

City University of New York (CUNY)

CUNY Academic Works

All Dissertations, Theses, and Capstone
Projects

Dissertations, Theses, and Capstone Projects

9-2017

Continuidad y cambio en los condicionantes lingüísticos y socio-demográficos del uso del pronombre sujeto en el español de la primera generación en Nueva York

Lorena García-Barroso
The Graduate Center, CUNY

[How does access to this work benefit you? Let us know!](#)

More information about this work at: https://academicworks.cuny.edu/gc_etds/2294

Discover additional works at: <https://academicworks.cuny.edu>

This work is made publicly available by the City University of New York (CUNY).
Contact: AcademicWorks@cuny.edu

CONTINUIDAD Y CAMBIO EN LOS CONDICIONANTES LINGÜÍSTICOS Y SOCIO-
DEMOGRÁFICOS DEL USO DEL PRONOMBRE SUJETO EN EL ESPAÑOL DE LA
PRIMERA GENERACIÓN EN NUEVA YORK

by

LORENA GARCIA BARROSO

A dissertation submitted to the Graduate Faculty in Latin American, Iberian and Latino Cultures
in partial fulfillment of the requirements for the degree of Doctor of Philosophy,
The City University of New York

2017

© 2017

LORENA GARCIA BARROSO

All Rights Reserved

Continuidad y cambio en los condicionantes lingüísticos y socio-demográficos del uso del
pronombre sujeto en el español de la primera generación en Nueva York

by

Lorena García Barroso

This manuscript has been read and accepted for the Graduate Faculty in Latin American, Iberian
and Latino Cultures in satisfaction of the dissertation requirement for the degree of
Doctor of Philosophy.

Date

Ricardo Otheguy

Chair of Examining Committee

Date

José del Valle

Executive Officer

Supervisory Committee:

Ricardo Otheguy

José del Valle

Beatriz Lado

THE CITY UNIVERSITY OF NEW YORK

ABSTRACT

Continuidad y cambio en los condicionantes lingüísticos y socio-demográficos del uso del pronombre sujeto en el español de la primera generación en Nueva York

by

Lorena García Barroso

Advisor: Ricardo Otheguy

In situations of bilingualism in the United States, it has been shown that speakers who alter their grammar (syntax or morphology) are mostly second generation, that is, Latinos born in the United States (Flores-Ferrán 2004; Montrul 2004 ; Silva-Corvalán, 1994). Research about Spanish in the US shows that second generation speakers have different features from those of the first generation who are Latinos born in Latin America. However, is it possible that all the speakers of the first generation behave grammatically in the same way due to the fact of being born in Latin America regardless of the amount of years living in the country?

This dissertation is a variationist study of Spanish in New York City that seeks to explore if the grammar of the newcomers and that of the established immigrants, who together make up the first generation, is modified like the grammar of the second generation speakers born in NYC. Through the study of the subject pronoun variable, this dissertation explores whether the hierarchies of variables and constraints are the same in these two groups coming to NYC from

Latin America.

For the purpose of this research, as in other sociolinguistics studies, first generation speakers, that is newcomers and established immigrants, are called exposure groups (to the city and English spoken in New York City).

In this dissertation, I study the presence or absence of the subject pronoun, making a hierarchical classification of the internal and external variables that condition its use, and of the factors in which they are subdivided. The corpus of this research is composed of 114 transcribed interviews with first generation immigrant speakers and newcomers from Puerto Rico, Cuba, the Dominican Republic, Mexico, Colombia and Ecuador residing in New York. The study takes into consideration the occurrences where the use of the pronoun is variable or optional. In the different national and exposure groups, I analyze the most important variables in terms of the predictive value of the pronoun appearance, and check if the same order of variables and constraints is given in the different groups.

The results show that the order and weight of the variables is different between the different groups. It is also observed that the internal variables are more important than the external ones to differentiate the immigrant and newcomers speakers in New York in the realization of the subject pronoun. The Caribbeans of both groups have the highest level of use of the pronoun. Immigrants are different from newcomers by the age they came to NYC and the level of English proficiency. The language contact hypothesis is only valid for the informants who have been in the city for the longest time and have been more exposed to the contact language. The internal variables that have greater prediction in the use of the pronoun are Reference and Person for the newcomers. For the immigrants they are Person and Reference. Finally, constraints hierarchies are different in the Person and Tense variables in the two

exposure groups. Despite these differences, the strength of Spanish continuity in Latin America in the pronominal distribution in Spanish of newcomers and immigrants is greater than the force of change.

Para mis queridos padres, Antonio y Lina

Para mi adorada bebé, Sofia

Agradecimientos

Me gustaría agradecer de manera muy especial la ayuda del profesor Ricardo Otheguy, mi director de tesis y mi Maestro. Gracias a tu sabiduría y tu excelente dirección he podido realizar este trabajo de sociolingüística variacionista. Ha sido un privilegio ser tu alumna y doctoranda durante todos estos años. En mi memoria siempre quedará el recuerdo de nuestras reuniones en tu oficina, tus lápices amarillos y las notas que guardo como un tesoro. Gracias por tus muestras de ánimo, tu generosidad, tu constancia y tu disponibilidad total.

Igualmente, quiero expresar mi gratitud al profesor José del Valle, otro Maestro excelente. Me siento afortunada de haber empezado y acabado mi tesis en un programa de doctorado dirigido por un académico como tú que combina tan perfectamente esos dos aspectos que a veces parecen irreconciliables en la academia: la profesionalidad y la cercanía. Siempre has estado dispuesto a enseñarme, a ayudarme, a guiarme.

Me gustaría sobresaltar la colaboración tan estrecha con las profesoras Beatriz Lado, miembro del comité de tesis, y Laura Callahan, miembro de mi segundo examen y propuesta de tesis. Estas dos profesoras siempre me retaron intelectualmente con sus diferentes áreas de investigación, contribuyendo a mi formación como lingüista. Disfruté mucho de vuestras clases.

También, agradezco la confianza de mi fiel y queridísimo amigo José Plácido, otro Maestro excelente. Te admiro por lo que haces y te doy las gracias por enseñarme y compartir conmigo tus magníficas lecciones. A mis amigas del doctorado Jelena, Carmen, Carolina, Ana Isabel porque casi todas pasaron con éxito este tramo final. A Jaime Atencio, Rocío Albéndiz, Mengji y Borja por sus productivas ideas. Y a Víctor, mi profesor de antropología, mi inspiración para no quedarme atrás, mi maestro de vida. Te quiero con locura y lo sabes. Nuestra bebé, nuestra mejor creación.

Finalmente, quiero mencionar a los más importantes: mis padres, Antonio y Lina. Gracias por todas las oportunidades que siempre nos habéis ofrecido a mí y a mis hermanas, Merche y Míriam. Esta tesis se la dedico a ellos. Soy muy, muy, muy afortunada de tener unos padres como ellos. A mi padre porque de él heredé el gusto por el estudio, la lectura y el afán de superación. A mi madre porque fue ella la que me enseñó a estudiar y a esforzarme por conseguir lo mejor siempre. Ella es la parte más fundamental de mi vida, la que le da sentido a todo. Me habría gustado estar pegada a ti toda la vida, Li.

A todas estas personas, mis más sinceras gracias por vuestras muestras de apoyo.

Índice

Lista de tablas.....	xi
1. Capítulo 1. Introducción.....	1
1. Introducción.....	1
2. Motivación: el contacto de lenguas.....	4
3. Preguntas de investigación.....	5
4. Revisión de la literatura.....	7
5. Hipótesis y previsión de los resultados.....	11
6. Resumen del estudio.....	13
7. Relevancia del estudio.....	13
2. Capítulo 2. Metodología.....	15
1. Introducción.....	15
2. La muestra.....	15
2.1 Las entrevistas.....	16
2.2 Composición de la muestra de este estudio.....	16
2.2.1 Género.....	16
2.2.2 Edad.....	17
2.2.3 Exposición a Nueva York.....	17
2.2.4 Origen nacional y regional.....	18
2.2.5 Clase social y educación.....	19
2.2.6 Competencia en inglés.....	19
3. El sobre de variación.....	20
4. Metodología.....	21

5. Las variables.....	22
5.1 Las variables dependientes.....	23
5.2 Las variables independientes externas.....	23
5.3 Las variables independientes internas.....	25
3. Capítulo 3. El español de los inmigrantes y los recién llegados a través de la tasa pronominal.....	29
1. Introducción.....	29
2. Distribución del uso del pronombre en la muestra entera y por grupos de exposición...32	
2.1 La tasa pronominal en la muestra entera.....	32
2.2 La tasa pronominal por grupos de exposición.....	33
3. La tasa pronominal por variables externas socio-demográficas en la muestra entera y por grupos de exposición.....	36
3.1 Género, edad, edad de llegada a EE.UU, educación y clase social en la muestra entera y por grupos de exposición.....	36
3.2 Estatus socio-económico (ESE) en la muestra entera y por grupos de exposición...40	
4. La tasa pronominal por variables externas de uso de la lengua en la muestra entera y por grupos de exposición.....	41
4.1 Competencia en inglés en la muestra entera y por grupos de exposición.....	42
5. Origen nacional y regional en la muestra entera y por grupos de exposición.....	43
6. Diferencias regionales en la tasa pronominal por características demográficas en la muestra entera y por grupos de exposición.....	50
7. Diferencias regionales en la tasa pronominal por variables externas de uso de la lengua en la muestra entera y por grupos de exposición.....	53

Resumen y conclusión.....	54
4. Capítulo 4. Análisis de las variables externas.....	60
1. Introducción.....	60
2. Las variables externas que definen la tasa pronominal.....	62
3. Regresiones lineales múltiples.....	67
3.1 Regresiones lineales en la muestra entera.....	70
3.2 Regresiones lineales en los grupos de exposición: recién llegados e inmigrantes.....	72
3.3 Regresiones lineales en los grupos regionales de exposición.....	76
4. Interpretación de los resultados de las regresiones multivariantes.....	80
4.1 El español en Nueva York en la primera generación.....	81
4.2 El español en Nueva York según los grupos de exposición.....	82
4.3 El español en Nueva York en los grupos regionales de exposición.....	85
Resumen y conclusión.....	88
5. Capítulo 5. Análisis de las variables internas.....	93
1. Introducción.....	93
2. Las variables gramaticales y discursivas que definen la ocurrencia del pronombre.....	97
3. Regresiones logísticas.....	99
4. Jerarquías de variables y factores.....	100
4.1 Jerarquías de variables y factores en la muestra entera.....	101
4.2 Jerarquías de variables y factores en los grupos de exposición: recién llegados e inmigrantes.....	105
4.3 Jerarquías de variables y factores en los grupos regionales de exposición.....	114
5. Interpretación de los resultados de las jerarquías de variables y factores.....	124

5.1 Continuidad en el español en Nueva York.....	125
5.2 Cambios en el español en Nueva York.....	129
Resumen y conclusión.....	132
6. Capítulo 6. Conclusiones.....	136
Bibliografía.....	144

Lista de tablas

Capítulo 2:

Tabla 1. Total de participantes por grupos de exposición.....	18
Tabla 2. Total de participantes por grupos de exposición y origen nacional.....	19
Tabla 3. Total de participantes por grupos de exposición y área regional.....	19

Capítulo 3:

Tabla 1. Porcentaje de verbos con pronombres sujeto.....	33
Tabla 2. ANOVA. Tasa pronominal por exposición a Nueva York.....	34
Tabla 3. ANOVA. Tasa pronominal por edad.....	37
Tabla 4. ANOVA. Tasa pronominal por edad de llegada a EE.UU.....	38
Tabla 5. ANOVA. Tasa pronominal por clase social, R.LL.....	38
Tabla 6. ANOVA. Tasa pronominal por edad de llegada a EE.UU, INM.....	39
Tabla 7. ANOVA. Tasa pronominal por competencia en inglés.....	42
Tabla 8. ANOVA. Tasa pronominal por competencia en inglés, INM.....	43
Tabla 9a. ANOVA. Tasa pronominal por país de origen.....	44
Tabla 9b. ANOVA. Post-hoc Tukey test de significación pair-wise.....	45
Tabla 10. ANOVA. Tasa pronominal por región de origen.....	46
Tabla 11. ANOVA. Tasa pronominal por país de origen, R.LL.....	47
Tabla 12. ANOVA. Tasa pronominal por país de origen, INM.....	48
Tabla 13. ANOVA. Tasa pronominal por región de origen, R.LL.....	49
Tabla 14. ANOVA. Tasa pronominal por región de origen, INM.....	49

Capítulo 4:

Tabla 1. Correlaciones Pearson con tasa pronominal en la muestra entera.....	65
--	----

Tabla 2. Correlaciones Pearson con tasa pronominal en los R.LL.....	66
Tabla 3. Correlaciones Pearson con tasa pronominal en los INM.....	66
Tabla 4a. Regresión múltiple para la tasa pronominal, muestra entera.....	70
Tabla 4b. Regresión múltiple para la tasa pronominal, muestra entera.....	70
Tabla 4c. Regresión múltiple para la tasa pronominal, muestra entera.....	71
Tabla 5a. Regresión múltiple para la tasa pronominal, R.LL.....	73
Tabla 5b. Regresión múltiple para la tasa pronominal, R.LL.....	73
Tabla 5c. Regresión múltiple para la tasa pronominal, R.LL.....	73
Tabla 6a. Regresión múltiple para la tasa pronominal, INM.....	74
Tabla 6b. Regresión múltiple para la tasa pronominal, INM.....	75
Tabla 6c. Regresión múltiple para la tasa pronominal, INM.....	75
Tabla 7.1a. Regresión múltiple para la tasa pronominal, R.LL. Caribeños.....	76
Tabla 7.2a. Regresión múltiple para la tasa pronominal, R.LL. Continentales.....	76
Tabla 7.1b. Regresión múltiple para la tasa pronominal, R.LL. Caribeños.....	77
Tabla 7.2b. Regresión múltiple para la tasa pronominal, R.LL. Continentales.....	77
Tabla 7.1c. Regresión múltiple para la tasa pronominal, R.LL. Caribeños.....	77
Tabla 7.2c. Regresión múltiple para la tasa pronominal, R.LL. Continentales.....	77
Tabla 8.1a. Regresión múltiple para la tasa pronominal, INM. Caribeños.....	79
Tabla 8.2a. Regresión múltiple para la tasa pronominal, INM. Continentales.....	79
Tabla 8.1b. Regresión múltiple para la tasa pronominal, INM. Caribeños.....	79
Tabla 8.2b. Regresión múltiple para la tasa pronominal, INM. Continentales.....	79
Tabla 8.1c. Regresión múltiple para la tasa pronominal, INM. Caribeños.....	79
Tabla 8.2c. Regresión múltiple para la tasa pronominal, INM. Continentales.....	79

Capítulo 5:

Tabla 1. Jerarquía de variables que predicen la Ocurrencia, muestra.....	101
Tabla 1a. Jerarquía de factores para la variable interna Persona, muestra entera.....	102
Tabla 1b. Jerarquía de factores para la variable interna Referencia, muestra entera.....	103
Tabla 1c. Jerarquía de factores para la variable interna Tiempo, muestra entera.....	103
Tabla 1d. Jerarquía de factores para la variable interna Léxico, muestra entera.....	104
Tabla 1e. Jerarquía de factores para la variable interna Reflexividad, muestra entera.....	104
Tabla 2. Jerarquía de variables que predicen la Ocurrencia, R.LL.....	106
Tabla 2a. Jerarquía de factores para la variable interna Referencia, R.LL.....	107
Tabla 2b. Jerarquía de factores para la variable interna Persona, R.LL.....	107
Tabla 2c. Jerarquía de factores para la variable interna Léxico, R.LL.....	108
Tabla 2d. Jerarquía de factores para la variable interna Tiempo, R.LL.....	108
Tabla 2e. Jerarquía de factores para la variable interna Reflexividad, R.LL.....	109
Tabla 3. Jerarquía de variables que predicen la Ocurrencia, INM.....	110
Tabla 3a. Jerarquía de factores para la variable interna Persona, INM.....	111
Tabla 3b. Jerarquía de factores para la variable interna Referencia, INM.....	111
Tabla 3c. Jerarquía de factores para la variable interna Tiempo, INM.....	112
Tabla 3d. Jerarquía de factores para la variable interna Reflexividad, INM.....	112
Tabla 3e. Jerarquía de factores para la variable interna Léxico, INM.....	113
Tabla 4.1a. Jerarquía de variables que predicen la Ocurrencia, R.LL. Caribeños.....	115
Tabla 4.2a. Jerarquía de variables que predicen la Ocurrencia, R.LL. Continentales.....	115
Tabla 4.1b. Jerarquía de factores para la variable interna Referencia, R.LL. Caribeños.....	116
Tabla 4.2b. Jerarquía de factores para la variable interna Referencia, R.LL. Continentales...	116

Tabla 4.1c. Jerarquía de factores para la variable interna Persona, R.LL. Caribeños.....	116
Tabla 4.2c. Jerarquía de factores para la variable interna Persona, R.LL. Continentales.....	116
Tabla 4.1d. Jerarquía de factores para la variable interna Léxico, R.LL. Caribeños.....	117
Tabla 4.2d. Jerarquía de factores para la variable interna Léxico, R.LL. Continentales.....	117
Tabla 4.1e. Jerarquía de factores para la variable interna Tiempo, R.LL. Caribeños.....	118
Tabla 4.2e. Jerarquía de factores para la variable interna Tiempo, R.LL. Continentales.....	118
Tabla 4.1f. Jerarquía de factores para la variable interna Reflexividad, R.LL. Caribeños.....	118
Tabla 4.2f. Jerarquía de factores para la variable interna Reflexividad, R.LL. Continentales..	118
Tabla 5.1a. Jerarquía de variables que predicen la Ocurrencia, INM. Caribeños.....	119
Tabla 5.2a. Jerarquía de variables que predicen la Ocurrencia, INM. Continentales.....	119
Tabla 5.1b. Jerarquía de factores para la variable interna Persona, INM. Caribeños.....	121
Tabla 5.2b. Jerarquía de factores para la variable interna Persona, INM. Continentales.....	121
Tabla 5.1c. Jerarquía de factores para la variable interna Referencia, INM. Caribeños.....	122
Tabla 5.2c. Jerarquía de factores para la variable interna Referencia, INM. Continentales.....	122
Tabla 5.1d. Jerarquía de factores para la variable interna Tiempo, INM. Caribeños.....	122
Tabla 5.2d. Jerarquía de factores para la variable interna Tiempo, INM. Continentales.....	122
Tabla 5.1e. Jerarquía de factores para la variable interna Reflexividad, INM.....	123
Tabla 5.2e. Jerarquía de factores para la variable interna Reflexividad, INM. Continentales...	123
Tabla 5.1f. Jerarquía de factores para la variable interna Léxico, INM. Caribeños.....	123
Tabla 5.2f. Jerarquía de factores para la variable interna Léxico, INM. Continentales.....	123

Capítulo 1

Introducción

1. Introducción

En situaciones de bilingüismo en Estados Unidos, se ha mostrado que los hablantes que alteran su gramática (en cuanto a su sintaxis o morfología) son mayormente los de *segunda generación*, es decir, los latinos nacidos en el país norteamericano (Bookhamer, 2012; Flores-Ferrán, 2004; Gutiérrez, 1995; Montrul, 2002, 2004, 2009; Silva-Corvalán, 1994, 1995a, 1995b; Travis, 2007). Muchos trabajos sobre el español en Estados Unidos han mostrado que la segunda generación tiene rasgos diferentes de la *primera generación*, que agrupa a los latinos nacidos en Latinoamérica. Por ejemplo, algunos autores han estudiado que hay modificaciones en el sistema verbal. Se observa que la segunda generación tiene menos usos del modo subjuntivo y más usos del indicativo en comparación con la primera generación (Bookhamer, 2013; Montrul, 2002; Silva-Corvalán, 1995). También en trabajos sobre la expresión de futuridad se ha comprobado que la segunda generación hace uso de la forma perifrástica mientras que la ocurrencia del futuro morfológico es casi inexistente (Gutiérrez, 1995; Zentella, 1997). Lo mismo ocurre con el condicional. Hay una progresiva desaparición de este tiempo verbal en la segunda generación (Silva-Corvalán, 1994). Otro tipo de trabajos ha analizado que la segunda generación usa de forma explícita el pronombre personal sujeto más que la primera generación (Flores-Ferrán, 2004; Montrul, 2004; Silva-Corvalán, 1982, 1994; Travis, 2007).

Entre otros estudios sobre el uso del pronombre personal, está el de Otheguy y Zentella (2012) sobre el español en Nueva York. Ellos estudiaron el uso variable del pronombre personal sujeto en hablantes de primera y segunda generación. En su estudio intergeneracional, los autores distinguen dentro de la primera generación dos grupos de hablantes en relación al nivel de

exposición al inglés que han tenido en la ciudad de Nueva York: *inmigrantes establecidos e inmigrantes recién llegados*. Esta distinción hecha en los grupos de exposición se basa en la edad de llegada y los años vividos en Nueva York. Los recién llegados son los informantes que llegaron con 17 años o más y han vivido en Estados Unidos por cinco años o menos. Los inmigrantes son aquellos que llegaron cuando tenían menos de 17 años y pasaron más de cinco años en el país norteamericano (Otheguy y Zentella, 2012: 33; 86). En su trabajo, analizaron las *tasas pronominales*, es decir, el porcentaje de verbos de forma personal que aparecen con un pronombre personal sujeto, en la primera generación: los recién llegados y los inmigrantes, y la segunda generación: los nacidos en Nueva York (Otheguy y Zentella, 2012: 68). También, realizaron en cada generación jerarquías estadísticas de variables y de factores, las cuales ordenan las variables independientes (i.e., las variables predictivas) según la magnitud de su influencia sobre el uso o la omisión del pronombre. Estos investigadores observaron que el contacto con el inglés se manifiesta en un aumento del número de pronombres y en diferencias de jerarquías de variables independientes internas, y también jerarquías de factores dentro de cada variable en la segunda generación. A través del análisis de las jerarquías de variables y de factores, detectaron que se acortan las diferencias de peso predictivo entre los factores en cada variable. Estos autores encontraron, por lo tanto, que la segunda generación es diferente en cuanto a la realización del pronombre.

Ahora bien, ¿es posible que todos los hablantes de primera generación se comporten gramaticalmente igual sólo por el hecho de haber nacido en Latinoamérica a pesar de la cantidad de años vividos en el país?, ¿es posible que los hablantes en EEUU inmigrantes y recién llegados no cambien su gramática a pesar del tiempo de exposición con la lengua de contacto y otras variedades orales de español? El propósito de este trabajo de sociolingüística variacionista es

determinar si la gramática de los inmigrantes y recién llegados, la primera generación, que viven en Nueva York se modifica de la misma manera que la gramática de los hablantes de segunda generación. Quiero estudiar si 1) los inmigrantes y recién llegados sufren cambios paralelos en la gramática como ocurre en la segunda generación y 2) determinar en qué momento se produce este cambio si lo hay. Para este análisis intrageneracional, estudiaré el rasgo morfosintáctico estudiado por Otheguy y Zentella, o sea, la alternancia entre la presencia y la ausencia del pronombre personal sujeto.

Aunque Otheguy y Zentella estudiaron diferencias intergeneracionales, lo que no hicieron en su trabajo sobre los pronombres fue ver si existen igualmente diferencias de tasa pronominal y de las jerarquías de variables y de factores dentro de la primera generación. La contribución del presente trabajo es establecer si se detectan cambios en los hablantes de primera generación en situación de bilingüismo y determinar en qué momento se establece que hay cambios en las jerarquías de variables y de factores.

Para conseguir este propósito y como señalan autores como Poplack y Levey (2010: 19) o Thomason (2001: 134), en estudios intergeneracionales e intrageneracionales, como en este caso, donde se analiza un rasgo de variación lingüística como por ejemplo la variación del pronombre personal sujeto, no es suficiente con estudiar su frecuencia de uso, es decir, la tasa pronominal. Es necesario seguir una metodología comparativa capaz de detectar el cambio y distinguir entre las causas internas y externas (Poplack y Levey, 2010: 19). El enfoque variacionista entiende que la gramática de los hablantes está formada por las jerarquías de factores. El método comparativo variacionista establecido por Poplack y otros autores (2012) sirve para comparar frecuencias de uso de estructuras y de patrones gramaticales a través de las jerarquías de factores de cada variable. De modo que en este estudio voy a realizar análisis de tasas pronominales, jerarquías de

variables independientes externas e internas y jerarquías de factores en todos los hablantes inmigrantes y recién llegados del corpus del español en Nueva York de Otheguy y Zentella (2012).

2. Motivación: el contacto de lenguas

Parte de la literatura sobre el contacto de lenguas señala que la gramática es poco permeable a las influencias externas (Appel y Muysken, 1987; Meillet, 1921; Müller, 1871; Sapir, 1921; Silva-Corvalán, 1994, 1995a y 2001; Weinreich, 1953). La idea es que características de la estructura lingüística de la lengua de contacto no se incorporan en la sintaxis o morfología del sistema lingüístico de la lengua de pre-contacto. Los efectos del inglés sobre el español hablado en Estados Unidos constituyen un buen ejemplo de influencia externa. Silva-Corvalán (1994: 4) distingue entre la transferencia directa y la indirecta. La transferencia directa ocurre cuando se incorporan nuevas estructuras en la variedad de contacto. La transferencia indirecta implica cambios en una característica existente en la variedad de pre-contacto. Silva-Corvalán dice que la transferencia directa sucede raramente en situaciones de contacto. Y señala que no existe en el español hablado en Estados Unidos. Para que la transferencia directa ocurra, tiene que haber un cambio a nivel sintáctico que no interactúa con la semántica, el léxico o la pragmática (Silva-Corvalán, 2001: 12). Para la autora, éste es el nivel de gramática que es “remarkably impermeable to foreign influence” (1994: 166).

El inglés y el español son dos lenguas que tienen pronombres personales sujetos. Pero la distribución pronominal es diferente. La expresión del pronombre en inglés es más frecuente en la mayoría de los casos mientras que en español es muy variable. Silva-Corvalán considera que hay un proceso de transferencia indirecta en el español de los bilingües de Los Ángeles. Señala que el contacto con el inglés no está incorporando ninguna nueva estructura al sistema

pronominal español. Tampoco se está produciendo un cambio a nivel sintáctico. El contacto con el inglés puede expandir los contextos lingüísticos en que los pronombres personales sujetos aparecen (1994: 154-165). Según esto, la gramática de los hablantes de primera generación, a pesar del contacto con otras lenguas y variedades dialectales, no se modifica.

No obstante, parece que los hablantes de primera generación que pasan más tiempo en la ciudad realizan más pronombres sujetos de forma explícita (Otheguy, Zentella y Livert, 2007; Otheguy y Zentella, 2012). Este hallazgo indica que aunque no haya transferencia directa, sí hay una influencia gramatical en el español que modifica sutilmente el sistema gramatical de estos hablantes. Existe un proceso de convergencia gramatical donde el inglés refuerza una estructura existente en el español (Silva-Corvalán, 1994: 4-5; Zapata, Sánchez y Toribio, 2005).

3. Preguntas de investigación

Estas preguntas sobre si hay diferencias entre los hablantes de primera generación se van a hacer basándose en los dos grupos de exposición que contempla el corpus del español en Nueva York de Otheguy y Zentella (2012): los recién llegados y los inmigrantes. Atendiendo a los criterios de clasificación propuestos por estos autores para diferenciar a estos hablantes, mis preguntas tratan de detectar si hay diferencias en la realización del pronombre personal sujeto entre los nacidos en Latinoamérica que acaban de llegar a la ciudad y los que están más establecidos porque llevan más tiempo.

Respecto a las diferencias dialectales entre los hispanohablantes de Nueva York, el corpus tiene hablantes de diferentes orígenes nacionales y regionales. Esto representa una oportunidad para investigar no solamente si el grupo de hispanohablantes de primera generación es uniforme o tiene diferenciación interna, sino investigar también si existen también, con el paso

del tiempo en Nueva York dentro de la primera generación, diferencias de índole dialectal. Las preguntas son las siguientes:

- ¿Hay cambios en el uso de los pronombres personales entre la primera generación recién llegados y la primera generación inmigrantes que vive en Nueva York?
- ¿El contacto con el inglés y otras variedades dialectales del español habladas en Nueva York influye en la modificación de la gramática de la primera generación inmigrantes como ocurre con la segunda generación?
- ¿Se mantiene igual la gramática de la primera generación recién llegados que la gramática de la primera generación inmigrantes?
- El cambio que pueda registrarse, ¿es igual en todos los hablantes de primera generación?
- ¿Hay diferencias en la tasa pronominal entre la primera generación recién llegados y la primera generación inmigrantes?
- ¿Cuáles son las variables externas e internas que producen un efecto estadístico predictivo en la realización del pronombre personal sujeto en la primera generación?
¿Son las mismas para cada grupo de exposición?
- ¿Hay diferencias en las jerarquías de variables entre la primera generación recién llegados y la primera generación inmigrantes?, ¿y en las jerarquías de factores?
- ¿En qué momento, o sea, a cuántos años de vivir en Nueva York, se establece que hay cambio en las jerarquías de variables y de factores, si los hay?
- ¿Incide llegar más joven o más mayor en el uso del pronombre en los dos grupos de exposición?
- ¿Hay diferencias respecto al país de origen en el uso del pronombre? ¿Y respecto al área regional?

4. Revisión de la literatura

En general, muchas de las investigaciones sobre los pronombres sujetos en español han estudiado su uso variable en el español en Estados Unidos, Latinoamérica y España. El interés por esta variable lingüística se debe a la existencia de diferencias dialectales y regionales en el español peninsular y latinoamericano. Se han escrito numerosos artículos, libros, capítulos de libros y tesis doctorales. Pero ninguno de ellos se pregunta si todos los hablantes de la primera generación en Nueva York muestran el mismo comportamiento en el uso del pronombre. La mayor parte de estos trabajos establece cuáles son los factores internos y externos que influyen y se correlacionan con el uso explícito del pronombre.

Se han realizado investigaciones sobre la alternancia entre la ausencia y la presencia del pronombre sujeto en casi todos los países de habla hispana incluyendo Estados Unidos (Morales, 1986; Bentivoglio, 1987; Silva-Corvalán, 1982, 1994, 1997; Cameron y Flores-Ferrán, 2004; Otheguy, Zentella y Livert, 2007; Ortiz, 2008; Otheguy y Zentella, 2007, 2012).

