 30%. En la variante *diferente turno de habla* el pronombre tiene un porcentaje de frecuencia importante tanto en los estudios de Bentivoglio 1987, Bentivoglio 1999, como en el de Morales 1982 (la presencia pronominal es siempre superior al 50%) menos en la muestra del habla actual de Caracas, en que el porcentaje de 31% es casi idéntico al obtenido por la variante *igual turno de habla*. Se debe acotar que esta variable no ha resultado significativa en la investigación de Torres-Caucollos y Travis 2008.

#### 4.1.6. Clase semántica del verbo

La sexta variable independiente es la CLASE SEMÁNTICA DEL VERBO que acompaña al pronombre sujeto; ha sido usada por diferentes investigadores (Bentivoglio 1987, 1999; Silva Corvalan 2003, entre otros), con clasificaciones más o menos parecidas. Con respecto a las construcciones perifrásticas, debo aclarar que en este análisis sólo he tomado en cuenta la forma no finita del verbo, ya que me ocuparé sólo del aspecto, pero no del significado modal.

Para hacer un análisis más preciso, decidí utilizar la clasificación de procesos ADESSE (Base de datos: Alternancias de Diátesis y Esquemas Sintáctico-Semánticos del Español).<sup>20</sup> Las variantes son los verbos de procesos: i) *mentales*; ii) *relacionales*; iii) *materiales*; iv) *verbales*; v) *existenciales*; y vi) de *modulación*,<sup>21</sup> según puede observarse en los ejemplos de (26) a (31), respectivamente:

- (26) /verbos de procesos mentales/  
 a. YO a veces *pienso* en eso (c31hf.06)  
 b. Ø *creo* que el único lugar donde no fuimos fue a Tucupita (c32ha.04)
- (27) /verbos de procesos relacionales/  
 a. YO *era* pasante de la Electricidad de Caracas (c32me.06)  
 b. y entonces Ø *no puedo* cerrar bien la mano como para pisar (c33he.08)
- (28) /verbos de procesos materiales/  
 a. YO *trabajaba* en compras (c32ma.04)  
 b. y entonces, bueno, me acuerdo que el golpe fue tan fuerte que Ø me *rompí* la cabeza (c33he.08)

---

<sup>20</sup> Esta clasificación, avalada por la Universidad de Vigo, puede consultarse en la página web: <http://adesse.uvigo.es/data/clases.php>, que se revisa y actualiza constantemente.

<sup>21</sup> Más detalle sobre esta clasificación puede verse en el Anexo 1.

- (29) /verbos de procesos verbales/  
 a. YO *insisto* en que no, porque yo soy muy necia (c31ma.04)  
 b. yo no, nunca Ø *dije* a mi hijo que estudiara Derecho (c32hf.08)
- (30) /verbos de procesos existenciales/  
 a. YO *nací* en el año cuarenta (c33me.07)  
 b. ahí Ø *transcurrí* dos años (c33hb.05)
- (31) /verbos de modulación/  
 a. bueno YO *ayudaba* en matemáticas (c31me.05)  
 b. Ø *evito tener* con... conflictos con ella (c31hc.05)

El cuadro 15 muestra los resultados obtenidos al analizar el corpus según la clase semántica del verbo:

Cuadro 15: Presencia o ausencia del pronombre de 1ª persona según la CLASE SEMÁNTICA DEL VERBO

Clase semántica del verbo		Presencia	Ausencia	Total	%
de procesos materiales	N	783	1867	2.650	74
	%	30	70		
de procesos mentales	N	128	246	374	10
	%	34	66		
de procesos relacionales	N	95	264	359	10
	%	26	74		
de procesos existenciales	N	22	65	87	3
	%	25	75		
de procesos verbales	N	36	46	82	2
	%	44	56		
de modulación	N	8	40	48	1
	%	17	83		
Total	N	1.072	2.528	3.600	100
	%	30	70		

$$\chi^2 = 18,10 \quad 5 \text{ g. d. l.} \quad p < 0,00282404$$

En el cuadro 15, la prueba del  $\chi^2$  indica que la CLASE SEMÁNTICA DEL VERBO está estadísticamente relacionada con la presencia~ausencia del pronombre de 1ª persona; el menor porcentaje de la presencia de pronombre de primera persona se debe a la clase semántica de los verbos de modulación (17%), y el mayor porcentaje (44%) a la

clase de procesos verbales. Las otras cuatro variantes tienen valores más o menos parecidos porque varían entre 25% y 34%.

En otros estudios, la clasificación semántica de los verbos ha sido diferente, y los resultados con respecto a la presencia del pronombre son los que se aprecian en el cuadro comparativo 16:

Cuadro 16: Comparación entre esta investigación y aquellas en las que se toma en cuenta la CLASE SEMÁNTICA DEL VERBO

Estudios	Clase semántica del verbo							
		de procesos materiales 30%	de procesos verbales 44%	de procesos mentales 34%	de procesos existenciales 25%		de procesos relacionales 26%	de modulación 17%
Torre-Caucollos y Travis 2008: Nuevo México 1992-1995				psicológico 52%				otros verbos 30%
Travis 2007: Colombia 1996			de habla 51%	psicológico 67%		copulativo 60%	de movimiento 31%	otros 39%
Travis 2007: Nuevo México 1992-1995			de habla 35%	psicológico 55%		copulativo 39%	de movimiento 20%	otros 25%
Samper <i>et alii</i> 2006: Norma culta de España	volición 30%		dicción 20%	cognición 40%				otros 25%
Bentivoglio 1999: Caracas 1968-1977		percepción 39%		de cognición 48%	existencial 55%	copulativo 37%	de movimiento 41%	
Blanco Canales 1999: Alcalá de Henares 1998	de deseo /voluntad 47%	percepción 60%	de lengua 0,343	de pensamiento 0,55				otros 0,49
Bentivoglio 1987: Caracas 1977	de volición 66%	percepción /sensación 57%	de dicción 55%	cognición/ mental 53%				otros 33%

#### 4.1.7. Conectividad discursiva

La séptima variable independiente es la CONECTIVIDAD DISCURSIVA, que “ha resultado ser el factor estadísticamente más significativo en todos los estudios” (Silva Corvalán, 2001:154). Para la presente investigación, utilicé la misma variable propuesta por Paredes Silva 2003, ya ampliamente descrita en la sección 2.2. de los antecedentes en la página; sin embargo, mientras hacía la codificación, me di cuenta de que dos casos no estaban contemplados en ninguno de los grados: i) las citas directas anteceditas por preguntas formuladas por el mismo entrevistado (ver ejemplo 32), y ii) los casos en los que se pierde la conectividad con el referente anterior (ver ejemplo 40). En el primer caso, como supuestamente esa cláusula proviene de un tercer participante no presente, decidí que estos casos podían formar parte del *grado 3*, en el que se considera el mantenimiento del mismo referente del sujeto, luego de la intervención de otro hablante. En cambio para el segundo caso, cuando se pierde toda conectividad, decidí crear un nuevo grado, el que he llamado *sin conexión*.

- (32) a.  $\emptyset_1$  llego a la casa oliendo a cigarrillo, ¿te fumaste una caja de cigarros?, y yo: no mamá,  $\emptyset_2$  me *fumé* uno (c31hc.05)

Las variantes de la variable conectividad son ocho (8), y se pueden apreciar en los ejemplos (33-40):

- (33) /grado 1: conexión óptima/  
 a.  $YO_1$  *estaba* preocupado porque  $YO_2$  *quería tener* la mejor pinta posible (c32hd.07)  
 b. luego  $\emptyset_1$  *estudié* en el Liceo Caracas y ahí  $\emptyset_2$  me *gradué* de bachiller (c32mb.05)

- (34) /grado 2: cambio de TAM/  
 a. Cuando Ø<sub>1</sub> *llegué* ahí, YO<sub>2</sub> *era* un muchacho de 18 años (c32hd.07)  
 b. Ø<sub>1</sub> *comencé* en el curso introductorio y Ø<sub>2</sub> *no era* competitivo (c31hf.06)
- (35) /grado 3: cambio de turno con mismo referente/  
 a. E.: ese tío ¿lo encontraron?  
 H.: sí, NOSOTROS lo *cuidamos* después (c33hf.09)  
 b. E.: ¿y eres la única hembra?  
 H.: no, no, Ø *tengo* una hermana menor (c32ma.04)
- (36) /grado 4: interrupción por cláusulas impersonales/  
 a. Ø<sub>1</sub> íbamos por la bajada de los perros, mas de ahí, se iba hasta Los Próceres, NOSOTROS<sub>2</sub> *teníamos que estar* chévere (c32me.06)  
 b. yo<sub>1</sub> lo que hacia era cuidarte, porque no había de (sic) otra, y pero sin embargo Ø<sub>2</sub> *creo*, yo creo que al final del cuento como yo soy mamá gallina (c32md.06)
- (37) /grado 5: interrupción por cláusulas personales/  
 a. Ø<sub>1</sub> *no sé hablar* el idioma que ellos hablan y, bueno, YO<sub>2</sub> *me voy* (c31me.05)  
 b. todo el mundo pensaba que Ø<sub>1</sub> iba a estudiar Comunicación Social, nadie se imaginó que era eso lo que Ø<sub>2</sub> iba a estudiar (c31mc.04)
- (38) /grado 6: el referente del sujeto cumple otra función en la cláusula previa/  
 a. ¡basta! no me hiciste caso, YO *tomo* el control (c31hf.06)  
 b. me gustan los periódicos también y esas cosas, las que Ø *he visto* en la universidad me han llamado la atención (c31he.05)
- (39) /grado 7: cambio de tópico/  
 a. más bien hay gente que quiere venir a Macarao, porque, por la misma cuestión de que es tranquilo, pues, más bien hay gente que, YO *llego* y llevo a una (sic) amigo para mi casa, de la universidad (c31ha.04)  
 b. los dos varones, uno militar y yo cura, [risas] en esa línea pues eh ... a pesar del distanciamiento, digamos de proyección eh ... profesional, sin embargo, somos hermanos, hermano, hermanos ¿no?, ciertamente eh ... cuando uno descubre muchas cosas uno va conociendo y va desarrollando, como Ø *te vuelvo a insistir* El Valle era eh ... lo que era realmente una comunidad, [en] El Valle todo el mundo se conocía (c33ha.05)

- (40) /grado 8: sin conexión/
- a. el programa que escoge CONATEL, → +10 cláusulas  
 para hacer las leyes, → 10 cláusulas  
 es Archivo Criminal, → 9 cláusulas  
 y bajo lupa lo examinaban, → 8 cláusulas  
**E.:** ¿qué tal?, → 7 cláusulas  
**H.:** porque les parecía sangriento, → 6 cláusulas  
 o porque de repente, le estaba mostrando al pueblo, → 5 cláusulas  
 que había demasiado asesinato, y demasiado acoso, → 4 cláusulas  
 entonces el, tampoco al gobierno le convenía mucho que, → 3 cláusulas  
 que vieran ese tipo de situaciones pues, → 2 cláusulas  
 que además que salen en la prensa, → 1 cláusula  
 porque NOSOTROS lo *escogíamos* así adrede, en la prensa (c32mg.08)
- b. pero igual es una comunidad como la vecindad del chavo, este, → +10 cláusulas  
 donde todo el mundo se conoce, → 10 cláusulas  
 los malandros están mosca → 9 cláusulas  
 de que no te roben tu carro, → 8 cláusulas  
 o sea no hay problema de ese tipo. → 7 cláusulas  
**E.:** ¿Tienes relación con los vecinos? → 6 cláusulas  
**H.:** Sí, claro. → 5 cláusulas  
**E.:** ¿Te llevas bien?, → 4 cláusulas  
 ¿Hay gente joven? → 3 cláusulas  
**H.:** de todo, o sea eso es una cosa del Banco Obrero → 2 cláusulas  
 que hicieron para ... → 1 cláusula  
 Ø *creo* que para el Ipasme, no sé (c32ha.04)

El cuadro 17 muestra los resultados obtenidos al analizar el corpus según la conectividad discursiva; nótese que los resultados se han ordenado jerárquicamente, del mayor porcentaje de casos obtenidos por cada grado, del grado 5 (35%) hasta el grado 7 (1%):



Cuadro 17: Presencia o ausencia del pronombre de 1ª persona según la CONECTIVIDAD DISCURSIVA

<b>Conectividad discursiva</b>		<b>Presencia</b>	<b>Ausencia</b>	<b>Total</b>	<b>%</b>
Grado 5	N	481	770	1.251	35
	%	38	62		
Grado 1	N	177	835	1.012	28
	%	18	82		
Grado 2	N	170	486	656	18
	%	26	74		
Grado 3	N	50	111	161	5
	%	31	69		
Grado 6	N	63	145	208	6
	%	30	70		
Grado 8	N	73	80	153	4
	%	48	52		
Grado 4	N	38	80	118	3
	%	32	68		
Grado 7	N	20	21	41	1
	%	49	51		
Total	N	1.072	2.528	3.600	100
	%	30	70		

$$\chi^2 = 153,84 \quad 7 \text{ g. d. l.} \quad p < 0,000000$$

Los resultados del cuadro 17 indican que la CONECTIVIDAD DISCURSIVA está estadísticamente relacionada con la presencia del pronombre. Hay variación en los porcentajes entre un mínimo de 18% (*grado 1*), lo que era esperable, hasta un máximo de 48-49% (*grados 8 y 7*). Los porcentajes de otros grados no se alejan mucho del promedio del 30% (26, 30, 31, 32, hasta 38%). El resultado arrojado por la prueba del  $\chi^2$  señala que se puede rechazar la hipótesis nula. La comparación de los resultados obtenidos en el último estudio de Paredes Silva (2003) con los de la presente investigación se reporta en el cuadro 18:

Cuadro 18: Comparación entre esta y la muestra 2000 de la investigación de Paredes Silva 2003

<b>Grados</b>	<b>Paredes Silva 2003</b>	<b>Pérez Brabandere 2010</b>
6	88%	30%
5	83%	38%
7	83%	49%
4	81%	32%
3	73%	31%
2	71%	26%
1	44%	18%

Los resultados de la comparación entre la presente investigación y la de Paredes Silva 2003 (corpus de 2000), muestran diferencias considerables que se deben indudablemente a las diferentes características de los corpus analizados (portugués vs. español). Sin embargo, permiten concluir que mientras mayor es el grado de CONECTIVIDAD DISCURSIVA menos probabilidades hay de que se exprese el pronombre. Estos resultados se asemejan a los de la presente investigación.

