

La influencia de la persona gramatical sobre la expresión del pronombre sujeto en el español del sur de Arizona

Hope Anderson
The University of Arizona

1.0 Introducción

La expresión del pronombre sujeto en español es un tema con mucha posibilidad para investigar y analizar. Como el español es un idioma con la característica *pro-drop*, el pronombre sujeto se utiliza con frecuencia en ambas de sus variantes—expresa y nula (el pronombre sujeto no aparece)—pero con diferentes matices sintácticos, pragmáticos y discursivos.

Aunque se han encontrado en la literatura muchos grupos de factores que promueven la expresión del sujeto, como el paralelismo y la conexión discursiva (Flores-Ferrán, 2004, 2010; Torres Cacoullós y Travis, 2010a, 2010b), un grupo de factores que se ve importante es el de la persona y el número (Bayley y Pease-Alvarez, 1997; Flores-Ferrán, 2004). En estudios anteriores, se han reportado promedios de expresión de sujeto diferentes (en función de la comunidad), pero en general, hay más expresión con el *yo* y menos expresión con el *nosotros* y el *ellos* (Orozco y Nunez, 2012; Carvalho y Child, 2011; Flores-Ferrán, 2004; Bayley & Pease-Alvarez, 1997). No sólo son las frecuencias que se diferencian; de hecho, dado que no hay manera léxica de expresar la primera y la segunda persona, es evidente que habrá diferencias en la expresión pronominal con estas personas, en comparación con la tercera persona. Estas distinciones no sólo existen en el español; son diferencias universales que aparecen en otros idiomas también (Dumont, 2006). En español, las diferencias son tan claras que algunos estudios han intentado probar si se distinguen también los patrones que se asocian con cada persona gramatical.

El estudio actual supone contribuir a esta línea de trabajo, usando datos del español del sur de Arizona. En este proyecto, se usó regresión de subida y bajada en Goldvarb X para investigar los grupos de factores que correlacionan con la expresión del pronombre (expreso o nulo). Se comparó la expresión de cada persona gramatical y de todas las personas juntas para investigar hasta qué punto se distinguen. El estudio sugiere que aunque todavía sería beneficioso comparar los resultados de cada persona gramatical dentro de un estudio, como lo hacen Otheguy et al. (2007), puede que los datos de cada persona gramatical no sean comparables, como en la perspectiva de Torres Cacoullós y Travis (2010a, 2010b) y Dumont (2006).

2.0 La variable

El ejemplo 1, tomado del corpus analizado para este estudio, muestra tanto el pronombre expreso como el pronombre nulo en una serie de tokens (en letra mayúscula) de la primera persona singular (el *yo*).

YO QUIERO AYUDARLE- ayudarle a la gente que tienen esa barrera de- de la idioma. Me gusta, como Ø HAGO un poco de eso en el trabajo, siempre Ø SIENTO como un- no sé, como (...) YO PUEDO, YO LE PUEDO AYUDAR a la gente con esa barrera de la idioma. (CESA005)

En el ejemplo 1, el primer token usa el pronombre expreso, aunque se ve el sujeto en la morfología también. Los siguientes tokens usan el nulo hasta después de una pausa, cuando regresa el pronombre expreso. Sin embargo, otros ejemplos del corpus muestran patrones muy distintos.

3.0 Estudios previos

Existen dos corrientes de pensamiento en cuanto al tema de la persona gramatical en estudios de la expresión del sujeto. Por un lado, Otheguy, Zentella y Livert (2007) estudiaron todas las personas gramaticales juntas, separándolas para análisis adicional sólo cuando se detectó una gran diferencia en los resultados para algunas de las personas. Por otro lado, Torres Cacoullós y Travis (2010a, 2010b) estudiaron sólo una persona gramatical, pensando que cada persona gramatical se comporta de una manera tan distinta de las otras que no se puede unir las en el mismo estudio.

3.1 Juntas

Bayley y Pease-Alvarez (1997) demostraron la importancia de la persona gramatical, ya que la persona y el número formaron la restricción más fuerte en la variación de la expresión del pronombre sujeto en narrativas personales de

niños viviendo en California pero nacidos en, y con padres nacidos en, o México o los Estados Unidos. Entre las personas gramaticales, la primera persona singular usó más el pronombre expreso (37%; peso de factor 0.691 en VARBRUL), seguida por la tercera singular (25%; 0.576), la tercera plural (20%; 0.433) y la primera plural (9%; 0.244). Encontraron también diferencias según el nivel de conexión que tenían los niños con los EEUU, encontrando que los niños con más enlaces a los EEUU usaron más el pronombre nulo que los nacidos en México; un patrón opuesto de lo que detectó Silva-Corvalán (1994).

En una comunidad puertorriqueña de Nueva York, Flores-Ferrán (2004) encontró la persona y el número como un factor muy relevante, con el mayor uso del pronombre expreso en la segunda persona singular no específica (60%), la segunda persona singular específica (53%), la primera persona singular (52%) y la tercera persona singular (48%). Estas frecuencias altas se diferenciaron de los plurales (la primera y la tercera persona), los cuales mostraron mucho menos uso (17% y 22%). Este patrón se pareció a los datos anteriores de San Juan, Puerto Rico, de un grado suficiente que no confirmó una hipótesis de efectos de contacto sobre la expresión. La conexión discursiva (*switch reference*) salió también como un factor importante en el uso de pronombres expresos.