Todos los estudios sobre los pronombres personales sujetos se han hecho sobre distintas variedades de español y sobre el español en contacto con otras lenguas. En el caso del español en Estados Unidos, la mayoría de los trabajos ha analizado este fenómeno de variación como un elemento diferenciador entre la primera y segunda generación. La segunda generación es diferente en la realización del pronombre en comparación con la primera generación. Las escasas investigaciones que analizan este rasgo lingüístico sólo en la primera generación no señalan si estos hablantes muestran también un comportamiento diferente en la expresión del pronombre (Otheguy, Zentella y Livert, 2007; Otheguy y Zentella, 2012). No existen trabajos que, estudiando la primera generación exclusivamente, determinen si la gramática de estos hablantes

se modifica en situaciones de lenguas en contacto como se ha mostrado con la segunda generación.

Barrenechea y Alonso (1969, 1977) y Pérez Sala (1973) fueron de los primeros investigadores que estudiaron el uso variable del pronombre personal sujeto en las hablas de Argentina y Puerto Rico. En ambos casos, los autores se limitaron a realizar análisis cuantitativos sobre la presencia del pronombre personal sujeto en discursos orales.

A partir de estos estudios, se comienzan a realizar investigaciones que analizan qué factores influyen más en la expresión del pronombre. El propósito de estos trabajos sociolingüísticos es determinar la importancia de los factores lingüísticos, sociales y estilísticos que intervienen en el hablante a la hora de expresar u omitir el pronombre personal sujeto (Cantero Sandoval, 1976; Morales, 1980, 1982, 1986, 1997; Enríquez, 1984; Bentivoglio, 1987; Cameron, 1992, 1993, 1995). Estos estudios ponen en evidencia la fuerte correlación entre los factores lingüísticos y el uso explícito del pronombre.

En general, las investigaciones sobre la variación del pronombre personal sujeto coinciden en señalar que los hablantes bilingües y monolingües tienden más a omitir el pronombre en el discurso oral y escrito que a usarlo explícitamente. No obstante, existe una serie de variables lingüísticas internas que condicionan el uso del pronombre (Cantero Sandoval, 1976; Morales, 1980, 1982, 1986; Enríquez, 1984; Bentivoglio, 1987; Cameron, 1992, 1993, 1995; Silva-Corvalán, 1994; Bayley y Pease-Álvarez, 1996, 1997; Flores y Toro, 2000; Flores-Ferrán, 2002, 2005; Cabré-Sans y Gavarró, 2006; Carvalho y Child, 2011; Otheguy y Zentella, 2012, Carvalho, Shin y Orozco, 2015).

En algunos estudios se observa que la presencia del pronombre aumenta con verbos conjugados en ciertas personas en relación a otras. La persona y el número del verbo es una de

las variables independientes internas que incide en la presencia del pronombre. Hay una tendencia a expresar la primera persona de singular *yo* frente a los demás pronombres. Sin embargo, los hablantes desfavorecen el uso del pronombre porque la mayoría de los verbos aparece sin pronombre (Shin y Otheguy, 2009, 2013; Silva-Corvalán, 1982).

Otros tipos de trabajo sobre los pronombres sujetos han analizado la influencia del contenido semántico del verbo en el uso explícito del pronombre. Los verbos que se refieren a actividades mentales (ej., *ella piensa*) o de estado (ej., *ella es profesora*) tienden a aparecer con un pronombre personal sujeto más que otro tipo de verbos (Travis, 2007; Carvalho y Child, 2011; Otheguy y Zentella, 2012).

En muchos de estos estudios se observa que la presencia del pronombre aumenta también cuando hay cambio de referencia en relación a cuando la referencia sigue siendo la misma. Se ha comprobado que el cambio de referente es un factor que sistemáticamente condiciona el uso explícito del pronombre personal (Morales 1986; Cameron 1992, 1995; Silva-Corvalán 1994; Bayley y Pease-Álvarez 1996; Flores-Ferrán 2002; Cameron y Flores-Ferrán, 2004; Otheguy y Zentella, 2012).

También, se ha prestado especial atención a examinar los factores que influyen en su posición preverbal o posverbal (Silva-Corvalán, 1982). Algunos autores se han interesado por el estudio de las causas de la aparición del sujeto pronominal en el español analizando las implicaciones de la hipótesis de compensación funcional (Hochberg, 1986; Morales, 1997; Erker, 2010). Además, se han realizado trabajos para estudiar el fenómeno de la llamada ‘perseveración’ y su importancia en la realización u omisión de los pronombres sujetos en español (Cameron, 1994; Cameron y Flores-Ferrán, 2004; Otheguy, 2014).

Aparte de las numerosas investigaciones que analizan las variables lingüísticas internas, existe otro tipo de trabajos. Hay estudios que dirigen su atención a los factores pragmáticos del discurso. Principalmente, se ha estudiado qué importancia tienen el tópico, el peso pragmático o el énfasis en la variación entre la presencia y la ausencia del pronombre (Montes Miró, 1986; Morales, 1986; Silva-Corvalán, 1994; Davidson, 1996; Stewart, 2003; Matos Amaral y Schwenter, 2005).

Otros trabajos analizan las variables sociales y estilísticas que afectan a este fenómeno de variación lingüística. Los resultados se muestran poco significativos respecto a la influencia en la alternancia del pronombre. Algunas de las variables estilísticas que se han analizado son el discurso formal y el discurso informal, el estilo narrativo y la conversación natural, y las formas de tratamiento formal e informal entre otras (Baumgardner, 1982; Moser, 2002; Hurtado, 2005; Bayona, 2006).

También, hay estudios que investigan si esta alternancia se relaciona con variables independientes externas o sociales. Se han analizado el sexo, la edad, el nivel de educación y cómo influyen en la expresión u omisión del pronombre (Barrenechea y Alonso, 1977; Enríquez, 1984; Morales, 1986; Bentivoglio, 1987; Cameron, 1992; Silva-Corvalán, 1994; Ávila-Jiménez, 1996; Bayley y Pease-Álvarez, 1996; Flores-Ferrán 2002, 2004, 2005; Hurtado, 2005; Otheguy y Zentella, 2012; Shin y Otheguy, 2009, 2013).

Una de las variables independientes externas que más atención ha recibido en los trabajos sobre el español en Estados Unidos y Puerto Rico es el contacto con el inglés y los efectos que tiene en la realización pronominal. Muchos autores parten de la noción de que el inglés requiere el uso expreso del pronombre personal sujeto en casi todos los contextos (Lapidus y Otheguy, 2005). La idea es que la exposición constante con el inglés se refleja en un mayor uso explícito

del pronombre personal en los hablantes de español. Muchas investigaciones hacen correlaciones de variables externas como el tiempo de exposición y el conocimiento de inglés, la edad de llegada a Estados Unidos, los años vividos en este país o la escolarización recibida en inglés (Pérez Sala, 1973; Morales, 1986; Ávila-Jiménez, 1996; Bayley y Pease-Álvarez, 1996, 1997; Silva-Corvalán, 1997, 1994; Flores y Toro, 2000; Flores-Ferrán, 2002, 2004, 2005; Otheguy y Zentella 2007; Shin y Otheguy, 2009; Otheguy y Zentella, 2012; Shin y Montes-Alcalá, 2014).

En otros trabajos sobre las diferencias dialectales se ha analizado que el español en el Caribe presenta rasgos distintos en cuanto al uso pronominal (Álvarez Nazario, 1990; Lipski, 1994). Estos estudios han comprobado que hay una mayor proporción de sujetos explícitos en las hablas latinoamericanas y que dentro de éstas las caribeñas son más propensas a expresar el pronombre sujeto (Cameron, 1992, 1994; Otheguy y Zentella, 2007, 2012; Posio, 2011).

En resumen, todos estos estudios investigan las variables lingüísticas internas, externas, estilísticas y pragmáticas que favorecen la expresión del pronombre personal sujeto. Algunos analizan diferentes variedades de español. Otros estudian el español en contacto con otras lenguas. Hay trabajos intergeneracionales e intrageneracionales sobre la variación del pronombre. Pero ninguna de estas investigaciones ha estudiado si todos los hablantes de primera generación en Nueva York presentan el mismo comportamiento a la hora de expresar el pronombre personal sujeto. Aún no se ha analizado si la primera generación también modifica su gramática en situaciones de lenguas en contacto.

5. Hipótesis y previsión de los resultados

Como señalé arriba, algunos estudios sobre el español en Estados Unidos han mostrado que los hablantes bilingües de segunda generación tienden a modificar su gramática en situaciones de lenguas en contacto (Bookhamer, 2012; Flores-Ferrán, 2004; Gutiérrez, 1995;

Montrul, 2002, 2004, 2009; Silva-Corvalán, 1994, 1995a, 1995b; Travis, 2007). Ninguno de estos trabajos ha examinado si el español hablado en Nueva York de la primera generación sufre también cambios en el sistema gramatical como consecuencia del contacto con el inglés y otras variedades de español. En esta tesis, se demuestra que los hablantes de primera generación en Nueva York, inmigrantes y recién llegados, exhiben diferentes pautas en la utilización del pronombre personal sujeto. La hipótesis es que estos dos grupos de hablantes, que se diferencian por el tiempo de exposición a la ciudad de Nueva York, muestran diferencias en el uso pronominal en gran parte por el incremento del uso y contacto con el inglés.

Los trabajos mencionados anteriormente obvian el hecho de que a pesar de pertenecer a la misma generación por haber nacido en Latinoamérica, estos hablantes no conforman un grupo homogéneo. Es importante distinguir, por ejemplo, entre un hablante que acaba de llegar a Estados Unidos y otro que lleva establecido desde hace mucho tiempo. Probablemente, existen variables y factores que afecten de manera distinta a cada uno de ellos.

En esta investigación hay dos líneas de hallazgos que derivan de la hipótesis de la nivelación dialectal y de la hipótesis de las lenguas en contacto. Mis análisis confirman que hay diferencias dialectales en el discurso de los recién llegados y los inmigrantes en Nueva York. He encontrado más presencia del pronombre en los caribeños que en los continentales como observaron Otheguy y Zentella (2012). También, he observado que, con el paso del tiempo en la ciudad, los hablantes de español cada vez más expuestos al inglés exhiben los modelos de uso del inglés en el uso pronominal. Esto conlleva a un aumento en la ocurrencia del pronombre.

Además, establezco que hay una fuerte correlación entre la tasa pronominal y pertenecer a unos de los dos grupos de exposición. Esto significa que los hablantes que usan más pronombres son los hablantes que han estado más expuesto al ambiente neoyorquino. En relación

a la exposición, la tesis muestra que hay un incremento importante del uso expreso del pronombre en informantes que llevan más años vividos aquí, especialmente si son caribeños e inmigrantes.

El propósito de la tesis es establecer que la primera generación sufre también modificaciones en el sistema gramatical. En términos generales, muestro que los inmigrantes de la primera generación presentan una gramática diferente a la de los recién llegados. Estos hablantes inmigrantes son diferentes de los recién llegados porque favorecen ciertas estrategias sobre otras en el uso del pronombre debido a la influencia del inglés.

6. Resumen del estudio

En el capítulo dos, presento la metodología utilizada en el estudio. Hablo sobre la muestra de informantes y explico cuales son las variables dependientes e independientes. El capítulo tres describe la tasa pronominal de la primera generación. Contiene distribuciones del uso del pronombre y analiza el impacto de varias variables sociodemográficas y de uso de la lengua. Los capítulos cuatro y cinco exploran qué variables independientes externas e internas afectan al uso del pronombre en los recién llegados y los inmigrantes. También, muestro los resultados de los análisis variacionistas hechos a través de las jerarquías de factores de las variables independientes internas. Finalmente, el capítulo seis resume los resultados encontrados y ofrece algunas conclusiones e interpretaciones de los datos analizados en los capítulos previos.

7. Relevancia del estudio

Este estudio es una contribución más a la inmensa literatura sobre el español en Estados Unidos y, sobre todo, en Nueva York. La aportación de este trabajo provee un mejor conocimiento del perfil sociolingüístico de los latinos en EE.UU así como del fenómeno del contacto lingüístico en general. Específicamente, el estudio dirige la atención sobre el hecho de

que los desarrollos diacrónicos que ocurren en la lengua como resultado del contacto lingüístico no siempre resultan en la perseverancia de modelos, sino en transformaciones detectables que tienen lugar a través de generaciones dentro de la misma comunidad. También, subraya la importancia de explorar explicaciones alternativas sobre los cambios gramaticales antes de concluir que la gramática de la primera generación no se modifica porque es poco permeable a las influencias externas en situaciones de contacto.

Capítulo 2

Metodología

1. Introducción

Este capítulo presenta la metodología y los materiales utilizados para responder a las preguntas de investigación. Describe la muestra usada para el estudio, incluyendo una descripción detallada de los participantes y las entrevistas orales. También, incluye información acerca del llamado sobre de variación (Labov, 1972) y del programa usado para realizar los análisis estadísticos que determinan la relación entre las variables predictivas y el uso del pronombre. Finalmente, explica cuáles son las variables dependientes e independientes y sus correspondientes factores.

2. La muestra

Los datos de este trabajo forman parte del corpus del español en Nueva York de Otheguy y Zentella (2012). El corpus está compuesto de 140 transcripciones de entrevistas sociolingüísticas realizadas a latinos que viven en Nueva York entre 2000 y 2005. La muestra de esta tesis está formada por 114 de esas entrevistas sociolingüísticas. El criterio para seleccionar este número de transcripciones se basa en que las 114 son entrevistas realizadas a informantes que conforman el grupo de *primera generación*, es decir, los latinos que nacieron en Latinoamérica y que son clasificados por Otheguy y Zentella (2012: 33; 86) según el nivel de exposición a la ciudad de Nueva York en *inmigrantes recién llegados* e *inmigrantes establecidos*. La muestra está estratificada por diferentes características socio-demográficas incluyendo región, origen nacional, género, edad, edad de llegada, años en Nueva York, educación, clase social y estatus socioeconómico. También por una característica del uso de la lengua: la competencia en inglés.

2.1 Las entrevistas

Cada entrevista dura aproximadamente entre 60 y 90 minutos. Fueron realizadas por entrevistadores latinos cuyo origen nacional era el mismo que el del entrevistado. Por ejemplo, si el informante era puertorriqueño, el entrevistador también. Los temas tratados eran abiertos. La idea era mantener una conversación con un estilo natural usando preguntas y técnicas de entrevista en sociolingüística (Labov, 1972). La mayoría de las veces, se hablaba sobre el país de origen, las razones del proceso de inmigración, las experiencias como inmigrantes en Estados Unidos, el primer día en el país de acogida, la vida sentimental, recuerdos de la infancia, etc. Todas las entrevistas fueron hechas en español, pero en muchas ocasiones los entrevistados usaron préstamos de palabras inglesas y hubo cambio de código al inglés.

2.2 Composición de la muestra de este estudio

La muestra fue elaborada en relación a unos criterios de selección con la finalidad de conseguir una muestra estratificada. La idea es conseguir un número equilibrado de hombres y mujeres y participantes de diferentes niveles socioeconómicos, educativos y áreas geográficas. Los criterios de estratificación son los siguientes: género, edad, origen nacional y regional, exposición a Nueva York, clase social, educación y competencia en inglés.

2.2.1 Género

El género es una variable muy analizada en estudios de sociolingüística. En general, se muestra que las mujeres son más conservadoras que los hombres y estos últimos son los que suelen promover el cambio lingüístico (Labov, 2001: 261). En esta muestra, hay 54 hombres (un 47 por ciento de la muestra) y 60 mujeres (un 53 por ciento de la muestra).

2.2.2 Edad

Hay diferentes rangos de edades. El participante más joven tiene 15 años y el más mayor tiene 80. La muestra contiene siete hablantes por debajo de los 20 años, 69 informantes entre los 20 y 40 años, 31 que tienen entre los 40 y 60 años y siete con más de 60 años.

2.2.3 Exposición a Nueva York

En la muestra, este criterio de estratificación es fundamental para nuestra investigación. Es importante considerar la existencia de dos grupos de latinoamericanos que se diferencian en el nivel de exposición a la ciudad de Nueva York: los *inmigrantes establecidos* e *inmigrantes recién llegados*. Esta distinción hecha en los grupos de exposición no sólo se basa en la edad de llegada y los años vividos en Nueva York, sino también en un criterio lingüístico. Es decir, más expuesto o menos expuesto no es solamente una cuestión de tiempo o lugar de nacimiento sino una medida bilingüe susceptible de cambiar. El grupo de exposición recién llegados agrupa a los informantes que llegaron con 17 años o más y han vivido en Estados Unidos por cinco años o menos (Otheguy y Zentella, 2012: 33; 86). Estos hablantes son menos susceptibles a la influencia del inglés y otras formas de español en Nueva York porque llegaron a una edad madura lingüísticamente y estuvieron poco expuestos al ambiente de la ciudad. Eso hace que sean más representativos del español de sus países y regiones de origen.

Los hablantes recién llegados de la muestra tienen como promedio 30 años y llevan menos de tres años en la ciudad. Representan el 34 por ciento de la muestra. El grupo de exposición inmigrantes reúne a aquellos que llegaron cuando tenían menos de 17 años y pasaron más de cinco años en el país norteamericano (Otheguy y Zentella, 2012: 33; 86). Este grupo es más susceptible a la influencia del inglés y a otras variedades de español. Forman el 66 por

ciento de la muestra. La siguiente tabla muestra el número total de informantes por cada grupo de exposición:

	N Participantes	Porcentaje
Recién llegados	39	34
Inmigrantes	75	66
Total	114	100

2.2.4 Origen nacional y regional

La procedencia geográfica del informante es también un rasgo muy estudiado en los trabajos de sociolingüística sobre el español en Latinoamérica (Lipski, 1994: 241; López Morales, 1992: 137; Ordoñez y Olarrea, 2001; Toribio 2000; Zogona, 2002; Goodall, 2004; Otheguy y Zentella, 2012: 143). Los participantes proceden de seis grupos nacionales diferentes: colombianos, cubanos, ecuatorianos, mexicanos, puertorriqueños y dominicanos. Estos seis grupos corresponden a los seis grupos nacionales de latinos más grandes en Nueva York. Además, estos seis grupos nacionales pueden ser agrupados en dos grupos regionales principales: los caribeños y los continentales. Esta distinción representa una de las líneas principales en dialectología que divide al español hablado en Latinoamérica. Los colombianos, ecuatorianos y mexicanos forman el grupo de los hablantes continentales. Los cubanos, dominicanos y puertorriqueños constituyen el grupo de caribeños. Las tablas 2 y 3 presentan la distribución de los hablantes inmigrantes (en la tabla *INM.*) y recién llegados (en la tabla *R.LL.*) según el origen nacional y el área regional.

Tabla 2. Total de participantes por grupos de exposición y origen nacional

	R.LL.	Porcentaje	INM.	Porcentaje
Colombianos	6	15	10	14
Cubanos	7	19	13	17
Ecuatorianos	8	21	12	16
Mexicanos	6	15	13	17
Puertorriqueños	6	15	12	16
Dominicanos	6	15	15	20
Total	39	100	75	100

Tabla 3. Total de participantes por grupos de exposición y área regional

	R.LL.	Porcentaje	INM.	Porcentaje
Caribeños	19	49	40	53
Continetales	20	51	35	47
Total	39	100	75	100

2.2.5 Clase social y educación

La clase social y la educación también son consideradas predictores en la variación en relación a muchos tipos de comportamiento lingüístico. Los hablantes de la muestra representan diferentes clases sociales y niveles de educación. Hay 57 informantes (el 50 por ciento de la muestra) que representan a la clase media-alta y otros 57 (el otro 50 por ciento) que se describen como la clase trabajadora. También el 55 por ciento (62 informantes de la muestra) va a la universidad o ya está graduado, mientras que el 45 por ciento (los otros 52 informantes) tiene un nivel elemental o está aún en la secundaria.

2.2.6 Competencia en inglés

Ya que la influencia del inglés sobre el español es una de las preguntas de investigación, este criterio de estratificación es importante. Las diferencias en la competencia en inglés son útiles para diferenciar a estos hablantes. Del total de los 114 informantes, 90 (el 79 por ciento de

la muestra) definen su competencia en inglés menos que excelente. El resto, 24 informantes que representan el 21 por ciento, dice que es excelente.

3. El sobre de variación

El estudio variable del uso del pronombre personal sujeto con verbos de forma personal requiere considerar todos los verbos que aparecen en las 114 entrevistas para distinguir aquellos que aparecen en contextos de variación entre la presencia y la ausencia del pronombre personal sujeto, incluidos en el estudio, y aquellos que aparecen en contextos no variables, es decir, contextos donde el pronombre está presente o ausente casi siempre. Los contextos, donde es posible la variación, se llaman variables. Los contextos, donde no es posible la variación, son categóricos.

El siguiente texto extraído de una entrevista contiene ejemplos de verbos encontrados en contextos variables (subrayados) y verbos encontrados en contextos categóricos (en cursiva).

1. [Cuenta su experiencia de cómo vivían en la casa familiar] Bueno, aquí, me acuerdo que cuando al principio yo creo que tenía muchas responsabilidades, porque mi papá *trabajaba* y mami *ayudaba* a papi, ah... vivíamos en el mismo departamento que vivimos ahora, el cual *estaba* ocupado por qué sé yo cuánta gente, ah... familia de papi, los cuales *se fueron, mudaron* poco a poco a sus propios apartamentos, pero al principio creo que *habían* sólo dos apartamentos, los cuales los *ocupaba* la familia Pérez. [319D]

El texto 1 contiene catorce verbos, siete que aparecen en contextos variables y que se incluyen en la muestra, y siete que están en contextos categóricos, que se excluyen del estudio. Los verbos incluidos son *me acuerdo*, *yo creo*, *tenía*, *vivíamos*, *vivimos*, *sé yo* y *creo*. Los verbos que no son elegibles para el estudio son *trabajaba*, *ayudaba*, *estaba*, *se fueron*, *mudaron*, *habían* y *ocupaba*. Los siete verbos que aparecen en contextos variables son elegibles porque ocurren con pronombre pero podrían haberse usado sin pronombre, o de lo contrario ocurren sin un pronombre pero podrían haberse utilizado con él. Siguiendo la terminología en sociolingüística variacionista, estos verbos están dentro del sobre de variación (Lavob, 1972; Tagliamonte, 2006:

86). Los siete verbos que se excluyen aparecen en contextos categóricos porque no van acompañados de un pronombre personal sujeto y porque, además, difícilmente podrían ocurrir con un pronombre sujeto. Tres de los verbos, *trabajaba*, *ayudaba* y *ocupaba*, no pueden llevar un pronombre personal sujeto porque ya tienen un sujeto nominal: *mi papá*, *mami* y *la familia Pérez*. Los otros cuatro verbos *estaba*, *se fueron*, *mudaron* y *habían* tampoco pueden tener un pronombre personal sujeto porque los tres primeros tienen como sujeto un pronombre de relativo y el cuarto verbo es impersonal. Ninguno de estos siete verbos son elegibles porque aparecen en tipos de contextos no variables. Estos verbos están fuera del sobre de variación.

En resumen, voy a considerar todas las ocurrencias donde hay un verbo conjugado en cualquier tiempo verbal que ocurre con o sin un pronombre personal sujeto. Los verbos elegibles de este trabajo son aquellos que aparecen con un pronombre personal sujeto pero podrían haber aparecido sin él o que ocurren sin el pronombre pero pueden aparecer con el mismo. En otras palabras, voy a seleccionar todas las ocurrencias de cada participante donde aparece un verbo conjugado y donde el uso del pronombre es variable. Es decir, donde es posible la ausencia o la presencia de un pronombre personal sujeto.

4. Metodología

Para el análisis estadístico de los datos usaré el *Paquete Estadístico para las Ciencias Sociales*, conocido comúnmente como *SPSS*, por sus siglas en inglés.

Recuerdo que este trabajo se centra en el estudio de las variables sociales y lingüísticas y de sus factores que motivan a los hablantes de los grupos de exposición inmigrantes establecidos e inmigrantes recién llegados a elegir entre la presencia o la ausencia del pronombre personal sujeto. Por lo tanto, realizaré análisis estadísticos para explorar el español en Nueva York de la primera generación.

Para estudiar el impacto de las variables externas en estos dos grupos, voy a hacer ANOVAS. Este método de análisis estadístico toma en consideración la *tasa pronominal* de los individuos en los grupos comparándolos y determina si las diferencias entre los grupos son mayores que las diferencias dentro de cada grupo. Por tasa pronominal entendemos la proporción o el porcentaje de verbos de forma personal que aparecen con un pronombre personal sujeto (Otheguy y Zentella, 2012: 68). También, haré regresiones lineales. Estas regresiones informan qué variables externas son más importantes.

Para analizar el orden de las variables independientes internas y el orden de los factores que afectan significativamente el uso de los pronombres personales sujetos en el español de la primera generación, haré uso de los análisis multivariantes. También, realizaré jerarquías de variables independientes internas y jerarquías de factores a través de regresiones logísticas que señalan qué variables internas y factores son significativos y qué orden tienen.

5. Las variables

La alternancia entre la presencia y la ausencia del pronombre personal sujeto está condicionada por una serie de variables lingüísticas y sociales. Este condicionamiento está sujeto a cambios en situaciones de lenguas en contacto (Klee y Lynch, 2009; Montrul, 2004; Silva-Corvalán, 1994; Toribio, 2004). A continuación, presento las variables que se analizan en el trabajo. Estas variables son variables independientes internas y externas (sociodemográficas y de uso de la lengua). Primero, definiré la variable dependiente del estudio. Seguidamente, haré una descripción de las variables externas e internas. Por último, definiré qué factores tiene cada una de estas variables.

5.1 Las variables dependientes

El rasgo morfosintáctico que se estudia es la presencia o la ausencia de los pronombres personales en el discurso oral. La variable dependiente para las variables independientes externas es la tasa pronominal. La variable consiste en el porcentaje de pronombres de cada informante. La variable dependiente para las variables independientes internas se llama <tasa pronominal>. La tasa pronominal es la proporción o el porcentaje de verbos de forma personal que aparecen con un pronombre personal sujeto (Otheguy y Zentella, 2012: 68). Tiene los siguientes factores:

1ª Factor: Presencia del pronombre, e.g. *yo canto, tú cantas, ella canta, etc.*

2ª Factor: Ausencia del pronombre, e.g. *canto, cantas, canta, etc.*

5.2 Las variables independientes externas

Las variables independientes externas se refieren a las causas externas a la estructura de la lengua. Algunas de las diferentes características sociodemográficas de la muestra fueron elegidas como variables externas. Es decir, muchos de los criterios usados para la selección de los participantes de la muestra se usaron como variables externas predictivas en el análisis. A continuación, presento una lista con las variables externas sociodemográficas y de uso de la lengua. Cada variable aparece junto a sus factores.

Origen nacional:

- Colombiano
- Cubano
- Dominicano
- Ecuatoriano
- Mexicano
- Puertorriqueño

Origen regional:

- Caribeño
- Continental

Género:

- Hombre
- Mujer

Edad:

- Jóvenes
- Mayores

Edad de llegada:

- Vino a EE.UU con o antes de 12 años
- Vino a EE.UU después de 12 años

Educación:

- Elemental e instituto
- Universidad y graduado

Clase social:

- Alta y media
- Trabajadora

Nivel de estatus socioeconómico (ESE):

- Bajo estatus socioeconómico
- Alto estatus socioeconómico

Competencia en inglés:

- Inglés menos que excelente

- Inglés excelente

Hay muchas razones por las que consideré estas características sociodemográficas y de uso de la lengua como variables predictivas. Algunas de las características sociodemográficas se seleccionaron específicamente para responder las preguntas de investigación como origen regional y nacional. Otras características de uso de la lengua como la competencia en inglés también se considera una variable predictiva para responder la pregunta sobre la influencia del inglés. El resto de las características sociodemográficas se incluyen como variables predictivas porque en estudios anteriores se ha demostrado que tienen una relación importante en el comportamiento lingüístico de los hablantes en modos muy significativos. La clase social y la educación son variables analizadas en muchos trabajos que demuestran una relación muy importante entre la variabilidad en español y el nivel de educación de los hablantes, su ocupación, sus ingresos económicos o su barrio (Alba, 1988; Cedgren, 1978; Lafford, 1986; Samper-Padilla, 1990). La edad se incluye como variable sociodemográfica predictiva, ya que algunos estudios indican que determinadas variables lingüísticas son condicionadas por la edad (Alba, 1988; Lafford, 1986; Samper-Padilla, 1990). También, el género es analizado como una variable predictiva por las diferencias cuantitativas y categóricas encontradas en el discurso de los hombres y las mujeres (Navas Sánchez-Elez, 1997; Rissel, 1989; Valdivieso y Magaña, 1988).

5.3 Las variables independientes internas

También, voy a tener en cuenta una serie de lo que se llaman en sociolingüística variables independientes internas. Las variables independientes internas se refieren a las causas internas que se relacionan con la estructura de la lengua. Son factores lingüísticos como la persona del verbo o el significado léxico. Las variables independientes internas de este estudio forman un

total de cinco. Éstas son las variables independientes internas: persona del verbo, tiempo, referencia, significado léxico y reflexividad. La razón de haber seleccionado éstas y no otras tiene que ver con sus propiedades morfológicas y semánticas, ideales para la variabilidad entre la presencia o ausencia del pronombre. En trabajos previos, se ha comprobado que estas variables influyen en la distribución de los pronombres (Morales, 1986; Bayley y Pease-Álvarez, 1996, 1997; Silva-Corvalán, 1982, 1994). Abajo presento una lista de las variables independientes internas con sus factores seguidas de ejemplos.

Persona del verbo:

- Primera persona singular

Ejemplo: *Yo vivo en Nueva York.*

- Segunda persona singular

Ejemplo: *Tú vives en Nueva York.*

- Tercera persona singular

Ejemplo: *Ella vive en Nueva York.*

- Primera persona plural

Ejemplo: *Nosotros vivimos en Nueva York.*

- Tercera persona plural

Ejemplo: *Ellos viven en Nueva York.*

Tiempo:

- Presente indicativo

Ejemplo: *Yo vivo en Nueva York.*

- Pasado indicativo

Ejemplo: *Yo viví/ vivía/he vivido en Nueva York.*

- Futuro indicativo

Ejemplo: *Yo voy a vivir/ viviré en Nueva York.*

- Condicional indicativo

Ejemplo: *Yo viviría en Nueva York.*

- Modo Subjuntivo

Ejemplo: *Espero que ella viva en Nueva York.*

Referencia:

- Cambio referencia

Ejemplo: *Yo vivo en Nueva York, pero ella viaja mucho.*

- Misma referencia

Ejemplo: *Yo vivo en Nueva York, pero yo no viajo mucho.*

Significado léxico:

- Verbo mental

Ejemplo: *Yo pienso que vivir en Nueva York está bien.*

- Verbo estimativo

Ejemplo: *Yo estimo que en Nueva York se vive bien.*

- Verbo de estado

Ejemplo: *Me parece bien vivir en Nueva York.*

- Verbo externo

Ejemplo: *Yo vivo en Nueva York.*

Reflexividad:

- Verbo reflexivo

Ejemplo: *Yo me llamo Lina y vivo en Nueva York.*

- Verbo no reflexivo

Ejemplo: *Yo soy Lina y vivo en Nueva York.*

Capítulo 3

El español de los inmigrantes y los recién llegados a través de la tasa pronominal

1. Introducción

En este capítulo estudio si los *grupos de exposición* que denominamos *inmigrantes establecidos e inmigrantes recién llegados*¹, muestran un comportamiento diferente en el uso del pronombre personal sujeto. Me refiero a estos dos modos de distinción entre inmigrantes establecidos e inmigrantes recién llegados como una diferencia en la *exposición* a la ciudad de Nueva York. Según ya se ha indicado en el capítulo 2 que trata de la metodología del presente estudio, esta distinción hecha en los grupos de exposición no sólo se basa en la edad de llegada y los años vividos en Nueva York, sino también en un criterio lingüístico. Es decir, más expuesto o menos expuesto no es una cuestión de tiempo o lugar de nacimiento sino una medida bilingüe susceptible de cambiar. Así el grupo de exposición de recién llegados agrupa a los informantes que llegaron con 17 años o más y han vivido en Estados Unidos por cinco años o menos. El grupo de exposición de inmigrantes reúne a aquellos que llegaron cuando tenían menos de 17 años y pasaron más de cinco años en el país norteamericano (Otheguy y Zentella, 2012: 33; 86).