#### 4.1.8. Efecto repetición

Con la octava variable independiente me propongo estudiar el EFECTO REPETICIÓN. Siguiendo a Cameron y Flores-Ferrán (2003:51), considero como realización previa del referente de primera persona la que está presente en las cláusulas anteriores, de las que tomo en cuenta desde la última hasta la décima. La variable está conformada por tres variantes: i) *no hay referencia de 1ª persona en ninguna de las diez cláusulas precedentes o no hay cláusula antecedente al caso en estudio*; ii) *hay un referente expreso de sujeto de 1ª persona en las diez cláusulas precedentes al caso en estudio*; iii) *hay un referente no expreso de sujeto de 1ª*

*persona en las diez cláusulas precedentes al caso en estudio. Las tres variantes se ilustran en (41-43):*

- (41) /no hay referente de 1<sup>a</sup> persona/  
 a. **[inicio grabación]**  
 E.: bueno, estamos grabando → 4 cláusulas  
 bueno cuéntame, → 3 cláusulas  
 dónde vives, → 2 cláusulas  
 el sitio donde vives, ¿en un apartamento?, ¿en una casa? → 1 cláusula  
 H.: e ...YO<sub>1</sub> *vivo* en El Paraíso (c31mb.04)
- b. mi papá hecho el gafito a quien, al único, como él no trabaja, → +10 cláusulas  
 a quien mantiene es a él únicamente [risas] entonces bueno, → 10 cláusulas  
 “mira, que se me acabó la tarjeta del celular”, → 9 cláusulas  
 “¿qué? ¿tan rápido?” [risas] entonces, bueno, él así, o sea, → 8 cláusulas  
 mi hermano siempre buscando las alternativas de, de, → 7 cláusulas  
 de no pedirle mucho dinero pero, ni modo, → 6 cláusulas  
 tiene que pedirle dinero → 5 cláusulas  
 porque no trabaja y, → 4 cláusulas  
 y él vive en la casa → 3 cláusulas  
 y vive es de mi papá pues, entonces, → 2 cláusulas  
 pero así son, → 1 cláusula  
 Ø<sub>1</sub> *somos* en la casa (c31md.04)
- (42) /referente anterior de 1<sup>a</sup> persona expreso /  
 a. cuando YO<sub>0</sub> *era* un muchacho joven de repente YO<sub>1</sub> me *sentaba* encima de un capó de alguien [...] y uno coye se ponía, rojo o apenado (c32hb.05)  
 b. YO<sub>0</sub> todo lo que *había comido* eran tres empanaditas, que Ø<sub>1</sub> me *comí* a las seis y pico de la mañana en Puerto Cabello (c33hc.06)
- (43) /referente anterior de 1<sup>a</sup> persona no expreso/  
 a. mientras Ø<sub>0</sub> *tenía* el de bachillerato YO<sub>1</sub> siempre *tuve* actividades extra-cátedra (c31hb.05)  
 b. Ø<sub>0</sub> *fui* e ... ayudante de misa del padre ¿verdad?, e ... Ø<sub>1</sub> *tuve* tiempo ayudando al padre a la misa (c33ha.04)

En el cuadro 19 puede observarse la distribución de los casos según la presencia o ausencia del EFECTO REPETICIÓN:

Cuadro 19: Presencia o ausencia del pronombre de 1ª persona según el EFECTO REPETICIÓN

<b>Efecto repetición</b>		<b>Presencia</b>	<b>Ausencia</b>	<b>Total</b>	<b>%</b>
Referente anterior de 1ª persona no expreso	N	546	1.845	2.391	66
	%	23	77		
Referente anterior de 1ª persona expreso	N	423	539	962	27
	%	44	56		
No hay referente de 1ª persona	N	103	144	247	7
	%	42	58		
Total	N	1.072	2.528	3.600	100
	%	30	70		

$$\chi^2 = 164,57 \quad 2 \text{ g. d. l.} \quad p < 0,0000$$

Como se observa en el cuadro 19, los porcentajes de presencia oscilan entre un máximo de 44%, cuando se trata de un *referente anterior de 1ª persona expreso*, y un mínimo de 23% cuando el *referente anterior de 1ª persona no es expreso*. La prueba del  $\chi^2$  señala que la variable EFECTO REPETICIÓN está estadísticamente relacionada con la expresión del pronombre. En el corpus caraqueño actual, así como en la investigación de Camerón y Flores Ferrán 2003, quienes incluyen entre sus variantes la de *no hay referente de 1ª persona*, al igual que en el presente estudio, se exhibe la misma jerarquización. En el cuadro 20 se reportan los resultados obtenidos en otros estudios con relación a la presencia del pronombre:

Cuadro 20: EFECTO REPETICIÓN en este y otros estudios

<b>Efecto repetición</b>	<b>Estudios</b>		<b>Presencia %</b>	<b>Ausencia %</b>
Referente anterior de 1ª persona expreso	<b>Pérez Brabandere 2010</b>	<b>Caracas 2004-2010</b>	<b>44</b>	<b>56</b>
	Travis 2007	Cali, Colombia 1996	55	45
	Travis 2007	Nuevo México 1992-1995	50	50
	Torre-Cacoullos y Travis 2008	Nuevo México 1992-1995	51	49
Referente anterior de 1ª persona no expreso	<b>Pérez Brabandere 2010</b>	<b>Caracas 2004-2010</b>	<b>23</b>	<b>77</b>
	Travis 2007	Cali, Colombia 1996	40	80
	Travis 2007	Nuevo México 1992-1995	23	77
	Torre-Cacoullos y Travis 2008	Nuevo México 1992-1995	23	77

Como se puede apreciar en el cuadro 20, en el presente estudio la variante que se usa más frecuentemente con un pronombre es la de *referente anterior de 1ª persona expreso*. Los valores en la comparación con otros estudios oscilan entre un mínimo de 44% (en el presente estudio) y un máximo de 55% (en el corpus colombiano que estudia Travis 2007), con valores intermedios en el corpus de Nuevo México, tanto en el analizado por Travis 2007 (50%), como en el analizado por Torre-Cacoullous y Travis 2008 (51%).

#### 4.1.9. Distancia referencial

Con la novena y última variable independiente estudio la distancia referencial, empleando las mismas variantes propuestas por Travis 2007. La variable está conformada por cinco variantes establecidas según el número de cláusulas (entre una y diez) que pueden hallarse entre las dos en las que se hallan los sujetos de 1ª persona analizados: a) *ninguna cláusula*; b) *una cláusula*; c) *entre dos y cuatro cláusulas*; y d) *entre cinco y diez cláusulas*, e) *no hay referente de 1ª persona* (44-48):

- (44) /no hay ninguna cláusula/<sup>22</sup>
- a. YO<sub>b</sub> no *conozco* la casa de un grupo de mis tíos y, sin embargo, YO<sub>a</sub> siempre los *veo* (c31hb.05)
  - b. ya Ø<sub>b</sub> *estoy cansada* de mi hermana, Ø<sub>a</sub> no la *quiero ver!* (c31me.05)
- (45) /hay una cláusula/
- a. Bueno, YO<sub>b</sub> *estaba* en un programa que eh, se llama ... como una bolsa de trabajo ¿no? y entonces, este → 1 cláusula  
YO<sub>a</sub> *daba* clases una hora al día durante los días de semana (c32ha.04)
  - b. Ø<sub>b</sub> No *he montado* mi casa de bolsa, el mundo ha cambiado un poco → 1 cláusula  
este, pero Ø<sub>a</sub> *me metí* en el mundo de las casa de cambio (c32hd.07)

<sup>22</sup> El caso en estudio está identificado con el subíndice “a”, y su referente previo con el subíndice “b”.

- (46) /hay entre dos y cuatro cláusulas/
- a. que YO<sub>b</sub> lo recuerdo  
 vivió tanto, → 4 cláusulas  
 fue tan fue... → 3 cláusulas  
 la familia de mi mamá son una gente longeva, → 2 cláusulas  
 él vivió tanto que ... que ... que → 1 cláusula  
 YO<sub>a</sub> lo recuerdo (c32ma.04)
- b. ese es mi horario que Ø<sub>b</sub> tengo comprendido de lunes a viernes,  
 eh sábado y domingo si es especial, → 2 cláusulas  
 algún cliente quiere un viaje, sobre todo viaje, → 1 cláusula  
 sobre todo viaje, Ø<sub>a</sub> salgo, si no no (c32hb.05)
- (47) /hay entre cinco y diez cláusulas/
- a. YO<sub>b</sub> soy muy familiar con ellos → 5 cláusulas  
 eso es → 4 cláusulas  
 lo que a mí me gusta, o sea mi familia → 3 cláusulas  
 y me encanta mi familia → 2 cláusulas  
 E.: ¿es grande tu núcleo familiar? → 1 cláusula  
 I.: bastante grande, o sea lo que pasa es que  
 donde YO<sub>a</sub> vivo, es, así todo como (c31ha.04)
- b. El primer día que Ø<sub>b</sub> me fui a entrevistar este ... en Sivensa → 7 cláusulas  
 la parte administrativa queda en la.. en la Torre América → 6 cláusulas  
 queda en la Avenida Venezuela, → 5 cláusulas  
 resulta ser que Sivensa → 4 cláusulas  
 ahí tiene... → 3 cláusulas  
 no es dueña de todo el edificio → 2 cláusulas  
 pero es prácticamente dueña de la mitad del edificio, → 1 cláusula  
 ahí tiene una serie de pisos  
 cuando Ø<sub>a</sub> llegué ahí, yo era un muchacho de 18 años (c32hd.07)

- (48) /no hay referente de 1ª persona/
- a. El Valle era lo que era realmente una comunidad, El Valle todo el mundo se conocía, → +10 cláusulas  
 en el Valle todo el mundo sabía → 10 cláusulas  
 quién era uno, → 9 cláusulas  
 en el Valle sabían quién era las cosas (*sic*) → 8 cláusulas  
 Por supuesto, mi familia es una familia bastante religiosa → 7 cláusulas  
 pero en el sentido de los que se llamaba la familia expresa religiosa ¿no?, de creencias, de profesión pero e... en el común de los normales de práctica → 6 cláusulas  
 en los tiempos más o menos que práctica una persona normalmente e... no todos los días, no todos los domingos, → 5 cláusulas  
 sino lo que era la Navidad, → 4 cláusulas  
 lo que era Semana Santa, las fiestas patronales, etc., etc., etc., → 3 cláusulas  
 el patrono del Valle es San Roque, una fiesta → 2 cláusulas  
 que se celebraba el dieciséis de agosto de todos los años, → 1 cláusula  
 al menos YO<sub>a</sub> *estuve* (c33hb.05)
- b. **H.:** sí por ejemplo cuando uno está en una está en una ingeniería, entonces hay las básicas que ve en la universidad, tú ves las básicas, todo el mundo ve matemática, física, química de las áreas de ciencias ¿no?, → +10 cláusulas  
 después tú caes a las básicas profesionales → 10 cláusulas  
 que son, como decir la la física aplicada, la dinámica estática, → 9 cláusulas  
 esas son profesionales pero un básico pues → 8 cláusulas  
 todo el que se dedique a la ingeniería → 7 cláusulas  
 va a ver esas → 6 cláusulas  
 si tú vas a estudiar otra área de las ciencias → 5 cláusulas  
 tú no vas a ver esas básicas, → 4 cláusulas  
 ves otras básicas por lo menos en ingeniería, → 3 cláusulas  
 tú ves unas básicas → 2 cláusulas  
 que son → 1 cláusula  
 Ø<sub>a</sub> no sé (c32hg.08)

En el cuadro 21 se muestran los resultados obtenidos al cruzar la variable dependiente con la variable DISTANCIA REFERENCIAL:

Cuadro 21: Presencia o ausencia del pronombre de 1ª persona según la DISTANCIA REFERENCIAL

<b>Distancia referencial</b>		<b>Presencia</b>	<b>Ausencia</b>	<b>Total</b>	<b>%</b>
Ninguna cláusula	N	336	1.301	1.637	46
	%	21	79		
2-4 cláusulas	N	259	439	698	19
	%	37	63		
1 cláusula	N	211	428	639	18
	%	33	67		
5-10 cláusulas	N	163	216	379	10
	%	43	57		
No hay referente de 1ª persona	N	103	144	247	7
	%	42	58		
Total	N	1.072	2.528	3.600	100
	%	30	70		

$$\chi^2 = 136,67 \quad 4 \text{ g. d. l.} \quad p < 0,0000$$

En el cuadro 21 se puede observar cómo los porcentajes oscilan de un máximo de 43% cuando la distancia oscila entre *5-10 cláusulas*, a un mínimo de 21%, cuando *no hay ninguna cláusula*. La prueba del  $\chi^2$  señala que se puede rechazar la hipótesis nula y relaciona estadísticamente la DISTANCIA REFERENCIAL con la presencia del pronombre.