Usando un corpus de entrevistas con 142 habitantes de las comunidades más grandes de hispanohablantes de la ciudad de Nueva York, Otheguy et al. (2007) compararon la expresión del pronombre sujeto de los hablantes de generaciones distintas y de regiones distintas de origen (el Caribe y el continente de América). Encontraron una subida en el uso de pronombres expresos como resultado de contacto con

el inglés, y también de un acercamiento de dialectos distintos en relación a la expresión del pronombre sujeto. Entre los recién llegados, la primera persona singular fue un contexto muy significativo para la expresión del pronombre de los del continente: estos hablantes produjeron casi el doble de pronombres explícitos en contextos de la primera persona singular, mientras esto no tuvo ningún efecto para los del Caribe. La tercera persona singular mostró el mismo efecto entre los recién llegados, con los del continente produciendo más del triple de pronombres expresos en este contexto, pero sin ningún efecto para los del Caribe.

3.2 Separadas: La primera persona

Torres Cacoullós y Travis (2010a) estudiaron la expresión del sujeto “yo” en entrevistas, conversaciones y narrativas de Nuevo México y concluyeron que el patrón de expresión del pronombre fue igual en esta variedad de español y en variedades sin contacto con el inglés. Se detectó más uso del pronombre expreso en discursos que contenían cambios de código, no como resultado de contacto en general, sino por el paralelismo con un sujeto expreso en el discurso anterior, sea ese discurso en inglés o en español, con el inglés promoviendo más expresión, ya que el inglés tiene más uso del pronombre expreso que el español.

Torres Cacoullós y Travis (2010b) compararon la frecuencia y patrones de expresión del pronombre de la primera persona singular entre hablantes de diferentes grados de bilingüismo en Nuevo México. Los resultados sugirieron que ni el contacto ni la presencia de cambios de código afecta estos patrones. Los factores que mejor correspondieron con la expresión pronominal fueron el paralelismo y *switch reference*. El factor de contenido lexical, aunque resultó significativo, fue debido a la

frecuencia de las frases “yo (no) sé” y “yo creo” (dos verbos psicológicos), y la ambigüedad morfológica fue un factor más débil.

3.3 Separadas: La segunda persona

Cameron (1993) empezó con todas personas juntas y encontró los mismos patrones en cuanto a la expresión pronominal en San Juan, Puerto Rico, y en Madrid, España. Sobre todo, salió una interacción entre dos factores importantes: la ambigüedad y *switch reference*. En su enfoque subsiguiente al pronombre *tú* en particular, mostró que, aunque los patrones de la expresión del sujeto del *tú* específico no variaron mucho entre los dos sitios (48% expreso en San Juan y 40% en Madrid), sí se encontró una gran diferencia en la expresión del *tú* no específico (69% expreso en San Juan y 19% en Madrid).

Usando datos de entrevistas con 41 puertorriqueños en la ciudad de Nueva York, Flores-Ferrán (2010) detectó una subida en el uso de pronombres expresos en narrativas de conflicto interpersonal, en comparación con narrativas de otros temas. También se mostraron diferencias notables por persona gramatical: la frecuencia del pronombre expreso por persona fue casi el revés en narrativas de conflicto (*tú* específico > él/ella > usted > ellos/ellas > yo > tú no específico > nosotros > ustedes) y sin conflicto (*ustedes* > nosotros > tú no específico > ellos/ellas > yo > usted > él/ella > tú específico). Resultó que la presencia de conflicto fue un factor tan robusto que los efectos se pudieron observar entre todas las edades del estudio y que pudo competir con *switch reference* como predictor de expresión pronominal.

3.4 Separadas: Tercera persona

Dumont (2006) estudió cómo las NPs completas contrastan con los pronombres

expresos y nulos en la expresión de la tercera persona singular y plural. Usando los datos de cinco entrevistas con hablantes de español en Nuevo México y Colorado, encontró una preferencia para pronombres nulos (58%), con un 22% para la NP completa y un 20% para el pronombre expreso. Los factores que mejor correspondieron con la expresión de la NP completa fueron principalmente la realización de la mención anterior y la distancia de la mención anterior, y en un menor grado la especificidad, la clase semántica del verbo (específicamente verbos de movimiento) y el número (es decir, la singularidad). Aunque el estudio presente no incluye la NP completa, los mismos factores se han asociado con la expresión del pronombre en otros estudios.

3.5 ¿Juntas o separadas?

Para empezar a resolver este debate, Orozco y Nunez (2012) compararon los patrones de expresión del *yo* con los patrones generales de expresión del pronombre sujeto, concluyendo que no había ventaja de separar las personas dentro de un estudio. Pero como encontraron otros factores que se diferenciaron entre las dos comunidades que estudiaron (colombianos en Barranquilla y colombianos inmigrantes en Nueva York), y efectos de edad y sexo también, dejaron abierta la posibilidad de que la persona gramatical sí tenga un efecto en la expresión del pronombre en el habla de otras comunidades.