La predicción es que los inmigrantes recién llegados deberían presentar una *tasa pronominal* más baja que la de los inmigrantes establecidos debido a la reducida exposición de los recién llegados a la ciudad y al escaso contacto con el inglés. La tasa pronominal, como ya hemos indicado en el capítulo 2 que trata de la metodología, es la proporción o el porcentaje de verbos de forma personal que aparecen con un pronombre personal sujeto (Otheguy y Zentella, 2012: 68). El siguiente pasaje de una de las transcripciones de la muestra ofrece un ejemplo de la tasa pronominal.

¹ A partir de ahora utilizaré inmigrantes para referirme a los inmigrantes establecidos y recién llegados para hablar sobre los inmigrantes recién llegados.

1a. [Habla de cómo cuidaba a sus hermanos en su niñez] **Yo creo** que **yo la asumía** yo sola, **yo no me acuerdo** que nadie me dijera: “**tienes** que ayudarlo, **tienes** que hacer esto”, aunque siempre había un aire de que cada quien aquí está ayudando a cada uno, porque **estamos** juntos y **tú tienes** hermanos y **yo soy** tu mamá y **yo ayudo** a tu papá, pero no me **decían**: “**tienes** que hacer esto”. [319D]

El texto, que corresponde a una informante de la República Dominicana del grupo de exposición inmigrantes establecidos, contiene 11 verbos elegibles, de los que seis aparecen con pronombre explícitamente y cinco lo omiten. La tasa pronominal de este extracto de entrevista es 6/11 o 55 por ciento.

El rasgo morfosintáctico que se estudia es la variación entre la presencia y la ausencia del pronombre personal sujeto español con verbos de forma personal (ejemplo: *creo ~ yo creo, estamos ~ nosotros estamos*). Esta característica del español en general, y específicamente del español de los latinoamericanos en Nueva York, hace que los verbos no se usen siempre con, o siempre sin, pronombres personales sujetos. Esta diferenciación lleva a la siguiente pregunta: ¿se pueden distinguir a los inmigrantes establecidos y los recién llegados por la frecuencia de uso del pronombre con verbos conjugados?

Los inmigrantes recién llegados, que han tenido un extenso y reciente contacto con sus países de origen, deberían constituir el grupo de latinos en Nueva York cuyo uso del pronombre es menos parecido al inglés y más similar al español. La idea es que el español hablado por los latinos bilingües puede estar influenciado por el inglés. Este contacto de lenguas induce al cambio en el sistema pronominal español. En situaciones de bilingüismo en Estados Unidos, se ha mostrado que los hablantes que alteran su gramática (en cuanto a su sintaxis o morfología) son mayormente los latinos que nacieron en este país y que conforman la *segunda generación* (Bookhamer, 2012; Flores-Ferrán, 2004; Gutiérrez, 1995; Montrul, 2002, 2004, 2009; Silva-Corvalán, 1994, 1995a, 1995b; Travis, 2007). En estudios anteriores sobre el pronombre personal sujeto, se observa que los hablantes latinos que nacieron en Estados Unidos, la *segunda*

generación, usan de forma explícita el pronombre personal sujeto más frecuentemente que los hablantes que nacieron en Latinoamérica (Flores-Ferrán, 2004; Montrul, 2004; Otheguy y Zentella, 2012; Silva-Corvalán, 1982, 1994; Travis, 2007). Voy a determinar, a través del estudio de su tasa pronominal, si la gramática de los hablantes inmigrantes y recién llegados que viven en Nueva York se modifica de la misma manera que la gramática de los hablantes latinos que nacieron en Estados Unidos. Trato de averiguar si 1) estos dos grupos de exposición dentro de la primera generación sufren cambios paralelos en la gramática como ocurre en la segunda generación y 2) de determinar en qué momento se produce este cambio si lo hay. Quiero ver si estas dos sub-comunidades de latinos de primera generación en la ciudad de Nueva York presentan diferencias gramaticales y quiero analizar si estas diferencias se deben al contacto con el inglés y otras variedades de español como se ha señalado en trabajos previos para la segunda generación (Otheguy y Zentella, 2012).

La variación en el uso pronominal, dentro de cada uno de los dos grupos de exposición, puede ser distinta en los informantes de la muestra según se relacione con diferentes variables externas como el área geográfica, el origen nacional, llegar más joven o más viejo a la ciudad, el género, la educación, etc. Las variables externas son características socio-demográficas externas a la estructura de la lengua. Esta diferenciación lleva a las preguntas siguientes: ¿hay variables externas que tienen un efecto estadístico más predictivo en el uso del pronombre en esta muestra?, ¿cuáles son las variables externas que influyen en la realización del pronombre en la muestra entera y en cada uno de estos dos grupos? Mis preguntas analizan si dentro de estos dos grupos de informantes lo que diferencia el uso del pronombre tiene que ver con si ser caribeño tiene más valor predictivo que ser puertorriqueño, si llegar más o menos joven a Estados Unidos tiene más valor predictivo que ser hombre, si tener un nivel alto de inglés tiene más valor

predictivo que tener un nivel alto de educación, etc. En definitiva, se trataba de ver si región tiene más fuerza predictiva que origen nacional, llegar más joven o más viejo tiene más fuerza predictiva que género, nivel de destreza en inglés tiene más fuerza predictiva que educación y así sucesivamente. Para responder estas cuestiones, es necesario examinar las diferencias y similitudes en la tasa pronominal de la muestra entera y de cada uno de los dos grupos de exposición, realizando análisis de variación con las variables externas demográficas y de uso de la lengua.

2. Distribución del uso del pronombre en la muestra entera y por grupos de exposición

Voy a determinar la tasa pronominal de todos los informantes de la muestra, que como ya hemos explicado, consiste en los 114 informantes de primera generación del corpus Otheguy-Zentella, para ver su distribución general. Seguidamente, voy a establecer la tasa pronominal por cada grupo de exposición. Es decir, analizaré qué porcentaje de pronombres utilizan explícitamente los hablantes inmigrantes y los recién llegados. Trato de observar qué grupo de individuos se comporta de manera más similar al de la muestra entera.

2.1 La tasa pronominal en la muestra entera

Investigo si los hablantes de español de esta muestra pueden ser distinguidos por la frecuencia de uso del pronombre. Busco diferencias significativas en relación a su tasa pronominal. Esto lleva a cuestionarnos qué cantidad de pronombres personales sujetos utilizan los hablantes hispanos de la muestra del presente estudio en la ciudad de Nueva York. La siguiente tabla indica el número de verbos que se usa con o sin pronombre.

<i>Tabla 1.</i> Porcentaje de verbos con pronombres sujeto		
	N Verbos	Porcentaje
Presencia	16,822	.33
Ausencia	35,053	.67
Total	51,875	.100

Los datos de la tabla 1 señalan que en esta muestra la mayor parte de los verbos ocurre sin pronombre, igual que sucede en el español en general. De todos los verbos en forma personal, dos tercios no usan pronombre y sólo una tercera parte aparece con el pronombre. La tasa pronominal de todos los informantes de la muestra es del 33 por ciento. Este resultado, que para nuestra muestra de primera generación es casi el mismo que el de Otheguy y Zentella (2012: 69) para todo el corpus Otheguy-Zentella, indica que el español en Nueva York de los inmigrantes y los recién llegados se caracteriza por la ausencia pronominal, como sucede en el español de todas partes.

Partiendo de la tasa pronominal de la muestra entera del presente estudio, es posible que existan diferentes sub-comunidades lingüísticas de español en Nueva York con un comportamiento pronominal distinto. El estudio de estas diferencias permitirá establecer grupos de latinos que se distingan por la ocurrencia del pronombre dentro de la primera generación.

2.2 La tasa pronominal por grupos de exposición

A partir de la tasa pronominal general, voy a estudiar si hay diferencias entre los grupos de exposición en el uso del pronombre. Espero encontrar que los inmigrantes usan más veces el pronombre personal sujeto que los recién llegados. La predicción es que los inmigrantes que llevan más tiempo en Nueva York tendrán mayor número de verbos con pronombres porque están más expuestos al inglés y a otras variedades de español que los recién llegados.

Para el estudio de esta característica particular, la ANOVA es el método de análisis estadístico elegido porque toma en consideración la tasa de los individuos en los grupos comparándolos y determina si las diferencias entre los promedios de tasa pronominal de los grupos son mayores que las diferencias dentro de cada grupo. El coeficiente de F indica las veces que la diferencia o variación entre los grupos es superior a la variación dentro de cada grupo. Cuanto mayor el coeficiente de F, más significativos serán los resultados. Por otro lado, el valor de *p* indica la significación estadística. Por convención, los resultados de $p < .05$ (o sea, *p* con un valor de menor que .05) se consideran significativos. Los resultados de entre $p < .05$ y $p < .10$ son marginalmente significativos. Los resultados de $p > .10$ no son significativos y suelen descartarse.

La tabla 2 presenta un análisis de varianza (ANOVA) que muestra la frecuencia general de uso del pronombre en recién llegados e inmigrantes.

Tabla 2. ANOVA. Tasa pronominal por exposición a Nueva York		
	N Hablantes	Tasa pronominal
Recién llegados	39	.29
Inmigrantes	75	.33
	114	
<i>F</i> = 3.30	<i>p</i> < .072	

En general, se observa una diferencia marginalmente significativa entre los recién llegados y los inmigrantes. Los recién llegados, que aparentemente se guían más por las normas de uso latinoamericanas y cuya influencia del inglés tiende a considerarse aún mínima, usan un pronombre personal sujeto en un 29 por ciento de los verbos en forma personal. Los inmigrantes, que tienen más exposición con el sistema gramatical inglés, realizan el pronombre el 33 por

ciento de las veces. Este grupo de exposición que muestra el mismo uso del pronombre que todos los informantes del corpus representa a la muestra entera. Además, esta distinción entre los inmigrantes y los recién llegados se apoya en que la varianza en el uso del pronombre es mayor más de tres veces ($F = 3.30$) entre los grupos que dentro de cada uno de los grupos. Este valor F que diferencia a los recién llegados de los inmigrantes registra un nivel marginal de significancia estadística ($p < .072$).

Los datos apoyan la idea de que cuanto mayor es la exposición de la población de habla hispana a la ciudad de Nueva York mayor es su tasa pronominal. En su estudio sobre el español de las primera y segunda generaciones de hispanos en Nueva York, Otheguy y Zentella (2012: 87-90) encontraron que el contacto con el inglés es una de las fuerzas que motiva el aumento del uso del pronombre en el español de estos hablantes. Si comparamos los resultados, ambos análisis muestran hallazgos que apuntan hacia la misma dirección: el aumento del uso del pronombre en relación a la exposición con el inglés. La diferencia es que en su estudio estos dos grupos presentan resultados estadísticamente distintos y más significativos. La varianza en el uso del pronombre es casi seis veces mayor ($F = 5.89$) y el valor de p es de $< .02$. La razón de esta diferencia es que fueron analizados junto a los latinos nacidos en Nueva York que conforman la segunda generación. Otheguy y Zentella (2012: 87) compararon a los recién llegados y los inmigrantes considerados dentro del grupo llamado primera generación por haber nacido en Latinoamérica con los latinos de segunda generación que nacieron en Nueva York. En el presente trabajo, en cambio, y como hemos aclarado, se estudia si los recién llegados y los inmigrantes son diferentes y para ello se estudia cómo es su tasa pronominal. O sea, que los datos de Otheguy y Zentella (2012) son intergeneracionales, mientras que los nuestros son intrageneracionales, comparando dos grupos dentro de la primera generación.

3. La tasa pronominal por variables externas socio-demográficas en la muestra entera y por grupos de exposición

Investigo si es adecuado agrupar a los informantes de la muestra entera y por grupos de exposición en base a su género, edad, edad de llegada a Estados Unidos, educación, clase social y estatus socioeconómico. También es importante analizar si la tasa pronominal permite agrupar a los hablantes en relación a su lugar de origen como sugiere la literatura sobre el español (Lipski, 1994: 241; López Morales, 1992: 137; Ordoñez y Olarrea, 2001; Toribio 2000; Zogona, 2002; Goodall, 2004; Otheguy y Zentella, 2012: 143). Esto se relaciona con la hipótesis establecida por Otheguy y Zentella (2012) acerca de la idea de que las diferencias y similitudes se deben al contacto entre inglés y español y a la nivelación dialectal originada por el contacto con distintas variedades dialectales de español en Nueva York.

3.1 Género, edad, edad de llegada a EE.UU, educación y clase social en la muestra entera y por grupos de exposición

En este apartado, comparo la tasa pronominal en relación a cinco variables sociales que comparten los informantes de la muestra entera y los inmigrantes y los recién llegados por separado. Respecto a la muestra completa, los análisis de varianza (ANOVA) señalan que la covariación entre la tasa pronominal y tres de estas variables no es estadísticamente significativa: género ($F = 0.84, p < .43$), educación ($F = 0.21, p < .64$) y clase social ($F = 1.41, p < .23$). Los valores bajos de F indican que estos hablantes no son distintos. En cada uno de estos casos, la varianza entre los grupos es más pequeña, o la misma, que dentro de cada grupo. Esto significa que la tasa pronominal no es suficiente para diferenciar a los informantes de la muestra tan sólo por el hecho de ser hombre o mujer, tener más o menos nivel de educación, o pertenecer a la

clase media o la clase trabajadora. Los datos muestran que los miembros de la muestra no son distintos con respecto a estas variables en base a su tasa pronominal.

Este resultado, similar al encontrado por Otheguy y Zentella (2012) en su trabajo, apoya la idea de Silva-Corvalán (2001: 133) acerca de la generalización de que las características personales no suelen impactar en la variabilidad de rasgos morfosintácticos. Sin embargo, el análisis revela que las variables externas edad y edad de llegada a Estados Unidos sí alcanzan niveles marginales de significación estadística en el uso explícito del pronombre. La siguiente tabla divide a los informantes según su edad. La categoría de los jóvenes agrupa a los hablantes que tienen entre 15 y 39 años. La categoría de los mayores reúne a aquellos que tienen entre 40 y 80 años.

<i>Tabla 3.</i> ANOVA. Tasa pronominal por edad		
	N Hablantes	Tasa pronominal
Jóvenes	76	.31
Mayores	38	.35
	114	
$F = 2.79$	$p < .097$	

La tabla 3 muestra que la diferencia en la tasa pronominal de estos informantes es estadísticamente significativa, a pesar de no lograr un nivel pleno de significación de $p < .05$. Cuanto más años tiene el informante, más uso del pronombre hace. Cuanto más joven es el informante, menos pronombres utiliza.

La tabla 4, que ponemos a continuación, muestra resultados con respecto a la edad a la que nuestros informantes de primera generación llegaron a Nueva York antes de los 12 años o con 12 años y después de esta edad.

Tabla 4. ANOVA. Tasa pronominal por edad de llegada a EE.UU		
	N Hablantes	Tasa pronominal
Antes de o con 12 años	19	.38
Después de 12 años	95	.31
	114	
$F = 6.18$	$p < .014$	

La tabla 4 informa que los hablantes que llegaron a Estados Unidos antes de los 12 años o con 12 años utilizan el pronombre más veces que aquellos que llegaron después de haber cumplido los 12 años. Los que llegaron a una edad más temprana realizan el pronombre el 38 por ciento de los casos. Los que vinieron siendo más mayores lo usan el 31 por ciento. A diferencia de los anteriores, estos resultados son altamente significativos, porque p está muy por debajo de $< .05$.

Ahora analizamos los datos de los análisis de co-varianza entre la tasa pronominal y las variables externas demográficas por grupos de exposición. Primero, muestro los resultados de cada una de estas variables en los recién llegados. Los resultados son diferentes. Las ANOVAS para los recién llegados señalan que hay cuatro variables que no son estadísticamente significativas: género ($F = 1.66, p < .20$), edad ($F = 1.48, p < .23$), edad de llegada a EE.UU ($F = 0.13, p < .71$) y educación ($F = 0.07, p < .78$). Los valores de estas ANOVAS son suficientes para demostrar que los recién llegados no se diferencian en el uso del pronombre por ser hombre o mujer, los años que tienen o el nivel de educación. Uno de los análisis revela que la clase social es significativa en los recién llegados. En la siguiente tabla *R.LL.* significa recién llegados.

Tabla 5. ANOVA. Tasa pronominal por clase social, R.LL.		
	N Hablantes	Tasa pronominal
Trabajadora	22	.32
Media-alta	17	.25
	39	
$F = 4.46$	$p < .042$	

La tabla 5 indica que las diferencias en las tasas pronominales entre las dos clases sociales son significativas para diferenciar a los recién llegados. Estos resultados ofrecen información acerca de los recién llegados a Nueva York. Hay mayor uso del pronombre cuando el informante se describe como miembro de la clase trabajadora.

En el grupo de exposición de inmigrantes, los análisis de co-varianza (ANOVAS) entre la tasa pronominal y las variables demográficas dan los resultados siguientes: género ($F = 0.30, p < .73$), edad ($F = 0.43, p < .51$), educación ($F = 1.39, p < .24$) y clase social ($F = 0.07, p < .77$). La tasa pronominal no diferencia a los inmigrantes de esta muestra por ser hombres o mujeres, más jóvenes o más mayores, tener más o menos nivel educativo o ser de clase trabajadora o media-alta.

No obstante, el análisis de una de las ANOVAS indica que la variable externa edad de llegada a EE.UU es significativa dentro del grupo de los inmigrantes. En la siguiente tabla *INM* se refiere a inmigrantes.

Tabla 6. ANOVA. Tasa pronominal por edad de llegada a EE.UU, INM.		
	N Hablantes	Tasa pronominal
Antes de o con 12 años	19	.38
Después de 12 años	56	.32
	75	
$F = 3.83$	$p < .054$	

Los datos de la tabla 6 señalan que los inmigrantes de la muestra se diferencian en el uso del pronombre por la edad de llegada a EE.UU. Aquellos informantes que llegaron antes de los 12 años usan más el pronombre que los que llegaron después de esa edad. La evidencia de estos

datos dicen que los inmigrantes en Nueva York no se distinguen por su uso del pronombre en relación a su género, edad, educación y clase social, pero sí por la edad a la que llegaron a EE.UU.

3.2 Estatus socio-económico (ESE) en la muestra entera y por grupos de exposición

Otro análisis más para estudiar la varianza de la tasa pronominal se ha hecho con el estatus socio-económico de los hablantes. Siguiendo a Otheguy y Zentella (2012: 71), esta variable es resultado de medir conjuntamente la educación y la ocupación de los informantes de la muestra. La razón es que estos autores observan que la pertenencia a la clase trabajadora, media o alta es subjetiva.

El resultado de este análisis muestra que hay una tendencia a aumentar el uso del pronombre en todos los informantes de la muestra de estatus socio-económico bajo que en los de estatus socio-económico más alto. La tasa pronominal es 34 por ciento para los de estatus bajo y 31 por ciento para los de nivel más alto. No obstante, los datos no muestran un resultado estadístico significativo ($F = 1.60, p < .20$). Estos datos apoyan los resultados obtenidos por Otheguy y Zentella (2012: 71).

Los análisis sobre la tasa pronominal y el ESE por cada grupo de exposición da resultados similares a los de todos los hablantes de la muestra. Los datos no son significativos estadísticamente en ninguno de los dos grupos. Los recién llegados no se diferencian por esta variable: ESE ($F = 1.30, p < .26$). Estos valores informan que esta medida es irrelevante para distinguir a los recién llegados en el uso del pronombre. En cambio, la tabla 5 de la sección anterior sobre la tasa pronominal por clase social en los recién llegados mostró que las diferencias entre las clases sociales trabajadora y media-alta son significativas para diferenciar a los recién llegados. Los resultados de la tabla 5 indicaron que hay mayor uso del pronombre cuando el informante se describe como miembro de la clase trabajadora. Los resultados del ESE

en los recién llegados también indican que los informantes de bajo estatus socioeconómico utilizan más el pronombre de manera explícita, aunque los datos no son significativos estadísticamente. Lo mismo ocurre con los inmigrantes: ESE ($F = 0.21, p < .64$). En ambos casos, los informantes recién llegados y los inmigrantes con un bajo ESE presentan mayores porcentajes en la tasa pronominal que los de mayor ESE.

4. La tasa pronominal por variables externas de uso de la lengua en la muestra entera y por grupos de exposición

Anteriormente estudié el impacto de las variables externas sociodemográficas en la muestra entera y en cada uno de los grupos de exposición. A continuación, analizo si hay alguna correlación entre las variables externas que se refieren al uso de la lengua y la utilización del pronombre personal sujeto. En particular, estudio qué efecto produce tener una competencia en inglés excelente o menos que excelente. La única variable lingüística externa que se analiza como variable del uso de la lengua es la competencia en inglés. El problema con este tipo de variables es la falta de criterio en la auto-evaluación del informante. En ocasiones, la competencia percibida no se corresponde con la competencia lingüística real. Se observa que los participantes a menudo sufren inseguridad lingüística y se auto-evalúan en relación a las normas de corrección en lugar de su competencia y fluidez reales (Otheguy y Zentella, 2012: 93).

Por tanto, las variables de uso de la lengua se basan en cuánto inglés usan los informantes de la muestra en los contextos profesional y familiar. Se espera que los hablantes que usan más inglés, tengan mayor competencia en esta lengua y una tasa pronominal más elevada como resultado del contacto de lenguas.

4.1 Competencia en inglés en la muestra entera y por grupos de exposición

Hemos visto que hay cambios en la tasa pronominal cuando se agrupa a los informantes por grupos de exposición. Los inmigrantes que llevan más tiempo en la ciudad presentan un incremento del uso del pronombre que los diferencia de los recién llegados. Estos resultados parecen indicar que los informantes que están más expuestos al contacto con el inglés tienen más posibilidades de que su gramática haya sido modificada por la gramática inglesa. Quiero saber si este aumento del uso del pronombre es sólo una consecuencia del estatus oficial del inglés o si tener una competencia alta en inglés se corresponde con un incremento de la tasa pronominal.

Para estudiar el efecto de una competencia alta en inglés en el español de Nueva York, los informantes están divididos en los que hablan inglés excelente y los que lo hablan menos que excelente. El objetivo es establecer si la competencia excelente en inglés motiva cambios en la gramática en español. El resultado se muestra a continuación:

	N Hablantes	Tasa pronominal
Inglés menos que excelente	88	.31
Inglés excelente	26	.36
	114	
$F = 2.88$	$p < .092$	

Los resultados sólo alcanzan un nivel marginal de significación, pero así y todo los datos de la tabla 7 reflejan que los informantes que tienen un inglés excelente presentan una tasa pronominal más elevada. Estos hablantes utilizan el 36 por ciento de la veces el pronombre de forma explícita. Los que tienen un inglés menos que excelente lo usan el 31 por ciento de los casos.

Estos resultados sugieren que es muy probable que el mayor uso del pronombre entre los hablantes de mayor exposición sea, en parte, debido al inglés.

Al analizar la competencia en inglés por separado en los dos grupos de exposición, los datos siguen siendo interesantes. Esta variable es significativa en los inmigrantes, pero no en los recién llegados.

Tabla 8. ANOVA. Tasa pronominal por competencia en inglés, INM.		
	N Hablantes	Tasa pronominal
Inglés menos que excelente	56	.32
Inglés excelente	19	.38
	75	
$F = 3.88$	$p < .053$	

La tabla 8 muestra que, dentro de la primera generación que estudiamos, son los inmigrantes que poseen un inglés excelente los que usan más el pronombre, haciéndolo en el 38 por ciento de los verbos. Los recién llegados no se diferencian por esta variable: la tasa pronominal de los recién llegados en relación a su competencia en inglés no arroja resultados significativos ($F = 0.42, p < .51$).

5. Origen nacional y regional en la muestra entera y por grupos de exposición

Después de haber analizado el impacto del género, de la edad, la educación, la clase social, del ESE y del nivel de inglés en el uso del pronombre en la muestra entera y por grupos de exposición, ahora estudio la importancia de otras variables externas: el origen nacional y regional. Me pregunto si pertenecer a uno de los países representados en la muestra o a una de las regiones, caribeña o continental, es significativo para diferenciar a estos hablantes en la

realización del pronombre personal como se ha señalado en trabajos previos (Otheguy, Zentella y Livert, 2007; Otheguy y Zentella, 2012; Raña Riso, 2013).

Los informantes de la muestra están divididos según su país de origen: colombianos, cubanos, dominicanos, ecuatorianos, mexicanos y puertorriqueños.

<i>Tabla 9a.</i> ANOVA. Tasa pronominal por país de origen		
	N Hablantes	Tasa pronominal
Dominicanos	21	.40
Puertorriqueños	18	.37
Cubanos	20	.36
Colombianos	16	.30
Ecuatorianos	20	.27
Mexicanos	19	.21
	114	.32
$F = 11.12$	$p < .001$	

La tabla 9a indica los porcentajes de la tasa pronominal en orden descendente para cada grupo de informantes por cada país. Los datos señalan que los dominicanos son los que utilizan más el pronombre con un promedio de 40 por ciento. Mientras que en los mexicanos sólo el 21 por ciento de los verbos llevan pronombres. El análisis de varianza ANOVA indica que la co-variación entre la tasa pronominal y el país de origen es muy significativo estadísticamente. Estos resultados para nuestra muestra de primera generación coinciden con los de Otheguy y Zentella (2012) para todo el corpus en el orden de países aunque no en los valores.

Para averiguar sobre la importancia de los datos, hay que analizar cuáles de las diferencias entre los países adyacentes son significativas. La pregunta es si la tasa pronominal de 40 por ciento en los dominicanos es estadísticamente distinta de la tasa pronominal de 37 por ciento de los puertorriqueños. La tabla 9b muestra los resultados del test de significación pair-wise. Los asteriscos marcan las intersecciones entre países adyacentes y muestran las diferencias

de las tasas pronominales que son estadísticamente significativas. Los países donde las intersecciones no se dan indican que las diferencias en la tasa pronominal no son significativas.

Tabla 9b. ANOVA. Post-hoc Tukey test de significación pair-wise

	Dom	PR	Cub	Col	Ecu	Mex
Dominicanos				*	*	*
Puertorriqueños					*	*
Cubanos					*	*
Colombianos	*					
Ecuatorianos	*	*	*			
Mexicanos	*	*	*			

Los asteriscos señalan una diferencia significativa en la tasa pronominal

Los datos presentados en la tabla 9b ayudan a justificar si efectivamente se puede caracterizar a los informantes por su procedencia basándose en la asociación entre la tasa pronominal y el país de origen. Vemos que el agrupamiento de país por país no se justifica. Lo que denotan los asteriscos es que los dominicanos son distintos de los mexicanos. Los puntos marcan una división entre dos grupos. Al igual que observaron Otheguy y Zentella (2012: 73) para el corpus entero, vemos que para nuestra muestra de primera generación el grupo de los caribeños está integrado por Cuba, República Dominicana y Puerto Rico. El grupo de países continentales lo forman Colombia, Ecuador y México.

La tabla 9b muestra que la diferencia de los tres puntos de porcentaje entre dominicanos, puertorriqueños y cubanos no es estadísticamente significativa. Por tanto, no hay ninguna diferencia significativa en la tasa pronominal en estos países. Más bien, los datos apoyan la idea de un único grupo: el de los caribeños. Lo mismo ocurre con los países continentales. Dentro de este grupo, hay una diferencia estadísticamente significativa entre los colombianos y los ecuatorianos y mexicanos. A pesar de esta distinción, estos tres países también constituyen un

grupo singular si observamos que sus porcentajes en la tasa pronominal son más próximos entre sí.

Los datos de las tablas 9a y 9b apoyan la idea que Otheguy y Zentella (2012) apuntaron sobre la existencia de un continuum nacional de las diferencias de la tasa pronominal en el español de los latinos en Nueva York. También, distinguen dos regiones: la caribeña y la continental.

La tabla 10 señala las diferencias en la tasa pronominal dividiendo a todos los informantes de la muestra en dos grandes zonas dialectales.

Tabla 10. ANOVA. Tasa pronominal por región de origen		
	N Hablantes	Tasa pronominal
Caribeños	59	.38
Continetales	55	.26
	114	
$F = 44.14$	$p < .001$	

Los datos indican que los caribeños usan el pronombre con mucha más frecuencia que los continentales como mostraron Otheguy y Zentella (2012: 73). En esta muestra, los 59 informantes del Caribe utilizan un pronombre con un verbo el 38 por ciento de las veces. Los 55 continentales lo usan sólo en el 26 por ciento de los verbos.

La tabla también informa que las diferencias en la tasa pronominal a través de los dos grupos son más grandes que dentro de cada grupo ($F = 44.14$). Las diferencias son estadísticamente muy significativas. Los datos de las tablas 9a, 9b y 10 reafirman la existencia de dos regiones: la caribeña y la continental.

Los siguientes textos de dos informantes de distintas regiones ilustran cualitativamente las diferencias entre los caribeños y los continentales. Las tasas pronominales de estos hablantes

se corresponden con la tasa promedio de uso del pronombre en cada región. El texto 2a corresponde a un puertorriqueño recién llegado que tiene una tasa pronominal de 38 por ciento. El texto 2b es de un mexicano también recién llegado que utiliza el pronombre el 26 por ciento de las veces.