Los resultados obtenidos en otros estudios en relación con la presencia del pronombre son los reportados en el cuadros 22:



Cuadro 22: DISTANCIA REFERENCIAL en este y otros estudios

<b>Distancia referencial</b>	<b>Estudios</b>		<b>Presencia %</b>	<b>Ausencia %</b>
5-10 cláusulas	<b>Pérez Brabandere 2010</b>	<b>Caracas 2004-2010</b>	<b>43</b>	<b>57</b>
	Travis 2007	Nuevo México 1992-1995	46	54
	Travis 2007	Cali, Colombia 1996	38	62
2-4 cláusulas	<b>Pérez Brabandere 2010</b>	<b>Caracas 2004-2010</b>	<b>37</b>	<b>63</b>
	Travis 2007	Cali, Colombia 1996	49	51
	Travis 2007	Nuevo México 1992-1995	39	61
1 cláusula	<b>Pérez Brabandere 2010</b>	<b>Caracas 2004-2010</b>	<b>33</b>	<b>67</b>
	Travis 2007	Cali, Colombia 1996	44	56
	Travis 2007	Nuevo México 1992-1995	34	66
Ninguna cláusula	<b>Pérez Brabandere 2010</b>	<b>Caracas 2004-2010</b>	<b>21</b>	<b>79</b>
	Travis 2007	Cali, Colombia 1996	55	45
	Travis 2007	Nuevo México 1992-1995	23	77

Los resultados de Travis 2007 no se asemejan mucho a los de la presente investigación, pero pese a que la investigadora no usó la variante *no hay referente de 1ª persona*, la jerarquización de las frecuencias relativas de las variantes de esta variable para la expresión del pronombre son prácticamente las mismas.

#### 4.1.9. Variables extralingüísticas

En los cuadros 23 y 25 se presentan los resultados de la tabulación cruzada y de la prueba del  $\chi^2$  aplicada a las dos variables sociales utilizadas en el presente estudio.

El cuadro 23 muestra los resultados del cruce de la variable GRUPO ETARIO con la variable presencia~ausencia del pronombre de 1ª persona:

Cuadro 23: Presencia o ausencia del pronombre de 1ª persona según GRUPO ETARIO

Grupo etario		Presencia	Ausencia	Total	%
Grupo 1 20-34 años	N	330	870	1.200	33
	%	28	72		
Grupo 2 35-54 años	N	360	840	1.200	33
	%	30	70		
Grupo 3 55 años o más	N	382	818	1.200	34
	%	32	68		
Total	N	1.072	2.528	3.600	100
	%	30	70		

$$\chi^2 = 5,43 \quad 2 \text{ g. d. l.} \quad p < 0,06618769$$

Los porcentajes de presencia del pronombre en los tres GRUPOS ETARIOS son muy similares, ya que oscilan entre 28% y 32% cuando el pronombre es explícito. El grupo de *tercera edad* (3) obtuvo el mayor porcentaje de aparición (32%). La prueba del  $\chi^2$  confirma lo que revelan los porcentajes, y señala que no puede rechazarse la hipótesis nula.

En el cuadro 24 se muestran los resultados que se obtuvieron en otros estudios en los que se incluyó el GRUPO ETARIO como variable extralingüística.

Cuadro 24: GRUPO ETARIO en este y otros estudios

<b>Grupo etario</b>	<b>Estudios</b>		<b>Presencia %</b>	<b>Ausencia %</b>
Grupo 1 20-34 años	<b>Pérez Brabandere 2010</b>	<b>Caracas 2004-2010</b>	<b>28</b>	<b>72</b>
	Samper, Hernández y Troya 2006	Las Palmas de Gran Canarias 1985	17	83
	Samper, Hernández y Troya 2006	Madrid 1985	22	78
	Samper, Hernández y Troya 2006	Salamanca 1985	28	72
	Samper, Hernández y Troya 2006	Sevilla1972-1973	26	74
Grupo 2 35-54 años	<b>Pérez Brabandere 2010</b>	<b>Caracas 2004-2010</b>	<b>30</b>	<b>70</b>
	Samper, Hernández y Troya 2006	Las Palmas de Gran Canarias 1985	33	67
	Samper, Hernández y Troya 2006	Madrid 1985	33	67
	Samper, Hernández y Troya 2006	Salamanca 1985	37	63
	Samper, Hernández y Troya 2006	Sevilla1972-1973	33	67
Grupo 3 55 años o más	<b>Pérez Brabandere 2010</b>	<b>Caracas 2004-2010</b>	<b>32</b>	<b>68</b>
	Samper, Hernández y Troya 2006	Las Palmas de Gran Canarias 1985	29	71
	Samper, Hernández y Troya 2006	Madrid 1985	28	72
	Samper, Hernández y Troya 2006	Salamanca 1985	32	68
	Samper, Hernández y Troya 2006	Sevilla1972-1973	28	72

Según se aprecia en el cuadro 24, del conjunto de corpus que estudiaron Samper *et alii* 2006, sólo en el de Las Palmas de Gran Canaria hay variación en la presencia del pronombre entre el grupo 1 (jóvenes entre 20 y 34 años) y los otros dos grupos 17% vs. 33% y 29%, respectivamente. Los resultados, al igual que los del presente estudio, muestran claramente que esta variable no influye en la variación presencia~ausencia del pronombre.

En el cuadro 25 se muestran los resultados de la presencia~ausencia del pronombre de primera persona de acuerdo con el factor SEXO:

Cuadro 25: Presencia o ausencia del pronombre de la 1ª persona según SEXO

Sexo		Presencia	Ausencia	Total	%
Femenino	N	580	1220	1800	50
	%	32	68		
Masculino	N	492	1308	1800	50
	%	27	73		
Total	N	1072	2528	3600	100
	%	30	70		

$\chi^2$  corregido por Yates = 10,05      p < 0,0015196

En el cuadro 25 puede apreciarse que los porcentajes de aparición del pronombre son muy similares tanto en el grupo de las mujeres (32%) como en el de los hombres (27%), con tan sólo cinco puntos porcentuales de diferencia. Sin embargo, los resultados de la prueba del  $\chi^2$  permiten rechazar la hipótesis nula.

En otros estudios en los que se incluyó el SEXO como variable extralingüística se obtuvieron los resultados que se muestran en el cuadro 26:

Cuadro 26: SEXO en este y otros estudios

Sexo	Estudios		Presencia %	Ausencia %
Femenino	<b>Pérez Brabandere 2010</b>	<b>Caracas 2004-2010</b>	<b>32</b>	<b>68</b>
	Samper, Hernández y Troya 2006	Las Palmas de Gran Canarias 1985	28	72
	Samper, Hernández y Troya 2006	Madrid 1985	28	72
	Samper, Hernández y Troya 2006	Salamanca 1985	31	69
	Samper, Hernández y Troya 2006	Sevilla1972-1973	30	70
	Bentivoglio 1999	Caracas 1968-1977	37	63
	Miró Vera y de Pineda 1990	Sevilla 1972-1973	27	73
	Bentivoglio 1987	Caracas 1977	40	60
Masculino	<b>Pérez Brabandere 2010</b>	<b>Caracas 2004-2010</b>	<b>27</b>	<b>73</b>
	Samper, Hernández y Troya 2006	Las Palmas de Gran Canarias 1985	28	72
	Samper, Hernández y Troya 2006	Madrid 1985	29	71
	Samper, Hernández y Troya 2006	Salamanca 1985	34	66
	Samper, Hernández y Troya 2006	Sevilla1972-1973	29	71
	Bentivoglio 1999	Caracas 1968-1977	47	53
	Miró Vera y de Pineda 1990	Sevilla 1972-1973	28	72
	Bentivoglio 1987	Caracas 1977	41	59

Los resultados que se observan en el cuadro 26 muestran que sólo en los estudios de Bentivoglio (1987 y 1999) sobre el habla de Caracas los porcentajes de presencia pronominal, son superiores al 40%; en todos los otros corpus la presencia varía entre 27% (corpus Caracas 2004-2010) y 34% (Salamanca)

De acuerdo con los resultados de las pruebas del  $\chi^2$  (ver parte inferior de los cuadros en los que se presentan los resultados de la presente investigación), de un total de once variables, nueve lingüísticas y dos extralingüísticas, tan sólo siete (NÚMERO GRAMATICAL DEL SUJETO, AMBIGÜEDAD DE LA FORMA VERBAL, CLASE SEMÁNTICA DEL VERBO, CONECTIVIDAD DISCURSIVA, EFECTO REPETICIÓN, DISTANCIA REFERENCIAL y SEXO) parecen tener una relación estadísticamente significativa con la alternancia entre presencia y ausencia del pronombre de 1ª persona.

#### 4.2. GoldVarb X

En la segunda etapa del análisis cualitativo utilicé el método directo (*one level*) y el análisis por pasos (*step-up and down*), pues en “los programas VARBRUL conviene realizar ambos, ya que la información que proporcionan es complementaria.” (Hernández Campoy y Almeida 2005:272).

En los análisis de regresión múltiple de GoldVarb X, los pesos probabilísticos superiores a 0,50 indican que una o más variantes de una variable contribuyen significativamente a una de las dos variantes de la variable dependiente, mientras que los inferiores a 0,50 no lo hacen. El análisis probabilístico del presente estudio está

dirigido a establecer la contribución de las diferentes variables a la variante *presencia del pronombre*.

Como señalé en la sección final de los antecedentes (21), LA CONECTIVIDAD DISCURSIVA, EL EFECTO REPETICIÓN y LA DISTANCIA REFERENCIAL surgen de la reflexión teórica en torno a la variable CAMBIO DE REFERENTE y, por lo tanto, algunas de sus variantes interactúan entre ellas. Debido a este hecho, haré dos análisis separados: uno con la variable CONECTIVIDAD, y el otro con las variables EFECTO REPETICIÓN y DISTANCIA REFERENCIAL, dado que, según Travis 2007, estas dos últimas son inversamente proporcionales y se complementan.

#### 4.2.1. Modelo con Conectividad

En el cuadro 27 se muestran las variables seleccionadas por GoldVarb X en el análisis binomial de subida y bajada (*binomial step up & down*), de acuerdo con el orden de importancia, que se manifiesta en el valor del rango.

Cuadro 27: Resultados del análisis binomial de subida y bajada con CONECTIVIDAD

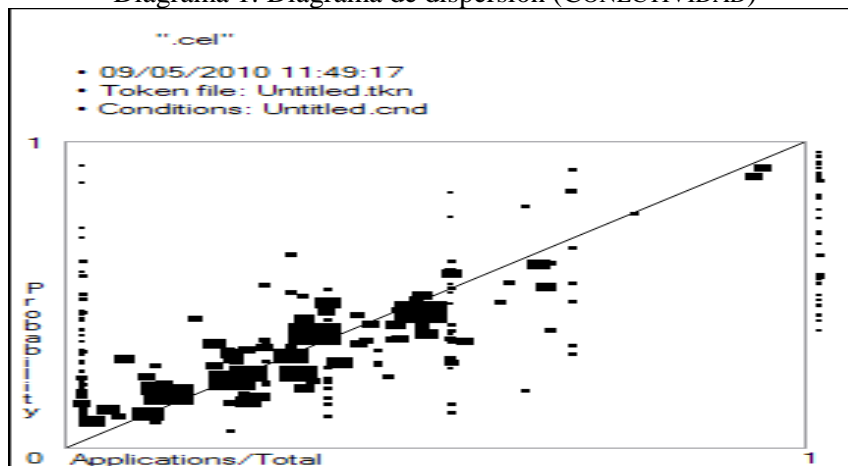
VARIABLES	Nº DE CASOS	% DE PRESENCIA	PESO PROBABILÍSTICO
AMBIGÜEDAD DE LA FORMA VERBAL			
Morfológica y contextualmente ambigua	82/96	85	<b>0,91</b>
Morfológicamente ambigua	181/439	41	<b>0,62</b>
No ambigua	809/3.065	26	0,46
<i>Rango</i>			45
CONECTIVIDAD			
Grado 8	73/153	48	<b>0,72</b>
Grado 7	20/41	49	<b>0,69</b>
Grado 5	481/1.251	38	<b>0,61</b>
Grado 4	38/118	32	<b>0,57</b>
Grado 3	50/161	31	<b>0,53</b>
Grado 6	63/208	30	<b>0,52</b>
Grado 2	170/656	26	0,45
Grado 1	177/1.012	18	0,35
<i>Rango</i>			37
CLASE SEMÁNTICA DEL VERBO			
Verbos de procesos verbales	36/82	44	<b>0,66</b>
Verbos de procesos mentales	128/374	34	<b>0,53</b>
Verbos de procesos materiales	783/2.650	30	0,50
Verbos de procesos relacionales	95/359	26	0,48
Verbos de procesos existenciales	22/87	25	0,47
Verbos de modulación	8/48	17	0,33
<i>Rango</i>			33
NÚMERO GRAMATICAL DEL SUJETO			
Singular	947/2.785	34	<b>0,55</b>
Plural	125/815	15	0,33
<i>Rango</i>			22
SEXO			
Femenino	580/1.800	32	<b>0,54</b>
Masculino	492/1.800	27	0,46
<i>Rango</i>			8
$\chi^2$ celdas = 1,1440    Total $\chi^2$ = 252,8136	N = 3.600	Input 0,277	Verosimilitud= - 1975,801

En los resultados del cuadro 27 se puede observar que el valor del  $\chi^2$  es superior al valor mínimo requerido en el modelo (40,113) para una probabilidad de 0,05, lo que indica que “disponemos de un modelo teórico que da cuenta de la variación”

(Hernández Campoy y Almeida, 2005:277). Cuando la probabilidad es inferior a 0,05 –el límite requerido para las ciencias sociales– la hipótesis nula puede rechazarse, es decir, los resultados no son azarosos, pues se ajustan al modelo creado. En este nivel de análisis, de las variables seleccionadas cuatro de las variantes tienen pesos probabilísticos muy superiores a 0,50: AMBIGÜEDAD *morfológica y contextualmente ambigua* tiene un peso de 0,91; CONECTIVIDAD *grado 8 y 7*, de 0,72 y 0,69, respectivamente; CLASE SEMÁNTICA, *verbos de procesos verbales* de 0,66; y AMBIGÜEDAD, *morfológicamente ambigua* 0,62).