El proyecto actual incorpora ambas corrientes, usando grupos de factores que correlacionan con la expresión del pronombre para comparar la expresión de cada persona gramatical, especialmente la primera persona singular, con todas las personas juntas. Puede contribuir a la obra de Orozco y Nunez (2012) para formar la base de estudios futuros que se podrían

enfocar específicamente en una persona gramatical o que podrían incluir todas las personas, según lo que sean sus resultados.

4.0 La pregunta de investigación

La pregunta de investigación es: En el español del sur de Arizona, ¿se distribuyen de maneras distintas las diferentes personas gramaticales en cuanto a la expresión del pronombre sujeto (el uso de pronombres sujeto nulos y expresos)?

5.0 Metodología

5.1 Colección de datos

Siete entrevistas sociolingüísticas, las cuales forman la parte inicial del Corpus de Español del Sur de Arizona (CESA), se hicieron por estudiantes de SPAN 584C de la Universidad de Arizona en octubre del 2012. Los siete entrevistados eran hablantes nativos del español o bilingües en inglés y español. Eran habitantes de Tucson o de sus alrededores, originalmente de México o de los Estados Unidos pero que habían vivido en el sur de Arizona por más de dos años.

Cada entrevista se transcribió por el/la entrevistador(a), y los tokens de pronombres sujeto nulos y expresos se codificaron por los mismos a lo largo de 12 grupos de factores: sexo del hablante, prácticas lingüísticas del hablante, ambigüedad morfológica, persona/número, tiempo/aspecto/modo, conexión discursiva (si el referente es igual que el anterior o no), tipo de cláusula, reflexividad (si el verbo es reflexivo o no), paralelismo (expresión del referente anterior), contenido lexical, turno y número del participante. Por eso, se podría detectar si había alguna diferencia en el comportamiento de las diferentes personas gramaticales en el habla de los siete entrevistados. La decisión de codificar por estos grupos de factores, y por los

factores individuales dentro de cada uno, en el sistema compartido de codificación del CESA fue influida por Otheguy et al. (2007), Travis (2007) y Carvalho y Child (2011). Aunque no se incluyeron los factores no lingüísticos en el análisis, por el número mínimo de hablantes en el corpus, estos se conservaron para análisis adicional en la Discusión.

5.2 Análisis

Los datos de las entrevistas se analizaron con regresión de subida y bajada en Goldvarb X. El método fue correr Goldvarb varias veces para detectar, usando el peso de factores, los factores lingüísticos que correlacionaron con la expresión del pronombre sujeto, y la interacción entre estos factores, y después, correr Goldvarb de nuevo con las personas gramaticales separadas. Este método puede ayudar a detectar cualquier cambio en el comportamiento de la expresión del pronombre sujeto para cada pronombre personal, con enfoque especial en la primera persona singular, ya que es la persona gramatical más estudiada y la que existe en más de la mitad de los tokens del corpus.

6.0 Resultados

6.1 Juntas

Como en los estudios previos, se mostró por persona una gran variedad de frecuencia de la expresión del pronombre en los datos del CESA, aunque la expresión pronominal de muchas de las personas salió más baja que en otros estudios como Orozco y Nunez (2012) y Carvalho y Child (2011). La comparación de frecuencias se puede ver en la Tabla 1.

* La segunda persona singular en los datos de Carvalho y Child (2011), la cual incluye el *tú*, el *usted* y el *vos*, se divide según la especificidad (43.6% para lo específico y 31.3% para lo no específico).

** La segunda persona singular en los datos del CESA incluye también el pronombre *usted*.

*** Los totales de Barranquilla y Nueva York incluyen el pronombre menos frecuente *uno*.

Se hicieron corridas en Goldvarb con y sin el tiempo/aspecto/modo (TAM), para obtener información más precisa de los grupos de factores más significativos en el análisis de todas las personas gramaticales. Había muchas interacciones entre los tiempos menos frecuen-

Tabla 1. Porcentaje de expresión del pronombre sujeto en cuatro comunidades, por persona gramatical

| | Barranquilla, Colombia (Orozco y Nunez 2012) | Nueva York (Orozco y Nunez 2012) | Rivera, Uruguay (Carvalho y Child 2011) | CESA |
|----------|--|----------------------------------|---|----------------------|
| yo | 44.5% | 58.7% | 38.0% | 24.5% (N = 482/1966) |
| tú | 32.5% | 40.0% | 43.6%/31.3%* | 11.1%** (N = 26/235) |
| él/ella | 38.0% | 43.3% | 41.4% | 23.0% (N = 100/435) |
| nosotros | 12.8% | 14.5% | 25.2% | 4.5% (N = 13/290) |
| ustedes | 38.9% | 20.8% | --- | 28.6% (N = 6/21) |
| ellos | 9.9% | 25.0% | 14.5% | 13.7% (N = 69/433) |
| todos | 34.2%*** | 42.9%*** | 35.3% | 20.2% (N = 696/3449) |

* La segunda persona singular en los datos de Carvalho y Child (2011), la cual incluye el *tú*, el *usted* y el *vos*, se divide según la especificidad (43.6% para lo específico y 31.3% para lo no específico).