2a. [Habla de su experiencia de ser puertorriqueño en Nueva York] O sea entonces **yo decía** “no, no **aplican**”. En Puerto Rico claro, entonces **yo decía** no, no, no, si es que aquí nunca. Porque **tú sabes** que el chisme es o la excusa es ... **Nosotros no podemos** ser racistas porque **somos** una mezcla de indios, africanos y españoles, que (347) es mierda. Pero con lo más interesante es que **yo me pongo** a re-evaluar mi ... la organización social que **yo tenía** cuando **estaba** en Puerto Rico. [435P]

2b. [Habla de cómo eligió sus estudios] **Estudié** el bachillerato, la preparatoria. Lo que pasa es que es el área, cuando **estás** en el penúltimo año **escoges** tu área. Cuando **estás** en el último año de prepa, ya **estás** en la área que **vas a estudiar** para la universidad, entonces **yo escogí** una área en donde **no sabía** las opciones que **tenían**. Cuando **salí** de la prepa se me cerraron las opciones y dentro de las carreras que **podía** escoger la que más me llamó la atención fue la de comunicaciones. [308M]

A continuación investigo si el origen nacional y regional se relaciona con el uso del pronombre en cada grupo de exposición. Específicamente, quiero analizar si hay sub-comunidades de latinos recién llegados y latinos inmigrantes que se distinguen por el país o la región de la que vienen y cómo su procedencia impacta en la tasa pronominal. Las siguientes tablas muestran la tasa pronominal por origen nacional de los recién llegados y los inmigrantes en orden descendiente.

Tabla 11 . ANOVA. Tasa pronominal por país de origen, R.LL.

	N Hablantes	Tasa pronominal
Dominicanos	6	.41
Puertorriqueños	6	.34
Cubanos	7	.32
Ecuatorianos	8	.26
Colombianos	6	.24
Mexicanos	6	.19
	39	.29
<i>F</i> = 5.44	<i>p</i> < .001	

<i>Tabla 12</i> . ANOVA. Tasa pronominal por país de origen, INM.		
	N Hablantes	Tasa pronominal
Dominicanos	15	.40
Puertorriqueños	12	.38
Cubanos	23	.38
Colombianos	10	.33
Ecuatorianos	12	.27
Mexicanos	13	.23
	75	.33
<i>F</i> = 6.59	<i>p</i> < .001	

La tabla 11 señala que en el grupo de recién llegados los dominicanos son los que tienen una mayor tasa pronominal y los mexicanos los que la tienen más baja. Los dominicanos usan el 41 por ciento de los verbos con un pronombre, mientras que los mexicanos lo utilizan el 19 por ciento de las veces. La ANOVA indica que la diferencia en la tasa pronominal es muy significativa. La tabla 12 informa que en el grupo de los inmigrantes también los dominicanos usan más pronombres y los mexicanos menos. El 40 por ciento de los verbos ocurre con un pronombre en los dominicanos y el 23 por ciento de los verbos tiene un pronombre en los mexicanos. La ANOVA también da un resultado muy significativo. Los resultados de las tablas 11 y 12 son muy similares. La única diferencia entre estos grupos de exposición es el orden de los países continentales. Mientras que en los recién llegados son los ecuatorianos los que tienen una tasa pronominal más alta, en los inmigrantes son los colombianos como en la muestra entera. Los inmigrantes presentan el mismo orden descendente en la tasa pronominal por país que en la muestra entera. Esto indica que los inmigrantes son más representativos para todos los informantes de la muestra.

En relación a la región de origen, los latinos recién llegados y los latinos inmigrantes son muy parecidos.

Tabla 13 . ANOVA. Tasa pronominal por región de origen, R.LL.		
	N Hablantes	Tasa pronominal
Caribeños	19	.36
Continetales	20	.23
	39	
<i>F</i> = 19.07	<i>p</i> < .001	

Tabla 14 . ANOVA. Tasa pronominal por región de origen, INM.		
	N Hablantes	Tasa pronominal
Caribeños	40	.39
Continetales	35	.27
	75	
<i>F</i> = 25.53	<i>p</i> < .001	

La tabla 13 informa que los 19 recién llegados caribeños tienen una tasa pronominal mayor (el 36 por ciento de los verbos ocurre con pronombre) que los 20 recién llegados continentales (el 23 por ciento de los verbos aparece con pronombre). La tabla muestra que estas diferencias son muy significativas. La tabla 14 indica que también los 40 caribeños inmigrantes utilizan más el pronombre que los 35 continentales inmigrantes. La tasa pronominal de los caribeños es del 39 por ciento, mientras que la de los continentales es del 27 por ciento. Estos datos reafirman la significación de esta diferencia. El grupo de exposición de los inmigrantes sigue representando más a todos los informantes de la muestra entera. Estos resultados que coinciden con los de la tabla 10 indican que la categorización de los hablantes por región de origen se justifica por los datos de la distribución pronominal en la muestra.

En definitiva, los datos indican que las diferencias por región, y también en cierto sentido por país, se mantienen firmes durante toda la primera generación, porque los inmigrantes siguen teniendo el mismo orden por región y por país que los recién llegados.

6. Diferencias regionales en la tasa pronominal por características demográficas en la muestra entera y por grupos de exposición

Ahora analizo la influencia de las variables externas de género, edad, edad de llegada a EE.UU, educación, clase social y estatus socioeconómico (ESE) en la tasa pronominal de los caribeños y los continentales de la muestra completa y de los caribeños y los continentales que se dividen en inmigrantes y recién llegados. Los análisis anteriores sobre la muestra entera desvelaron que no hay diferencias en la tasa pronominal entre grupos definidos por el género, la educación y la clase social. Únicamente cuando los informantes son divididos según los años que tienen y la edad a la que llegaron a Estados Unidos, el comportamiento en el uso del pronombre difiere.

Las variables externas no son diferentes de forma significativa respecto a la tasa pronominal de los caribeños, a excepción de la edad de llegada a EE.UU. Los resultados de las ANOVA de estas variables externas para los caribeños son: género ($F = 0.97, p < .38$), edad ($F = 1.01, p < .31$), educación ($F = 2.17, p < .14$), clase social ($F = 1.09, p < .30$) y ESE ($F = 0.03, p < .95$). Sin embargo, casi todas estas variables sí distinguen a los continentales en el uso pronominal. Los resultados para los continentales son: género ($F = 7.00, p < .01$), edad ($F = 9.60, p < .01$), edad de llegada a EE.UU ($F = 0.12, p < .72$), educación ($F = 3.28, p < .07$), clase social ($F = 4.44, p < .04$) y ESE ($F = 4.84, p < .03$). Estos datos indican que los caribeños no se distinguen en la tasa pronominal en casi ninguna de estas características demográficas. Sólo se muestran diferentes según la edad de llegada a EE.UU ($F = 5.79, p < .01$). Así que ser caribeño y

llegar después de los doce años se correlaciona con el aumento del pronombre. A pesar de esta diferencia, se observa una tendencia a la uniformidad en el uso del pronombre.

En general, los resultados indican que los caribeños de la muestra conforman un grupo más homogéneo internamente. Los informantes continentales se distinguen por casi todas las variables externas analizadas. Los datos informan que son las mujeres quienes usan más el pronombre, los informantes de más edad, los que poseen un nivel más bajo de educación, los que pertenecen a la clase trabajadora y aquellos con un SES más bajo.

Los análisis de co-varianza entre la tasa pronominal y las variables externas demográficas por grupos de exposición en los caribeños y continentales dan resultados diferentes. Las ANOVAS para los caribeños recién llegados indican que la mayoría de las variables externas no son significativas en el uso del pronombre. Los resultados son: género ($F = 0.00, p < .93$), edad ($F = 0.17, p < .67$), edad de llegada a EE.UU ($F = 0.14, p < .70$), educación ($F = 3.03, p < .09$), clase social ($F = 0.79, p < .38$) y ESE ($F = 0.62, p < .43$). Únicamente, el nivel de educación diferencia a los caribeños recién llegados. Aquellos con mayor educación usan más el pronombre. Las ANOVAS para los caribeños inmigrantes ofrecen los siguientes resultados: género ($F = 0.99, p < .38$), edad ($F = 3.44, p < .07$), edad de llegada a EE.UU ($F = 21.98, p < .01$), educación ($F = 0.69, p < .40$), clase social ($F = 0.08, p < .76$) y ESE ($F = 0.18, p < .66$). Los caribeños inmigrantes se distinguen por los años que tienen y la edad a la que llegaron a Estados Unidos. Dentro de este grupo los que son más jóvenes y llegaron a Estados Unidos después de los doce años utilizan más el pronombre.

Las ANOVAS para los continentales recién llegados son: género ($F = 7.90, p < .01$), edad ($F = 0.17, p < .67$), edad de llegada a EE.UU ($F = 0.11, p < .74$), educación ($F = 1.57, p < .22$), clase social ($F = 2.41, p < .13$) y ESE ($F = 3.23, p < .09$). Los datos muestran que el género y el

ESE son significativos para distinguir a los informantes de este grupo. Los continentales recién llegados que son mujeres de bajo estatus socioeconómico utilizan más el pronombre. Los resultados para los continentales inmigrantes son: género ($F = 1.22, p < .27$), edad ($F = 6.15, p < .01$), edad de llegada a EE.UU ($F = 0.75, p < .39$), educación ($F = 0.62, p < .43$), clase social ($F = 1.59, p < .21$) y ESE ($F = 1.09, p < .30$). Estos hablantes no son diferentes por ninguna de estas características demográficas, excepto por la edad. Los continentales inmigrantes más mayores presentan una tasa pronominal alta.

En resumen, se observa que hay diferencias en la tasa pronominal entre los informantes de la muestra cuando se analizan por el origen geográfico así como cuando se analizan por el origen geográfico junto con la exposición a la ciudad de Nueva York. Los continentales y caribeños son distintos cuando se definen por características socio-demográficas. Los caribeños se diferencian por la edad de llegada a EE.UU. Aquellos que vinieron después de los doce años realizan más el pronombre. Los continentales muestran diferencias en el uso pronominal en casi todas las variables: el género, la edad, la educación, la clase social y el ESE. Los continentales que son hombres o mujeres, jóvenes o mayores, que tienen un nivel de educación elemental o universitaria, que se consideran clase trabajadora o clase media-alta y que tienen un estatus socioeconómico bajo o alto se diferencian por su tasa pronominal.

El resultado varía cuando se divide a los informantes de la muestra además de por el origen geográfico, por la exposición a Nueva York. Los caribeños inmigrantes y recién llegados son diferentes en la tasa pronominal. Mientras que la educación es la única variable externa que diferencia a los caribeños recién llegados en el uso del pronombre, la edad y la edad de llegada a EE.UU son las características demográficas que distinguen a los caribeños inmigrantes. Los continentales recién llegados también son diferentes del resto por el género y el ESE. Las mujeres

de estatus socioeconómico bajo son las que usan el pronombre explícitamente la mayoría de las veces. Los continentales inmigrantes sólo se distinguen por la edad. Cuantos más años tienen, más alta es su tasa pronominal.

7. Diferencias regionales en la tasa pronominal por variables externas de uso de la lengua en la muestra entera y por grupos de exposición

La única variable externa de uso de la lengua analizada es la competencia en inglés. Investigo la influencia de esta variable en la tasa pronominal de los caribeños y los continentales y de los caribeños y los continentales inmigrantes y recién llegados. Es decir, observo si hay correlación entre la competencia en inglés y el origen geográfico y entre la competencia en inglés y el origen geográfico y la exposición a Nueva York.

Los resultados de las ANOVAS de esta variable externa para los caribeños es: competencia en inglés ($F = 4.12, p < .04$) y para los continentales: competencia en inglés ($F = 0.29, p < .58$). Esta variable es significativa para los caribeños pero no para los continentales. Los datos sugieren que la competencia en inglés diferencia a los caribeños de la muestra de los continentales. Los caribeños que tienen una competencia excelente en inglés usan más el pronombre. Por el contrario, el coeficiente de F en los continentales señala que hay más variación dentro de los dos grupos (más y menos inglés) que entre ellos, y, el valor de la p confirma que la diferencia entre los dos grupos obviamente no es significativa.

Cuando los informantes se dividen según el origen geográfico y la exposición, los datos siguen siendo parecidos. Para los caribeños recién llegados, la ANOVA informa que la competencia en inglés es ($F = 0.74, p < .39$). Para los caribeños inmigrantes, el resultado de la competencia en inglés es ($F = 5.94, p < .02$). Se observa que dentro del grupo de los caribeños son los inmigrantes los que se distinguen en su tasa pronominal por tener un determinado nivel

en inglés. La ANOVA para los continentales recién llegados en competencia en inglés es ($F = 0.67, p < .42$) y para los continentales inmigrantes es ($F = 0.14, p < .70$). En cualquiera de los casos, esta variable no es significativa para diferenciar a los continentales.

En resumen, el contacto con el inglés se relaciona con el aumento de la tasa pronominal en los caribeños de la muestra. Específicamente esta variable es significativa para los caribeños inmigrantes. Los informantes que son caribeños, inmigrantes y con un inglés excelente utilizan más el pronombre. Sin embargo, tener un inglés excelente o menos que excelente no es importante para diferenciar a los continentales inmigrantes y recién llegados en el uso del pronombre.

Resumen y conclusión

He establecido qué grupos de informantes tienen diferencias significativas en su tasa pronominal. Para conseguir el propósito, he investigado el impacto de nueve variables externas, ocho sociodemográficas y una de uso de la lengua.

He analizado si las variables externas de género, edad, edad de llegada, educación, clase social, estatus socioeconómico y competencia en inglés pueden ayudar a distinguir a estos dos grupos por las posibles diferencias de su tasa pronominal. También, he estudiado si la región de procedencia y el origen nacional influyen de la misma forma en el uso del pronombre. Y si el origen geográfico en relación a la tasa pronominal es importante para agrupar uniformemente a los hablantes como se ha demostrado en otros trabajos anteriores sobre el español (Lipski, 1994: 241; López Morales, 1992: 137; Ordoñez y Olarrea, 2001; Toribio 2000; Zogona, 2002; Goodall, 2004; Otheguy y Zentella, 2012: 143).

He observado si algunas de estas variables externas diferencia a los recién llegados de los inmigrantes. Mis preguntas analizaban si dentro de estos dos grupos de informantes lo que

diferencia el uso del pronombre tiene que ver con si ser caribeño tiene más valor predictivo que ser puertorriqueño, si llegar más o menos joven a Estados Unidos tiene más valor predictivo que ser hombre, si tener un nivel alto de inglés tiene más valor predictivo que tener un nivel alto de educación, etc. En definitiva, se trataba de ver si región tiene más fuerza predictiva que origen nacional, llegar más joven o más viejo tiene más fuerza predictiva que género, nivel de destreza en inglés tiene más fuerza predictiva que educación y así sucesivamente.

A continuación, presento algunos de los hallazgos encontrados en la muestra entera sobre los 114 informantes:

- La mayor parte de los verbos ocurre sin pronombre, igual que sucede en el español en general.
- Las variables externas sociodemográficas más significativas son: origen regional, origen nacional y edad de llegada a Estados Unidos.
- Las variables externas sociodemográficas y de uso de la lengua que presentan niveles marginales de significación estadística en el uso explícito del pronombre son edad y competencia en inglés.

El primer dato relevante sobre los 114 hablantes de la muestra es que la mayoría de los verbos ocurren sin el pronombre personal sujeto como sucede en el español en cada país hispanohablante. Los resultados analizados en el capítulo en la muestra entera además indican la importancia de varias variables externas que predicen la tasa pronominal en los 114 informantes. Dentro de las variables externas sociodemográficas, los resultados señalaron que las correlaciones con la tasa pronominal que son estadísticamente muy significativas o significativas incluyen las variables origen regional, origen nacional y edad de llegada. Las variables más predictivas en el uso del pronombre en estos informantes es la región y el país de origen. Con

niveles muy alto de significancia estadística ($p < .001$), estas variables representan la continuidad con el español de Latinoamérica como afirman en su trabajo Otheguy y Zentella (2012: 129). Parece que de los 114 informantes de la muestra los que son caribeños y han llegado antes de o con 12 años realizan más el pronombre.

En general, los datos sobre el origen nacional y regional también reflejan que los caribeños de la muestra conforman un grupo más homogéneo internamente. Los informantes continentales se distinguen por casi todas las variables externas analizadas. Los resultados hallados informan que son las mujeres quienes usan más el pronombre, los informantes de más edad, los que poseen un nivel más bajo de educación, los que pertenecen a la clase trabajadora y aquellos con un SES más bajo.

La variable externa sociodemográfica edad y la variable externa de uso de la lengua competencia en inglés también son importantes para distinguir a estos 114 hablantes en la realización del pronombre, aunque en menor medida debido a que alcanzaron niveles marginales de significación estadística. Aun así, cuantos más años tiene el informante y más nivel de inglés, más uso del pronombre hace.

En contraste, los análisis de co-variación entre la tasa pronominal y las variables estatus socioeconómico (ESE), clase social, género y educación no son significativas. Es decir, pertenecer a un nivel socioeconómico alto o bajo, a la clase media-alta o trabajadora, ser hombre o mujer y tener un nivel de educación elemental o universitario no son relevantes para diferenciar a los 114 hablantes de la muestra en la realización del pronombre personal sujeto.

Cuando los informantes de la muestra entera se dividen en inmigrantes y recién llegados, se observa lo siguiente:

- Los inmigrantes utilizan más veces el pronombre personal sujeto que los recién llegados, a pesar de que la significación estadística para diferenciar a estos dos grupos de exposición es marginalmente significativa.
- Las variables externas sociodemográficas significativas para los recién llegados son: origen regional, origen nacional y clase social.
- Las variables externas sociodemográficas y de uso de la lengua significativas para los inmigrantes son: origen regional, origen nacional, edad de llegada a Estados Unidos y competencia en inglés.

Estos resultados obtenidos son una buena aproximación del español en Nueva York de los inmigrantes y recién llegados. Para empezar se puede decir que son los inmigrantes los que usan más el pronombre explícitamente. Los datos presentados en secciones anteriores apoyan la idea de que cuanto mayor es la exposición de la población de habla hispana a la ciudad de Nueva York mayor es su tasa pronominal. Al estudiar los grupos por separado, vemos que la clase social es importante en los recién llegados. Estos usan más el pronombre cuando el informante se describe como miembro de la clase trabajadora. Los inmigrantes se distinguen por la edad de llegada a Estados Unidos y por la competencia en inglés. Los que llegaron antes de los 12 años y los que poseen un inglés excelente usan más el pronombre explícitamente.

También se puede hablar sobre la distinción ya hecha entre caribeños y continentales (López Morales, 1992; Otheguy y Zentella, 2012). Hay una diferencia clara en el uso del pronombre entre las dos regiones. Los caribeños, ya sean inmigrantes o recién llegados, presentan una tasa pronominal más alta que los continentales. Los datos presentados previamente indican que las diferencias por región, y también por país, se mantienen firmes durante toda la

primera generación, porque los inmigrantes siguen teniendo el mismo orden por región y por país que los recién llegados.

Si además miramos a cada grupo de exposición según el origen geográfico, vemos que:

- Los caribeños inmigrantes y los continentales inmigrantes utilizan más veces el pronombre personal sujeto.
- La variable externa sociodemográfica significativa para los caribeños recién llegados es: educación.
- Las variables externas sociodemográficas significativas para los continentales recién llegados son: género y ESE.
- Las variables externas sociodemográficas y de uso de la lengua significativas para los caribeños inmigrantes son: edad, edad de llegada y competencia en inglés.
- La variable externa sociodemográfica significativa para los continentales inmigrantes es: edad.

Los caribeños inmigrantes y recién llegados y los continentales inmigrantes y recién llegados son diferentes en la tasa pronominal. Los caribeños inmigrantes realizan más el pronombre que los caribeños recién llegados. Lo mismo ocurre en el caso de los continentales inmigrantes que tienen más pronombres que los continentales recién llegados.

Mientras que la educación es la única variable externa que diferencia a los caribeños recién llegados en el uso del pronombre, la edad, la edad de llegada a EE.UU y la competencia en inglés son las características demográficas y lingüísticas que distinguen a los caribeños inmigrantes. En el grupo de los caribeños son los inmigrantes los que más se distinguen en su tasa pronominal por diferentes variables externas. Los continentales recién llegados también son diferentes del resto por el género y el ESE. Las mujeres de estatus socioeconómico bajo son las

que usan el pronombre explícitamente la mayoría de las veces. Los continentales inmigrantes sólo se distinguen por la edad. Cuantos más años tienen, más alta es su tasa pronominal.

En resumen, cuando las dos regiones se dividen en inmigrantes y recién llegados, los resultados varían. Las variables externas que influyen en la tasa pronominal de los caribeños inmigrantes y recién llegados no son las mismas. Lo mismo sucede con los continentales inmigrantes y recién llegados. Estos datos parecen demostrar que no es suficiente con agrupar a todos los hablantes según el área geográfica.

Capítulo 4

Análisis de las variables externas

1. Introducción

En el capítulo tres, establecimos que los grupos de hablantes de español de la primera generación que se distinguen en Nueva York por el origen regional, la edad, la edad de llegada y la competencia en inglés presentan tasas pronominales distintas. También, observé que cuando estos hablantes son clasificados según el nivel de exposición a la ciudad, los recién llegados son diferentes en la tasa pronominal dependiendo del origen regional y la clase social, mientras que los inmigrantes se diferencian no sólo por el origen regional, sino también por la edad de llegada y la competencia en inglés, pero no por la clase social. Cuando, además, la primera generación se divide por el criterio regional, los resultados muestran que el uso pronominal en los caribeños se correlaciona con la edad de llegada y la competencia en inglés y en los continentales con el género, la edad, la educación, la clase social y el ESE. Finalmente, los datos revelaron que si agrupamos a la primera generación en relación al origen regional y a la exposición a la ciudad, las variables externas predictivas en la tasa pronominal son la educación para los caribeños recién llegados; la edad, la edad de llegada y la competencia en inglés para los caribeños inmigrantes; el género y el estatus socioeconómico (ESE) para los continentales recién llegados; y, por último, la edad para los continentales inmigrantes. A pesar de los hallazgos encontrados, la pregunta de cuál de estas diferencias entre los hablantes informa más sobre la expresión del pronombre personal sujeto en la primera generación no está resuelta.

El objetivo al considerar la importancia de cada una de estas variables externas en el análisis de la variable tasa pronominal en la primera generación en Nueva York es ofrecer una visión acertada sobre el uso del español de estos hablantes en la ciudad. Algunas de estas

variables externas sociodemográficas, como la región, reflejan la continuidad entre el español hablado en Latinoamérica y el español en la ciudad. Otras variables externas del uso de la lengua, como la competencia en inglés, indican los nuevos factores que afectan el español en Estados Unidos por el contacto de lenguas. Una de las preguntas de esta tesis tiene que ver con si la variación pronominal en el español de los inmigrantes y los recién llegados en Nueva York representa la continuidad con modelos de variabilidad de Latinoamérica o si por el contrario, representa cambios influenciados por fuerzas de contacto lingüístico local y dialectal. Recordemos que el objetivo es analizar si estas dos sub-comunidades de latinos *dentro* de la primera generación en la ciudad de Nueva York presentan diferencias gramaticales y si estas diferencias se deben al contacto con el inglés y otras variedades de español como Otheguy y Zentella (2012) ya mostraron en su trabajo *intergeneracional*.

Las diferencias en la tasa pronominal entre los dos grupos de primera generación (los inmigrantes y los recién llegados) llevan a cuestionarnos si se puede considerar la primera generación como un grupo homogéneo. Podemos pensar que el impacto de diferentes variables externas en cada uno de los grupos nos indica que estos informantes presentan distintos modelos de uso del pronombre. Por ejemplo, señalé anteriormente que los recién llegados se diferencian entre sí por el origen regional y la clase social, mientras que los inmigrantes difieren en su tasa pronominal por el origen regional, la edad de llegada y la competencia en inglés. Ahora quiero saber cuáles de estas variables externas son las más importante para entender la tasa pronominal de los inmigrantes y los recién llegados y cuáles son las menos relevantes y predictivas. Este capítulo responde estas preguntas a través del análisis de regresión múltiple. Primero, hablaré sobre los criterios para considerar las variables externas sociodemográficas y de uso de la lengua como variables predictivas. Seleccionaré las variables basándome en los resultados ya

encontrados en el capítulo tres en las correlaciones bivariantes. Luego, realizaré análisis de regresiones multivariantes en la muestra entera y en cada grupo de exposición, clasificados según el origen geográfico, primero, y el origen geográfico y la exposición, segundo, para estudiar comparativamente si hay diferencias entre los dos grupos que indiquen cambios en el sistema gramatical.

2. Las variables externas que definen la tasa pronominal

Los análisis sobre las ANOVAS en el capítulo anterior revelaron que de las nueve variables externas estudiadas, ocho sociodemográficas y una de uso de la lengua, ocho son significativas para entender la variación en el uso pronominal de los inmigrantes y recién llegados. Recuerdo que no hay una justificación válida para mantener la variable origen nacional en el estudio, ya que los resultados informaron sobre la inexistencia de este criterio con respecto a la tasa pronominal. Como vimos, los resultados informan, por el contrario, sobre la existencia de un continuum nacional de las diferencias de la tasa pronominal en el español de los latinos en Nueva York y la distinción de dos regiones: la caribeña y la continental.

Estas ocho variables, las cuales de ahora en adelante tratamos como nombres propios y escribimos con mayúscula inicial (ej. *Competencia en inglés*) tienen un valor predictivo diferente sobre los informantes según estos son clasificados atendiendo a criterios diversos. Por ejemplo, la variable de uso de la lengua Competencia en inglés es significativa para los informantes en la muestra entera, para los inmigrantes, para los caribeños y para los caribeños inmigrantes. Pero la variable sociodemográfica ESE distingue sólo a los continentales y a los continentales recién llegados.

Esta sección trata sobre la variable dependiente, llamada tasa pronominal, y sobre el impacto de las variables independientes externas. En términos estadísticos, la variable

dependiente es una variable continua que consiste en un número que se refiere a la tasa pronominal de cada informante. Las variables independientes externas son variables binarias nominales que tienen dos factores. A continuación, muestro una lista con las variables predictivas y sus factores:

Origen regional:

- Caribeño
- Continental

Género:

- Hombre
- Mujer

Edad:

- Jóvenes
- Mayores

Edad de llegada:

- Vino a EE.UU con o antes de 12 años
- Vino a EE.UU después de 12 años

Educación:

- Elemental e instituto
- Universidad y graduado

Clase social:

- Alta y media
- Trabajadora

Nivel de estatus socioeconómico (ESE):

- Bajo estatus socioeconómico
- Alto estatus socioeconómico

Competencia en inglés:

- Inglés menos que excelente
- Inglés excelente

Al analizar la tasa pronominal en el capítulo anterior, observamos que estas variables externas diferencian a los recién llegados de los inmigrantes. Ahora, voy a averiguar si en su capacidad de predecir la tasa pronominal de estos dos grupos de exposición lo que diferencia el uso del pronombre tiene que ver con si Región tiene más valor predictivo que Género, si Edad de llegada a Estados Unidos tiene más valor predictivo que Clase social, si Competencia en inglés tiene más valor predictivo que la Educación, etc. Empiezo de manera simple con diferentes correlaciones Pearson entre la tasa pronominal y todas las variables externas que fueron significativas en el capítulo anterior. Estas correlaciones bivariantes miden la relación entre el uso del pronombre y cada una de las variables externas. Las variables aparecen por orden de fuerza de su correlación con la tasa pronominal.

Las correlaciones son positivas o negativas por una cuestión simple de codificación que no impacta en el análisis. Las correlaciones negativas que aparecen con el signo menos fueron codificadas con el número 0, las correlaciones positivas fueron codificadas con el 1. Ahora bien, en las variables continuas, o en las que la codificación sigue una lógica relacionada con el contenido de lo que se ha codificado (como por ejemplo, Edad y Edad de llegada) el valor positivo o negativo de la correlación no tiene que ver de forma arbitraria con la forma de la codificación, sino que tiene un significado real. La tabla siguiente presenta los resultados de cada correlación e indica la importancia de las variables que predicen la tasa pronominal en los 114

informantes. (Recordemos que cuando hablamos de “la muestra entera” nos referimos a la muestra entera del presente estudio que, como sabemos, consiste en la totalidad de los informantes de primera generación --114 hablantes nacidos en Latinoamérica-- del corpus Otheguy-Zentella.)

Tabla 1. Correlaciones Pearson con tasa pronominal en la muestra entera

Ranking	Variable	N Hablante	r	P
1	Región	114	-.53	**
2	Edad de llegada	114	-.23	*
3	Competencia inglés	114	.16	a
4	Edad	114	.15	a
5	ESE	114	-.12	
6	Clase social	114	.11	
7	Género	114	.10	
8	Educación	114	.04	

** = p <.01
 * = p <.05
 a = p <.10

Como se esperaba, los resultados de la tabla 1 señalan que las correlaciones con la tasa pronominal que son estadísticamente muy significativas, significativas o marginalmente significativas incluyen las variables Región, Edad de llegada, Competencia en inglés y Edad. En contraste, las correlaciones de la tasa pronominal con las variables ESE, Clase social, Género y Educación no son significativas. En la tabla 1, el valor negativo de Edad de llegada, -.22, dice que según disminuye la edad, aumenta la tasa pronominal (o sea, dice que la correlación es inversa, cuanto más joven el hablante, más pronombres usa). Igual pasa con Competencia en inglés. El valor positivo de Competencia en inglés, .16, señala que según aumenta el conocimiento del inglés (según se acerca a 1), aumenta también la tasa pronominal; a más conocimiento del inglés, más pronombres en español.

Cuando la muestra entera es dividida según la exposición a la ciudad de Nueva York, los resultados siguen siendo los mismos. A partir de ahora, en las tablas *R.LL.* significa recién

llegados e *INM.* se refiere a inmigrantes.

Tabla 2. Correlaciones Pearson con tasa pronominal en los R.LL.

Ranking	Variable	N Hablante	r	P
1	Región	39	-.58	**
2	Clase social	39	.31	*
3	Género	39	.20	
4	Edad	39	.20	
5	ESE	39	-.19	
6	Competencia inglés	39	-.11	
7	Edad de llegada	39	.06	
8	Educación	39	-.04	

** = p < .01
 * = p < .05
 a = p < .10

Tabla 3. Correlaciones Pearson con tasa pronominal en los INM.

Ranking	Variable	N Hablante	r	P
1	Región	75	-.51	**
2	Competencia inglés	75	.24	*
3	Edad de llegada	75	-.22	*
4	Educación	75	.13	
5	Edad	75	.08	
6	Clase social	75	.04	
7	ESE	75	-.03	
8	Género	75	.01	

** = p < .01
 * = p < .05
 a = p < .10

Las tablas 2 y 3 también corroboran los hallazgos del capítulo previo. Para los recién llegados, las correlaciones estadísticamente significativas con la tasa pronominal incluyen región y clase social. Para los inmigrantes, las correlaciones que dieron valores importantes incluyen región, competencia en inglés y edad de llegada.