En el primer nivel de análisis del GoldVarb, el programa genera un gráfico denominado *Diagrama de dispersión*, en el que “se ponen en relación los valores observados (eje vertical) con los esperados (eje horizontal) para cada una de las celdas. Cuanto más se ajusten los valores representados a la línea de regresión, más adecuado será el modelo que aquí se presenta al modelo ideal” (Hernández Campoy y Almeida, 2005:284).

Diagrama 1: Diagrama de dispersión (CONECTIVIDAD)





En el diagrama de dispersión 1 muchos de los valores observados se acercan al eje o línea de regresión.

#### 4.2.2. Modelo con Efecto repetición y Distancia referencial

En el cuadro 28 se muestran las variables seleccionadas por GoldVarb X en el análisis binomial de subida y bajada (*binomial step up & down*), de acuerdo con el orden de importancia de cada variable, orden que se manifiesta en el rango.

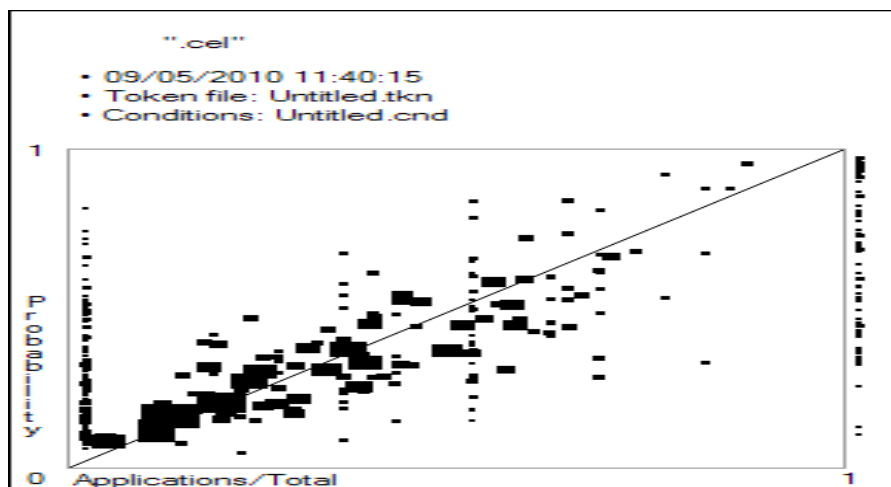
Cuadro 28: Resultados del análisis binomial de subida y bajada con la variable EFECTO REPETICIÓN y DISTANCIA REFERENCIAL

VARIABLES	Nº DE CASOS	% DE PRESENCIA	PESO PROBABILÍSTICO
AMBIGÜEDAD DE LA FORMA VERBAL			
Morfológica y contextualmente ambigua	82/96	85	<b>0,91</b>
Morfológicamente ambigua	181/439	41	<b>0,62</b>
No ambigua	809/3.065	26	0,47
<i>Rango</i>			<i>44</i>
CLASE SEMÁNTICA			
Verbos de procesos verbales	36/82	44	<b>0,66</b>
Verbos de procesos mentales	128/374	34	<b>0,52</b>
Verbos de procesos materiales	783/2.650	30	<b>0,50</b>
Verbos de procesos relacionales	95/359	26	0,47
Verbos de procesos existenciales	22/87	25	0,47
Verbos de modulación	8/48	17	0,30
<i>Rango</i>			<i>36</i>
DISTANCIA REFERENCIAL			
De cinco a diez cláusulas	163/379	43	<b>0,68</b>
De dos a cuatro cláusulas	259/698	37	<b>0,60</b>
Una cláusula	211/639	33	<b>0,56</b>
No hay referente de 1ª persona	103/247	42	<b>0,53</b>
No hay ninguna cláusula	336/1.637	21	0,39
<i>Rango</i>			<i>29</i>
NÚMERO GRAMATICAL DEL SUJETO			
Singular	947/2.785	34	<b>0,55</b>
Plural	125/815	15	0,33
<i>Rango</i>			<i>22</i>
EFECTO REPETICIÓN			
No hay referente de 1ª persona	103/247	44	<b>0,65</b>
Referente anterior de 1ª persona expreso	423/962	42	<b>0,64</b>
Referente anterior de 1ª persona no expreso	546/2.391	23	0,43
<i>Rango</i>			<i>22</i>
CAMBIO DE TURNO			
Igual turno de habla	1.010/3.399	30	<b>0,51</b>
Diferente turno de habla	62/201	31	0,41
<i>Rango</i>			<i>10</i>
SEXO			
Femenino	580/1.800	32	<b>0,54</b>
Masculino	492/1.800	27	0,46
<i>Rango</i>			<i>8</i>
Total $\chi^2 = 350,0468$ $\chi^2$ celdas = 1,0480	N = 3.600	Input 0,272	Verosimilitud = -1925,934

En los resultados del cuadro 28 se puede observar que el valor del  $\chi^2$  es superior al valor mínimo requerido: de las variables seleccionadas por GoldVarb cinco de las variantes tienen pesos probabilísticos muy superiores a 0,50: la variante *morfológica y contextualmente ambigua*, de la variable AMBIGÜEDAD, obtiene un peso de 0,91; la *de cinco a diez cláusulas* de la variable DISTANCIA REFERENCIAL, 0,68; la *de procesos verbales* de la variable CLASE SEMÁNTICA, 0,66; las variantes *no hay referente de 1ª persona* y *referente anterior de 1ª persona expreso*, en la variable EFECTO REPETICIÓN, obtienen pesos de 0,65 y 0,64, respectivamente; la *morfológicamente ambigua*, de la variable AMBIGÜEDAD, 0,62; y la *de dos y cuatro cláusulas*, de la variable DISTANCIA REFERENCIAL, 0,60.

A continuación se presenta el diagrama de dispersión de este modelo, que como se dijo en la página 68, es un gráfico generado en el primer nivel de análisis del GoldVarb.

Diagrama 2: Diagrama de dispersión (EFECTO REPETICIÓN y DISTANCIA REFERENCIAL)



En el diagrama de dispersión 2 muchos de los valores observados se acercan al eje o línea de regresión.

#### 4.2.3. Comparación de resultados de los dos modelos con estudios previos

La variable más significativa en ambos modelos es la AMBIGÜEDAD DE LA FORMA VERBAL, la cual obtuvo casi los mismos pesos en el modelo con CONECTIVIDAD (MC) y en el modelo con EFECTO REPETICIÓN/DISTANCIA REFERENCIAL (MEF): 0,91 si los casos son *morfológica y contextualmente ambiguos*; 0,62, si los casos son *morfológicamente ambiguos*, y 0,46 (MC) / 0,47 (MEF) si no hay ninguna *ambigüedad*. Esta variable ha resultado significativa –aunque no siempre la más significativa– en todos los estudios donde se ha utilizado. Para Bentivoglio 1987 es el tercer factor después de CORREFERENCIALIDAD y NÚMERO GRAMATICAL DEL SUJETO: la variante *ambigua* arroja un peso de 0,59. En Bentivoglio 1999, la AMBIGÜEDAD sigue siendo el tercer factor relevante, pero esta vez luego de la CORREFERENCIALIDAD y CAMBIO DE TURNO; la variante más influyente continúa siendo la *ambigüedad* con un peso de 0,66. Samper *et alii* 2006 también encuentran en todos los dialectos (Las Palmas de Gran Canaria, Madrid, Salamanca, Sevilla) analizados conjuntamente con GoldVarb que la variante que favorece la aparición del pronombre es la de *morfológica y contextualmente ambigua* (0,62). En cambio, para Blanco Canales 1999, y Torres-Caucoullós y Travis 2008 la AMBIGÜEDAD es la variable más desfavorecedora.

En el MC la CONECTIVIDAD es el segundo factor más significativo. Al detallar los resultados se ven las diferencias respecto de los estudios de Paredes Silva (1993, 2003), pues en la presente investigación los pesos probabilísticos de los diferentes grados de conectividad no están ordenados jerárquicamente, pues el *grado 6*, que debería situarse inmediatamente debajo del *grado 7* (con peso 0,69) y antes del *grado 5* (con peso 0,61), por el contrario, aparece después del *grado 3* (con peso 0,53). En efecto, los pesos de los *grado 3* y *6* son casi idénticos: 0,53 y 0,52, respectivamente; lo esperado habría sido que el *grado 6* tuviera un peso probabilístico entre 0,69 (del *grado 7*) y 0,61 (del *grado 5*).

La CLASE SEMÁNTICA DEL VERBO es la tercera variable por orden de rango en el MC y la segunda en el MEF: en ambos modelos los procesos que contribuyen mayormente a la presencia del pronombre son los verbales (MC y MEF: 0,66), los mentales (MC: 0,53 / MEF: 0,52), y los materiales (MC y MEF: 0,50). En estudios anteriores, se había estudiado la CLASIFICACIÓN SEMÁNTICA DE LOS VERBOS y aunque las categorías no son las mismas que las del presente estudio, sería interesante señalar las variantes significativas: en Bentivoglio 1987 y en Blanco Canales 1999 los *verbos de percepción* (0,59) y (0,60); en Bentivoglio 1999, los *existenciales* (0,64), los *de movimiento* (0,53), y los *de cognición* (0,53). En Samper *et alii* 2006 resultaron significativos los que expresan *cognición* (0,62), *percepción* (0,61), y *volición* (0,52). Travis 2007, por su parte, señala como significativos los *psicológicos* (0,70), los *copulativos* (0,55), y los *de lengua* (0,53). Para Torres-Caucoullós y Travis 2008 los *psicológicos* obtienen un peso de 0,67.

La variable NÚMERO GRAMATICAL DEL SUJETO es la cuarta por rango de significación tanto en MC como en MEF, y la variante *singular* es la que contribuye más a la presencia del pronombre con un peso de 0,55. En los análisis de Bentivoglio 1987 y Blanco Canales 1999 el número *singular* tiene un peso de 0,68 y 0,59, respectivamente.

En el MEF, la DISTANCIA REFERENCIAL resultó ser la tercera variable en cuanto a rango. La DISTANCIA REFERENCIAL había sido mencionada por Cameron y Flores Ferrán 2003, pero es Travis 2007 quien la codifica y analiza. En su análisis del habla de Nuevo México, *5+ cláusulas* (0,63); *2-4 cláusulas* (0,58); *1 cláusula* (0,53); *0 cláusulas* (0,39); y para Cali (Colombia): *5+ cláusulas* (0,58); *2-4 cláusulas* (0,51); *1 cláusula* (0,45); *0 cláusulas* (0,42). Llama la atención el hecho de que los valores de Travis 2007 se asemejan a los obtenidos en el presente estudio, pese que en este último hay una variante más: *5-10 cláusulas* (0,68); *2-4 cláusulas* (0,60); *1 cláusula* (0,56); *no hay referente de 1ª persona* (0,53); *0 cláusulas* (0,39).

La variable EFECTO REPETICIÓN también resultó significativa: *no hay referente de 1ª persona* (0,65) y *referente anterior de 1ª persona expreso* (0,64). La investigación de Cameron y Flores Ferrán 2003 es la única en la que se emplean para el EFECTO REPETICIÓN las mismas variantes: *referente anterior de 1ª persona expreso* (0,64), *no hay referente de 1ª persona* (0,54), y *referente anterior de 1ª persona no expreso* (0,37); sus resultados parecen coincidir con los míos. Incluso si comparamos los resultados de la presente investigación con los de Travis 2007 y los de Torres-

Caucoullous y Travis 2008, que sólo tienen dos variantes, se aprecian las semejanzas en los resultados.

La variable CAMBIO DE TURNO sólo fue usada en el MEF, pues se solapa con la variante *grado 3* de la variable CONECTIVIDAD. CAMBIO DE TURNO ha sido objeto de estudio en investigaciones previas y más de una vez estuvo entre las variables más relevantes; en cambio, en el presente análisis, la variante que más contribuye a la presencia del pronombre es *igual turno de habla* con un peso de 0,51, que apenas está por encima del umbral de 0,50, valor mínimo requerido por el programa GoldVarb. En las dos investigaciones de Bentivoglio (1987, 1999), la variante *cambio de turno de habla* resultó ser relevante con pesos de 0,55 y 0,72, respectivamente. Quizás esta diferencia se deba al hecho de que la presente investigación se basa en un corpus mucho más amplio que el usado por la investigadora en esos estudios.

Por último, la variable sociolingüística SEXO es la última seleccionada como significativa por GoldVarb X en ambos modelos (MC y MEF): la variante *sexo femenino* contribuye a la expresión del pronombre con un peso de 0,54. En estudios previos el SEXO también se había revelado como una variable relevante para la expresión o no del pronombre. Bayley y Pease-Alvárez 1996 señalan que las mujeres favorecen la expresión del pronombre con un valor de  $\pi$ : 0,545. En algunos estudios se ha señalado en cambio que es la variante *sexo masculino* la más importante [ver Bentivoglio 1987 (0,55) y 1999 (0,69); Blanco Canales 1999 (0,56)]. Pero también se dan casos como los de la investigación de Samper *et alii* 2006 para quienes el SEXO no resulta ser una variable significativa.