** La segunda persona singular en los datos del CESA incluye también el pronombre *usted*.

*** Los totales de Barranquilla y Nueva York incluyen el pronombre menos frecuente *uno*.

tes (futuro, subjuntivo, condicional y tiempos compuestos), las personas gramaticales menos ambiguas (todos menos primera y tercera persona singulares) y los factores del paralelismo (especialmente el tercer token en una serie, el cual es el tipo menos frecuente en el corpus), y también entre el paralelismo y la conexión discursiva, resultando en knockouts que se tuvieron que resolver antes de hacer el análisis. Los knockouts son pares de factores que no tienen ningún ejemplo del token. (Por ejemplo, en este estudio, no había tokens del pronombre *nosotros*, ni expreso ni nulo, en el subjuntivo o en el condicional.) Los factores con knockouts se tienen que combinar de maneras lógicas para poder continuar con la regresión de subida y bajada. Además, para muchas de las personas gramaticales, sólo el presente salió con un número adecuado de tokens para el análisis; por eso, con la lista entera de tokens y para cada persona individualmente, se hizo una corrida sin el TAM para estudiar mejor los otros grupos de factores, pero no afectó mucho los resultados, como se puede ver en las tablas. Por eso, el resto del análisis se basó en la corrida con el TAM.

Con todas las personas juntas, como se ve en la Tabla 2, el paralelismo salió como el grupo más significativo (rango = 0.508),

seguido por el TAM (0.453), la persona y el número (0.403), el contenido lexical (0.260), la conexión discursiva (0.148), el tipo de cláusula (0.122) y la reflexividad (0.113). Goldvarb eliminó la ambigüedad y la posición del token en el turno, aunque la ambigüedad salió como el factor menos significativo en la corrida sin el TAM.

Esta clasificación de los grupos de factores se parece a las que se encontraron por Orozco y Nunez en Barranquilla y en Nueva York, aunque esos autores detectaron la conexión discursiva como más importante y el paralelismo como menos importante, y también incluyeron en su codificación otros factores, incluso los que no fueron lingüísticos. Sus resultados se presentan en la Tabla 3.

Los resultados del CESA son aún más parecidos a los obtenidos por Carvalho y Child (2011) en Rivera, Uruguay, y por Otheguy et al. (2007) con inmigrantes de Ecuador, Colombia y México en Nueva York, aunque el TAM no salió significativo en Rivera y el paralelismo no lo fue en Nueva York, como se ve en la Tabla 4.

Dentro de cada grupo de factores en el CESA, se pueden ver (en la Tabla 5) tanto los patrones esperados como los inesperados. En cuanto al paralelismo, el token anterior ex-

Tabla 2. Grupos de factores significativos en el análisis de todas personas gramaticales en el CESA
Con tiempo/aspecto/modo: Sin tiempo/aspecto/modo:

| Grupo de factores | Rango |
|------------------------|-------|
| 1. Paralelismo | 0.508 |
| 2. Tiempo/Aspecto/Modo | 0.453 |
| 3. Persona y Número | 0.403 |
| 4. Contenido Lexical | 0.260 |
| 5. Conexión Discursiva | 0.148 |
| 6. Tipo de Cláusula | 0.122 |
| 7. Reflexividad | 0.113 |

Log likelihood = -1541.356 Significance = 0.019

| Grupo de factores | Rango |
|------------------------|-------|
| 1. Paralelismo | 0.501 |
| 2. Persona y Número | 0.381 |
| 3. Contenido Lexical | 0.279 |
| 4. Conexión Discursiva | 0.155 |
| 5. Tipo de Cláusula | 0.128 |
| 6. Reflexividad | 0.125 |
| 7. Ambigüedad | 0.067 |

Log likelihood = -1555.661 Significance = 0.039

Tabla 3. Grupos de factores significativos en el análisis de tokens de todas personas gramaticales en dos comunidades (Orozco y Nunez 2012)

| Barranquilla | |
|------------------------------------|-------|
| Grupo de factores | Rango |
| 1. Persona y número | 0.46 |
| 2. Continuidad del sujeto | 0.32 |
| 3. Tiempo/aspecto/modo | 0.30 |
| 4. Sujeto anterior (Paralelismo) | 0.26 |
| 5. Clase del verbo | 0.20 |
| 6. Realización del sujeto anterior | 0.17 |
| 7. Contenido lexical | 0.09 |
| 8. Sexo/edad del hablante | 0.09 |
| 9. Tipo de cláusula | 0.09 |
| 10. Presencia de tercera partes | 0.07 |

| Nueva York | |
|------------------------------------|-------|
| Grupo de factores | Rango |
| 1. Tiempo/aspecto/modo | 0.53 |
| 2. Persona y número | 0.35 |
| 3. Educación | 0.28 |
| 4. Continuidad del sujeto | 0.24 |
| 5. Sujeto anterior (Paralelismo) | 0.19 |
| 6. Presencia de tercera partes | 0.19 |
| 7. Realización del sujeto anterior | 0.18 |
| 8. Clase del verbo | 0.15 |
| 9. Contenido lexical | 0.15 |
| 10. Tipo de cláusula | 0.08 |
| 11. Distancia del sujeto anterior | 0.07 |

preso promueve la expresión; aún más cuando es el tercero de una serie. Dentro del TAM, es lógico que promuevan la expresión del pronombre el subjuntivo y el imperfecto, ya que el subjuntivo normalmente aparece en cláusulas subordinadas, y el imperfecto cuenta con ambigüedad en algunas de sus formas: las formas que eran las más comunes dentro del corpus (aunque la ambigüedad en sí no salió como un factor importante, y las cláusulas subordinadas no promovieron la expresión del pronombre, en esta codificación).