Las tablas 1, 2 y 3 proveen una indicación preliminar de la importancia relativa de las variables externas que predicen la tasa pronominal. La variable más predictiva en el uso del pronombre en la muestra entera, en los recién llegados y en los inmigrantes es la región. Esta

variable representa la continuidad con el español de Latinoamérica como afirman en su trabajo Otheguy y Zentella (2012: 129). La variable que refleja el contacto de lenguas en Nueva York, la competencia en inglés, sólo es significativa, según es de esperar, en la muestra entera y en los inmigrantes. La edad de llegada es uno de los predictores más fuertes nuevamente para la muestra entera y los inmigrantes. La variable Clase social es la única variable sociodemográfica que afecta a los recién llegados mientras que la variable Edad es marginalmente significativa para todos los informantes de la muestra.

3. Regresiones lineales múltiples

Los análisis bivariantes de las tablas anteriores no responden realmente a la pregunta relativa a la fuerza de las variables externas en el uso pronominal. Estas correlaciones Pearson consideran el impacto de cada una de las variables por separado sobre la tasa pronominal. Es decir, analizan el impacto de Región sobre la tasa pronominal, el impacto de Competencia de inglés sobre la tasa pronominal, etc. Para examinar el efecto de varias variables predictivas sobre la variable dependiente, hay que realizar regresiones múltiples que consideran el efecto de todas las variables externas simultáneamente. Los análisis de regresiones múltiples informan cuál de las variables produce mayor variación, indicando qué variable externa es la que tiene mayor efecto predictivo sobre la tasa pronominal. Usando los resultados de las regresiones múltiples, se puede establecer una jerarquía de variables externas basadas en los efectos sobre la variable dependiente que refleja la relación real entre las variables predictivas y la tasa pronominal. A través de los análisis de regresión, conoceré la relativa importancia de las variables sociodemográficas y de las fuerzas regionales y de contacto que operan en el español de los inmigrantes y recién llegados.

En total, voy a realizar 21 regresiones lineales múltiples: tres regresiones para la muestra entera, tres regresiones por cada grupo de exposición y seis regresiones por cada grupo de exposición clasificados además según el origen regional. A pesar de que sólo algunas variables externas dieron resultados significativos en las correlaciones bivariantes anteriores, voy a incluirlas todas en las regresiones para tener datos más fiables y posteriormente, poder comparar los resultados.

Para que las regresiones múltiples den resultados fiables, las variables predictivas deben ser medidas conceptual y estadísticamente de forma independiente. Si las variables no son independientes o si sus factores se solapan, la regresión es menos efectiva. Este problema estadístico de solapamiento es conocido comúnmente como colinealidad (Newton and Rudestam, 1999: 264) o multicolinealidad (Field, 2009: 257). Si en nuestro análisis tuviéramos las variables Nación y Región (o sea, origen nacional y origen regional), no sería coherente incluirlas en la misma regresión debido a que estas dos variables están relacionadas conceptualmente. Igualmente, ocurre con las variables Edad de llegada y Competencia en inglés. A pesar de que estas variables son distintas conceptualmente, no es conveniente agruparlas en la misma regresión, ya que la competencia en inglés aumenta a una edad más temprana de llegada a Nueva York. Por razones similares, tampoco se incluyen en la misma regresión las variables Edad de llegada y Edad. Para evitar el problema de colinealidad ninguna de las regresiones realizadas incluyen estas variables juntas.

La regresión lineal calcula la cantidad de variación de la que da cuenta el modelo estadístico del modelo estadístico y la expresa en el valor de R^2 (R cuadrada). La R^2 es una medida del grado de cobertura del modelo en cuanto al total de variación que hay en los datos. Cuanto mayor la R^2 , mayor es la proporción de la variabilidad que explica la regresión y, por lo

tanto, mejor es el modelo. Incluida en el output de la regresión hay una ANOVA que prueba si el modelo entero es estadísticamente significativo. La regresión también calcula el peso relativo o la fuerza de asociación entre cada variable independiente y la variable dependiente usando una Beta estandarizada (β_1). El valor de la Beta estandarizada da información sobre el orden de las variables independientes en relación a otras (Field, 2009: 274; Newton y Rudestam, 1999: 267). Como resultado, el valor más alto de β_1 indica cuál es la variable predictiva con más fuerza. Así, se puede establecer una jerarquía de variables independientes externas. Además de calcular el peso de cada variable en la regresión, la β_1 , que puede ser positiva o negativa, permite comprender la relación entre cada una de las variables externas y la variable dependiente. Por ejemplo, el valor negativo de una β_1 para Región, codificada como 0 = caribeños y 1 = continentales, indica que los que más usan el pronombre de manera explícita en Nueva York son los caribeños. De forma similar, el valor positivo de una β_1 para Competencia en inglés, codificada como 0 = inglés menos que excelente y 1 = inglés excelente, señala que los informantes que saben mejor inglés son los que más utilizan el pronombre. La regresión también calcula un valor de p para cada β_1 . Este valor de p indica la significación estadística de cada variable independiente externa cuando es analizada simultáneamente con todas las variables en la misma regresión.

Por último, cada valor de R^2 y β_1 va seguido de uno o dos asteriscos o de una a que indican respectivamente $** = p < .01$, $* = p < .05$ y $a = p < .10$. En el caso de R^2 , estos valores informan si el modelo entero es estadísticamente significativo. En el caso de β_1 , estos valores nos dicen si cada variable es estadísticamente significativa.

En la siguiente sección, presento los resultados de las regresiones para la muestra entera. Luego, presento los resultados de las regresiones para cada uno de los grupos de exposición.

Finalmente, muestro las regresiones de los grupos de exposición clasificados según el origen regional.

3.1 Regresiones lineales en la muestra entera

En las primeras tres regresiones, las variables estudiadas son Región, Género y ESE que aparecen en los tres modelos alternativamente con Edad de llegada, Competencia en inglés y Edad. Las variables Educación y Clase social son excluidas de la regresión porque son parecidas conceptualmente a la variable ESE y porque de las tres éstas dos tuvieron una correlación más baja, a pesar de que ninguna de ellas fue significativa. A continuación, presento las tablas de las regresiones con los resultados para la muestra entera, o sea, para todos los hablantes de primera generación.

Tabla 4a. Regresión múltiple para la tasa pronominal, muestra entera		
Variable	β_1	p
Región	-.54	**
Edad de llegada	-.21	**
ESE	-.11	
Género	.11	
R2 = .359**		N = 114
** = p < .01		
* = p < .05		
a = p < .10		

Tabla 4b. Regresión múltiple para la tasa pronominal, muestra entera		
Variable	β_1	p
Región	-.52	**
Competencia inglés	.14	a
Género	.13	a
ESE	-.13	a
R2 = .331**		N = 114
** = p < .01		
* = p < .05		
a = p < .10		

Tabla 4c. Regresión múltiple para la tasa pronominal, muestra entera		
Variable	β_1	p
Región	-.54	**
Género	.12	
ESE	-.09	
Edad	.04	
R2 = .314**		N = 114
** = $p < .01$		
* = $p < .05$		
a = $p < .10$		

Los valores de R2 de las tablas 4a, 4b y 4c indican que hay una gran proporción de variabilidad. Los resultados de estos modelos son muy fiables ($p < .01$). La columna de la β_1 señala la fuerza de la relación entre cada variable predictiva y la variable dependiente. Esto permite establecer una jerarquía de variables independientes externas. En las tres tablas la variable que aparece en primer lugar es Región. Todas las regresiones señalan que los resultados para Región son fiables, ya que esta variable es estadísticamente muy significativa. Las variables Género (tabla 4a), ESE (tabla 4b) y Edad (tabla 4c) se muestran en la última posición de cada regresión respectivamente. Sólo Edad de llegada aparece segunda en la tabla 4a, Competencia en inglés ocupa la segunda posición en la tabla 4b y también Género en la tabla 4c es segunda.

El valor de la p en la columna derecha informa sobre la significación de cada variable cuando se mide simultáneamente con las demás variables del modelo. Los resultados indican que la única variable que es significativa en todos los modelos es Región. Este dato señala que podemos tener fiabilidad en estos resultados. De las demás variables, en el primer modelo (tabla 4a) Edad de llegada también es estadísticamente muy significativa. Competencia en inglés, Género y ESE tienen unos resultados marginalmente significativos en la regresión de la tabla 4b. Mientras que Género, ESE y Edad no son significativos según el modelo de la tabla 4c. Los datos sobre Género y ESE de la tabla 4b no parecen fiables, ya que en ninguna de las otras regresiones son marginalmente significativos.

Estos resultados parecen similares a los resultados de la correlación Pearson para la muestra entera, donde Región y Edad de llegada son muy significativos o significativos y el resto de las variables presentan valores marginales de significancia o ninguna significación estadística. La dirección de estos resultados es similar a la de los análisis bivariados del capítulo anterior: los hablantes de primera generación que presentan una tasa pronominal más elevada son los caribeños, que han llegado a Estados Unidos antes de los 12 años o con 12 años, y tienen un inglés excelente.

3.2 Regresiones lineales en los grupos de exposición: recién llegados e inmigrantes

En esta sección, analizo las mismas variables que consideré para las regresiones realizadas en la muestra entera, pero esta vez examino por separado la incidencia estadística que tienen las variables externas sobre la tasa pronominal en los dos grupos de exposición, recién llegados e inmigrantes.

Primero, muestro los resultados para los recién llegados. Para cada una de las tres regresiones, incluyo las variables Región, Género y Clase social alternativamente en los tres modelos con Edad de llegada, Competencia en inglés y Edad. Las variables Educación y ESE se excluyen de la regresión porque Educación obtuvo en la correlación un valor muy bajo y la variable ESE es conceptualmente igual a la variable Clase social que sí resultó ser significativa para este grupo de exposición en los análisis del capítulo anterior. Las siguientes tablas presentan los resultados de cada correlación.

Tabla 5a. Regresión múltiple para la tasa pronominal, R.LL.

Variable	$\beta 1$	p
Región	-.59	**
Género	.25	a
Clase social	.04	
Edad de llegada	-.01	
R2 = .408**		N = 39
** = $p < .01$		
* = $p < .05$		
a = $p < .10$		

Tabla 5b. Regresión múltiple para la tasa pronominal, R.LL.

Variable	$\beta 1$	p
Región	-.61	**
Género	.24	a
Competencia en inglés	-.12	
Clase social	.01	
R2 = .420**		N = 39
** = $p < .01$		
* = $p < .05$		
a = $p < .10$		

Tabla 5c. Regresión múltiple para la tasa pronominal, R.LL.

Variable	$\beta 1$	p
Región	-.58	**
Género	.23	
Edad	.04	
Clase social	.04	
R2 = .409**		N = 39
** = $p < .01$		
* = $p < .05$		
a = $p < .10$		

Este grupo de exposición difiere en significación y cantidad de variación según los resultados de las regresiones. Los valores de R2 de las tablas 5a, 5b y 5c indican que los modelos dan cuenta de una gran proporción de la variabilidad. Los resultados de todos los modelos son estadísticamente muy significativos ($p < .01$). Las jerarquías de variables externas expresadas a través de la $\beta 1$ permiten establecer que la variable más importante para los recién llegados en la tasa pronominal es la Región. Ninguna de las demás variables independientes externas muestra resultados significativos, a excepción de Género que en los dos primeros modelos presenta

resultados marginalmente significativos. Pero este resultado no es totalmente fiable, ya que la tercera regresión no la contempla como una variable significativa. Estos resultados difieren ligeramente de los resultados mostrados en las ANOVAS del capítulo anterior y la correlación Pearson de arriba. En ambos análisis, la variable Clase social demuestra valores de significación estadística. Las regresiones lineales, por el contrario, la sitúan en la penúltima y última posiciones en las tablas 5a, 5b y 5c y en ninguno de los modelos resulta significativa. En su lugar, las regresiones muestran que la variable Género es marginalmente significativa en los modelos de las tablas 5a y 5b como señalé previamente. En los análisis bivariantes y las correlaciones el predictor Género no es significativo, a pesar de que ocupa una tercera posición. Los datos sobre la clase social y el género no son muy fiables, porque en algunos análisis resultan significativos y en otros no. Lo que sí tienen en común los resultados es coincidir en señalar la importancia de la variable Región en la tasa pronominal de los recién llegados.

A continuación, presento los resultados para los inmigrantes. En las tres regresiones, incluyo las variables Región, Género y Educación de forma alternativa con Edad de llegada, Competencia en inglés y Edad. Las variables Clase social y ESE se excluyen de la regresión porque tuvieron valores más bajos que Educación y también porque se evita el problema de colinealidad. Las tablas siguientes muestran las regresiones para los inmigrantes.

Tabla 6a. Regresión múltiple para la tasa pronominal, INM.		
Variable	β_1	p
Región	-.52	**
Edad de llegada	-.24	*
Educación	-.05	
Género	.03	
R2 = .315**		N = 75
** = $p < .01$		
* = $p < .05$		
a = $p < .10$		

Tabla 6b. Regresión múltiple para la tasa pronominal, INM.		
Variable	β_1	p
Región	-.50	**
Competencia en inglés	.19	a
Género	.05	
Educación	-.04	
R2 = .293**		N = 75
** = $p < .01$		
* = $p < .05$		
a = $p < .10$		

Tabla 6c. Regresión múltiple para la tasa pronominal, INM.		
Variable	β_1	p
Región	-.50	**
Género	.05	
Educación	.03	
Edad	.02	
R2 = .262**		N = 75
** = $p < .01$		
* = $p < .05$		
a = $p < .10$		

En general, los resultados de las tablas 6a, 6b y 6c indican que todas las regresiones son estadísticamente significativas. Los valores de la R2 muestran que dan cuenta de una gran proporción de la variabilidad. Basándonos en la columna de la β_1 , observamos varias cosas sobre las variables. Todos los modelos sitúan la variable Región en primer lugar por ser muy significativa. La regresión de la tabla 6a informa además sobre la significación de la variable Edad de llegada para los inmigrantes. La regresión de la tabla 6b también señala que la variable Competencia en inglés es marginalmente significativa. Ninguna de las variables Género, Educación y Edad resultó ser importante para este grupo de exposición. Estos datos son muy similares a los obtenidos a través de las ANOVAS y las correlaciones. Parece que estos resultados son fiables. Las variables que diferencian a los inmigrantes en la tasa pronominal son la Región, la Edad de llegada y la Competencia en inglés.

3.3 Regresiones lineales en los grupos regionales de exposición

Ahora analizo qué variables externas son más significativas en cuanto a la tasa pronominal para los grupos de exposición recién llegados e inmigrantes divididos según el área regional. Voy a utilizar las variables estudiadas en las regresiones previas. La variable Región, que ha sido considerada en todos los análisis y que ha sido la que ha obtenido resultados más significativos, es eliminada porque no resulta relevante cuando se estudia cada región de forma separada.

Las regresiones siguientes son para los recién llegados caribeños y los recién llegados continentales. En cada modelo, incluyo las variables Género y Educación alternativamente con Edad de llegada, Competencia en inglés y Edad. Las variables Clase social y ESE se excluyen de la regresión porque obtuvieron valores más bajos que Educación en las ANOVAS del capítulo anterior. Las tablas de la izquierda describen los datos de las regresiones para los recién llegados caribeños y las tablas de la derecha presentan las regresiones para los recién llegados continentales.

Tabla 7.1a. Regresión múltiple para la tasa pronominal, R.LL. Caribeños		
Variable	$\beta 1$	p
Educación	.40	
Género	-.03	
Edad de llegada	-.00	
R2 = .153		N = 19
** = p < .01		
* = p < .05		
a = p < .10		

Tabla 7.2a. Regresión múltiple para la tasa pronominal, R.LL. Continentales		
Variable	$\beta 1$	p
Género	.52	*
Educación	-.10	
Edad de llegada	.06	
R2 = .183		N = 20
** = p < .01		
* = p < .05		
a = p < .10		

Tabla 7.1b. Regresión múltiple para la tasa pronominal, R.LL. Caribeños		
Variable	β_1	p
Educación	.57	*
Competencia inglés	-.46	a
Género	-.21	
R2 = .312		N = 19
** = p < .01		
* = p < .05		
a = p < .10		

Tabla 7.2b. Regresión múltiple para la tasa pronominal, R.LL. Continentales		
Variable	β_1	p
Género	.55	*
Competencia inglés	-.20	
Educación	-.02	
R2 = .349 ^a		N = 20
** = p < .01		
* = p < .05		
a = p < .10		

Tabla 7.1c. Regresión múltiple para la tasa pronominal, R.LL. Caribeños		
Variable	β_1	p
Educación	.42	a
Edad	.20	
Género	-.12	
R2 = .183		N = 19
** = p < .01		
* = p < .05		
a = p < .10		

Tabla 7.2c. Regresión múltiple para la tasa pronominal, R.LL. Continentales		
Variable	β_1	p
Género	.51	*
Edad	.11	
Educación	-.03	
R2 = .320 ^a		N = 20
** = p < .01		
* = p < .05		
a = p < .10		

Los resultados de las tablas que van de la 7.1a a la 7.2c señalan que los recién llegados caribeños y continentales varían en significación y variabilidad según demuestran las jerarquías de variables externas. Las tablas informan que las variables en la regresión tienen mayor variación en los recién llegados continentales que en los caribeños. Los valores de R2 apoyan este dato. Dos de las tres regresiones para los continentales son marginalmente significativas, mientras que ninguno de los tres modelos es significativo para los caribeños. En los recién llegados continentales, la única regresión que no da valores de significación es la que incluye edad de llegada, aunque está próxima a niveles marginales de significación.

Los recién llegados caribeños y los recién llegados continentales son diferentes respecto a la relevancia de las variables. Para los recién llegados caribeños, la β_1 señala en las tablas 7.1b y 7.1c que Educación y Competencia en inglés son las variables más importantes. La tabla 7.1a

también sitúa Educación en primer lugar a pesar de no tener valor significativo. Este resultado sobre la incidencia de la variable Educación en la tasa pronominal de los recién llegados caribeños coincide con el resultado de las ANOVAS del capítulo 3. Respecto a la variable Competencia en inglés, hay que señalar que su significación es marginal. Estos datos no son muy fiables porque ninguno de los modelos resultó ser estadísticamente significativo.

Las jerarquías de variables externas, representadas también por la β_1 , son diferentes en los recién llegados continentales. La variable que aparece en primer lugar en todas las regresiones y que es estadísticamente significativa en todas las tablas de los continentales es Género. Las demás variables no tienen ningún impacto. También, este resultado es similar al encontrado en los análisis bivariantes sobre la tasa pronominal. Los datos son más fiables. No obstante, las regresiones de la tabla 7.2a no son significativas y las de las tablas 7.2b y 7.2c son marginalmente significativas.

Las tablas que presento abajo representan las regresiones realizadas para el grupo de inmigrantes caribeños e inmigrantes continentales. Cada regresión es exactamente igual que las realizadas para los recién llegados caribeños y continentales. Se incluyen las variables Género y Educación de forma alternada con Edad de llegada, Competencia en inglés y Edad. Nuevamente, las variables Clase social y ESE se excluyen de los modelos para evitar la colinealidad. Las tablas de la izquierda indican los datos de las regresiones para los inmigrantes caribeños y las tablas de la derecha muestran los resultados de las regresiones para los inmigrantes continentales.

Tabla 8.1a. Regresión múltiple para la tasa pronominal, INM. Caribeños

Variable	$\beta 1$	p
Edad de llegada	-.65	**
Educación	-.12	
Género	-.01	
R2 = .378**		N = 40
** = p < .01		
* = p < .05		
a = p < .10		

Tabla 8.2a. Regresión múltiple para la tasa pronominal, INM. Continentales

Variable	$\beta 1$	p
Género	.21	
Edad de llegada	.18	
Educación	-.10	
R2 = .080		N = 35
** = p < .01		
* = p < .05		
a = p < .10		

Tabla 8.1b. Regresión múltiple para la tasa pronominal, INM. Caribeños

Variable	$\beta 1$	p
Competencia inglés	.37	*
Educación	.03	
Género	.00	
R2 = .146		N = 40
** = p < .01		
* = p < .05		
a = p < .10		

Tabla 8.2b. Regresión múltiple para la tasa pronominal, INM. Continentales

Variable	$\beta 1$	p
Género	.18	
Educación	-.11	
Competencia inglés	-.04	
R2 = .052		N = 35
** = p < .01		
* = p < .05		
a = p < .10		

Tabla 8.1c. Regresión múltiple para la tasa pronominal, INM. Caribeños

Variable	$\beta 1$	p
Edad	-.33	
Educación	-.08	
Género	-.03	
R2 = .088		N = 40
** = p < .01		
* = p < .05		
a = p < .10		

Tabla 8.2c. Regresión múltiple para la tasa pronominal, INM. Continentales

Variable	$\beta 1$	p
Edad	.38	*
Género	.13	
Educación	.00	
R2 = .175		N = 35
** = p < .01		
* = p < .05		
a = p < .10		

Las tablas 8.1a a la 8.2c informan que los inmigrantes caribeños y los inmigrantes continentales muestran diferentes niveles de significación para las variables y para la cobertura de la variación en los diferentes modelos. En general, los valores de la R^2 son bajos, así que hay menor grado de cobertura de la variabilidad. Los datos no son muy fiables. Sólo una de las regresiones (tabla 8.1a) alcanzó significación en los inmigrantes caribeños. Ninguna de las regresiones fue significativa en los inmigrantes continentales.

En relación a las jerarquías de variables para los inmigrantes caribeños, la columna de la β_1 indica que la variable Edad de Llegada es muy significativa en la tabla 8.1a y la Competencia en inglés es significativa en la tabla 8.1b. Las demás variables pierden fuerza en las regresiones. Estos datos son semejantes a los resultados de los análisis bivariados del capítulo anterior, que manifestaban la importancia de estas dos mismas variables. La variable Edad, que resultó ser marginalmente significativa en los análisis previos, ahora ocupa el primer lugar de la tercera regresión y aunque está próxima a alcanzar significación pierde peso estadístico.

El ranking de variables externas para los inmigrantes continentales sitúa a Género y Edad en primeras posiciones. La variable Género no es significativa en las regresiones de las tablas 8.2a y 8.2b, mientras que la variable Edad sí lo es en la regresión de la tabla 8.2c. También, estos resultados coinciden con los de las ANOVAS. La variable Edad es la única significativa que incide en la tasa pronominal de los inmigrantes continentales.

4. Interpretación de los resultados de las regresiones multivariadas

Los resultados de las correlaciones y las regresiones multivariadas mostrados a lo largo del capítulo desde la tabla 1 hasta la tabla 8.2c permiten hacer algunas generalizaciones sobre el español en Nueva York de en la primera generación, que abarca a los recién llegados y a los inmigrantes. Primero, voy a interpretar los resultados obtenidos para la muestra entera de los

hablantes de primera generación. Luego, voy a considerar los datos sobre los recién llegados y los inmigrantes. Finalmente, interpretaré los hallazgos encontrados en los grupos regionales de exposición.

4.1 El español en Nueva York en la primera generación

Los resultados de las regresiones sobre la muestra entera revelan que hay continuidad entre los modelos de uso del pronombre personal en Latinoamérica y la distribución de los pronombres en Nueva York en la primera generación. El factor más importante que da cuenta de las diferencias de uso del pronombre personal sujeto en Nueva York es la región de Latinoamérica de donde vienen los hablantes. Este resultado muestra que el español de la primera generación es una continuación del uso regional del español en el Caribe y del español en el continente. Otheguy y Zentella (2012: 139) en su estudio también mostraron que la región de donde proviene el hablante era el predictor más fuerte en la ocurrencia del pronombre. La diferencia es que estos autores estudiaron, para llegar a esa conclusión, la primera y segunda generaciones juntas y en mi estudio sólo analizo los informantes de la primera.

Otro descubrimiento importante para entender la variación del uso del pronombre en la primera generación es la importancia de la edad que tenía el hablante cuando llegó a Nueva York. Este dato señala que venir a la ciudad a partir de una edad concreta es determinante en la tasa pronominal. Llegar a Estados Unidos antes de los doce años significa un aumento del pronombre. El resto de las características sociodemográficas no son importantes. Es decir, ser hombre o mujer, tener una edad determinada, pertenecer a una clase social específica y poseer un nivel de educación específico no son rasgos que definan a estos hablantes en la realización del pronombre. La falta de relevancia de estas características se explica desde el planteamiento de Silva-Corvalán (2001: 133) de que las características sociales tienen un impacto mínimo en la

variación de características morfosintácticas. Sin embargo, el uso del pronombre en el español de los latinos de primera generación en Nueva York no puede entenderse considerando sólo la procedencia regional del hablante.

Otro dato importante es que el español de la primera generación no es una simple continuación del español latinoamericano. Hay otros elementos que están propiciando cambios y que apoyan la hipótesis del contacto de lenguas. La influencia del inglés contribuye a dar cuenta de la variación en el uso del pronombre en los hablantes de nuestra muestra de primera generación. Cuanto mayor es la competencia en inglés, más cantidad de pronombres utiliza el informante. Este hallazgo nos lleva a interpretar que está dándose una tendencia a realizar explícitamente el pronombre en los hablantes de primera generación que poseen un inglés excelente, como también demostraron en su estudio Otheguy y Zentella (2012: 140) cuando estudiaban la primera y segunda generaciones. La diferencia entre su trabajo y éste es que la influencia del inglés no es uno de los rasgos más fundamentales para entender el uso pronominal en los hablantes de la primera generación. Otheguy y Zentella (2012: 141) observaron que la influencia del inglés en la ocurrencia del pronombre era más relevante que otras características como, por ejemplo, el lugar de nacimiento o el tiempo de exposición en la ciudad. En mi trabajo, a pesar de que hay una relación entre tener un nivel de competencia alto en inglés y el aumento del pronombre, la influencia de la lengua de contacto no es de las características más determinantes en los hablantes de la variación pronominal de la primera generación.

4.2 El español en Nueva York según los grupos de exposición

La comparación de los grupos de exposición recién llegados e inmigrantes permite que se entienda mejor la hipótesis de la continuidad y del cambio en el español hablado por los latinos

nacidos en el Caribe y Latinoamérica. Los dos grupos de exposición muestran semejanzas y diferencias.

Los recién llegados siguen representando en las normas de uso del pronombre la continuidad con el español de Latinoamérica. Es decir, las regiones de donde proceden los hablantes marcan las pautas de uso pronominal de este grupo de hablantes. La realización del pronombre en el español de los recién llegados en Nueva York está condicionada básicamente por la continuidad del origen regional del hablante.

Respecto al impacto de otras características sociodemográficas en la tasa pronominal de este grupo de exposición, se pueden proponer varias interpretaciones. La primera tiene que ver con la pertenencia a una clase social determinada. De acuerdo a este criterio y como quedó demostrado en el capítulo anterior, la clase social del hablante diferencia a los recién llegados entre sí en cuanto al uso del pronombre. Los informantes que se identifican como de clase trabajadora utilizan más los pronombres. Sin embargo, este dato pierde importancia en relación al género del informante. Parece que ser hombre o mujer es lo que más distingue a los recién llegados en la tasa pronominal después de la región de procedencia. Las mujeres son las que utilizan más explícitamente el pronombre.

La influencia del inglés sobre el uso del pronombre no se encuentra entre los recién llegados. Este dato no apoya la hipótesis del contacto de lenguas en el español en Nueva York de los recién llegados. Para que el inglés influya el uso del pronombre de manera más evidente en los latinos de la ciudad, tiene que haber pasado más tiempo, como sucede en el caso de los inmigrantes de nuestra muestra.

Los inmigrantes muestran también la continuidad con los modelos de uso pronominal en Latinoamérica. A pesar de que estos informantes llevan más tiempo en Estados Unidos, la

superioridad de la procedencia regional refleja que hay un continuum con los modelos latinoamericanos que marca la tasa pronominal de este grupo de exposición. Entre los inmigrantes, el ser caribeño sigue determinando un mayor uso del pronombre.

Aparte de este dato de continuidad, en que los caribeños en Nueva York usan más pronombres que los continentales, reflejando el hecho de que en Latinoamérica se use más el pronombre en el Caribe que en las otras áreas, hay otra serie de hallazgos que evidencian cambios. Tener un nivel excelente de inglés en el español de los inmigrantes se corresponde con un incremento de la tasa pronominal.

Otra de las características sociodemográficas que también es definitoria del uso del pronombre en los inmigrantes: la edad a la que llegaron los inmigrantes a la ciudad. Los datos sugieren que llegar antes de los 12 años o llegar después de los 12 años a Nueva York afecta la tasa pronominal de este grupo. Los inmigrantes que llegaron antes de los 12 años usan más el pronombre. Hay un proceso de cambio que se manifiesta en llegar más joven o más mayor. La edad de llegada antes o después de los 12 años es predictiva de la tasa pronominal cuando se mira la muestra entera y esto sigue siendo así cuando se toman en cuenta sólo los inmigrantes.

El resto de las características externas no diferencia a los inmigrantes entre sí. Ni ser hombre o mujer, ni venir de una clase social alta o baja, ni tener una cierta edad son significativos en cuanto al nivel de la tasa pronominal. Los datos encontrados para los inmigrantes son más similares a los de la muestra entera. Las interpretaciones de los hallazgos coinciden en señalar a este grupo de exposición como el más representativo de la muestra. Tal vez, la explicación es que hay más informantes inmigrantes que recién llegados.

4.3 El español en Nueva York en los grupos regionales de exposición

La comparación de las regiones caribeña y continental dentro de los recién llegados y los inmigrantes permite obtener un mejor entendimiento acerca de la continuidad y del cambio en el español hablado por los latinos nacidos en Latinoamérica.

Las interpretaciones de los resultados para los recién llegados indican que los recién llegados caribeños y los recién llegados continentales conforman un grupo heterogéneo en que sus informantes usan el pronombre de distinta manera. Estos dos grupos regionales son diferentes en el rasgo que más impacta en la tasa pronominal. La única característica que distingue a los recién llegados caribeños entre sí es el nivel de educación. Los recién llegados caribeños son distintos del resto porque usan más el pronombre cuando tienen mayor educación. Esta interpretación sobre este dato coincide con el resultado del capítulo anterior sobre la tasa pronominal.

Respecto al nivel de competencia en inglés, los hallazgos son sorprendentes. Parece que los recién llegados caribeños que tienen menos competencia en inglés son los que presentan una tasa pronominal más elevada. Si esto es cierto, la hipótesis del contacto de lenguas en este grupo no es válida. Parece que este grupo regional de hablantes que todavía no lleva mucho tiempo en la ciudad de Nueva York no es susceptible a la influencia del inglés. Las demás características sociodemográficas como ser hombre o mujer, la edad a la que llegó a la ciudad y los años que tenía en ese momento no son rasgos que diferencien a los recién llegados caribeños en el uso del pronombre.