#### 4.2.4. Ajuste del modelo

Para decidir cuál de los dos modelos se ajusta mejor a los datos en estudio, se toma en cuenta lo que dice Tagliamonte (2006:133)

How do we find a set of values which best accounts for the observed data? In statistics, how well a model with given factor effects fits a data set can be measured by several criteria.

Variable rule analysis uses the likelihood criterion because it can account for the extreme distributional imbalances, including contrasting full vs. near-empty cells, in corpus-based data. [...]

The likelihood measure indicates how likely it is that a particular set of data has been generated by the model which has the given values for the factor effects. [...] “The principle of maximum likelihood provides a means to choose the set of values which is most likely to have generated the data” (Sankoff 1988c: 990).<sup>23</sup>

Según lo anteriormente dicho, aunque la diferencia sea muy leve, el modelo que mejor se ajusta a nuestro estudio es el MEF pues, su nivel de verosimilitud es de **-1925,934**, mientras que el de MC es de **-1975,801**.

---

<sup>23</sup> Mi traducción es:

¿Cómo encontramos un conjunto de valores que dé mejor cuenta de los datos observados? En estadística, la manera en que un modelo se ajusta a un conjunto de datos puede medirse de acuerdo con varios criterios.

Los análisis de reglas variables usan el criterio de verosimilitud porque este puede dar cuenta de los desequilibrios distribucionales extremos, incluyendo el contrastar celdas llenas y casi vacías, en datos basados en corpus. [...]

El valor de la verosimilitud indica cuán probable es que un conjunto particular de datos haya sido generado por el modelo que tiene los valores dados para los efectos de un factor. [...] “El principio de verosimilitud máxima provee un factor para elegir el conjunto de valores que más probablemente hayan generado los datos” (Sankoff 1988c: 990).



#### 4.3. Frecuencia del verbo

En este apartado, se presenta una lista de los verbos encontrados en el corpus analizado, con el fin de ver cuáles son los más frecuentes y con cuáles se expresa más el pronombre, según puede apreciarse en el cuadro 30:<sup>24</sup>

Cuadro 30: Verbos más frecuentes en el corpus analizado

Forma léxica	Clase semántica	Cantidad	Con pronombre expreso
Tener	Relacional	391	96
Ser	Relacional	236	97
Saber	Mental	175	42
Decir	Verbal	141	69
Creer	Mental	137	72
Ir	Material	130	30
Estar	Material	97	18
Ver	Mental	78	31
Recordar	Mental	59	22
Estudiar	Mental	57	21
Vivir	Existencial	53	29
Acordar	Mental	51	22
Estar	Relacional	49	15
Hacer	De modulación	49	4
Conocer	Mental	46	14
Llegar	Material	46	13
Salir	Material	30	7
Hacer	Material	25	4
Entrar	Material	23	9
Quedar	Material	23	7
Querer	Mental	22	12
Comer	Material	21	7
Trabajar	Material	19	9
Graduar	Material	19	6
Venir	Material	19	5

<sup>24</sup> La lista de todos los verbos del corpus (no sólo de los más frecuentes) se presenta en el Anexo 4.

Forma léxica	Clase semántica.	Cantidad	Pronombre Expreso
Empezar	Existencial	19	1
Dar	Relacional	18	9
Pensar	Mental	18	9
Jugar	Material	18	4
Terminar	Existencial	18	2
Mudar	Material	17	2
Sentir	Mental	16	6
Nacer	Existencial	15	8
Hablar	Verbal	15	2
Tratar	De modulación	13	2
Pasar	Existencial	13	1
Pasar	Material	13	1
Leer	Mental	12	6
Meter	Material	12	4
Entender	Mental	11	6
Comprar	Relacional	11	4
Agarrar	Material	11	3
Casar	Material	11	3
Regresar	Material	11	3
Llevar	Material	10	4
Comenzar	Existencial	10	3
Necesitar	Relacional	10	3
Poder	Relacional	10	3
Querer	Mental	10	1

Los dos primeros verbos más frecuentes –‘tener’ y ‘ser’– son de *procesos relacionales*; los otros tres verbos con más de cien casos son de *procesos mentales* (‘saber’, ‘creer’) y *verbales* (‘decir’). Si tomamos en cuenta no sólo los cinco (5) primeros verbos sino todos aquellos con más de diez (10) casos cada uno, puede apreciarse que la clase semántica de *procesos materiales* es la que tiene mayor variedad de verbos (19). Siguen, por orden de aparición, los verbos de *procesos*

*mentales (13), de procesos relacionales (7), existenciales (6), verbales (2), y de modulación (2).*

## 5. CONCLUSIONES

Como se señaló en el apartado de los objetivos, el propósito principal del presente estudio era determinar los factores sintáctico-semántico-pragmáticos que influyen en la expresión variable del sujeto pronominal de primera persona en el español hablado actualmente en Caracas. Para esto se extrajeron los cien (100) primeros casos de las grabaciones de 36 hablantes, con grado de instrucción alto, del corpus PRESEEA-Caracas 2004-2010.

En investigaciones similares acerca de otras variedades del español se observa cómo la presencia de sujetos explícitos varía entre un mínimo de 23% en Buenos Aires y un máximo de 42% en el corpus de la norma culta de Caracas, con valores intermedios en Caracas 1977 (40%), Los Angeles (39%), Alcalá de Henares (35%), Santiago de Chile (30%), y con 30% en las varias ciudades estudiadas por Samper *et alii* 2006. En la muestra aquí analizada, la variación oscila entre un 30% de presencia pronominal y un 70% de ausencia.

En la presente investigación, en la búsqueda de un enfoque actualizado del fenómeno en el corpus actual caraqueño, empleé variables usadas en investigaciones similares, pero observé que ciertas variantes interactuaban entre ellas. Por este motivo me vi en la obligación de realizar dos análisis estadísticos separados a los que llamé: *Modelo con Conectividad* (MC) y *Modelo con Efecto repetición y Distancia referencial* (MEF).

En el MC el programa estadístico GoldVarb X señala que las variables que contribuyen significativamente a la expresión de los sujetos de primera persona en el habla de Caracas son jerárquicamente: en AMBIGÜEDAD, *morfológica y contextualmente ambigua* (0,91); en CONECTIVIDAD, *grado 8 y 7* (0,72 y 0,69, respectivamente); en CLASE SEMÁNTICA, *de procesos verbales* (0,66); en AMBIGÜEDAD, *morfológicamente ambigua* (0,62); en CONECTIVIDAD, *grado 4* de (0,57); en NÚMERO GRAMATICAL, *singular* (0,55); en SEXO, *femenino* (0,54); en CONECTIVIDAD, *grado 3* (0,53); y en CLASE SEMÁNTICA, *de procesos mentales* (0,53).

GoldVarb X también indica que las variables que contribuyen significativamente a la expresión de los sujetos de primera persona en el habla de Caracas en el MEF son, en orden jerárquico: en AMBIGÜEDAD, *morfológica y contextualmente ambigua* (0,91); en DISTANCIA REFERENCIAL, *de cinco a diez cláusulas* (0,68); en CLASE SEMÁNTICA, *de procesos verbales* (0,66); en EFECTO REPETICIÓN, *no hay referente de 1ª persona, y referente anterior de 1ª persona expreso* (0,65 y 0,64, respectivamente); en AMBIGÜEDAD, *morfológicamente ambigua* (0,62); en DISTANCIA REFERENCIAL, *de dos a cuatro cláusulas, y una cláusula* (0,60 y 0,56, respectivamente); en NÚMERO GRAMATICAL, *singular* (0,55); en SEXO, *femenino* (0,54); en DISTANCIA REFERENCIAL, *no hay referente de 1ª persona* (0,53); en CLASE SEMÁNTICA, *de procesos mentales* (0,52); en CAMBIO DE TURNO, *igual turno de habla* (0,51); y en CLASE SEMÁNTICA, *de procesos materiales* (0,50).

Se nota así que las variantes que resultan significativas suelen serlo en ambos modelos y además coinciden con las variables que han señalado otros autores como

más significativas para el fenómeno (ver el apartado 4.2.3., en el que se comparan los resultados de los dos modelos con estudios previos). Al analizar dos modelos surge obviamente la pregunta: ¿cuál de ellos se ajusta mejor a nuestro estudio? Para poder responder esta interrogante tuve que recurrir a Tagliamonte 2006, quien señala que es el valor de verosimilitud el que sirve para estos casos, y a pesar de que la diferencia entre ambos valores fue muy leve, el modelo que mejor se ajusta al presente estudio es el MEF.

También realicé un conteo de todos aquellos verbos con más de diez (10) casos cada uno, para apreciar así qué clase semántica tiene más variedad de verbos. Los resultados indican que hay mayor variedad en los verbos *de procesos materiales* (19), seguidos, en orden de aparición por los *de procesos mentales* (13), *procesos relacionales* (7), *existenciales* (6), *verbales* (2), y *de modulación* (2).

Por último, sólo queda aconsejar que en futuras investigaciones se comparta una metodología similar a la del presente estudio, para poder hacer comparaciones posteriores.

## 6. REFERENCIAS BIBLIOGRÁFICAS

- Alarcos Llorach, Emilio. 1994. *Gramática de la lengua española*. Madrid: Espasa Calpe.
- Ariel, Mira. 1990. *Accessing noun-phrase antecedents*. London: Routledge.
- Barrenechea, Ana María y Alicia Alonso. 1977. Los pronombres personales sujetos en el español hablado en Buenos Aires. En Juan Lope Blanch (ed.), *Estudios sobre el español hablado en las principales ciudades de América*, 333-49. México: Universidad Autónoma de México.
- Bayley, Robert y Lucinda Pease-Álvarez. 1996. Null and expressed pronoun variation in Mexican-descent children's Spanish. En Jennifer Arnold, Renée Blake, Brad Davidson, Scott Schwenter y Julie Solomon (eds.), *Sociolinguistic variation data, theory, and analysis. Selected papers from NWAV23 at Stanford*. 85-111. Stanford: Stanford University.
- Bentivoglio, Paola. 1980. Why "canto" and not "yo canto"? The problem of first person subject pronoun in spoken Venezuelan Spanish. Tesis de M.A. inédita, University of California, Los Angeles.
- Bentivoglio, Paola. 1987. *Los sujetos pronominales de primera persona en el habla de Caracas*. Caracas: Universidad Central de Venezuela.
- Bentivoglio, Paola. 1999. Recursos estadísticos aplicados al análisis sintáctico: una propuesta metodológica. En Joseph H. Matluck y Carlos A. Solé (eds.), *Simposio internacional de la lengua española: pasado, presente y futuro actas*. 225-236. Austin: Texas University-Press.
- Bentivoglio, Paola e Irania Malaver. 2006. La lingüística de corpus en Venezuela: un nuevo proyecto. *Lingua Americana*, 19. 37-46.
- Bentivoglio, Paola y Mercedes Sedano. 1993. Investigación sociolingüística: sus métodos aplicados a una experiencia venezolana. *Boletín de Lingüística* 8. 3-35.
- Blanco Canales, Ana. 1999. Presencia/ausencia de sujeto pronominal de primera persona en español. *Español Actual* 72. 31-39.
- Blanco Canales, Ana. 2004. *Estudio sociolingüístico de Alcalá de Henares*. Alcalá de Henares: Servicio de Publicaciones de la Universidad de Alcalá.
- Bright, William. 1974. Las dimensiones de la sociolingüística. En Paul Garvin y Yolanda Lastra (eds.), *Antología de estudios de etnolingüística y sociolingüística*, 197-202. México: Universidad Nacional Autónoma de México.

- Brown, Penelope y Stephen C. Levinson. 1987. *Politeness. Some universals in language usage*. Cambridge: Cambridge University Press.
- Calsamiglia Blancafort, Helena y Amparo Tusón Valls. 1999. *Las cosas del decir. Manual de análisis del discurso*. Barcelona: Ariel Lingüística.
- Cantero, Jorge Gustavo. 1976. Peculiaridades en el empleo del pronombre personal *yo* en el habla culta de la ciudad de México. *Anuario de Letras*, Volumen XIV. 233-237.
- Cameron, Richard. 1992. Pronominal and null subject variation in Spanish: Constraints, dialects, and functional compensation. Tesis de Ph. D. inédita, University of Pennsylvania, Philadelphia, Estados Unidos.
- Cameron, Richard. 1997. Accessibility theory in a variable syntax of Spanish. *Journal of Pragmatics* 28. 29-67.
- Cameron, Richard y Nydia Flores-Ferrán. 2003. Perseveration of subject expression across regional dialects of Spanish. *Spanish in Context* 1. 41-65.
- Cifuentes, Hugo. 1980-81. Presencia y ausencia del pronombre personal en el habla culta de Santiago de Chile. *Boletín de Filología de la Universidad de Chile. Homenaje a Ambrosio Rabanales* 31.743-52.
- D'Introno, Francesco. 1989. Empty and full pronouns in Spanish. *Hispanic Linguistics* 3. 27-47.
- de Paiva, Maria da Conceição y Maria Eugênia Lamoglia Duarte. 2003. Introdução: A mudança lingüística em curso. En Maria da Conceição de Paiva y Maria Eugênia Lamoglia Duarte (Organización), *Mudança lingüística em tempo real*. 97-114. Rio de Janeiro: Contra Capa.
- de Pineda, Miguel Ángel. 1983. *Sociolingüística andaluza 2. Material de encuestas para el estudio del habla urbana culta de Sevilla*. Sevilla: Secretariado de publicaciones de la Universidad de Sevilla.
- Enríquez, Emilia V. 1984. *El pronombre personal sujeto en la lengua española hablada en Madrid*. Madrid: Consejo Superior de Investigaciones Científicas.
- Enríquez, Emilia V. 2000. El sistema pronominal del español. En Manuel Alvar (dir.), *Introducción a la lingüística española*, 307-329. Barcelona: Ariel.
- Esgueva, Manuel y Margarita Cantarero. 1981. *El habla de la ciudad de Madrid. Materiales para su estudio*. Madrid: Consejo Superior de Investigaciones Científicas.
- Fernández Soriano, Olga. 1999. Expresión y omisión del pronombre personal. En Ignacio Bosque y Violeta Demonte (eds.), *Gramática descriptiva de la lengua española*, tomo 1, 1209-1273. Madrid: Espasa-Calpe.