En cuanto a la persona y el número, un factor muy significativo como encontraron

también Bayley & Pease-Alvarez (1997) y Flores-Ferrán (2004), los pesos de factores se parecen a las frecuencias explicadas previamente, con el *yo* más expresado y el *nosotros* menos expresado. En el contenido lexical, los verbos psicológicos y los que tienen que ver con el habla más promueven la expresión, seguidos por las cópulas, los verbos de movimiento y otros verbos.

Dentro de la conexión discursiva, *switch reference* obviamente es el factor más asociado con la expresión, ya que hay menos expresión con el mismo referente. La importancia de este factor también se detectó por

Tabla 4. Grupos de factores significativos en el análisis de tokens de todas personas gramaticales en dos comunidades

| Rivera, Uruguay (Carvalho y Child, 2011) | |
|---|-------|
| Grupo de factores | Rango |
| 1. Paralelismo | |
| 2. Persona | |
| 3. Conexión discursiva | |
| 4. Contenido lexical | |
| 5. Reflexividad | |

| Inmigrantes de Ecuador, Colombia y México en Nueva York (Otheguy et al., 2007) | |
|---|-------|
| Grupo de factores | Rango |
| 1. Persona | |
| 2. Conexión discursiva | |
| 3. Frase hecha | |
| 4. Tiempo/aspecto/modo | |
| 5. Tipo de cláusula | |
| 6. Contenido lexical | |
| 7. Reflexividad | |

Tabla 5. Clasificación de factores, dentro de los grupos de factores, que correlacionan con la expresión del pronombre en todas personas gramaticales, en el CESA (en el análisis con tiempo/aspecto/modo)

| Paralelismo | Peso | Tiempo/Aspecto/Modo | Peso | Persona y Número | Peso |
|--------------|-------|---------------------|-------|---------------------|-------|
| EE_ | 0.885 | future | 0.741 | yo | 0.583 |
| E_ | 0.638 | subjuntivo | 0.579 | él/ella | 0.569 |
| NE_ | 0.628 | imperfect | 0.574 | ellos/ellas/ustedes | 0.419 |
| primer token | 0.531 | presente | 0.521 | tú/usted | 0.328 |
| EN_ | 0.448 | condicional | 0.514 | nosotros | 0.180 |
| N_ | 0.406 | pretérito | 0.407 | | |
| NN_ | 0.377 | compuesto | 0.288 | | |

| Contenido Lexical | Peso | Conexión Discursiva | Peso |
|-------------------|-------|-----------------------------------|-------|
| psicológico | 0.675 | switch reference | 0.582 |
| de habla | 0.529 | primer token | 0.500 |
| copula | 0.479 | el mismo referente; TAM diferente | 0.437 |
| movimiento | 0.461 | el mismo referente y TAM | 0.434 |
| otro | 0.415 | | |

| Tipo de Cláusula | Peso | Reflexividad | Peso |
|------------------|-------|--------------|-------|
| coordinada | 0.570 | no reflexivo | 0.509 |
| independiente | 0.496 | reflexivo | 0.396 |
| relativa | 0.450 | | |
| otra subordinada | 0.448 | | |

Flores-Ferrán (2004, 2010) y Cameron (1993). Lo más interesante de esta parte del análisis es del tipo de cláusula: se esperaba que las cláusulas subordinadas se asociaran con la expresión, como en Carvalho y Child (2011), pero salió el revés. Unas explicaciones posibles se dan en la Discusión. La reflexividad fue el último factor de todos los factores significativos, con los verbos no reflexivos promoviendo la expresión.

6.2 Separadas

En los resultados de la corrida de la primera persona, los cuales se ven en la Tabla 6, el grupo de factores más significativo fue el paralelismo (0.564), semejante a Torres Cacoullos y Travis (2010a, 2010b); fue seguido por el

TAM (0.518), el contenido lexical (0.283), el tipo de cláusula (0.183), la reflexividad (0.168) y la conexión discursiva (0.159). Sin el TAM, los otros grupos fueron casi iguales, con la apariencia adicional de dos otros grupos significativos: la ambigüedad y la posición del token en el turno.

La diferencia más grande entre estos resultados y los de todas las personas juntas es la conexión discursiva. Resultó menos significativo en la primera persona singular que en todas las personas, mientras el tipo de cláusula y la reflexividad salieron un poco más importantes en el *yo* que en los resultados generales.

Dentro de muchos de los grupos de factores, como se ve en la Tabla 7, no había una

Tabla 6. Grupos de factores significativos en el análisis de tokens de la primera persona singular en el CESA

| Con tiempo/aspecto/modo: | | Sin tiempo/aspecto/modo: | |
|--------------------------|-------|--------------------------|-------|
| Grupo de factores | Rango | Grupo de factores | Rango |
| 1. Paralelismo | 0.564 | 1. Paralelismo | 0.561 |
| 2. Tiempo/Aspecto/Modo | 0.518 | 2. Contenido Lexical | 0.307 |
| 3. Contenido Lexical | 0.283 | 3. Tipo de Cláusula | 0.193 |
| 4. Tipo de Cláusula | 0.183 | 4. Reflexividad | 0.177 |
| 5. Reflexividad | 0.168 | 5. Conexión Discursiva | 0.170 |
| 6. Conexión Discursiva | 0.159 | 6. Ambigüedad | 0.132 |
| | | 7. Turno | 0.077 |

Log likelihood = -953.076 Significance = 0.012

Log likelihood = -966.752 Significance = 0.019

gran diferencia en el efecto entre el *yo* y todas personas juntas, pero en el TAM, el subjuntivo fue relativamente más importante en el *yo* que en todas juntas.