Los recién llegados continentales no se distinguen entre sí según el nivel de educación, como hemos visto que sucede entre los caribeños. Los continentales se distinguen entre sí únicamente por ser hombre o mujer. Los datos sugieren que el género del informantes es lo único

relevante para entender el uso pronominal en los recién llegados continentales. Son las mujeres las que más utilizan el pronombre. Este hallazgo nos lleva a proponer la siguiente interpretación: la mujer en los países continentales de Latinoamérica presenta una tasa pronominal más elevada que la de un hombre.

Respecto al impacto de otros rasgos sociales, los recién llegados continentales no se diferencian entre sí por ninguno. Poseer un nivel excelente o menos que excelente de inglés, tener una edad determinada, haber llegado antes o después de los doce años o tener educación elemental o superior no son características que distingan a estos hablantes en la realización del pronombre.

Los hallazgos encontrados sobre la tasa pronominal de los inmigrantes indican que estos hablantes pueden ser divididos en un grupo caribeño más heterogéneo diferenciado por distintas características sociales y un grupo continental homogéneo, más estable.

Una de las características sociodemográficas y otra asociada con el contacto de lenguas son las que diferencian a los inmigrantes caribeños. Lo más significativo para los inmigrantes caribeños es la edad a la que llegó el informante. Este rasgo sociodemográfico es lo que más define la tasa pronominal de estos hablantes. Llegar a Estados Unidos antes o con doce años de edad es el criterio más relevante para diferenciar a los inmigrantes caribeños en la realización del pronombre.

También, hay otro rasgo que distingue a estos hablantes del resto. La relación entre el uso pronominal y el conocimiento de la lengua de contacto es relevante. Aquellos informantes inmigrantes del Caribe que tienen un inglés excelente, presentan un incremento en el uso del pronombre. Este dato evidencia la influencia del sistema gramatical inglés sobre el español y confirma la hipótesis del contacto de lenguas. En el caso de los inmigrantes caribeños que ya por

su procedencia regional presentan una tasa pronominal elevada, el conocimiento del inglés refuerza el uso del pronombre.

Los hallazgos señalan también que los años del informante, el nivel de educación y el género no tienen ninguna repercusión en la variación pronominal. Es decir, tener veinte o cuarenta años, poseer un título elemental o universitario y ser hombre o mujer no diferencian a los inmigrantes caribeños en la realización del pronombre. Los datos presentados coinciden también con los datos del capítulo anterior sobre la tasa pronominal de los inmigrantes caribeños.

Mirando ahora a los inmigrantes continentales, hemos visto que constituyen un grupo más homogéneo y estable. Solamente los años del informante diferencian a este grupo de exposición regional. Este dato confirma que lo que condiciona el uso del pronombre en los inmigrantes continentales es tener una edad determinada, por ejemplo, tener veinte, treinta, cuarenta, cincuenta años o más. Es decir, ser más joven o más mayor. Los inmigrantes continentales que tienen más años presentan una tasa pronominal más elevada.

Otro hallazgo es el género del hablante, a pesar de que no es definitorio. Este dato parece señalar que también las mujeres continentales, en este caso las mujeres inmigrantes, utilizan el pronombre de forma explícita más frecuentemente que los hombres como ocurre con las mujeres recién llegadas continentales. Este dato nos lleva a proponer nuevamente que en Latinoamérica las mujeres utilizan más los pronombres.

En relación a los demás rasgos externos, los inmigrantes continentales no son distintos por ninguna otra característica sociodemográfica. Tampoco, se diferencian por el nivel de inglés que tienen. El conocimiento de la lengua de contacto no ejerce ninguna presión en el aumento de la tasa pronominal de este grupo regional de inmigrantes continentales a pesar de que estos hablantes han estado más expuestos al inglés porque llevan más años en Estados Unidos.

Resumen y conclusión

Los datos analizados en este capítulo permiten confirmar varios hallazgos sobre el uso del pronombre personal sujeto en los recién llegados y los inmigrantes en Nueva York. Primero, hay diferencias en la tasa pronominal entre los dos grupos de exposición y entre los grupos regionales de exposición. Segundo, la fuerza de la continuidad del español en Latinoamérica en la distribución pronominal en el español de estos dos grupos de informantes es mayor que la fuerza de cambio.

Los diferentes análisis, a través de las correlaciones y las regresiones múltiples, han revelado la importancia de las variables sociodemográficas y de las fuerzas regionales y de contacto que definen el español de los recién llegados y los inmigrantes. Hemos analizado cuál de las variables produce mayor variación y hemos observado qué variables externas son las que tienen mayor efecto predictivo sobre la tasa pronominal, examinando que estos dos grupos de exposición se distinguen.

En la muestra entera, éstas son algunas diferencias significativas encontradas:

- Las variables externas que diferencian a todos los hablantes de primera generación son Origen Regional, Edad de llegada a Estados Unidos y Competencia en inglés.
- Hay continuidad en los modelos de uso del pronombre personal en Latinoamérica.
- Existen cambios por el contacto de lenguas que contribuyen a la variación en el uso del pronombre.

Las regresiones mostraron que la variable más importante que distingue entre sí a todos los hablantes de primera generación es la Región. Los análisis indican que los caribeños de la muestra entera exhiben una tasa pronominal más elevada que la de los continentales. El origen geográfico es un criterio válido para agrupar de forma uniforme a los hablantes (Cameron, 1995:

8; Lipski, 1994: 241; López Morales, 1992: 137; Otheguy y Zentella, 2012: 146). El hecho de que Región sea la variable más significativa implica una continuación del uso regional de las normas de utilización del pronombre del español caribeño y del español continental. Es decir, estos hablantes utilizan el pronombre de manera similar a como se usa en el Caribe y Latinoamérica.

El segundo dato interesante es que la variable Edad de llegada también distingue a estos informantes en la tasa pronominal. Hay una fuerte correlación entre llegar a Estados Unidos antes de los doce años y el aumento del pronombre. Los informantes que vinieron a la ciudad antes de haber cumplido doce años son los que más usan el pronombre explícitamente. No hay otra variable sociodemográfica que impacte en la variación pronominal. Ser hombre o mujer, tener un título de la secundaria o de la universidad, pertenecer a la clase trabajadora o media no son rasgos definitorios en el uso del pronombre.

También, los resultados confirman que hay evidencia del contacto de lenguas a través de la influencia del inglés. Los informantes que tienen un nivel de competencia excelente en esta lengua realizan más el pronombre. La variable de uso de la lengua con unos valores de significación marginales nos lleva a considerar que el español de la primera generación está condicionado por los patrones de uso del sistema pronominal inglés.

En los grupos de exposición recién llegados e inmigrantes, ocurre lo siguiente:

- Los grupos son iguales por el impacto del origen regional sobre la tasa pronominal.
- Los grupos son diferentes por la incidencia de otros factores externos en la tasa pronominal.
- La hipótesis de la continuidad con el español latinoamericano sigue siendo más fuerte que la hipótesis del cambio.

Los recién llegados y los inmigrantes son iguales en cuanto a la región de procedencia en el uso del pronombre. Esta variable externa sigue siendo la que ejerce mayor fuerza predictiva en la tasa pronominal. Son los caribeños de ambos grupos los que presentan el mayor nivel de uso del pronombre. Este dato sugiere que el aumento del uso del pronombre en el español de los recién llegados y los inmigrantes en Nueva York está condicionado por la continuidad del origen regional del hablante. Nuevamente, los resultados reafirman la fuerza de la hipótesis de la continuidad.

Sin embargo, los recién llegados y los inmigrantes sí se distinguen los unos de los otros por la variable de contacto y otras de tipo sociodemográfico. Los dos grupos de exposición se muestran distintos en la influencia del inglés, del género y de la edad de llegada. Los recién llegados son diferentes únicamente por el género del hablante. Son las mujeres recién llegadas las que presentan una tasa pronominal alta. Este resultado sitúa a las mujeres caribeñas y latinoamericanas como las promotoras en la variación pronominal en Nueva York. En cambio, los inmigrantes son distintos por la edad a la que llegaron y el nivel de competencia en inglés. Los inmigrantes que llegaron antes de los 12 años a Estados Unidos y que tienen un inglés excelente utilizan más el pronombre. La hipótesis del contacto de lenguas sólo toma fuerza con los inmigrantes.

En los grupos de exposición recién llegados caribeños y continentales e inmigrantes caribeños y continentales, los resultados añaden más cosas:

- Los recién llegados caribeños y los recién llegados continentales conforman un grupo heterogéneo diferenciado por el impacto de una variable distinta en cada grupo regional.

- Los inmigrantes caribeños y los inmigrantes continentales conforman un grupo heterogéneo también. Los inmigrantes caribeños son más cambiantes y los continentales son más estables.
- Los cambios por el contacto de lenguas que influyen en la variación en el uso del pronombre sólo ocurren en los inmigrantes caribeños.

Los análisis muestran que la variable externa que tiene mayor fuerza predictiva en la tasa pronominal es distinta en cada grupo de exposición regional. Los recién llegados caribeños se diferencian por el nivel de educación que tienen. Utilizan más el pronombre cuando tienen un nivel más alto de educación. Los recién llegados continentales son distintos por el género del informante. Las mujeres son las que mayor tasa pronominal presentan. Los inmigrantes caribeños se distinguen por la edad a la que vinieron a Estados Unidos y el nivel de competencia en inglés. Los que tienen menos de 12 años y saben un inglés excelente son más propensos a usar el pronombre. Los inmigrantes continentales sólo son diferentes por los años que tienen. Tener más edad se correlaciona con el aumento del pronombre. Los análisis también señalan que la fuerza del cambio por el contacto de lenguas es mínima en tres de los grupos. Únicamente, la hipótesis del contacto con el inglés adquiere sentido en los inmigrantes que son caribeños.

En definitiva, las jerarquías de variables externas confirman los datos del capítulo previo y establecen la primacía de la continuidad regional en la tasa pronominal de los recién llegados y los inmigrantes en Nueva York. El constante flujo de inmigrantes latinoamericanos y caribeños contribuye al mantenimiento de las normas de habla de sus países. También, las regresiones corroboran que la hipótesis del contacto de lenguas sólo es válida para los informantes que llevan más tiempo en la ciudad y han estado más expuestos a la lengua de contacto. Es decir, los que tienen más conocimiento del inglés son los inmigrantes y, específicamente, los que vienen del

Caribe. Los recién llegados presentan una cierta resistencia a la influencia del inglés en el uso pronominal probablemente debido a la poca exposición con esta lengua y a que su nivel de conocimiento es menor.

La variación pronominal de los recién llegados y los inmigrantes es una combinación de las fuerzas de la continuidad con el español de Latinoamérica y las presiones del cambio por el contacto con el inglés.

Capítulo 5

Análisis de las variables internas

1. Introducción

Este capítulo analiza la importancia de los factores internos en la variación en el uso pronominal de los inmigrantes y recién llegados. Los factores internos son diferentes de los factores externos o sociodemográficos, estudiados anteriormente. Estos revelaron que la continuidad con las normas latinoamericanas y el contacto entre lenguas (español-inglés) son fuerzas independientes que inciden de manera distinta en el uso pronominal. Las variables sociales se referían al origen regional, la edad de llegada a EE.UU, la edad, la educación, la clase social, el estatus socio-económico (ESE) y la competencia en inglés.

Los factores internos que investigo se relacionan con los verbos en forma personal que aparecen con o sin un pronombre personal sujeto. Estos factores internos se basan en tres clases diferentes de propiedades: morfológicas, sintácticas y discursivas. Las propiedades morfológicas son aquellas que tienen que ver con el número y la persona del verbo (ejemplo: *(yo) hablo ~ (tú) hablas*) o con el tiempo verbal (*(yo) hablo ~ (yo) hablé*). Las propiedades sintácticas se refieren, entre otras cosas, a la reflexibilidad o no reflexibilidad del verbo (ejemplo: *(yo) me visto ~ (yo) visto a los niños*). Las propiedades discursivas tienen relación con la continuidad o el cambio de referente. Si el referente del pronombre personal sujeto del verbo es el mismo que el del verbo anterior hay continuidad con la referencia. Un ejemplo de misma referencia es *Antonio dijo que (él) llegaría tarde*. Si el referente es distinto, entonces hay cambio de referencia. Un ejemplo de cambio de referencia es *Antonio dijo que (tú) llegarías tarde*.

Los factores internos de este capítulo representan una combinación de la coocurrencia del verbo y de las características denotativas del sujeto del verbo. En los ejemplos anteriores, la

flexión que indica persona y número en el verbo distingue entre la primera persona del singular (terminación en *-o*) y la segunda persona del singular (terminación en *-as*). Los pronombres sujetos en español suelen distribuirse de manera desigual con respecto a los contextos lingüísticos y tipos de denotaciones asociados con los verbos con los que coexisten. En el corpus, la mayoría de los pronombres ocurre más con verbos en singular que en plural. Igualmente, hay más pronombres explícitos en situaciones de cambio de referencia (con pronombre: *Antonio dijo que tú llegarías tarde*) que en los casos donde la referencia es la misma (sin pronombre: *Antonio dijo que llegaría tarde*). Estas condiciones diferentes en las que aparecen los verbos son las *variables independientes* que co-varían en diferentes grados con la presencia y la ausencia de los pronombres. Igual que las variables independientes centradas en el hablante del capítulo 4, como Región, Edad de llegada a EE.UU, Género, etc, en adelante estudio las variables independientes internas o gramaticales como Persona del verbo o Cambio de referencia (las cuales de ahora en adelante tratamos como nombres propios y escribimos con mayúscula inicial, ej. *Persona* y *Referencia*).

Las variables independientes internas difieren de manera importante de las variables independientes externas antes estudiadas, tales como Región o Edad de llegada a EE.UU, que se refieren a características que distinguen a los grupos de hablantes que mostraron distintas tasas pronominales. Las variables externas no son parte integral de la gramática de los hablantes (Otheguy y Zentella, 153: 2012). Ser caribeño o continental y haber llegado más joven o más mayor no son aspectos que configuran la gramática. Sin embargo, las variables internas como Persona o Referencia deben ser analizadas como elementos que influyen en la gramática a la hora de utilizar o no un pronombre personal sujeto (Cedergren y Sankoff, 1974; Clark, 2005; Guy 2005, 2007; Newmeyer, 2006; Poplack 2000).

En sociolingüística variacionista, una variable independiente como Persona o Referencia se llama grupo de factores (ing. “factor group”). A cada uno de los diferentes niveles de la variable independiente o del grupo de factores se le llama “factor”. Por ejemplo, la variable independiente o el grupo de factores que llamamos Persona tiene cinco factores que son las cinco desinencias de los verbos en español (primera, segunda y tercera personas del singular y primera y tercera personas del plural). La variable independiente o el grupo de factores que llamamos Referencia tiene dos factores que son el cambio de referencia o la misma referencia. A partir de ahora, utilizaré los términos *variable independiente* (en lugar de grupo de factores) y *factores* por ser los más usados en la investigación sociolingüística.

Analizaré las variables independientes y sus factores para estudiar la importancia relativa de la continuidad con las normas de uso del español latinoamericano y el cambio asociado con el contacto de lenguas español-inglés en Nueva York. Las variables se comparan respecto a su fuerza o grado de influencia a la hora de usar un pronombre u omitirlo. Quiero averiguar si Persona es una variable independiente más fuerte o influyente sobre la decisión de usar u omitir un pronombre sujeto que, por ejemplo, Referencia. Primero, hacemos este análisis en la muestra entera y en hablantes pertenecientes a lo que hemos llamado diferentes grupos de exposición (en el sentido de largo del tiempo de exposición a la vida en Nueva York), y segundo, en hablantes clasificados según el origen geográfico, y el origen geográfico y la exposición. Para el análisis de los factores, las comparaciones de fuerza o influencia predictiva se miden en términos de la probabilidad de que un pronombre sea usado en la presencia de un factor determinado. Por ejemplo, la probabilidad de que un pronombre ocurra con un verbo con cambio de referencia es mayor que con un verbo con la misma referencia. Voy a estudiar si los verbos con cambio de referencia que favorecen más el uso explícito del pronombre se comportan del mismo modo en la

muestra entera y en los diferentes grupos de exposición, diferenciados también regionalmente. O si, por el contrario, ha habido alguna modificación originada por el contacto de lenguas. Quiero determinar si las consideraciones de carácter interno (ejemplo, las variables internas y sus factores) se pueden usar para estudiar la continuidad y el cambio inducido por el contacto en el español en Nueva York, como ocurrió con las consideraciones de tipo social en los capítulos anteriores.

Los rankings de variables y factores se hacen a través de análisis de regresión estadística cuyos resultados producen jerarquías de variables y jerarquías de factores. Al analizar si Persona es más fuerte o más débil que Referencia en un grupo de hablantes en el uso del pronombre, analizo si todos los grupos de hablantes que estudiamos comparten la misma jerarquía de variables. Lo mismo sucede cuando investigo si dos grupos de hablantes, como por ejemplo los inmigrantes recién llegados y los inmigrantes establecidos, presentan el mismo ranking para los verbos con cambio de referencia que para los verbos con misma referencia. Analizo si estos dos grupos comparten la misma jerarquía de factores para la variable Referencia.

La comparación y el análisis de las jerarquías de variables y factores se conoce como sociolingüística comparativa (Tagliamonte, 2002: 729). En este tipo de trabajos, como los estudios de Poplack (2000), Poplack y Tagliamonte (1999) y Poplack y Sankoff (1987), las variables y los factores que condicionan la variación son más relevantes para la identificación de los grupos y de sus relaciones. Igualmente ayudan a distinguir más entre la continuidad y el cambio que el estudio de la tasa pronominal realizado en capítulos anteriores (Poplack, 2000: 14).

La otra gran diferencia en este capítulo con respecto a los anteriores es el cambio de la variable dependiente que utilizamos para comparar grupos de informantes. Anteriormente,

estudié una característica que tiene que ver con el informante y a la que llamamos Tasa Pronominal. Ahora, analizo una característica relacionada con el verbo, que tiene que ver con la ausencia o la presencia de un pronombre personal sujeto. Por este motivo, la co-variación de las variables independientes de los capítulos 3 y 4 se basaba en las propiedades sociodemográficas y personales de los informantes (Región, Edad de llegada, Género, etc), mientras que la co-variación de las variables independientes que se estudian ahora son propiedades gramaticales y discursivas de los verbos (Persona, Referencia, Reflexividad, etc). O sea, que el análisis estadístico de las variables sociodemográficas y de la lengua se hizo al nivel del informante. Mientras que el análisis estadístico de las variables gramaticales se realiza al nivel de cada verbo. Finalmente, la variable Tasa Pronominal era una variable continua, es decir, un porcentaje. La variable dependiente ahora se llama Ocurrencia y es una variable categórica, que es la presencia o la ausencia del pronombre con un verbo particular. Esta variable es categórica porque cada verbo puede aparecer con o sin un pronombre.

2. Las variables gramaticales y discursivas que definen la ocurrencia del pronombre

Una vez establecida la variable dependiente Ocurrencia, presento las cinco variables independientes internas y sus factores con abreviaciones. He seleccionado estas variables porque la literatura muestra que son las más influyentes en la distribución del pronombre (Barrenechea y Alonso, 1977; Bayley y Pease-Álvarez, 1996; Cameron, 1996; Enríquez 1984; Morales, 1997; Silva-Corvalán, 1982, 1988, 1994).

La lista de variables independientes y de sus factores es muy clara. Por ejemplo, la variable Tiempo es una variable morfológica que se basa en las terminaciones del tiempo verbal para diferenciar entre presente, pasado, futuro, etc. También, Referencia es una variable discursiva y comunicativa que se basa en si el referente del verbo es el mismo o diferente

respecto al verbo anterior. A continuación, presento una lista con la variable dependiente, las variables independientes gramaticales y discursivas y sus factores:

Variable dependiente:

Ocurrencia del pronombre personal (Ocurrencia)

- Ausencia del pronombre
- Presencia del pronombre

Variables independientes gramaticales y discursivas:

Persona del verbo (Persona)

- Primera persona singular (1sg)
- Segunda persona singular (2sg)
- Tercera persona singular (3sg)
- Primera persona plural (1pl)
- Tercera persona plural (3pl)

Tiempo y modo del verbo (Tiempo)

- Presente de indicativo
- Pasado de indicativo (incluye pretéritos indefinidos e imperfectos)
- Futuro
- Condicional
- Presente perfecto
- Modo subjuntivo

Reflexividad del verbo (Reflexividad)

- Verbo está en construcción reflexiva
- Verbo no está en construcción reflexiva

Significado semántico y léxico del verbo (Léxico)

- Verbo mental o estimativo
- Verbo de estado
- Verbo externo

Referencia del sujeto (Referencia)

- Misma referencia
- Cambio de referencia

3. Regresiones logísticas

Como la variable dependiente Ocurrencia es dicotómica (presencia o ausencia del pronombre) el tipo de análisis estadístico que uso es la regresión logística. En el capítulo anterior, la variable dependiente Tasa Pronominal era continua y el análisis estadístico fue la regresión lineal. Según Menard (2002) y Pampel (2000), el análisis apropiado para medir una variable dependiente dicotómica es la regresión logística.

La regresión logística predice el resultado de una variable dependiente nominal usando algunas variables predictivas nominales o continuas y determina cuál de ellas representa la mayor variación. Los resultados de las regresiones logísticas se usan para establecer una jerarquía de variables predictivas basadas en su efecto sobre la ocurrencia (la presencia o ausencia) del pronombre personal sujeto. Y se usan para establecer, dentro de cada variable, una jerarquía de factores predictivos, que nos indican cuáles de ellos favorece o desfavorece la ocurrencia del pronombre, y en qué orden.

Una regresión logística calcula la cantidad de variación representada por las variables en la regresión, expresada como un valor R^2 (R cuadrada). La regresión también calcula el peso relativo o la fuerza de asociación entre cada variable independiente y la variable dependiente

usando un valor de *Wald* (Menard, 2002: 43). Las jerarquías de variables son clasificadas en términos del valor de Wald de cada variable independiente. La variable con el valor de Wald más alto representa la mayor fuerza o influencia estadística sobre la ocurrencia del pronombre, cuando se compara con todas las demás variables introducidas en la regresión.

Un segundo valor, el valor exponencial *B*, también nos permite determinar tanto el peso como la dirección de la relación entre cada uno de los factores de las variables independientes y la variable dependiente. De esta manera, podemos crear dos jerarquías, una jerarquía de variables basada en los valores de Wald y una jerarquía de factores basada en los valores exponenciales *B*. Las jerarquías de factores, que derivan de los mismos modelos de regresión logística usados para crear jerarquías de variables, clasifican la probabilidad de ocurrencia del pronombre personal sujeto con cada uno de los factores en una variable.

4. Jerarquías de variables y factores

Ahora investigo cuál de estas variables independientes internas es la más relevante en relación a la variable dependiente Ocurrencia. Trato de estudiar qué consideraciones lingüísticas influyen en los hablantes inmigrantes y los recién llegados a la hora de usar u omitir el pronombre personal sujeto.

Para medir la fuerza de las diferentes variables independientes, uso el mismo análisis multivariable del capítulo anterior con el objetivo de comparar las variables independientes pero en esta ocasión lo hago a través del análisis de la regresión logística. Esta vez, no haré asociaciones bivariadas para medir la fuerza entre Ocurrencia y Persona, Ocurrencia y Tiempo, Ocurrencia y Referencia, etc. Este tipo de comparación bivariable puede distorsionar los resultados. Para evitar este problema, realizo análisis de regresiones logísticas que están más estandarizados en la investigación sociolingüística (Tagliamonte, 2002, 2006). Además en los

resultados señalados en Otheguy y Zentella (2012: 158-160), estas cinco variables independientes internas fueron las más significativas.

En los siguientes apartados, presento los resultados de cinco regresiones logísticas. La primera para la muestra entera. La segunda y la tercera regresiones para cada uno de los grupos de exposición. La cuarta y quinta son las regresiones de los grupos de exposición clasificados según el origen regional.

4.1 Jerarquías de variables y factores en la muestra entera

La jerarquía de variables para la primera regresión logística incluye todos los verbos de la muestra, en total, 50,626. El objetivo es investigar cuál de estas variables internas predice mejor la ocurrencia del pronombre en la muestra entera. Los resultados aparecen en la tabla 1.

Tabla 1. Jerarquía de variables que predicen la Ocurrencia, muestra entera			
Orden	Variable	Wald	<i>p</i>
1	Persona	1736.61	**
2	Referencia	1479.46	**
3	Tiempo	189.46	**
4	Léxico	121.81	**
5	Reflexividad	93.90	**
R2 = .139		N verbos = 50,626	
** = $p < .01$			
* = $p < .05$			
a = $p < .10$			

La *N* en la tabla 1 indica el número de verbos analizados en este modelo de regresión. En total, son 50,626. La tabla también señala, en el coeficiente de *R2* (R cuadrada), la cantidad de variación entre la presencia y la ausencia del pronombre personal sujeto que contabiliza este modelo de regresión logística. En este caso, hay casi un catorce por ciento de variación. El valor de Wald asignado a cada variable tiene su significación estadística anotada con asteriscos en la forma usual. Este valor indica la significación estadística de cada variable independiente interna cuando es analizada simultáneamente con todas las variables en la misma regresión. Así, cada

valor va seguido de uno o dos asteriscos o de una a que indican respectivamente $** = p < .01$, $* = p < .05$ y $a = p < .10$. El modelo de regresión para todos los hablantes de primera generación nos dice que la variable Persona es el predictor más fuerte en la realización del pronombre seguida de Referencia, Tiempo, Léxico y Reflexividad.

Estos datos reflejan que la consideración más importante que guía el uso o la omisión del pronombre en la primera generación es la persona del verbo; seguida por la referencia, es decir, si el sujeto del verbo es el mismo que el sujeto del verbo anterior o si hay un cambio de referencia; seguida por el tiempo y modo verbales si es presente, pasado, modo subjuntivo, etc; seguida por el tipo de verbo si es mental, de estado, etc; y por último, seguida por la reflexividad del verbo si el verbo va acompañado de un pronombre reflexivo o no.

La jerarquía de variables de la tabla 1 nos informa acerca de la importancia de todas estas variables independientes internas, pero no nos dice cuáles son los factores que favorecen más o menos, dentro de cada una de estas variables, la realización del pronombre. Para saber qué factores predicen la aparición del pronombre, necesitamos hacer jerarquías de factores para cada una de las variables internas. Las regresiones proporcionan un valor exponencial B para cada uno de los factores. Siguiendo la metodología utilizada para codificar los datos, si el valor exponencial B es mayor que 1.00 significa que el factor predice la aparición de un pronombre personal. Los valores menores que 1.00 indican que el factor desfavorece el uso del pronombre personal. A continuación, presento cinco tablas con la jerarquía de factores de cada variable.

Tabla 1a. Jerarquía de factores para la variable interna Persona, muestra entera			
Orden	Factor	Exp (B)	p
1	2ª persona singular	42.23	**
2	3ª persona singular	37.48	**
3	1ª persona singular	35.00	**
4	3ª persona plural	0.91	
5	1ª persona plural	0.46	**
N verbos = 50,626			

En el caso de la variable Persona, observamos que sólo las personas del singular favorecen el uso explícito del pronombre personal. Los tres primeros factores de la variable Persona presentan valores por encima de 1.00. Hay diferencias entre los tres primeros factores y los dos últimos. La jerarquía de factores muestra que los hablantes de primera generación prefieren los pronombres con las formas singulares. Las personas en plural desfavorecen el uso pronominal. Usan el pronombre más con la segunda persona del singular seguida por la tercera persona del singular y por la primera persona del singular. Además, los tres primeros factores son muy significativos desde el punto de vista estadístico, mientras que los dos últimos no lo son.

Tabla 1b. Jerarquía de factores para la variable interna Referencia, muestra entera

Orden	Factor	Exp (B)	<i>p</i>
1	Cambio referencia	1.50	**
2	Misma referencia	.66	**
N verbos = 50,626			

En la variable Referencia, ocurre que los verbos cuyos sujetos tienen una referencia diferente a la del verbo anterior en el discurso tienden a aparecer con un pronombre personal sujeto. Los verbos que tienen la misma referencia que el verbo precedente desfavorecen el uso del pronombre.

Tabla 1c. Jerarquía de factores para la variable interna Tiempo, muestra entera

Orden	Factor	Exp (B)	<i>p</i>
1	Condicional	2.13	**
2	Presente	1.55	**
3	Pasado	1.52	**
4	Presente perfecto	1.27	**
5	Modo subjuntivo	0.87	**
6	Futuro	0.73	**
N verbos = 50,626			

La tabla 1c informa que en la primera generación los tiempos que más favorecen la utilización del pronombre son el condicional, el presente, el pasado y el presente perfecto. La jerarquía de factores indica que los valores de los primeros cuatro factores están por arriba de 1.00. El tiempo y el modo verbales con los que mayor probabilidad de uso tiene el pronombre son el condicional, seguido por el presente, el pasado y el presente perfecto. El único modo y tiempo que en la primera generación no favorece al pronombre es el modo subjuntivo y el futuro. La jerarquía también nos informa de que todos los factores son muy significativos.

Tabla 1d. Jerarquía de factores para la variable interna Léxico, muestra entera			
Orden	Factor	Exp (B)	<i>p</i>
1	Mental-estimativo	1.19	**
2	Estado	.96	
3	Externo	.87	
N verbos = 50,626			

La jerarquía de factores de la tabla 1d nos dice que sólo hay un valor mayor que 1.00. Los verbos de tipo mental-estimativo son los que tienen mayor probabilidad de llevar un pronombre. A pesar de los valores exponenciales *B* próximos a 1.00, los verbos de estado y los externos no tienden a ir acompañados del pronombre.

Tabla 1e. Jerarquía de factores para la variable interna Reflexividad, muestra entera			
Orden	Factor	Exp (B)	<i>p</i>
1	No reflexivo	1.17	**
2	Reflexivo	.85	
N verbos = 50,626			

El último modelo de la tabla 1e informa que la primera generación usa el pronombre con mayor probabilidad con los verbos que no son reflexivos. Si el verbo está acompañado de un pronombre reflexivo desfavorece la presencia del pronombre personal sujeto.

En términos generales, todas las variables internas, Persona, Referencia, Tiempo, Léxico y Reflexividad, tienen influencia en la ocurrencia del pronombre en la primera generación. Más explícitamente, la presencia del pronombre es favorecida con las tres personas del singular, con el cambio de referencia, con los tiempos verbales del indicativo a excepción del futuro y del modo subjuntivo, con los verbos de contenido semántico mental-estimativo y con los verbos que no llevan un pronombre reflexivo.