- García, Erica. 1985. Shifting variation. *Lingua* 67.189-224.
- Gili Gaya, Samuel. 1976. *Curso superior de sintaxis española*. Barcelona: Biblograf.
- Givón, Talmy. 1983. Topic continuity in discourse: An introduction. En Talmy Givón (ed.), *Topic continuity in discourse: A quantitative cross-language study*, 1-41. Amsterdam/Philadelphia, PA: Benjamins.
- Hernández Campoy, Juan y Manuel Almeida. 2005. Metodología de la investigación sociolingüística. Málaga: Editorial Comares
- Hochberg, Judith G. 1986. Functional compensation for /s/ deletion in Puerto Rican Spanish. *Language* 62.609-621.
- Hurtado, Luz Mercedes. 2001. La variable expresión del sujeto en el español de los colombianos y Colombo-americanos residentes en el condado de Miami-Dade. Tesis doctoral inédita, University of Florida, Miami.
- Labov, William. 1966. *The social stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistic.
- Labov, William. 1969. Contraction, deletion and inherent variability of the English copula. *Language* 45.715-762.
- Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania.
- Labov, William. 1975. Empirical foundations of linguistic theory. En Robert Austerlitz (ed.), *The scope of American Linguistics*, 77-133. Lisse: de Ridder.
- Lamoglia Duarte, Maria Eugênia. 2003. A evolução na representação do sujeito pronominal em dois tempos. En Maria da Conceição de Paiva y Maria Eugênia Lamoglia Duarte (Organização), *Mudança lingüística em tempo real*. 115-128. Rio de Janeiro: Contra Capa.
- Lavandera, Beatriz. 1978. Where does the sociolinguistic variable stop? *Language in society* 7. 171-138.
- López Morales, Humberto. 2004. *Sociolingüística*. Madrid: Gredos.
- Luján, Marta. 1999. Expresión y omisión del pronombre personal. En Ignacio Bosque y Violeta Demonte (eds.), *Gramática descriptiva de la lengua española*, tomo 1, 1275-1315. Madrid: Espasa-Calpe.
- Martín Rojo, Luisa y Michael Meewis. 1993. Referentes del sujeto pronominales y tácitos en la conversación en español: un enfoque pragmático. En Henk Haverkate, Kees Hengeveld y Gijs Mulder (eds.), *Aproximaciones pragmlingüísticas al español*, 87-118. Amsterdam/Atlanta: Editions Rodopi

- Martín Zorraquino, María Antonia y José Portolés Lázaro. 1999. Los marcadores del discurso. En Ignacio Bosque y Violeta Demonte (eds.), *Gramática descriptiva de la lengua española*, tomo 3, 4051-4073. Madrid: Espasa-Calpe.
- Meneses de Santana, Maria Zelma. 2008. O sujeito nulo na fala urbana itabiense: alguns fatores favorecedores. Ponencia presentada en el XV Congreso Internacional de la asociación de Lingüística y Filología de América Latina (ALFAL), del 18 al 21 de agosto 2008, Montevideo, Uruguay.
- Miró Vera, Ramona y Miguel Ángel de Pineda. 1990. Determinación sociolingüística de la presencia/ausencia del pronombre personal sujeto. En María Teresa Palet Plaja (ed.), *Sociolingüística andaluza 5: Habla de Sevilla y hablas americanas*, 37-44. Sevilla: Universidad de Sevilla.
- Morales, Amparo. 1982. La expresión de sujeto pronominal, primera persona, en el español de Puerto Rico. *Boletín de la Academia Puertorriqueña de la Lengua Española* 8. 91-102.
- Morales, Amparo. 1986. Algunos aspectos de gramática en contacto: la expresión del sujeto en el español de Puerto Rico. *Anuario de Letras*. Volumen XXIV Universidad Nacional Autónoma de México. 71-85.
- Morales, Amparo. 1997. La manifestación del sujeto en español: tendencias universales y variación dialectal. En Robert Hammond y Marguerite MacDonald (eds.), *Linguistic Studies in Honor of Bohdan Saciuk*, 87-103. West Lafayette, IN: Learning Systems, Inc.
- Morales, Amparo y María Vaquero. 1990. *El habla culta de San Juan*. Río Piedra, Puerto Rico: Editorial de la Universidad de Puerto Rico.
- Otheguy, Ricardo y Ana Cecilia Zentella. 2007. Apuntes preliminares sobre el contacto lingüístico y dialectal en el uso pronominal del español en Nueva York. En Richard Cameron y Kim Powski (eds.), *Spanish in contact: Policy, social and linguistic inquiries*, 275-295. Amsterdam/Philadelphia: John Benjamins.
- Paredes Silva, Vera Lúcia. 1993. Subject omission and functional compensation: Evidence from written Brazilian Portuguese. *Language Variation and Change* 5. 35-49.
- Paredes Silva, Vera Lúcia. 2003. Motivações funcionais no uso do sujeito pronominal: uma análise em tempo real. En Maria da Conceição de Paiva y Maria Eugênia Lamoglia Duarte (Organização), *Mudança lingüística em tempo real*, 97-114. Rio de Janeiro: Contra Capa.
- Pelly, María Elena. 1984. *La primera persona del plural en Cuba*. La Habana: Editorial de Ciencias Sociales.

- Quijada Bringtown, Milagros. 2006. *Ø cantaba... ¿él, ella o yo? Sujetos de tercera persona en el habla de Caracas*. Trabajo de grado, Maestría en Lingüística, Universidad Central de Venezuela, Caracas.
- Real Academia Española (RAE). 2009. *Nueva gramática de la lengua española*. Madrid: Espasa-Calpe.
- Romaine, Suzanne. 1981. The status of variable rules in sociolinguistic theory. *Journal of Linguistic* 17. 93-119.
- Romaine, Suzanne. 1984. On the problem of syntactic variation and pragmatic meaning in sociolinguistic theory. *Folia lingüística* 18. 409-437.
- Sankoff, David. 1978. *Linguistic variation: Models and methods*. New York: Academia Press.
- Sankoff, David. 1988. Sociolinguistics and syntactic variation. En Fritz Newmeyer (ed.), *Language: The sociocultural context*, 140-161. New York: Cambridge University Press.
- Sankoff, David, Sali Tagliamonte, Eric Smith. 2005. Goldvarb X: A variable rule application for Macintosh and Windows. Department of Linguistics, University of Toronto. Disponible en:  
[http://individual.utoronto.ca/tagliamonte/Goldvarb/GV\\_index.htm](http://individual.utoronto.ca/tagliamonte/Goldvarb/GV_index.htm) [consultada 22/09/09]
- Samper Padilla, José Antonio, Clara Eugenia Hernández Cabrera y Magnolia Troya Déniz. 2006. Presencia/ausencia del sujeto pronominal de primera persona singular en la norma culta de España. En Concepción Company Company (ed.), *El español en América. Diatopía, diacronía e historiografía*, 87-109. México: Universidad Nacional Autónoma de México.
- Sedano, Mercedes. En preparación. Manual de gramática del español. Con especial referencia al español de Venezuela. Caracas: Universidad Central de Venezuela.
- Seco, Manuel. 2001. *Gramática esencial del español*. Madrid: Espasa Calpe.
- Silva-Corvalán, Carmen. 1977. *A discourse study of word order in the Spanish spoken by Mexican-Americans in West Los Angeles*. Tesis de Master of Arts inédita, University of California, Los Angeles.
- Silva-Corvalán, Carmen. 1982. Subject expression and placement in Mexican-American Spanish. En Jon Amastae y Lucía Elías-Olivares (eds.), *Spanish in the United States: Sociolinguistic aspects*, 93-120. New York: Cambridge University Press.
- Silva-Corvalán, Carmen. 2001. *Sociolingüística y pragmática del español*. Washington, DC: Georgetown University Press.

- Silva-Corvalán, Carmen. 2003. Otra mirada a la expresión del sujeto como variable sintáctica. En Francisco Moreno Fernández, Francisco Gimeno Menéndez, José Antonio Samper, María Luz Gutiérrez Araus, María Vaquero y César Hernández (eds.), *Lengua, variación y contexto. Homenaje a Humberto López Morales*, Vol. II, 849-860. Madrid: Arco/Libros.
- Stewart, Miranda. 2001. Los 'hedges' y el uso del 'yo' en la interacción cara-a-cara. *Mestizaje lingüístico*. Actas del I Congreso de la Asociación de Lingüística y Filología de América Latina (ALFAL) región noroeste de Europa. Universidad de Groningen (Holanda).
- <http://elies.rediris.es/elies13/stewart.htm> [consulta 30/08/09]
- Tagliamonte, Sali A. 2006. *Analysing sociolinguistic variation*. Cambridge: Cambridge University Press.
- Travis, Catherine E. 2007. Genre effects on subject expression in Spanish: Priming in narrative and conversation. *Language Variation and Change* 19, 2.101-135.
- Torre-Cacoullós, Rena y Catherine E. Travis. 2008. Variable *yo* expression in New Mexico: English influence? En Susana Rivera-Mills y Daniel Villa (eds.), *Spanish of the Southwest: A Language in transition*, 1-16. Frankfurt: Iberoamericana /Vervuert.
- Universidad Central de Venezuela. 1979. *El habla culta de Caracas. Materiales para su estudio*. Caracas: Universidad Central de Venezuela.
- Universidad Nacional Autónoma de México. 1971. *El habla de la ciudad de México. Materiales para su estudio*. México: Universidad Nacional Autónoma de México.
- Universidad Nacional de Buenos Aires. 1987. *El habla culta de la ciudad de Buenos Aires. Materiales para su estudio*. Buenos Aires: Universidad Nacional de Buenos Aires.
- Vázquez Rozas, Victoria. 2004. Algunas reflexiones sobre el cálculo de la distancia referencial. D.E.L.T.A. 20,1. 27-47.

# Anexos

## Anexo 1

Tipos de procesos reconocidos en ADESSE (lista provisional)<sup>25</sup>

<b>1 Mental</b>	Verbos-ejemplo
<b>11 Sensación</b>	<i>Sentir</i>
<b>111 Volición</b>	<i>gustar, temer</i>
<b>12 Percepción</b>	<i>querer, desear</i>
<b>13 Cognición</b>	<i>ver, escuchar</i>
<b>131 Conocimiento</b>	<i>pensar, entender</i>
<b>132 Creencia</b>	<i>saber1, recordar</i>
<b>14 Elección</b>	<i>creer, opinar</i>
<b>2 Relacional</b>	<i>decidir, elegir</i>
<b>21 Atribución</b>	<i>ser, parecer</i>
<b>211 Relación</b>	<i>equivaler, depender</i>
<b>212 Propiedad</b>	<i>estar2, quedar2</i>
<b>2121 Medida</b>	<i>medir, costar</i>
<b>2122 Apariencia</b>	<i>oler2, saber2</i>
<b>213 Denominación</b>	<i>nombrar2, llamar2</i>
<b>22 Posesión</b>	<i>tener, necesitar</i>
<b>221 Adquisición</b>	<i>conseguir, recibir</i>
<b>222 Transferencia</b>	<i>dar, pagar</i>
<b>3 Material</b>	
<b>31 Espacio</b>	<i>mover, girar</i>
<b>311 Desplazamiento</b>	<i>ir, andar, llevar</i>
<b>312 Localización</b>	<i>poner, pender</i>
<b>313 Postura-posición</b>	<i>sentar, agachar</i>
<b>314 Orientación</b>	<i>señalar, apuntar</i>
<b>315 Manera movimiento</b>	<i>aletear, mecer</i>
<b>316 Unión</b>	<i>juntar, añadir</i>
<b>32 Cambio</b>	<i>pintar, cocinar</i>
<b>321 Creación</b>	<i>crear, fabricar</i>
<b>322 Modificación</b>	<i>romper, secar</i>
<b>3221 Cuidado corporal</b>	<i>lavar, cepillar</i>
<b>323 Destrucción</b>	<i>destruir, eliminar</i>
<b>3231 Consumo</b>	<i>comer, gastar</i>
<b>33 Otros hechos</b>	
<b>331 Meteorología</b>	<i>llover, atardecer</i>
<b>332 Emisión</b>	<i>brillar, emitir</i>
<b>333 Contacto</b>	<i>golpear, tocar</i>

<sup>25</sup> Consulta 16/04/2009

334 <b>Control</b>	<i>agarrar, coger</i>
335 <b>Uso</b>	<i>emplear, usar</i>
336 <b>Sustitución</b>	<i>Reemplazar, suplir</i>
337 <b>Actividad</b>	<i>trabajar, actuar</i>
338 <b>Competición</b>	<i>luchar, vencer</i>
34 <b>Comportamiento</b>	<i>portar2, comportar2</i>
342 <b>Fisiología</b>	<i>llorar, orinar</i>
343 <b>Relaciones sociales</b>	<i>acoger, casar2</i>
<b>4 Verbal</b>	
41 <b>Comunicación</b>	<i>decir, hablar</i>
411 <b>Mandato</b>	<i>ordenar1, pedir</i>
42 <b>Valoración</b>	<i>criticar, felicitar</i>
<b>5 Existencial</b>	
51 <b>Existencia</b>	<i>haber, ocurrir</i>
52 <b>Fase-tiempo</b>	
521 <b>Fase</b>	<i>empezar, acabar</i>
522 <b>Tiempo</b>	<i>durar, retrasar</i>
53 <b>Vida</b>	<i>vivir, matar</i>
<b>6 Modulación</b>	
60 <b>Causación</b>	<i>hacer2</i>
601 <b>Inducción</b>	<i>incitar, ayudar</i>
602 <b>Permiso</b>	<i>permitir, dejar2</i>
603 <b>Obligación</b>	<i>obligar, forzar1</i>
61 <b>Disposición</b>	<i>tratar, atreverse</i>
62 <b>Aceptación</b>	<i>aceptar, rechazar</i>
64 <b>Verbos de apoyo</b>	<i>hacer-0, dar-0</i>

## Anexo 2

### Sobre el $\chi^2$

El test del  $\chi^2$  hace una comparación global de grupos de frecuencias para determinar si la frecuencia observada de un fenómeno es significativamente igual a la frecuencia teórica prevista, o sí, por el contrario, estas dos frecuencias acusan una diferencia significativa para, por ejemplo, un nivel de significación del 0,05.