En el análisis del *tú/usted*, después de combinar algunos factores del paralelismo y quitar los factores menos frecuentes del TAM para resolver los knockouts, el único grupo de factores que salió importante fue la especificidad del referente, con el referente específico promoviendo la expresión (peso de factores: específico: 0.762; no específico: 0.399; rango: 0.363). La apariencia de la especificidad como factor

importante se parece al patrón encontrado por Cameron (1993) del *tú* en la comunidad madrileña (y el opuesto de lo que se encontró en San Juan), aunque la frecuencia de la expresión en el CESA es más baja en ambos casos, como se ve en la Tabla 8. También se hace eco con los resultados de Otheguy et al. (2007), los cuales encontraron más expresión del pronombre en la segunda persona específica que en la no específica en ambas comunidades inmigrantes estudiadas.

Para analizar el resto de las personas gramaticales, se tuvieron que combinar algunos

Tabla 7. Clasificación de factores, dentro de los grupos de factores, que correlacionan con la expresión del pronombre en la primera persona singular, en el CESA (en el análisis con tiempo/aspecto/modo)

| Paralelismo | Peso | Tiempo/Aspecto/Modo | Peso | Contenido Lexical | Peso |
|--------------|-------|---------------------|-------|-------------------|-------|
| EE_ | 0.927 | subjuntivo | 0.804 | psicológico | 0.679 |
| NE_ | 0.757 | futuro | 0.707 | de habla | 0.545 |
| E_ | 0.670 | imperfecto | 0.615 | movimiento | 0.437 |
| primer token | 0.534 | condicional | 0.563 | cópula | 0.420 |
| EN_ | 0.413 | presente | 0.532 | otro | 0.396 |
| N_ | 0.381 | pretérito | 0.371 | | |
| NN_ | 0.363 | compuesto | 0.286 | | |

| Tipo de Cláusula | Peso | Reflexividad | Peso | Conexión Discursiva | Peso |
|----------------------|-------|--------------|-------|---------------------|-------|
| coordinada | 0.598 | no reflexivo | 0.512 | switch reference | 0.617 |
| independiente | 0.506 | reflexivo | 0.344 | primer token | 0.467 |
| subordinada/relativa | 0.415 | | | el mismo referente | 0.458 |

Tabla 8. Porcentaje de expresión de la segunda persona singular en tres comunidades, por especificidad

| | tú en San Juan (Cameron, 1993) | tú en Madrid (Cameron, 1993) | tú y usted en el CESA |
|---------------|--------------------------------|------------------------------|-----------------------|
| Específico | 48% | 40% | 24.6% |
| No específico | 69% | 19% | 6.3% |

factores del paralelismo y de la conexión discursiva y quitar unos tiempos del TAM para resolver los knockouts. Después de eso, la conexión discursiva fue el único grupo de factores que salió importante para la tercera persona singular (rango: 0.330) y para la primera persona plural (rango: 0.509), como se ve en la Tabla 9.

En la segunda y tercera persona plural (*ustedes/ellos/ellas*), no se encontró ningún grupo de factores significativo para la expresión del pronombre sujeto, pero en el análisis de la tercera persona plural individualmente (sin *ustedes*), el paralelismo salió significativo (0.296). Este último resultado es semejante a Dumont (2006), la cual encontró el paralelismo como más significativo en la expresión de la tercera persona con NP completa (singular y plural juntas).

7.0 Discusión

7.1 Juntas

Aunque la frecuencia de expresión en el CESA fue menor que en muchos de los estudios anteriores, los patrones de todas las personas juntas se parecen a los de mucha de la literatura, como ya se mencionó. Sin embargo, hay que considerar que había mucha variación

entre participantes en cuanto a su expresión, como se puede ver en la tabla. A diferencia de los estudios anteriores, los cuales específicamente dividieron a los participantes en grupos según su generación o conexión a los Estados Unidos, no se separaron en este estudio los participantes del CESA, dado que sólo hay siete participantes en el corpus. Por ejemplo, Bayley & Pease-Alvarez (1997), Otheguy et al. (2007) y Silva-Corvalán (1994) detectaron diferencias por generación (aunque en direcciones opuestas); esto no fue posible en el estudio actual.

Con la excepción de un participante, CESA004, las prácticas lingüísticas, es decir, el idioma más usado por cada participante en su vida diaria, coinciden con la frecuencia de su expresión del pronombre, con los que usan más el inglés expresando el pronombre con más frecuencia que los que usan más el español, como se ve en la Tabla 10. Es posible que la separación por prácticas lingüísticas u otros factores no lingüísticos, cuando sea más grande el corpus, pueda revelar patrones distintos en la expresión del pronombre, hasta más factores significativos en los datos de las otras personas gramaticales.