4.2 Jerarquías de variables y factores en los grupos de exposición: recién llegados e inmigrantes

En los capítulos previos sobre las variables externas y su incidencia en la tasa pronominal, quedó establecido que son los inmigrantes, a diferencia de los hablantes de más reciente arribo a Nueva York, los que presentan un mayor aumento en el uso del pronombre porque llevan más tiempo en EE.UU y por su competencia en inglés. Quiero saber qué más diferencia a estos dos grupos de exposición entre sí, en cuanto a las variables internas que nos ocupan en este capítulo.

En esta sección, analizo las mismas variables internas consideradas en la muestra entera, pero tomando en consideración a cada grupo de exposición. Investigo si estas variables y sus factores tienen el mismo impacto en la ocurrencia del pronombre en los recién llegados y los inmigrantes.

La regresión logística permite comparar las jerarquías de variables y factores a través de estos dos grupos con el fin de determinar si el condicionamiento de la variable dependiente es

similar en los dos grupos. Voy a hacer jerarquías de variables y factores para cada grupo de exposición.

En primer lugar, muestro los resultados obtenidos para los hablantes recién llegados en la tabla 2 sobre la importancia y el orden de las variables internas en la ocurrencia del pronombre personal sujeto.

Tabla 2. Jerarquía de variables que predicen la Ocurrencia, R.LL.			
Orden	Variable	Wald	<i>p</i>
1	Referencia	524.77	**
2	Persona	481.30	**
3	Léxico	73.25	**
4	Tiempo	72.13	**
5	Reflexividad	29.65	**
R2 = .143		N verbos = 17,857	
** = $p < .01$			
* = $p < .05$			
a = $p < .10$			

El número total de verbos en los recién llegados es de 17,857. El coeficiente de R^2 (R cuadrada), la cantidad de variación entre la presencia y la ausencia del pronombre personal sujeto, es de un catorce por ciento de variación. La jerarquía de variables de la tabla 2 señala que todas las variables son muy significativas estadísticamente ($p < .01$). El orden de las variables en los recién llegados es diferente si lo comparamos con el resultado de la muestra entera. En este modelo, la variable que más peso tiene sobre el pronombre es la Referencia, seguida por la Persona, el Léxico, el Tiempo y la Reflexividad. De todas ellas, las variables Referencia y Persona se distancian del resto de variables por sus altos valores de Wald.

Los datos de la regresión logística también informan sobre la importancia de los factores. Ahora, presento las regresiones logísticas para las jerarquías de factores en los recién llegados.

Tabla 2a. Jerarquía de factores para la variable interna Referencia, R.LL.			
Orden	Factor	Exp (B)	<i>p</i>
1	Cambio referencia	1.51	**
2	Misma referencia	.66	**
N verbos = 17,857			

La tabla 2a muestra el mismo resultado que la jerarquía de factores para la variable Referencia en la muestra entera. Sólo el cambio de referencia presenta un valor exponencial *B* mayor que 1.00, aunque los dos factores son muy significativos. El cambio de referencia en el discurso aumenta el uso del pronombre. Mientras que la continuidad de referencia hace que tienda a omitirse el pronombre.

Tabla 2b. Jerarquía de factores para la variable interna Persona, R.LL.			
Orden	Factor	Exp (B)	<i>p</i>
1	2ª persona singular	37.74	**
2	3ª persona singular	31.50	**
3	1ª persona singular	25.87	**
4	3ª persona plural	0.68	
5	1ª persona plural	0.39	
N verbos = 17,857			

La tabla 2b sobre la variable Persona también es igual que la jerarquía de factores para esta variable en la muestra entera. Las personas del verbo en singular incrementan el uso pronominal y las personas en plural lo desfavorecen. Los tres primeros factores tienen valores por encima de 1.00 y son muy significativos. La jerarquía de factores muestra que los recién llegados utilizan el pronombre con las formas singulares del verbo que con las plurales. El orden es la segunda persona del singular, la tercera persona del singular y la primera persona del singular.

Tabla 2c. Jerarquía de factores para la variable interna Léxico, R.LL.			
Orden	Factor	Exp (B)	<i>p</i>
1	Mental-estimativo	1.31	**
2	Estado	.90	*
3	Externo	.84	
N verbos = 17,857			

La jerarquía de factores de la tabla 2c señala que sólo los verbos de tipo mental-estimativo favorecen el uso del pronombre. Los otros dos factores presentan valores exponenciales por debajo de 1.00. Los verbos mentales-estimativos son muy significativos y los de estado también son significativos. De nuevo, los recién llegados comparten la misma jerarquía de factores para Léxico que todos los hablantes de la muestra.

Tabla 2d. Jerarquía de factores para la variable interna Tiempo, R.LL.			
Orden	Factor	Exp (B)	<i>p</i>
1	Pasado	1.77	**
2	Presente	1.76	**
3	Condicional	1.64	
4	Presente perfecto	0.95	
5	Futuro	0.50	
6	Modo subjuntivo	0.39	
N verbos = 17,857			

La tabla 2d indica que los recién llegados favorecen el pronombre con los tiempos en pasado, presente y condicional. En la jerarquía de factores se observa que los tres primeros factores presentan valores exponenciales que están por encima de 1.00. Sólo el pasado y el presente son muy significativos. Los tiempos verbales con los que este grupo de exposición muestra probabilidades de uso del pronombre son, por orden, el pasado, el presente y el condicional. El presente perfecto, el futuro y el modo subjuntivo desfavorecen el pronombre. Esta jerarquía de

factores para la variable Tiempo hasta el momento es la que presenta un orden distinto en la jerarquía de factores para la misma variable con respecto a los hablantes de la muestra entera.

Tabla 2e. Jerarquía de factores para la variable interna Reflexividad, R.LL.			
Orden	Factor	Exp (B)	<i>p</i>
1	No reflexivo	1.18	**
2	Reflexivo	.01	
N verbos = 17,857			

En la tabla 2e el factor que tiene un valor exponencial *B* mayor que 1.00 es el factor no reflexivo que además es muy significativo. Este dato, que también coincide con el de la muestra entera, señala que los recién llegados tienen una mayor probabilidad de usar el pronombre con los verbos que no son reflexivos.

En resumen, el uso del pronombre en los recién llegados se caracteriza por el impacto significativo de todas las variables internas en el siguiente orden: Referencia, Persona, Léxico, Tiempo y Reflexividad. En concreto, los recién llegados tienden a utilizar el pronombre cuando hay cambio de referencia más que cuando no lo hay; con cualquier persona verbal que esté en singular y de todas ellas, más con la segunda persona del singular; con los verbos que son de tipo mental-estimativo más que con otros tipos de verbos; con los tiempos verbales del pasado, el presente y el condicional más que con los otros tiempos; y con los verbos que no son reflexivos más que con los verbos reflexivos. Estos resultados muestran similitudes y diferencias con los de la muestra entera. Si nos fijamos en las jerarquías de variables, el orden de las variables en los recién llegados es diferente al de la muestra entera. Pero cuando observamos las jerarquías de factores, los recién llegados son iguales en el orden de los factores que más favorecen el

pronombre para las variables Referencia, Persona, Léxico y Reflexividad. Son distintos en la jerarquía de factores de la variable Tiempo.

Ahora, presento los resultados obtenidos de las regresiones logísticas para los inmigrantes. Las siguientes tablas muestran la jerarquía de variables independientes internas y las jerarquías de factores de cada variable.

Tabla 3. Jerarquía de variables que predicen la Ocurrencia, INM.			
Orden	Variable	Wald	<i>p</i>
1	Persona	1271.20	**
2	Referencia	949.79	**
3	Tiempo	116.00	**
4	Reflexividad	64.83	**
5	Léxico	55.01	**
R2 = .163		N verbos = 32,769	
** = $p < .01$			
* = $p < .05$			
a = $p < .10$			

La tabla 3 muestra que el número de verbos en los inmigrantes es de 32,769. El coeficiente de R^2 (R cuadrada) informa que las variables representan un dieciséis por ciento de variación en la ocurrencia del pronombre personal sujeto. La jerarquía de variables señala que todas las variables son estadísticamente muy significativas con valores de p inferiores a .01. El orden de las variables en los inmigrantes es diferente del resultado de los recién llegados y similar al resultado de la muestra entera. Esta regresión logística sitúa en primer lugar, como la variable que más peso tiene en la aparición del pronombre, a la Persona con un valor de significación mucho más alto que el de las demás. A continuación, se encuentra la Referencia que tiene también un valor de Wald alto, seguida por el Tiempo, la Reflexividad y el Léxico.

Esta jerarquía de variables internas nos informa de que estos dos grupos de exposición, los recién llegados y los inmigrantes, no son iguales internamente. Es decir, las variables lingüísticas y discursivas que guían a ambos grupos en la ocurrencia del pronombre son las

mismas, pero el impacto de cada una de estas variables en los recién llegados y los inmigrantes es diferente.

Ahora, quiero ver si los inmigrantes también son diferentes de los recién llegados en sus jerarquías de factores. Las tablas que van de la 3a a la 3e presentan las regresiones logísticas de cada variable.

Tabla 3a. Jerarquía de factores para la variable interna Persona, INM.			
Orden	Factor	Exp (B)	<i>p</i>
1	2ª persona singular	44.23	**
2	1ª persona singular	39.86	**
3	3ª persona singular	30.80	**
4	3ª persona plural	0.67	
5	1ª persona plural	0.34	
N verbos = 32,769			

La jerarquía de factores para Persona en los inmigrantes dice que las tres primeras personas del singular favorecen el pronombre. Al igual que en el caso de los recién llegados, hay una preferencia por las personas en singular que son muy significativas. Superan en valores exponenciales *B* a las personas del plural que no alcanzan 1.00. El orden de la jerarquía de factores es diferente del orden para los recién llegados. Ambos grupos coinciden en la posición del primer factor, excepto en los factores que ocupan la segunda y tercera posiciones. Para los inmigrantes, la primera persona del singular se usa más con el pronombre. Para los recién llegados, los verbos en tercera persona del singular requieren más pronombres explícitos que los verbos en primera persona del singular.

Tabla 3b. Jerarquía de factores para la variable interna Referencia, INM.			
Orden	Factor	Exp (B)	<i>p</i>
1	Cambio referencia	1.50	**
2	Misma referencia	.66	
N verbos = 32,769			

La tabla 3b expone el mismo resultado que hemos observado hasta ahora. Los verbos con cambio de referencia siguen siendo los que tienen la mayor probabilidad de utilización del pronombre. Los verbos cuya referencia es la misma tienden a desfavorecer el uso pronominal. Los inmigrantes y los recién llegados no se diferencian por la jerarquía de factores de la variable Referencia.

Tabla 3c. Jerarquía de factores para la variable interna Tiempo, INM.			
Orden	Factor	Exp (B)	<i>p</i>
1	Condicional	2.50	**
2	Presente	1.50	**
3	Pasado	1.45	
4	Presente perfecto	0.44	
5	Modo subjuntivo	0.36	
6	Futuro	0.23	
N verbos = 32,769			

El modelo de la tabla 3c presenta un orden de factores para la variable Tiempo en los inmigrantes distinto al de los recién llegados. De los siete factores, tres siguen favoreciendo el uso del pronombre en ambos grupos de exposición, aunque el orden es diferente. Para los inmigrantes, el condicional es el tiempo con el que hay más probabilidad de utilizar el pronombre. Seguidamente, encontramos el presente y el pasado con valores exponenciales *B* próximos. El condicional y el presente son muy significativos con valores de *p* inferiores a .01. Los demás tiempos verbales y el modo subjuntivo desfavorecen el pronombre. Esta jerarquía de factores es igual en el orden que la que presenta la muestra entera, a excepción de que los hablantes de primera generación también favorecen el pronombre con el presente perfecto.

Tabla 3d. Jerarquía de factores para la variable interna Reflexividad, INM.			
Orden	Factor	Exp (B)	<i>p</i>
1	No reflexivo	1.17	**
2	Reflexivo	.85	
N verbos = 32,769			

Esta regresión logística no añade ninguna información novedosa. El valor del factor “no reflexivo” por encima de 1.00 indica que los inmigrantes también favorecen el pronombre con verbos que no están acompañados de un pronombre reflexivo. La tabla 3d muestra unos datos similares en relación a la jerarquía de factores para Reflexividad en los recién llegados.

Tabla 3e. Jerarquía de factores para la variable interna Léxico, INM.

Orden	Factor	Exp (B)	<i>p</i>
1	Mental-estimativo	1.13	**
2	Estado	1.03	**
3	Externo	.88	
N verbos = 32,769			

La última de las jerarquías de factores para los inmigrantes da detalles sobre la variable interna Léxico. La tabla informa de que los verbos de tipo mental-estimativo y de estado suelen favorecer el pronombre personal sujeto. El orden de los factores es igual que en los recién llegados. Sin embargo en los inmigrantes, el segundo factor “verbos de estado” tiene un valor exponencial *B* que está ligeramente por arriba de 1.00. También, los dos primeros factores son muy significativos. Esto indica que este grupo de exposición usa el pronombre con dos tipos de verbo de contenido semántico diferente.

En conclusión, estas cinco variables internas influyen en el uso del pronombre en los inmigrantes. El orden es: Persona, Referencia, Tiempo, Reflexividad y Léxico. Este grupo de exposición favorece el pronombre con las personas verbales en singular, habiendo una preferencia por la segunda persona en singular; cuando el referente es distinto al del verbo anterior en el discurso; con tres de los tiempos verbales del indicativo, siendo el condicional el tiempo verbal que requiere más pronombres; con los verbos que no llevan un pronombre reflexivo; y con los verbos que semánticamente son mentales-estimativos y de estado.

A través de las jerarquías de variables internas, observamos que los recién llegados y los inmigrantes presentan un comportamiento gramaticalmente diferente. Hemos visto que las cinco variables son muy significativas en los dos grupos. Pero el orden de las variables es diferente. Mientras que Referencia es la variable que guía el uso del pronombre en los recién llegados, Persona lo es en los inmigrantes. En cuanto a las jerarquías de variables, las posiciones de Tiempo, Léxico y Reflexividad son diferentes en ambos grupos. Si, además, prestamos atención a la información de las jerarquías de factores de los dos grupos de exposición, veremos que únicamente las jerarquías de factores para las variables Referencia y Reflexividad son iguales en los recién llegados y los inmigrantes. El resto de jerarquías (Persona, Referencia, Tiempo), nuevamente diferencia a los hablantes de los dos grupos de exposición.

4.3 Jerarquías de variables y factores en los grupos regionales de exposición

Ahora investigo qué variables internas son más significativas en cuanto a la ocurrencia del pronombre en los grupos de exposición recién llegados e inmigrantes divididos según el área regional. Voy a utilizar las cinco variables estudiadas en las regresiones logísticas previas. La investigación dialectal ha señalado que los hablantes del Caribe tienden a usar más el pronombre personal sujeto (Cameron, 1995: 8; Lipski, 1994: 241; López Morales, 1992: 137; Otheguy y Zentella, 2012: 146). El capítulo sobre la tasa pronominal también corroboró este dato. Ahora hay que analizar si los recién llegados caribeños y continentales y los inmigrantes caribeños y continentales son iguales o diferentes en las jerarquías de variables y de factores.

Las regresiones logísticas siguientes son para los recién llegados caribeños y los recién llegados continentales. Las tablas 4.1a y 4.2a presentan las jerarquías de variables internas para cada una de las regiones dentro del grupo de exposición recién llegados. Las tablas proporcionan una comparación entre los recién llegados caribeños y los continentales.

Tabla 4.1a. Jerarquía de variables que predicen la Ocurrencia, R.LL. Caribeños

Orden	Variable	Wald	<i>p</i>
1	Referencia	366.08	**
2	Persona	293.53	**
3	Léxico	70.01	**
4	Tiempo	66.00	**
5	Reflexividad	9.96	**
R2 = .187		N verbos = 8,889	
** = <i>p</i> < .01			
* = <i>p</i> < .05			
a = <i>p</i> < .10			

Tabla 4.2a. Jerarquía de variables que predicen la Ocurrencia, R.LL. Continentales

Orden	Variable	Wald	<i>p</i>
1	Persona	160.96	**
2	Referencia	159.03	**
3	Léxico	25.89	**
4	Reflexividad	20.60	**
5	Tiempo	13.20	**
R2 = .144		N verbos = 8,958	
** = <i>p</i> < .01			
* = <i>p</i> < .05			
a = <i>p</i> < .10			

El número de verbos en los recién llegados caribeños (8,889) es un poco inferior al número de verbos en los recién llegados continentales (8,958). Los resultados señalan que las variables principales tienen un orden diferente en ambas regiones. Todas son muy significativas. Los coeficientes de *R*² (*R* cuadrada) indican que las variables representan un dieciocho por ciento de variación en la ocurrencia del pronombre personal sujeto en los caribenos y un catorce por ciento en los continentales. Para los recién llegados caribeños, el orden es el siguiente: Referencia, Persona, Léxico, Tiempo y Reflexividad. Para los recién llegados continentales es: Persona, Referencia, Léxico, Reflexividad y Tiempo. En ambos grupos regionales, Referencia y Persona son las variables que ocupan las primeras posiciones mientras que Tiempo y Reflexividad aparecen en los últimos puestos. La única variable que aparece en la misma posición en los dos grupos es Léxico, pero con distinto valor de Wald. Las jerarquías de variables informan que los recién llegados caribeños y los recién llegados continentales se distinguen por la región de procedencia.

La segunda comparación en los recién llegados caribeños y continentales se basa en el estudio del orden y de la fuerza de los factores dentro de cada variable. Las tablas de la izquierda muestran los datos de las jerarquías de factores para los recién llegados caribeños y las tablas de

la derecha describen las jerarquías de factores para los recién llegados continentales. Comparamos a los caribeños y los continentales en el estudio de cada factor.

Tabla 4.1b. Jerarquía de factores para la variable interna Referencia, R.LL. Caribeños			
Orden	Factor	Exp (B)	<i>p</i>
1	Cambio referencia	1.55	**
2	Misma referencia	.65	
N verbos = 8,889			

Tabla 4.2b. Jerarquía de factores para la variable interna Referencia, R.LL. Continentales			
Orden	Factor	Exp (B)	<i>p</i>
1	Cambio referencia	1.48	**
2	Misma referencia	.67	
N verbos = 8,958			

La primera jerarquía de factores es para la variable Referencia. Ésta es una de las variables en la que los recién llegados de ambas regiones comparten el mismo orden de factores. El cambio de referencia, con unos valores exponenciales *B* muy similares y una alta significación, es lo que favorece el uso pronominal. El factor “misma referencia” no se acerca a 1.00 en ninguna de las jerarquías. Esto significa que cuando la referencia es la misma que la del verbo anterior no favorece el uso del pronombre. Los recién llegados caribeños y continentales no se diferencian por la jerarquía de factores de Referencia.

Tabla 4.1c. Jerarquía de factores para la variable interna Persona, R.LL. Caribeños			
Orden	Factor	Exp (B)	<i>p</i>
1	2ª persona singular	1.85	**
2	1ª persona singular	1.41	**
3	3ª persona singular	.86	
4	1ª persona plural	.11	*
5	3ª persona plural	.08	*
N verbos = 8,889			

Tabla 4.2c. Jerarquía de factores para la variable interna Persona, R.LL. Continentales			
Orden	Factor	Exp (B)	<i>p</i>
1	3ª persona singular	1.85	**
2	1ª persona singular	1.12	**
3	2ª persona singular	.51	
4	3ª persona plural	.33	
5	1ª persona plural	.23	
N verbos = 8,958			

Las tablas 4.1c y 4.2c presentan la jerarquía de factores de la variable Persona. Observamos que los dos primeros factores en cada grupo (segunda persona del singular y primera persona del singular en los recién llegados caribeños y tercera persona del singular y primera persona del singular en los recién llegados continentales) presentan la diferencia. En los dos grupos

regionales, el orden de los factores es distinto y los dos primeros factores son muy significativos. Los caribeños favorecen el pronombre más con la segunda persona del singular mientras que los continentales lo hacen con la tercera. La primera persona del singular es el segundo de los factores que favorece la ocurrencia del pronombre y ocupa el segundo lugar en los dos grupos. Los demás factores (entre los que se incluyen las personas en plural) tienen un valor exponencial B por debajo de 1.00, así que desfavorecen el pronombre. Estos resultados, que reflejan el favorecimiento del pronombre en las personas en singular, diferencian a los recién llegados caribeños y continentales en el orden diferente de los factores.

Tabla 4.1d. Jerarquía de factores para la variable interna Léxico, R.LL. Caribeños			
Orden	Factor	Exp (B)	p
1	Mental-estimativo	1.26	**
2	Estado	.98	
3	Externo	.80	
N verbos = 8,889			

Tabla 4.2d. Jerarquía de factores para la variable interna Léxico, R.LL. Continentales			
Orden	Factor	Exp (B)	p
1	Mental-estimativo	1.40	**
2	Externo	.85	
3	Estado	.84	
N verbos = 8,958			

Las regresiones logísticas de las tablas 4.1d y 4.2d también coinciden en posicionar en primer lugar el factor “mental-estimativo” en la variable Léxico. Es el único de los factores que es muy significativo en las dos jerarquías. Sólo si el verbo es semánticamente un verbo mental o estimativo el pronombre es favorecido en este grupo de exposición en las dos regiones. Los demás factores aparecen en distinto orden para estos grupos de hablantes pero ninguno de ellos llega a ser significativo. Esta jerarquía de factores sobre el Léxico dice que los recién llegados caribeños y continentales son iguales.

Tabla 4.1e. Jerarquía de factores para la variable interna Tiempo, R.LL. Caribeños

Orden	Factor	Exp (B)	<i>p</i>
1	Presente perfecto	1.97	**
2	Presente	1.91	**
3	Pasado	1.74	**
4	Condicional	0.66	
5	Modo subjuntivo	0.39	
6	Futuro	0.34	
N verbos = 8,889			

Tabla 4.2e. Jerarquía de factores para la variable interna Tiempo, R.LL. Continentales

Orden	Factor	Exp (B)	<i>p</i>
1	Pasado	1.57	**
2	Condicional	1.48	**
3	Presente	1.40	*
4	Presente Perfecto	0.35	
5	Modo subjuntivo	0.23	
6	Futuro	0.21	
N verbos = 8,958			

El resultado de las jerarquías de factores para la variable Tiempo diferencia a estos dos grupos en el orden de los tres primeros factores que son muy significativos. Cada jerarquía de factores sitúa en distinto orden los tiempos verbales. Los recién llegados caribeños favorecen más el pronombre con el presente perfecto, el presente y el pasado. Los recién llegados continentales favorecen el pronombre con los tiempos en pasado, condicional y presente. Los demás tiempos y modo verbales desfavorecen el pronombre en las dos regiones. Los dos grupos regionales en los recién llegados se diferencian en el orden de los factores de la variable Tiempo.

Tabla 4.1f. Jerarquía de factores para la variable interna Reflexividad, R.LL. Caribeños

Orden	Factor	Exp (B)	<i>p</i>
1	No reflexivo	1.09	**
2	Reflexivo	.91	
N verbos = 8,889			

Tabla 4.2f. Jerarquía de factores para la variable interna Reflexividad, R.LL. Continentales

Orden	Factor	Exp (B)	<i>p</i>
1	No reflexivo	1.29	**
2	Reflexivo	.77	
N verbos = 8,958			

Las tablas 4.1f y 4.2f muestran las jerarquías de factores de la variable Reflexividad. De los dos factores, la “no reflexividad” favorece el uso del pronombre. Por el contrario, si el verbo es reflexivo lo desfavorece. Estos resultados indican que estos hablantes no se distinguen por la jerarquía de esta variable.

En definitiva, los recién llegados caribeños y continentales son diferentes a nivel interno en el orden de las variables y, concretamente, en el orden de los factores de las variables Persona

y Tiempo. Las distintas regresiones logísticas señalan que ambos grupos regionales en los recién llegados utilizan el pronombre influenciados por las mismas variables internas. La diferencia se encuentra en el orden y el peso de las variables. En el grupo de exposición regional recién llegados caribeños, la Persona es lo primero que influye la ocurrencia del pronombre, mientras que en el grupo regional de exposición recién llegados continentales es la Referencia. Además de eso, todas las variables presentan unos valores de Wald distintos que indican que los pesos de cada variable son diferentes. En cuanto a las jerarquías de factores, el orden de los factores diferente en las variables Persona y Tiempo diferencia a los recién llegados caribeños de los continentales. Las jerarquías de factores para las variables Referencia, Léxico y Reflexividad son iguales en el orden en los dos grupos, aunque ligeramente diferentes en los valores exponenciales *B*.

Las tablas que presento abajo son las regresiones logísticas realizadas para los grupos regionales de exposición inmigrantes caribeños e inmigrantes continentales. La tabla 5.1a de la izquierda indica los resultados de la jerarquía de variables para los inmigrantes caribeños y la tabla 5.2a de la derecha muestra los datos de la jerarquía de variables para los inmigrantes continentales.

Tabla 5.1a. Jerarquía de variables que predicen la Ocurrencia, INM. Caribeños			
Orden	Variable	Wald	<i>p</i>
1	Persona	854.52	**
2	Referencia	687.49	**
3	Tiempo	96.60	**
4	Léxico	71.86	**
5	Reflexividad	34.07	**
R2 = .183		N verbos = 17,020	
** = <i>p</i> < .01			
* = <i>p</i> < .05			
a = <i>p</i> < .10			

Tabla 5.2a. Jerarquía de variables que predicen la Ocurrencia, INM. Continentales			
Orden	Variable	Wald	<i>p</i>
1	Persona	403.72	**
2	Referencia	291.22	**
3	Reflexividad	26.68	**
4	Tiempo	24.84	**
5	Léxico	2.46	**
R2 = .138		N verbos = 15,743	
** = <i>p</i> < .01			
* = <i>p</i> < .05			
a = <i>p</i> < .10			

El número de verbos en los inmigrantes caribeños (17,020) es superior al número de verbos en los inmigrantes continentales (15,743). Los coeficientes de R^2 (R cuadrada) indican que las variables representan un dieciocho por ciento de variación en la ocurrencia del pronombre personal sujeto en los inmigrantes caribeños y casi un catorce por ciento en los inmigrantes continentales. Los resultados de estos modelos indican varias cosas. Primero, todas las variables son muy significativas en la ocurrencia del pronombre en los dos grupos. Segundo, los inmigrantes caribeños y continentales son iguales en el orden de las dos primeras variables. En los dos grupos regionales de exposición el ranking de variables está liderado por Persona seguida de Referencia. Esto quiere decir que la ocurrencia del pronombre personal sujeto es favorecida por la persona del verbo, primero, y por el cambio de referencia en el discurso, segundo. El orden es el mismo en las dos variables, pero los valores de Wald son distintos en cada grupo. Los inmigrantes caribeños presentan unos valores más altos que los continentales. Tercero, el orden en el resto de las variables varía de un grupo a otro. En los inmigrantes caribeños, el tercer, cuarto y quinto puestos, los ocupan Tiempo, Léxico y Reflexividad. En los inmigrantes continentales son Reflexividad, Tiempo y Léxico. Las jerarquías de variables de las tablas 5.1a y 5.2a señalan que estas tres variables son tan significativas como las dos primeras en la probabilidad de ocurrencia del pronombre. La diferencia en el orden y los valores de Wald indican que estas tres variables diferencian a los dos grupos. Según las jerarquías de variables, los inmigrantes caribeños y continentales son distintos en el orden de los tiempos verbales, el contenido semántico del verbo y la reflexividad.

Ahora, muestro los resultados de las regresiones logísticas realizadas para cada variable en los inmigrantes caribeños y continentales. De nuevo, las tablas de la izquierda 5.1b a 5.1f presentan los datos de las jerarquías de factores de cada variable interna para los inmigrantes

caribeños y las tablas de la derecha 5.2b a 5.2f muestran las jerarquías de factores en los inmigrantes continentales.

Tabla 5.1b. Jerarquía de factores para la variable interna Persona, INM. Caribeños			
Orden	Factor	Exp (B)	<i>p</i>
1	2ª persona singular	20.82	**
2	1ª persona singular	13.41	**
3	3ª persona singular	8.43	
4	1ª persona plural	0.10	
5	3ª persona plural	0.04	
N verbos = 17,020			

Tabla 5.2b. Jerarquía de factores para la variable interna Persona, INM. Continentales			
Orden	Factor	Exp (B)	<i>p</i>
1	3ª persona singular	2.94	**
2	1ª persona singular	2.48	**
3	2ª persona singular	.86	
4	3ª persona plural	.69	
5	1ª persona plural	.51	
N verbos = 15,743			

Las tablas 5.1b y 5.2b presentan las jerarquías de factores para Persona en los inmigrantes caribeños y continentales. Una de las diferencias en estos datos es que el orden de los factores es diferente en cada jerarquía. Los inmigrantes caribeños favorecen el pronombre más con la segunda persona del singular. Los inmigrantes continentales lo favorecen con la tercera persona del singular. El único factor que coincide en el segundo orden en las dos jerarquías es la primera persona del singular que también favorece el pronombre. Estos dos primeros factores son muy significativos en las dos jerarquías. Los inmigrantes caribeños favorecen el pronombre con las tres personas del singular, porque también la tercera persona de singular favorece el pronombre. Los inmigrantes continentales lo favorecen sólo con dos personas del singular. Los demás factores (las personas del plural) se presentan en distinto orden y desfavorecen el pronombre en las dos regiones. Otra diferencia es que los inmigrantes caribeños presentan los factores con valores exponenciales *B* por encima de 1.00 y todos superiores a los que presentan los continentales. Los inmigrantes caribeños favorecen el pronombre con la segunda, primera y tercera personas del singular. Los continentales lo favorecen con la tercera y primera persona del singular. Este resultado en ambas jerarquías diferencia a los inmigrantes caribeños y continentales.

Tabla 5.1c. Jerarquía de factores para la variable interna Referencia, INM. Caribeños

Orden	Factor	Exp (B)	<i>p</i>
1	Cambio referencia	1.55	**
2	Misma referencia	.64	
N verbos = 17,020			

Tabla 5.2c. Jerarquía de factores para la variable interna Referencia, INM. Continentales

Orden	Factor	Exp (B)	<i>p</i>
1	Cambio referencia	1.48	**
2	Misma referencia	.67	
N verbos = 15,743			

Las regresiones logísticas para Referencia son iguales en los dos grupos. Las tablas 5.1c y 5.2c señalan que el cambio de referencia es el factor que afecta más el uso pronominal. Ocupa el primer lugar en las dos jerarquías de factores. Tiene un valor exponencial *B* por encima de 1.00 y es muy significativo con un valor de *p* inferior a .01. Este dato, que coincide con los datos anteriores, indica que cuando el referente de la oración cambia con respecto al del verbo que lo precede se favorece el pronombre. Cuando la referencia es la misma, se omite el pronombre.