### Anexo 3

#### Sobre GoldVarb

Los programas VARBRUL: estos programas estadísticos se han venido utilizando sobre todo desde los años 70 del siglo XX y fueron diseñados por David Sankoff especialmente para el análisis de la variación (socio)lingüística. Precisamente, el nombre que los designa, VARBRUL, deriva de la expresión inglesa *Variable rule* ‘regla variable’. La función principal de dichos programas es la de llevar a cabo un análisis de regresión de variables cualitativas; los procedimientos de los que se dispone para llevar a cabo este análisis son: el método directo (*one-level*) y el análisis por pasos (*step-up and down*).

El análisis estadístico de primer nivel, o método directo, proporciona la siguiente información: i) el número de interacciones para hallar la verosimilitud; ii) el *input* que se obtiene al aplicar el modelo logístico; iii) el efecto de cada uno de los factores de cada grupo sobre el factor dependiente (peso); iv) la división del número de aplicaciones de cada factor entre el número de casos posibles; v) la probabilidad resultante de combinar el peso y el *input*; vi) los factores utilizados en la creación de cada celda y el número total de datos por celda; vii) el valor esperado por celda; proporción de error por celda; viii) el ji-cuadrado ( $\chi^2$ ) total; ix) el ji-cuadrado ( $\chi^2$ ) dividido por número de celdas; x) el logaritmo de la función de verosimilitud; y xi) el diagrama de dispersión, en el que se compara la proporción de casos de aplicación del factor estudiado en cada una de las celdas con las probabilidades calculadas en el análisis.

En el procedimiento por pasos el programa ofrece información sobre frecuencias absolutas y relativas y sobre la estructura y número de celdas. En el nivel 0 del análisis de subida el programa ofrece la probabilidad global de aparición de la variante en estudio en el corpus analizado; de allí se pasa al siguiente nivel de análisis, el nivel 1, en el que se obtienen las probabilidades de los factores de cada variable, el grado de significación y el logaritmo de verosimilitud. En el nivel 2 el programa trata de comprobar si el modelo tiene una mejor capacidad de predicción que en el nivel anterior; en este nivel el programa “va añadiendo nuevos factores al análisis hasta conseguir el mejor ajuste de todas las combinaciones” (280), es decir, el que tenga diferencias significativas después de incorporar el mayor número de factores. En el nivel 3 se analiza el efecto de la combinación de todos los factores en conjunto; en este nivel el logaritmo de verosimilitud alcanza su valor más próximo a 0. Cuando finaliza el análisis de subida, se ejecuta el análisis de bajada, *stepdown*, que comienza con las combinaciones más complejas y se detiene en el nivel 2 del análisis. Al terminar el análisis de subida y bajada el programa indica qué factores son significativos y cuáles no lo son.

## Anexo 4

## Todos los verbos del corpus

CANTIDAD	PRONOMBRE EXPRESO	CLASE SEMÁNTICA	FORMA LÉXICA
1	-	312	abandonar
3	1	322	abrir
1	-	311	acabar de venir
1	1	62	aceptar
1	-	311	acercar
1	-	322	acomodar
1	-	311	acompañar a salir
2	2	41	aconsejar
51	22	131	acordar
5	1	313	acostar
1	1	601	acostumbrar
1	-	322	activar
1	-	211	adaptar
1	-	311	adelantar
2	-	11	admirar
6	2	11	adorar
1	1	313	agachar
11	3	334	agarrar
3	2	41	agradecer
1	-	11	aguantar
2	1	221	alcanzar
2	1	11	amar
1	1	131	analizar
1	-	21	andar
4	3	311	andar
1	-	316	añadir
1	-	222	aportar
2	-	11	apreciar
1	-	12	apreciar
9	-	131	aprender

CANTIDAD	PRONOMBRE EXPRESO	CLASE SEMÁNTICA	FORMA LÉXICA
1	-	12	aprender a descubrir
1	-	131	aprender a estudiar
2	-	41	aprender a hablar
1	-	12	aprender a leer
2	-	335	aprovechar
1	-	343	aprovechar de visitar
1	-	321	armar
4	1	311	arrancar
1	1	211	asociar
1	1	131	asumir
1	-	11	asustar
1	-	338	atacar
1	-	61	aventurar
1	-	131	averiguar
1	-	41	avisar
3	3	601	ayudar
1	-	315	bailar
2	1	337	bailar
7	6	311	bajar
1	-	12	bajar a ver
1	1	322	barrer
1	-	211	basar
1	1	343	bautizar
2	-	311	brincar
8	1	12	buscar
1	-	312	caber
1	-	212	caer
2	-	311	caer
1	-	41	caer a gritos
1	-	322	calar
1	1	213	calificar
1	-	311	cambiar
3	-	336	cambiar
1	-	11	cansar

CANTIDAD	PRONOMBRE EXPRESO	CLASE SEMÁNTICA	FORMA LÉXICA
4	2	41	cantar
11	3	343	casar
3	-	343	castigar
1	-	334	cazar
1	1	64	celebrar
1	-	3231	cenar
1	-	322	cerrar
1	-	333	chocar
1	-	41	codificar
2	-	334	coger
1	-	51	coincidir
1	1	337	colaborar
1	-	312	colocar
1	-	316	combinar
2	-	41	comentar
10	3	521	comenzar
1	1	131	comenzar a estudiar
2	1	64	comenzar a hacer
1	-	53	comenzar a vivir
21	7	3231	comer
1	-	211	comparar
9	1	22	compartir
11	4	221	comprar
1	1	41	comunicar
1	-	131	confundir
46	14	131	conocer
8	1	221	conseguir
1	-	602	consentir
6	1	22	conservar
8	6	13	considerar
1	1	41	contactar
8	1	41	contar
2	-	521	continuar
2	-	343	contratar

CANTIDAD	PRONOMBRE EXPRESO	CLASE SEMÁNTICA	FORMA LÉXICA
1	-	312	converger
3	1	41	conversar
2	-	343	convivir
2	2	322	corregir
1	1	311	correr
2	-	61	correr el riesgo
9	3	322	crecer
137	72	132	creer
4	2	343	criar
2	1	311	cruzar
3	1	12	cuidar
3	-	521	cumplir
18	9	222	dar
1	-	333	dar un beso
7	1	64	darse cuenta
2	-	211	deber
2	-	222	deber (deuda)
1	-	311	deber asistir
1	-	41	deber decir
1	-	521	deber empezar
1	1	321	deber escribir
1	-	131	deber formar
1	-	521	deber terminar
1	1	11	decepcionar
1	-	14	decidir
2	-	601	decidir ayudar
1	-	311	decidir entrar
1	-	337	decidir jugar
141	69	41	decir
1	1	213	decir (llamarlos)
2	1	222	dedicar
1	-	337	dedicar
8	1	312	dejar
1	-	521	dejar

CANTIDAD	PRONOMBRE EXPRESO	CLASE SEMÁNTICA	FORMA LÉXICA
1	-	337	dejar de jugar
1	-	521	dejar de subir
3	1	22	dejar de tener
1	-	322	dejar desordenado
1	-	522	demorar
1	-	521	desarrollar
2	-	311	desbarrancar
1	1	322	descuartizar
4	1	12	descubrir
1	1	111	desear
2	1	342	desmayar
1	-	3221	desnudar
1	-	316	despegar
1	-	342	despertarse
1	-	51	destacar
1	-	4	discutir
2	-	321	diseñar
7	-	11	disfrutar
2	-	343	divorciar
1	1	213	doblar (subtitular)
3	2	342	dormir
1	-	64	echar
1	-	131	educar
1	-	343	egresar (graduar)
1	-	332	emitir
1	-	11	emocionar
1	1	343	empatar (noviazgo)
19	1	521	empezar
1	1	41	empezar a cantar
1	-	334	empezar a coger
2	-	338	empezar a competir
1	-	12	empezar a descubrir
1	-	131	empezar a estudiar
1	-	221	empezar a ganar

CANTIDAD	PRONOMBRE EXPRESO	CLASE SEMÁNTICA	FORMA LÉXICA
1	-	41	empezar a hablar
3	-	64	empezar a hacer
1	-	321	empezar a hacer
1	-	521	empezar a hacer
1	-	11	empezar a joder
1	-	312	empezar a montar
1	-	337	empezar a participar
1	1	13	empezar a pensar
1	-	212	empezar a poner
1	-	342	empezar a respirar
4	2	337	empezar a trabajar
1	-	335	empezar a usar
1	-	311	empezar a zumar (lanzar)
2	1	21	empezar siendo
3	-	11	enamorar
2	1	337	encargar
1	-	312	encerrar
5	-	312	encontrar
2	1	338	enfrentar
1	-	322	enredar
1	1	131	enseñar
11	6	13	entender
3	1	131	enterar
23	9	311	entrar
2	1	131	entrar a estudiar
1	1	222	entregar
2	1	322	entrenar
1	1	311	escapar
1	1	14	escoger
3	1	312	esconder
1	1	321	escribir
4	1	12	escuchar
3	-	111	esperar
3	3	521	esperar

CANTIDAD	PRONOMBRE EXPRESO	CLASE SEMÁNTICA	FORMA LÉXICA
1	-	343	esperar graduar
49	15	212	estar
97	18	312	estar
1	-	333	estar accidentado
1	-	131	estar acordando
4	3	601	estar acostumbrado
1	1	11	estar aguantando
1	1	314	estar apuntando
1	1	522	estar apurado
1	1	322	estar arreglando
1	-	41	estar asesorado
1	-	601	estar ayudando
1	-	322	estar azarado (estresada, complicada)
1	1	337	estar bailando
2	-	343	estar casado
2	-	337	estar colaborando
2	-	41	estar comentando
1	-	3231	estar comiendo
1	1	312	estar como tratando de estar
1	-	22	estar compartiendo
2	-	131	estar confundida
1	-	321	estar construyendo
1	-	41	estar contando
1	-	337	estar contando
1	-	132	estar convencido
3	1	222	estar dando
4	1	41	estar diciendo
1	1	337	estar encargando
1	1	12	estar enfocado
1	1	131	estar enseñando
1	-	322	estar entonado (emborrachar)
3	1	521	estar esperando
6	1	131	estar estudiando
1	-	11	estar fastidiado



CANTIDAD	PRONOMBRE EXPRESO	CLASE SEMÁNTICA	FORMA LÉXICA
7	-	41	estar hablando
5	-	64	estar haciendo
1	-	337	estar jugando
1	-	311	estar llegando
1	-	337	estar llevando
1	1	61	estar metido
4	-	312	estar metido
1	-	53	estar muriendo
1	-	41	estar nombrando
1	-	337	estar ocupado
2	-	212	estar pendiente
1	1	11	estar preocupado
1	-	322	estar rascados (emborrachar)
3	-	337	estar realizando
1	-	11	estar relajado
1	-	316	estar reunido
1	-	337	estar sacando (preparar)
1	-	311	estar saliendo
1	1	521	estar terminando
8	2	337	estar trabajando
1	-	311	estar viajando
2	1	21	estar viendo
6	-	53	estar viviendo
2	-	522	estar viviendo
1	-	311	estar yendo
1	-	342	estornudar
57	21	131	estudiar
1	-	61	evitar
1	-	42	exagerar
3	1	601	familiarizar
1	1	11	fastidiar
1	-	131	formar
2	-	321	formar
1	-	322	fracturar

CANTIDAD	PRONOMBRE EXPRESO	CLASE SEMÁNTICA	FORMA LÉXICA
1	-	343	frecuentar
1	-	322	frenar
6	-	3231	fumar
2	1	221	ganar
1	1	3231	gastar
19	6	343	graduar
1	1	64	guardar
1	-	14	haber seleccionado
15	2	41	hablar
1	-	60	hacer
49	4	64	hacer
3	-	212	hacer
25	4	321	hacer
4	-	132	imaginar
3	1	321	imaginar
1	-	343	inaugurar
1	-	316	incluir
1	-	132	inculcar
3	-	337	incursionar
1	-	316	ingresar
1	1	316	inscribir
2	1	41	insistir
1	1	61	intentar
1	-	321	inventar
2	1	131	investigar
1	-	343	invitar
3	2	60	involucrar
3	3	212	ir
130	30	311	ir
1	1	334	ir a agarrar
2	2	32	ir a anotar
1	-	322	ir a cambiar
2	1	41	ir a cantar
1	1	334	ir a cazar