Es necesario analizar más los patrones inesperados de tipos de cláusula, ya que se supone

Tabla 9. Clasificación de factores, dentro de la conexión discursiva, que correlacionan con la expresión del pronombre en la tercera persona singular y en la primera persona plural, en el CESA

| Conexión discursiva (él/ella) | Peso | Conexión discursiva (nosotros) | Peso |
|-----------------------------------|-------|---|-------|
| <i>switch reference</i> | 0.582 | <i>switch reference</i> | 0.805 |
| primer token | 0.648 | otros (primer token o el mismo referente) | 0.296 |
| el mismo referente y TAM | 0.318 | | |
| el mismo referente; TAM diferente | 0.447 | | |

Tabla 10. Porcentaje de expresión del pronombre sujeto, número total de tokens y prácticas lingüísticas por participante del CESA

| Participante | Porcentaje expresos | Tokens (Total) | Prácticas lingüísticas |
|----------------|---------------------|----------------|---------------------------------|
| CESA005 | 35.6% | 640 | inglés |
| CESA006 | 28.4% | 475 | inglés |
| CESA001 | 18.7% | 444 | inglés |
| CESA008 | 16.2% | 551 | español |
| CESA002 | 15.5% | 715 | español |
| CESA003 | 9.4% | 341 | español |
| CESA004 | 6.4% | 283 | tanto el inglés como el español |
| Totales | 20.2% | 3449 | --- |

que las cláusulas subordinadas, donde existe más el *switch reference*, tendrían más expresión, en vez de menos. Es notable que las cláusulas coordinadas se asocian con la expresión del pronombre: puede que fuera un efecto de los tokens del tipo EE_ en el corpus (es decir, los que fueron el tercero de una serie, precedidos por dos pronombres expresos), entre los cuales se vio una fuerte correlación con la expresión. Por otro lado, como hay suficientes ejemplos de cláusulas coordinadas que inhiben la expresión, no puede ser sólo por esto, como se ve en el ejemplo 2.

Entonces ÉL SIEMPRE, POR EJEMPLO, Ø NO ES alguien que no:- que Ø NO PU- EDES (...) um, (...) e-es muy ay, cómo lo digo, Ø ES muy honesto, Ø ES muy duro a- a la misma vez, Ø ES alguien muy fuerte, no- Ø NO DEJA que la gente l- lo- se aprovechan de él, pero a la misma vez Ø SÍ LO DEJA. (CESA005)

En general, los resultados de todas las personas juntas se parecen a los patrones vistos en la literatura previa, y, con la habilidad de separar a los participantes por factores no lingüísticos, es probable que se parezcan aún más.

7.2 Separadas

Se detectó una distribución parcialmente

diferente de factores y grupos de factores entre el *yo* y todas las personas juntas, especialmente con la conexión discursiva y el subjuntivo. En estos resultados, las diferencias no son tan grandes para concluir que sea imposible analizar todas las personas juntas. Sin embargo, hay que considerar que dado que el 57% de los tokens son de la primera persona singular, en realidad lo que se están comparando en este análisis son los patrones del *yo* con los patrones del *yo* y de otras personas juntas. Se tienen que comparar los resultados del *yo* con los resultados de cada otra persona para ver si hay grandes diferencias que podrían sugerir la necesidad para análisis individual, y sí las hay: en cada otra persona gramatical, sólo la especificidad, la conexión discursiva o el paralelismo resultó significativo, a diferencia de los resultados de la primera persona.

Antes de continuar con las otras personas gramaticales, es importante comentar sobre una de las tabulaciones cruzadas que se hizo con la primera persona. En algunas entrevistas, por ejemplo CESA001 y CESA005, el contenido lexical exhibió un efecto más grande sobre la expresión de la primera persona singular que sobre todos los pronombres en general, debido al efecto de las frases frecuentes “yo sé”, “yo pienso” y “yo creo”, que se usaron fre-

cuentemente con el pronombre expreso, como en Torres Cacoullós y Travis (2010b), pero este efecto se neutralizó en los datos completos. Los patrones individuales, entonces, tuvieron efectos que posiblemente puedan afectar los resultados de factores no lingüísticos en la primera persona también.

En el análisis del *tú/usted*, además de la especificidad ya mencionada, la reflexividad también tuvo un efecto, aunque no se lo pudo mostrar por Goldvarb a causa de un knockout. Es decir, todos los 8 tokens reflexivos fueron del pronombre nulo, mientras los 227 no reflexivos mostraron una frecuencia de 11.5% expresos y 88.5% nulos. También puede ser que el TAM tuviera un efecto, ya que los 17 tokens que se quitaron del análisis del TAM a causa de los knockouts (futuro, condicional, subjuntivo y tiempos compuestos) usaron 0% el pronombre expreso (100% el pronombre nulo), mientras los otros 218 tokens mostraron patrones distintos por tiempo y aspecto, como se presenta en la Tabla 11. Con más tokens de los otros TAMs que no sean el presente, si siguen el mismo patrón, puede haber un efecto del TAM en la segunda persona singular.