CANTIDAD	PRONOMBRE EXPRESO	CLASE SEMÁNTICA	FORMA LÉXICA
1	-	64	ir a celebrar
1	-	521	ir a comenzar
1	-	3231	ir a comer
1	1	41	ir a confesar
4	-	41	ir a contar
1	-	521	ir a cumplir
1	-	12	ir a curiosear
2	2	222	ir a dar
1	-	333	ir a dar
6	-	41	ir a decir
1	-	521	ir a dejar
1	-	342	ir a dormir
1	-	311	ir a echar
2	1	312	ir a estar
7	2	131	ir a estudiar
1	1	41	ir a explicar
3	2	32	ir a grabar
2	-	41	ir a hablar
5	2	64	ir a hacer
1	-	337	ir a ir a rumbear
2	-	343	ir a jubilar
1	-	337	ir a jugar
1	-	311	ir a llegar
1	-	32	ir a llevar a laquear
1	1	312	ir a meter
2	-	53	ir a morir
1	-	311	ir a pasar
1	1	13	ir a pensar
2	-	312	ir a poner
1	-	338	ir a poner a discutir
1	1	338	ir a poner a pelear
1	1	41	ir a preguntar
2	1	12	ir a probar
1	-	312	ir a quedar

CANTIDAD	PRONOMBRE EXPRESO	CLASE SEMÁNTICA	FORMA LÉXICA
1	1	221	ir a quitar
1	-	131	ir a recordar
1	-	321	ir a redactar
5	1	41	ir a referir
1	-	342	ir a reír
1	1	41	ir a rezar
1	-	311	ir a salir
3	1	21	ir a ser
2	1	22	ir a tener
1	-	32	ir a tener que anotar
1	-	13	ir a tomar
1	-	34	ir a tratar de afrontar
1	-	322	ir a vender
1	-	311	ir a venir
2	-	12	ir a ver
3	-	343	ir a visitar
6	1	53	ir a vivir
1	-	316	ir acoplado
1	-	313	ir agachado
2	-	311	ir caminando
1	1	311	ir corriendo
1	-	64	ir dando cuenta
1	-	337	ir dedicando
2	-	12	ir descubriendo
5	-	13	ir entendiendo
1	-	311	ir escalando
2	-	312	ir quedando
1	-	316	ir separando
1	-	333	ir tocando
1	-	333	jamaquear (batir)
18	4	337	jugar
3	-	41	jurar
1	-	42	justificar
2	-	311	lanzar

CANTIDAD	PRONOMBRE EXPRESO	CLASE SEMÁNTICA	FORMA LÉXICA
12	6	12	leer
3	1	313	levantar
1	1	322	limpiar
8	3	41	llamar
5	3	213	llamar
46	13	311	llegar
1	1	131	llegar a aprender
1	1	51	llegar a pasar
1	-	12	llegar oliendo
1	1	322	llenar
5	1	211	llevar
10	4	311	llevar
10	-	343	llevar
1	-	342	llorar
4	-	221	lograr
1	-	11	lograr disfrutar
1	-	312	lograr poner
1	-	338	luchar
4	1	335	manejar
2	-	12	manifestar
1	-	333	manosear
3	1	521	mantener
1	-	53	matar
3	-	41	mencionar
1	1	11	menospreciar
1	1	61	meter
12	4	311	meter
9	2	312	meter
2	1	131	meter a estudiar
2	-	337	meter a trabajar
2	2	11	molestar
1	-	337	montar
2	-	333	morder
1	-	53	morir

CANTIDAD	PRONOMBRE EXPRESO	CLASE SEMÁNTICA	FORMA LÉXICA
2	1	31	mover
17	2	311	mudar
15	8	53	nacer
10	3	22	necesitar
1	-	334	necesitar agarrar
1	-	64	necesitar hacer
1	-	41	nombrar
2	2	12	notar
1	-	337	obedecer
1	1	12	observar
3	-	221	obtener
1	1	131	obviar (ignorar)
2	1	12	oír
1	-	131	olvidar
1	-	41	opinar
1	-	14	optar
2	-	213	ordenar
1	-	322	organizar
9	1	222	pagar
5	3	12	parar (prestar atención, atender)
2	-	312	pararse
1	-	313	pararse
3	2	342	parir
1	-	322	partir
13	1	51	pasar
5	-	64	pasar
13	1	311	pasar
5	1	521	pasar
8	1	522	pasar
1	-	64	pasar haciendo
1	-	41	pasar llamando
1	-	311	patinar
2	-	411	pedir
5	1	338	pelear

CANTIDAD	PRONOMBRE EXPRESO	CLASE SEMÁNTICA	FORMA LÉXICA
18	9	13	pensar
7	1	22	perder
1	-	335	perder
1	1	602	permitir
1	-	311	perseguir
1	1	22	pertenecer
1	-	32	pintar
1	1	41	plantear
10	3	21	poder
1	1	221	poder adquirir
1	-	322	poder alargar
1	-	32	poder anotar
1	-	337	poder bailar
1	1	61	poder brindar
2	2	41	poder cantar
1	-	334	poder cargar (sostener)
2	1	132	poder confiar
2	-	41	poder contar
1	-	334	poder controlar
2	-	343	poder convivir
1	-	222	poder dar
4	1	41	poder decir
1	1	312	poder dejar
1	-	521	poder dejar
1	-	311	poder echar
1	1	311	poder entrar
1	1	312	poder estar
1	1	64	poder estar haciendo
1	1	131	poder estudiar
1	-	41	poder hablar
3	1	64	poder hacer
2	1	321	poder hacer
1	-	132	poder imaginar
1	-	311	poder ir

CANTIDAD	PRONOMBRE EXPRESO	CLASE SEMÁNTICA.	FORMA LÉXICA
3	-	311	poder llegar
1	-	337	poder llevar a cabo (realizar)
1	-	521	poder mantener
1	1	53	poder morir
1	-	12	poder oír
1	-	131	poder olvidar
1	1	311	poder pasar
1	-	333	poder pisar
1	-	312	poder quedar
2	-	41	poder quejar
1	1	131	poder recordar
1	-	521	poder seguir
3	2	313	poder sentar
2	1	334	poder soportar
2	-	22	poder tener
1	-	337	poder trabajar
4	3	12	poder ver
1	-	53	poder vivir
5	-	32	poner
4	3	41	poner
3	-	212	poner
6	2	312	poner
2	1	337	poner
1	1	603	poner
1	-	321	poner a dibujar
1	-	41	poner a hablar
1	-	342	poner a llorar
1	1	13	poner a pensar
1	-	337	poner a trabajar
1	-	12	poner a ver
1	-	603	poner en práctica
2	-	34	portar
3	-	64	practicar
1	1	14	preferir



CANTIDAD	PRONOMBRE EXPRESO	CLASE SEMÁNTICA	FORMA LÉXICA
8	5	41	preguntar
1	-	32	prender
4	1	12	presentar
1	-	343	presidir
1	-	335	probar
1	-	321	programar
1	-	338	proteger
1	1	211	provenir
1	-	132	proyectar
1	-	22	quedar
3	-	212	quedar
23	7	312	quedar
1	-	131	quedar a estudiar
1	-	62	quedar admitido
1	-	337	quedar trabajando
1	1	12	quedar viendo
1	-	41	quejar
10	1	11	querer
22	12	111	querer
1	-	334	querer agarrar
1	1	337	querer bailar
1	-	12	querer buscar
4	2	211	querer decir
1	1	311	querer desplazar
1	1	312	querer estar
8	3	131	querer estudiar
2	1	212	querer hacer
1	-	321	querer hacer
1	1	34	querer imitar
5	4	311	querer ir
1	-	213	querer llamar
1	-	311	querer llegar
1	-	312	querer meter
1	1	131	querer que estudiaran

CANTIDAD	PRONOMBRE EXPRESO	CLASE SEMÁNTICA	FORMA LÉXICA
1	-	312	querer quedar
1	-	41	querer referir
1	-	221	querer sacar (obtener)
1	1	311	querer seguir
2	2	521	querer seguir
6	2	21	querer ser
3	2	22	querer tener
1	1	343	querer terminar de graduar
2	1	337	querer trabajar
4	1	12	querer ver
1	1	11	querer vivir
4	1	312	quitar
1	-	131	racionalizar
1	-	41	ratificar
1	-	322	reactivar
3	-	337	realizar
2	1	343	recibir
1	1	41	recitar
1	1	41	recomendar
4	2	12	reconocer
2	-	62	reconocer
59	22	131	recordar
1	-	311	recorrer
1	-	322	recubrir
1	-	221	recuperar
11	3	311	regresar
1	-	311	regresar caminando
3	2	342	reír
1	-	322	remodelar (decorar)
1	-	41	repetir
1	-	11	respetar
1	1	334	retener
1	-	311	retirarse
2	-	521	retomar

CANTIDAD	PRONOMBRE EXPRESO	CLASE SEMÁNTICA	FORMA LÉXICA
2	-	316	reunir
1	-	12	revisar
1	-	322	romper
175	42	131	saber
1	-	322	saber alimentar
1	-	337	saber bailar
1	-	3231	saber comer
1	-	34	saber comportar
1	-	41	saber hablar
1	-	132	saber que esperar
5	2	311	sacar
5	2	221	sacar (obtener)
1	-	335	sacar (usar)
1	1	337	sacar cuentas
3	-	212	salir
30	7	311	salir
1	-	3231	salir a comer
1	-	337	salir a rumbear (bailar)
1	1	337	salir a trabajar
2	-	311	salir corriendo
4	-	521	seguir
1	-	41	seguir contando
1	-	131	seguir estudiando
1	-	521	seguir manteniendo
1	-	222	seguir pagando
3	-	21	seguir siendo
3	-	337	seguir trabajando
1	1	12	seguir viendo
4	3	313	sentar
1	1	3231	sentar a tomar
16	6	12	sentir
1	-	12	sentir atraído
236	97	21	ser
1	-	21	ser (estar)

CANTIDAD	PRONOMBRE EXPRESO	CLASE SEMÁNTICA	FORMA LÉXICA
1	1	11	ser apegado
1	-	343	ser bautizado
1	-	42	ser descalificado
1	1	212	servir
1	-	53	sobrevivir
1	-	601	soler
2	2	334	soportar
1	-	12	sorprenderse
1	1	311	subir
1	-	311	subir volando
2	-	11	sufrir
2	2	211	superar
6	-	132	suponer
2	1	322	tapar
2	-	522	tardar
1	-	34	tender
1	1	41	tender a decir
391	96	22	tener
1	-	22	tener comprendido
1	-	222	tener dando
1	-	13	tener entendido
1	1	211	tener equiparado
1	-	337	tener previsto (prever)
1	-	334	tener que abandonar
1	-	34	tener que afrontar
1	-	311	tener que andar
1	-	32	tener que anotar
1	-	131	tener que aprender
1	-	34	tener que centrarse
1	1	12	tener que chequear (revisar)
2	1	41	tener que contestar
1	-	322	tener que cortar
1	1	12	tener que cuidar
1	-	521	tener que empezar

CANTIDAD	PRONOMBRE EXPRESO	CLASE SEMÁNTICA	FORMA LÉXICA
1	-	131	tener que enseñar
1	-	22	tener que estar
3	2	212	tener que estar
1	-	131	tener que estudiar
1	-	41	tener que hablar
2	-	64	tener que hacer
1	1	212	tener que hacer
1	-	321	tener que hacer
1	1	337	tener que intervenir
4	2	311	tener que ir
1	-	41	tener que llamar
1	-	22	tener que llevar
1	-	521	tener que mantener
1	-	61	tener que meter
1	-	51	tener que pasar
1	-	311	tener que pasar
1	-	522	tener que pasar
1	-	42	tener que perdonar
1	-	603	tener que poner
1	-	312	tener que quedar
2	-	11	tener que relajar
1	-	521	tener que retomar
1	1	131	tener que saber
1	-	61	tener que superar (abandonar)
3	-	22	tener que tener
1	-	311	tener que venir
2	-	311	tener que viajar
1	-	53	tener que vivir
1	-	53	tener viviendo
18	2	521	terminar
1	-	313	terminar acostándose
1	-	337	terminar de trabajar
1	-	11	tirar
1	-	311	tirar

CANTIDAD	PRONOMBRE EXPRESO	CLASE SEMÁNTICA	FORMA LÉXICA
2	-	332	tocar
2	1	221	tomar
5	2	334	tomar
3	-	3231	tomar
1	1	12	topar
19	9	337	trabajar
6	1	311	traer
2	-	3231	tragar
1	-	322	trancar (cerrar)
1	-	522	transcurrir
1	1	322	transformar
1	-	311	trasladar
13	2	61	tratar
1	-	22	tratar de conservar
1	1	335	tratar de manejar
1	-	316	unir
1	-	335	usar
1	-	335	utilizar
1	-	2121	valorar
19	5	311	venir
1	-	12	venir a ver
1	-	311	venir corriendo
1	-	41	venir diciendo
1	-	221	venir ganando
1	1	343	venir graduada
1	-	12	venir oyendo
1	-	521	venir terminando
78	31	12	ver
1	-	343	ver que atiende
8	-	311	viajar
1	1	343	visitar
3	2	11	vivir
53	29	53	vivir
1	-	321	vivir inventando

CANTIDAD	PRONOMBRE EXPRESO	CLASE SEMÁNTICA	FORMA LÉXICA
2	-	11	vivir pendiente
1	-	314	voltear
1	-	212	volver
5	-	311	volver
2	-	41	volver a decir
2	-	12	volver a encontrar
1	-	41	volver a insistir
1	-	311	volver a ir
2	1	12	volver a ver
2	1	14	votar