Parte del problema en detectar la importancia de factores como el paralelismo y la conexión discursiva en las otras personas que no fueran el *yo* es que no había tantas series tan largas de discurso con estas personas como las había con el *yo*. Sólo había 235 tokens de la segunda persona singular, en contraste con 1966 del *yo* (y 3449 total del corpus), y no se distribuyeron sobre las mismas funciones discursivas y narrativas como el *yo*. La oración en

el ejemplo 3, que sólo tiene dos tokens del *tú* en el medio de una serie del *nosotros* es representativa del uso de los tokens del *tú*:

En veces NOS ENCONTRAMOS para almorzar juntos, y para decir, ah, lo que ESTÁS HACIENDO hoy, o lo que HICISTES, o, sí, SIEMPRE NOS- NOS QUEDAMOS en contacto. (CESA005)

El problema se alivia un poco con la tercera persona singular, con 435 tokens y más narrativas, pero como hay más *switch reference* entre dos sujetos en narrativas de la tercera persona, todavía no se puede obtener series tan largas para análisis del paralelismo y de la conexión discursiva. Por eso, hay necesidad de obtener más datos para análisis futuros, especialmente series más largas para personas que no sean el *yo*.

Para contestar la pregunta de investigación, sería difícil concluir que las personas gramaticales se distribuyan de maneras claramente distintas en cuanto a la expresión del pronombre sujeto, dado que los patrones del *yo* y de todos los pronombres personales son parecidos, igual que la conclusión de Orozco y Nunez (2012). Sin embargo, sí hay evidencia que sugiere que los datos de cada persona gramatical no son comparables, como en la perspectiva de Torres Cacoullós y Travis (2010a, 2010b) y Dumont (2006), por tres razones. Primero, menos grupos de factores surgieron como significativos en el análisis de las otras personas, insinuando a patrones distintos para cada una. Segundo, hay diferentes restricciones semánticas sobre la expresión de la tercera persona y sobre la de las otras, como

explica Dumont (2006). Tercero, no existen en los corpus los mismos tipos de narrativas para cada persona gramatical, especialmente el *tú*, que existen para el *yo*.

Por lo tanto, la mejor conclusión de este estudio es que todavía sería beneficioso comparar los resultados de cada persona gramatical dentro de un estudio, como lo hacen Otheguy et al. (2007) y otros, aunque hay que notar que para llegar a resultados aún más interpretables y confiables, puede que sea necesario separarlas como Torres Cacoullós y Travis (2010a, 2010b) y estudiarlas individualmente.

Referencias

- Bayley, R., y Pease-Alvarez, L. (1997). Null pronoun variation in Mexican-descent children's narrative discourse. *Language Variation and Change*, 9(3), 349-371.
- Cameron, R. (1993). Ambiguous agreement, functional compensation, and nonspecific *tú* in the Spanish of San Juan, Puerto Rico, and Madrid, Spain. *Language Variation and Change*, 5(3), 305-334.
- Carvalho, A.M., & Child, M. (2011). Subject pronoun expression in a variety of Spanish in contact with Portuguese. En Jim Michnowicz & Robin Dodsworth (Eds.), *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics* (pág. 14-25). Somerville, MA: Cascadilla Proceedings Project.
- Dumont, J. (2006). Full NPs as subjects. En N. Sagarra & A.J. Toribio (Eds.), *Selected Proceedings of the 9th Hispanic Linguistics Symposium* (pág. 286-296). Somerville, MA: Cascadilla Proceedings Project.
- Flores-Ferrán, N. (2004). Spanish subject personal pronoun use in New York City Puerto Ricans: Can we rest the case for English contact? *Language Variation and Change*, 16(1), 49-73.
- Flores-Ferrán, N. (2010). *¿Tú no me hables!* Pronoun expression in conflict narratives. *International Journal of the Sociology of Language*, 203(1), 61-82.
- Orozco, R., y Nunez, B. (2012). *¿Todos juntos o cada uno por separado? 'All (SPPs) together or each one separately?': A pronominal expression study.* NWAV41 [Congreso], Bloomington, Indiana, 26 de octubre del 2012.
- Otheguy, R., Zentella, A.C., y Livert, D. (2007). Lan-

- guage and dialect contact in Spanish in New York: Toward the formation of a speech community. *Language*, 83(4), 770-802.
- Sankoff, D., Tagliamonte, S.A. y Smith, E. (2005). Goldvarb X: A variable rule application for Macintosh and Windows. Departamento de lingüística, Universidad de Toronto.
- Silva-Corvalán, C. (1994). *Language contact and change: Spanish in Los Angeles*. Oxford: Clarendon.
- Torres Cacoullós, R., y Travis, C. (2010a). Testing convergence via code-switching: Priming and the structure of variable subject expression. *International Journal of Bilingualism*, 15, 241-267.
- Torres Cacoullós, R., y Travis, C. (2010b). Variable *yo* expression in New Mexico: English influence? En S. Rivera-Mills & D. Villa (Eds.), *Spanish of the U.S. Southwest: A language in transition* (pp. 185-206). Madrid: Iberoamericana.
- Travis, C. (2007). Genre effects on subject expression in Spanish: Priming in narrative and conversation. *Language Variation and Change*, 19, 101-135.

Tabla 11. Expresión de la segunda persona singular en el CESA, por TAM:

| | Expreso | Nulo |
|---------------------|---------|-------|
| Pretérito (N = 13) | 30.8% | 69.2% |
| Presente (N = 193) | 10.9% | 89.1% |
| Imperfecto (N = 12) | 8.3% | 91.7% |