Tabla 5.1d. Jerarquía de factores para la variable interna Tiempo, INM. Caribeños

Orden	Factor	Exp (B)	<i>p</i>
1	Condicional	2.59	**
2	Presente	1.65	**
3	Presente Perfecto	1.49	
4	Pasado	0.99	
5	Modo subjuntivo	0.26	
6	Futuro	0.16	
N verbos = 17,020			

Tabla 5.2d. Jerarquía de factores para la variable interna Tiempo, INM. Continentales

Orden	Factor	Exp (B)	<i>p</i>
1	Condicional	2.35	**
2	Pasado	1.58	
3	Presente Perfecto	1.52	
4	Presente	0.41	
5	Modo subjuntivo	0.40	
6	Futuro	0.19	
N verbos = 15,743			

Los modelos de las tablas 5.1d y 5.2d nos dicen que los inmigrantes caribeños y continentales favorecen el pronombre de igual manera en los tiempos del condicional, primer lugar, y presente perfecto, tercer lugar. El condicional es el tiempo verbal que más favorece el pronombre en los dos grupos. Los tiempos del pasado y presente que también favorecen el uso pronominal señalan que estos hablantes son diferentes en el orden. Es decir, los inmigrantes caribeños realizan más pronombres con el presente y los inmigrantes continentales realizan más el pronombre con el

pasado. Los demás tiempos verbales y el modo subjuntivo desfavorecen el pronombre en los dos grupos de hablantes. Esta jerarquía de factores diferencia a estos grupos regionales.

Tabla 5.1e. Jerarquía de factores para la variable interna Reflexividad, INM. Caribeños			
Orden	Factor	Exp (B)	<i>p</i>
1	No reflexivo	1.17	**
2	Reflexivo	.85	
N verbos = 17,020			

Tabla 5.2e. Jerarquía de factores para la variable interna Reflexividad, INM. Continentales			
Orden	Factor	Exp (B)	<i>p</i>
1	No reflexivo	1.17	**
2	Reflexivo	.86	
N verbos = 15,743			

Las tablas 5.1e y 5.2e presentan las mismas jerarquías de factores de la variable Reflexividad. Los resultados afirman que los verbos que no llevan un pronombre reflexivo tienden a estar acompañados de un pronombre personal sujeto. Estos dos grupos se comportan igualmente en el uso del pronombre cuando el verbo es reflexivo.

Tabla 5.1f. Jerarquía de factores para la variable interna Léxico, INM. Caribeños			
Orden	Factor	Exp (B)	<i>p</i>
1	Mental-estimativo	1.18	**
2	Estado	.97	
3	Externo	.87	
N verbos = 17,020			

Tabla 5.2f. Jerarquía de factores para la variable interna Léxico, INM. Continentales			
Orden	Factor	Exp (B)	<i>p</i>
1	Mental-estimativo	1.09	**
2	Estado	1.03	*
3	Externo	.89	
N verbos = 15,743			

Las jerarquías de factores sobre el Léxico diferencian a los dos grupos. La tabla 5.1f dice que los inmigrantes caribeños favorecen los pronombres únicamente con verbos de tipo mental-estimativo. La tabla 5.2f señala que los inmigrantes continentales favorecen los pronombres con verbos que son de tipo mental-estimativo y de estado. Los verbos de tipo externo en los dos grupos desfavorecen el pronombre. Estos resultados confirman que los inmigrantes caribeños y continentales son iguales por el factor “mental-estimativo”, pero distintos por el factor “estado” en el uso pronominal.

En resumen, los inmigrantes caribeños y continentales presentan similitudes y diferencias en las jerarquías de variables y en las jerarquías de factores. Todas las variables son muy significativas como se aprecia en las tablas 5.1a y 5.2a. Según el orden, estos grupos regionales de exposición muestran un comportamiento parecido en las variables Persona y Referencia que son las que influyen más en la probabilidad de ocurrencia del pronombre. La probabilidad del pronombre en los dos grupos de hablantes está guiada primero por la persona verbal y segundo, por el estatus del referente como consabido o nuevo. Sin embargo, las diferentes posiciones de Tiempo, Reflexividad y Léxico en ambas jerarquías de variables demuestran que la probabilidad de que ocurra un pronombre es distinta en los inmigrantes caribeños y continentales. El orden diferente en las jerarquías de variables indica que los inmigrantes del Caribe y del Continente son distintos en la probabilidad de utilizar u omitir un pronombre personal sujeto.

Las jerarquías de factores también revelan que hay diferencias y similitudes en estos hablantes. El orden de los factores en Persona y Tiempo es diferente en cada jerarquía para la mayoría de los factores. Además, los valores de significación también son muy distintos. Así que los factores de Persona y Tiempo afectan de forma desigual a los inmigrantes caribeños y continentales en el favorecimiento del pronombre. Por el contrario, las jerarquías de factores para las variables Referencia, Reflexividad y Léxico, que contemplan dos y tres factores, indican que los dos grupos son internamente iguales a la hora de favorecer el pronombre.

5. Interpretación de los resultados de las jerarquías de variables y factores

Los resultados de las regresiones logísticas realizadas en este capítulo desde la tabla 1 hasta la tabla 5.2f nos llevan a establecer algunas generalizaciones sobre la gramática de los hablantes de primera generación en Nueva York, concretamente la gramática de los recién llegados y los inmigrantes.

5.1 Continuidad en el español en Nueva York

Uno de los hallazgos tiene que ver con que el nivel de continuidad de las normas de uso del español de Latinoamérica ocurre en el español de los recién llegados en Nueva York. Recuerdo que los recién llegados conforman el grupo de exposición más representativo dentro de la primera generación de los modos de hablar latinoamericanos porque es el grupo que lleva menos tiempo en la ciudad y el menos expuesto al inglés. Las jerarquías de variables de los recién llegados, traídas desde Latinoamérica, son iguales en los recién llegados, como grupo de exposición, y los recién llegados caribeños, como grupo regional de exposición. El orden de las variables que más favorece el pronombre es el mismo para los recién llegados en total y los recién llegados que vienen del Caribe. Las variables que más influencia tienen en el uso pronominal son la Referencia, la Persona, el Léxico, el Tiempo y la Reflexividad.

La excepción dentro de los recién llegados son los hablantes que vienen del Continente. Los recién llegados continentales, al contrario, presentan un orden diferente al de los recién llegados caribeños en la jerarquía de variables. Para estos hablantes, las variables que influyen la ocurrencia del pronombre son Persona, Referencia, Léxico, Reflexividad y Tiempo. Este dato señala que las dos regiones son diferentes en Latinoamérica. Dentro de los recién llegados, los que vienen del Caribe y los que vienen del Continente muestran unas jerarquías de variables diferentes en el orden. Esto indica que las variables que tienen mayor probabilidad en el uso del pronombre en los latinoamericanos del Caribe y del Continente son distintas. La posición de las variables Persona y Referencia en los recién llegados continentales coincide en el mismo lugar en los rankings de la primera generación (en total, incluyendo a los dos grupos de exposición) y los inmigrantes.

Para entender el orden de las variables en las jerarquías, hay que comprender que la función de los pronombres personales sujetos es la de apoyar las inflexiones verbales en la tarea de identificar a los referentes. Existen teorías del lenguaje que explícitamente asignan contenido semántico a palabras con función gramatical. Un ejemplo es el caso de la deixis codificada en los pronombres con esta función de identificación (García, 1975). Otras teorías formales que diferencian entre módulos gramaticales internos, afirman que la comprensión del uso de los pronombres se basa en la relación entre la semántica y la sintaxis (Sorace, 2005). De lo que no hay duda es que los pronombres personales sujetos ayudan a identificar a los referentes en el discurso.

Teniendo en cuenta esta característica esencial del pronombre personal sujeto, es comprensible que Referencia y Persona aparezcan en las primeras posiciones de las jerarquías de variables de los recién llegados en su totalidad en la tabla 2 y de los recién llegados caribeños en la tabla 4.1a. La Referencia es el predictor más fuerte en la presencia o ausencia del pronombre, porque si el referente es el mismo, o cambia del referente del verbo anterior, los hablantes activan o no la función de identificación del pronombre. La explicación para la variable Persona se encuentra también en la identificación de los referentes asociados a los sujetos del verbo. Las inflexiones persona y número de los verbos españoles dan información básica respecto a si el participante directamente más asociado al evento descrito en el verbo es el hablante, el interlocutor, etc (por ejemplo, usando la primera persona del singular, la segunda o la tercera). Queda perfectamente clara la razón por la que Referencia y Persona aparecen en las primeras posiciones de las jerarquías de variables. Por su función de identificar el rol del pronombre. Las posiciones en las jerarquías de las otras variables son más difíciles de explicar. Aunque en general, las variables más conectadas directamente con la función expresiva, como Persona,

Referencia y Tiempo, parecen influir el uso del pronombre más que otras variables donde esta conexión es débil y donde la definición de la variable responde más a relaciones sintácticas y semánticas como Reflexividad y Léxico.

Las jerarquías de variables de las tablas 2 y 4.1a nos dan información de lo que guía a los recién llegados en su totalidad y a los recién llegados caribeños en la probabilidad de usar u omitir el pronombre. Esto incluye un interés básico en la identificación y especificación del referente (de ahí que Referencia y Persona ocupen las primeras posiciones), una preocupación por las propiedades semánticas y sintácticas del verbo (Léxico y Reflexividad están en tercer y cuarto lugar), y un interés adicional en las propiedades asociadas con el lugar de los verbos a través del tiempo (en quinta posición Tiempo).

Usando la jerarquía de variables de los recién llegados en su totalidad en la tabla 2 como grupo de referencia con las normas de uso latinoamericanas, el hallazgo central es que únicamente los recién llegados caribeños comparten la misma jerarquía. Pero no sabemos si esto es simplemente porque hay más caribeños en la tabla 2, por lo tanto sus usos particulares tienen mayor impacto en el uso del grupo en su totalidad. Este hallazgo puede estar relacionado con la existencia de distintas tasas de ocurrencia del pronombre (estudiadas en capítulos previos para comparar a los grupos de exposición). La diferencia en las jerarquías de variables entre los recién llegados y los inmigrantes podía ser predecible. Los dos grupos de exposición difieren no sólo en la tasa pronominal y en el impacto de las variables externas (como se observó en los capítulos anteriores), sino también en el orden de las variables internas.

Cuando los grupos de exposición se dividen según el criterio regional, hemos visto que hay también diferencias. Las jerarquías de variables son distintas en los recién llegados continentales y en los recién llegados caribeños en relación a las jerarquías de variables de las

tablas 4.1a y 4.2a. El diferente orden de las jerarquías de variables en los grupos regionales de exposición nos lleva a establecer la existencia previa de diferencias gramaticales en la decisión de omitir o usar el pronombre personal sujeto en estos hablantes de la primera generación.

Otro de los resultados apunta que la continuidad también se refleja en las jerarquías de factores en los dos grupos de exposición. El orden de los factores en tres de las jerarquías, Referencia, Léxico y Reflexividad, es el mismo en los recién llegados y en los inmigrantes, ya sean caribeños o continentales. Los rankings de factores en las jerarquías de los recién llegados en las tablas de 2a a 2e (sin contar 2b y 2d) y de 4.1b a 4.2f (menos 4.1c, 4.2c, 4.1e, y 4.2e) y en los inmigrantes en las tablas de 3b a 3e (quitando 3c) y de 5.1c a 5.2f (a excepción de 5.1d y 5.2d) son resultados esperados y fáciles de explicar.

Todas las jerarquías de factores para la variable Referencia señalan que los contextos donde hay cambio de referencia favorecen el uso del pronombre. La explicación tiene que ver con la función de alertar al interlocutor e identificar el referente del verbo al que se refiere el nuevo pronombre personal (Silva-Corvalán, 1982: 196). Cuando se produce en el discurso un cambio de referencia, se presta más atención a la información nueva que a la que se repite y es la misma.

Respecto a la variable Léxico, las jerarquías de factores sitúan siempre en la primera posición a los verbos de contenido mental-estimativo. Los enunciados que expresan verbos que denotan estados mentales, actitudes, emociones, deseos, valores, juicios e, incluso, conocimiento, aparecen con pronombres explícitos (Enríquez, 1984; Morales, 1997). Este dato tiene que ver con la idea discursiva de dejar claro quien es el que realiza la acción del verbo. De nuevo, se trata de identificar al referente de la oración.

No obstante, se observa el cambio en las jerarquías de factores de los inmigrantes y, concretamente, de los inmigrantes continentales, ya que el segundo factor “verbos de estado” en estos hablantes también favorece la ocurrencia del pronombre.

Finalmente, el orden de los factores en la variable Referencia apoya la idea de la continuidad con el español de Latinoamérica. Los resultados confirman que los verbos en contextos reflexivos tienen la tendencia a desfavorecer la utilización del pronombre. Son los verbos que no llevan el pronombre reflexivo los que favorecen la presencia del pronombre personal sujeto. La explicación tiene que ver con el deseo del hablante de hacer identificable el sujeto del verbo.

5.2 Cambios en el español en Nueva York

Los elementos que influyen a los hablantes de primera generación en la ocurrencia del pronombre son permeables a las fuerzas del contacto. Tanto es así que pueden causar cambios en las tasas del uso de los pronombres como vimos en capítulos anteriores. El estudio de las jerarquías de variables, que representan el funcionamiento de la gramática de los hablantes recién llegados e inmigrantes, nos permite observar qué tipo de alteraciones gramaticales existen en los grupos de exposición recién llegados e inmigrantes. Para ver si hay cambio, no basta con estudiar las tasas de ocurrencia pronominal, sino también el orden de las variables y sus factores (Poplack y Levey, 2010).

Hay partes de la gramática de hablantes de la primera generación que han empezado a cambiar. Las jerarquías reflejan o constituyen el cambio. Así que el punto de referencia para detectar el cambio se encuentra en las diferencias entre los recién llegados y los inmigrantes establecidos en el orden de las variables y de los factores de las jerarquías de variables y de factores.

En la primera generación, los cambios comienzan a observarse en los inmigrantes establecidos cuando los comparamos con los recién llegados. En el capítulo 3 sobre la tasa pronominal, quedó demostrado que el español de los inmigrantes en Nueva York es distinto del español de los recién llegados por la tasa pronominal. En el capítulo 4 sobre el impacto de las variables externas, se reafirmó este dato al observar que las variables socio-demográficas y del uso de la lengua condicionan la realización del pronombre de forma diferente en los dos grupos de exposición. Ahora, el resultado de las diferencias entre recién llegados e inmigrantes establecidos en sus jerarquías de variables y de factores constituye los cambios.

El primer cambio que se observa es que los inmigrantes presentan un orden diferente al de los recién llegados en las jerarquías de variables. Las variables internas que causan mayor probabilidad de uso de los pronombres en los inmigrantes son las mismas que en los recién llegados, pero ocupan diferentes posiciones. En la tabla 3, la variable que ejerce más probabilidad de aparición del pronombre es Persona, seguida por Referencia, Tiempo, Reflexividad y Léxico. Si nos fijamos en la tabla 1 sobre la muestra entera, veremos que el orden es el mismo en las tres primeras variables. Parece que estos hablantes, los inmigrantes, representan más al grupo de la primera generación. No sabemos si esto es así porque hay en la muestra entera hay más inmigrantes que recién llegados, y por lo tanto sus usos particulares tienen mayor impacto en el uso del grupo en su totalidad. También esto ocurre cuando miramos las jerarquías de variables de los inmigrantes caribeños y continentales (tablas 5.1a y 5.2a). Estos grupos tienen el mismo orden para las dos primeras variables (que son Persona y Referencia).

Este primer cambio, distinto orden en las jerarquías de variables, confirma la predicción de que los inmigrantes se comportan de manera diferente a los recién llegados. Es más, indica que incluso dentro de los inmigrantes los que vienen del Continente y los que son del Caribe han

empezado a modificar su gramática como consecuencia de la nivelación dialectal y del contacto con el inglés. Este cambio es un efecto evidente del contacto, derivado de la competencia en inglés de los inmigrantes. Si comparamos este resultado con el encontrado en Otheguy y Zentella (2012: 160), observamos lo siguiente. Los inmigrantes presentan la misma jerarquía de variables que los hablantes de segunda generación. El cambio también se nota cuando un informante de primera generación se parece al de segunda. Los inmigrantes de la primera generación comparten el mismo orden en las variables que influyen el uso pronominal con los nacidos en EE.UU que con los recién llegados.

El segundo cambio se observa en las jerarquías de factores para las variables Persona y Tiempo. En ambas jerarquías, los factores cambian el orden en los grupos de exposición. Es cierto que en todas las jerarquías de factores para la variable Persona, los verbos en singular favorecen más pronombres y los verbos en plural los desfavorecen. Este hecho se relaciona con el significado de la deixis, mencionado anteriormente, que tiene que ver con la focalización inherente de los singulares más que en los plurales. Debido a esto, las primeras posiciones siempre están ocupadas por las personas del singular. Pero el orden de los factores es distinto. Los inmigrantes presentan un orden diferente al de los recién llegados en el segundo y tercer factores. Para estos hablantes, lo que favorece el uso del pronombre en segunda posición es la primera persona del singular y en tercer lugar, la tercera persona del singular. También, hay un cambio en el orden de los factores si nos fijamos en los grupos regionales de exposición. Los recién llegados caribeños (tabla 4.1c) y los inmigrantes caribeños (tabla 5.1b) presentan la misma jerarquía de factores en las tres primeras posiciones. Lo mismo sucede con los recién llegados continentales (tabla 4.2c) y los inmigrantes continentales (tabla 5.2b). Parece que estos cambios se deben a los efectos de la nivelación dialectal.

Las jerarquías de factores de la variable Tiempo también ofrecen datos interesantes. Hay cambios en las posiciones de los factores en los grupos de exposición cuando comparamos recién llegados con inmigrantes. Los recién llegados favorecen el pronombre más con el pasado, el presente y el condicional (tabla 2d). Los inmigrantes con el condicional, el presente y el pasado (tabla 3c). Los tiempos del pasado (el imperfecto, sobre todo) y el condicional (que comparte las terminaciones del imperfecto) son tiempos verbales que presentan ambigüedad en varias de las inflexiones verbales. Es comprensible que aparezcan en primeras posiciones y favorezcan más los pronombres. El cambio en estos grupos de exposición tiene que ver con el orden de estos factores. Cuando analizamos a los grupos regionales de exposición, se observa que los inmigrantes caribeños y continentales son iguales en la primera posición del condicional, pero luego varían en el segundo factor. Todos estos cambios son producidos por el efecto del contacto con otras variedades de español.

Por último, respecto a los grupos regionales de exposición, es de particular interés el resultado de la variable Referencia. Los continentales, cuando se comparan con los caribeños, hacen menos distinción entre la misma referencia y el cambio de referencia (tablas 4.1b, 4.2b, 5.1c y 5.2c). Dada la importancia de esta variable por aparecer siempre en las primeras posiciones en las jerarquías de variables y su aplicabilidad pan-hispánica, ésta es una diferencia importante regional. Esta reducción, estudiada por Flores-Ferrán (2002) y Shin y Otheguy (2009), apoya el mantenimiento de los modelos donde el cambio de referencia favorece los pronombres y la misma referencia los desfavorece.

Resumen y conclusión

Hemos tratado de documentar, de manera precisa, qué queda igual en las gramáticas que guían el uso de los pronombres personales en el español en Nueva York de los grupos de

exposición y qué cambia bajo el impacto del contacto con el inglés y otras variedades de español. El resultado es que los recién llegados han transmitido partes intactas de su gramática a los inmigrantes, pero también los inmigrantes han modificado su gramática como consecuencia del contacto con otras variedades hispanas y el inglés.

Hemos observado que todas las variables internas son muy significativas en los grupos de exposición y en los grupos regionales. El cambio se detecta en el orden diferente de las variables y de los factores. Esto puede deberse, en parte, a que el número de inmigrantes (75) en la muestra es más elevado que el de los recién llegados (39). Como consecuencia, los inmigrantes representan casi siempre la mayor cantidad de variación en la ocurrencia del pronombre. También, los inmigrantes conforman el grupo de hablantes más representativo dentro de la primera generación. No sabemos si esto es por el número elevado de inmigrantes en la muestra. Siempre han sido los que han mostrado órdenes similares a los de la muestra entera dentro de las jerarquías de variables y de factores.

La fuerza de la continuidad del español en Latinoamérica en la gramática interna en el español de los recién llegados y los inmigrantes es más fuerte que la fuerza de cambio. Los recién llegados y los inmigrantes comparten características comunes. Todas las variables internas del estudio son muy significativas en el impacto del uso del pronombre. No obstante, al comparar las jerarquías de variables y de factores el cambio se empieza a hacer visible en los inmigrantes que presentan órdenes diferentes al de los recién llegados.

Los informantes de esta muestra, a pesar de las diferencias en género, edad, edad de llegada, clase social, y competencia en inglés, utilizan los pronombres en modos significativos que reafirman en parte su identidad como miembros de la misma comunidad de habla hispana en Nueva York. Pero la sensibilidad de los inmigrantes mostrada en el orden diferente de algunas

variables y algunos factores refleja que se está produciendo una adaptación lingüística como consecuencia del contacto que tienen con otros grupos de hablantes (por ejemplo, de segunda generación), y por tanto, de la exposición con el inglés y otras variedades del español.

En los inmigrantes, la evidencia del impacto con el inglés comienza a emerger en la debilidad de algunos de los modelos que comparten con los recién llegados. Esto confirma que la continuidad y el cambio son fuerzas muy significativas dentro de la primera generación.

Los diferentes análisis, a través de las regresiones múltiples, han revelado la importancia de las variables internas y de las fuerzas regionales y de contacto que definen el español de los recién llegados y los inmigrantes. Hemos analizado cuál de las variables y de los factores produce mayor variación y hemos observado qué variables internas y qué jerarquías de factores son las que tienen mayor efecto predictivo sobre la ocurrencia pronominal, examinando que estos dos grupos de exposición se distinguen.

A modo de conclusión, éstas son algunas de las diferencias y similitudes significativas encontradas en las jerarquías de variables y factores:

- Las jerarquías de variables diferencian a los recién llegados de los inmigrantes. El orden cambia en los dos primeros factores en todas las jerarquías. Para los recién llegados es Referencia, Persona. Para los inmigrantes es Persona, Referencia.
- Las jerarquías de factores son las mismas en las variables Referencia, Léxico y Reflexividad en los recién llegados y en los inmigrantes, ya sean caribeños o continentales.
- Las jerarquías de factores son diferentes en las variables Persona y Tiempo en los grupos de exposición. En Persona, los recién llegados caribeños y los inmigrantes caribeños presentan la misma jerarquía de factores en las tres primeras posiciones. Igual pasa con los

recién llegados continentales y los inmigrantes continentales. En Tiempo, los inmigrantes caribeños y continentales son iguales en la primera posición del condicional. Los recién llegados caribeños son diferentes en las posiciones primeras cuando se comparan con los recién llegados continentales.

Esto nos lleva a decir que:

- Hay continuidad en los modelos de uso del pronombre personal en Latinoamérica.
- Existen cambios por el contacto de lenguas que contribuyen a la variación en el uso del pronombre.
- La hipótesis de la continuidad con el español latinoamericano sigue teniendo más fuerza frente a la hipótesis del cambio.

Capítulo 6

Conclusiones

Esta tesis ha presentado resultados que muestran diferencias en la utilización del pronombre personal sujeto en español correlacionadas con factores externos e internos en el habla de la primera generación migratoria en la ciudad de Nueva York. Llamamos variables externas a las que tienen que ver con características sociodemográficas del hablante, como por ejemplo su nacionalidad. Llamamos variables internas a las que tienen que ver con características del entorno lingüístico del uso del pronombre sujeto, como por ejemplo el tiempo del verbo con el que se usa el pronombre. Dividimos los hablantes de español de esta primera generación en dos grupos de lo que hemos llamado *exposición* a la vida en la ciudad de Nueva York y, consecuentemente, exposición al inglés. Los hablantes a quienes hemos denominado *recién llegados* y los hablantes a quienes hemos denominado *inmigrantes* (o *inmigrantes establecidos*) exhiben diferentes pautas en la realización del pronombre. El estudio de la variabilidad de los pronombres personales sujetos en estos dos grupos de exposición revela la forma en que el español de la primera generación migratoria neoyorquina sigue las normas de uso del español latinoamericano y, al mismo tiempo, manifiesta la influencia del inglés.

Este estudio desarrolla, expande y extiende lo que ya se ha hecho sobre los pronombres personales sujetos en Estados Unidos en los trabajos de Silva-Corvalán (1982, 1994, 1997); Cameron (1992, 1993, 1995); Flores y Toro (2000); Flores-Ferrán (2002, 2004, 2005); Cameron y Flores-Ferrán (2004); Lapidus y Otheguy (2005); Otheguy, Zentella y Livert (2007); Otheguy y Zentella (2007, 2012); Shin y Otheguy (2009, 2013); Erker (2010); Raña-Risso (2013); y Shin y Montes-Alcalá (2014). Hemos sacado nuestros datos, al igual que algunas de estas otras

investigaciones, del *Corpus Otheguy-Zentella del español en Nueva York*, cuyas características se describen en Otheguy y Zentella (2007, 2012).

Ninguna de las investigaciones basadas en el Corpus Otheguy-Zentella ha analizado si existen diferencias de tasa pronominal y de jerarquías de variables y de factores dentro de la primera generación. La contribución del presente trabajo es documentar los cambios que se detectan en los hablantes de primera generación en la situación de bilingüismo neoyorquina, y determinar que, efectivamente, hay cambios en las jerarquías de variables y de factores en los inmigrantes establecidos cuando se les compara con los recién llegados. Proponemos, siguiendo lo que se expone en Otheguy y Zentella (2012), que los cambios que detectamos en la primera generación pueden plausiblemente atribuirse al contacto con el inglés y a la nivelación dialectal que se da en Nueva York debido a la presencia de latinoamericanos originarios de áreas donde el pronombre sujeto se usa de maneras distintas.

En el trabajo, hemos analizado más de 50,000 verbos en una muestra de 114 informantes que vienen de seis países diferentes y de dos regiones de Latinoamérica. Hemos realizado diferentes tipos de análisis para mostrar que nuestros dos grupos de exposición (recién llegados e inmigrantes) son efectivamente diferentes en cuanto al uso del pronombre. Como señalan Poplack y Levey (2010), los cambios en las tasas pronominales no son necesariamente de por sí evidencia del cambio y del contacto. Más bien, los cambios relacionados con el contacto de lenguas se demuestran comparando las jerarquías de factores. Por este motivo, hemos realizado análisis bivariantes de las tasas de ocurrencia pronominal de los recién llegados y los inmigrantes y su covariación con variables sociodemográficas y de uso de la lengua; análisis de regresión multivariable para medir el impacto de cada variable externa; y análisis multivariantes de las

variables internas que identifican qué partes de la gramática del español de la primera generación en Nueva York permanecen iguales y qué partes se modifican debido al contacto y la nivelación.

En el capítulo 3, establecimos qué grupos de informantes presentaban diferencias significativas respecto a la tasa pronominal. Investigamos el impacto de nueve variables externas, ocho sociodemográficas y una de uso de la lengua. Analizamos cómo algunas de las variables externas de género, edad, edad de llegada, educación, clase social, estatus socioeconómico y competencia en inglés distinguen a estos dos grupos por las diferencias de su tasa pronominal. También, estudiamos que la región de procedencia y el origen nacional influyen de la misma forma en el uso del pronombre en los dos grupos de exposición.

Algunos de los hallazgos encontrados que diferencian a los recién llegados y a los inmigrantes en la tasa pronominal señalan que los inmigrantes usan el pronombre con más frecuencia. Cuanto mayor es la exposición de la población de habla hispana a la ciudad de Nueva York, mayor es su tasa pronominal. Al estudiar los grupos por separado, vimos que la clase social es importante en los recién llegados. Estos usan más el pronombre cuando el informante se describe como miembro de la clase trabajadora. Los inmigrantes se distinguían por la edad de llegada a Estados Unidos y por la competencia en inglés. Los que llegaron antes de los 12 años y los que poseen un inglés excelente usan el pronombre con más frecuencia.

También observamos una diferencia en el uso del pronombre que distingue las dos regiones en las que hemos dividido a nuestros informantes, a quienes hemos llamado *continentales* (con orígenes en Colombia, Ecuador, México) y *caribeños* (con orígenes en Cuba, Dominicana, Puerto Rico). Nuestra observación concuerda con la de trabajos anteriores de López Morales (1992) y de Otheguy y Zentella (2012). Los caribeños inmigrantes y recién llegados y los continentales inmigrantes y recién llegados son diferentes en la tasa pronominal. Los

caribeños, ya sean inmigrantes o recién llegados, presentan una tasa pronominal más alta que los continentales. Dentro de los caribeños, los inmigrantes realizan más el pronombre que los recién llegados. Lo mismo sucede en el caso de los continentales inmigrantes que tienen más pronombres que los continentales recién llegados. Mientras que la educación fue la única variable externa que diferencia a los caribeños recién llegados en el uso del pronombre, la edad, la edad de llegada a EE.UU y la competencia en inglés fueron las características demográficas y lingüísticas que distinguen a los caribeños inmigrantes. En el grupo de los caribeños fueron los inmigrantes los que más se distinguen en su tasa pronominal por diferentes variables externas. Los continentales recién llegados también fueron diferentes del resto por el género y el ESE. Las mujeres de estatus socioeconómico bajo son las que usan el pronombre explícitamente la mayoría de las veces. Los continentales inmigrantes sólo se distinguen por la edad. Cuantos más años tienen, más alta es su tasa pronominal. Así que cuando las dos regiones se dividen en inmigrantes y recién llegados, las variables externas que influyen en la tasa pronominal de los caribeños inmigrantes y recién llegados y los continentales inmigrantes y recién llegados no son las mismas.

Una vez establecidas las diferencias existentes en la tasa pronominal entre los dos grupos de exposición y entre los grupos regionales de exposición, en el capítulo 4 analizamos cuál de las variables producía mayor variación y observamos qué variables externas fueron las que tuvieron mayor efecto predictivo sobre la tasa pronominal. Uno de los resultados reveló que la fuerza de la continuidad del español en Latinoamérica en la distribución pronominal en el español de estos dos grupos de exposición es mayor que la fuerza de cambio.

La variable externa Región fue la que ejerció mayor fuerza predictiva en la tasa pronominal en los recién llegados y los inmigrantes. Son los caribeños de ambos grupos los que presentan el

mayor nivel de uso del pronombre. El aumento del uso del pronombre en el español de los recién llegados y los inmigrantes en Nueva York está condicionado por la continuidad del origen regional del hablante.

No obstante, los recién llegados y los inmigrantes se distinguieron por distintas variables de tipo sociodemográfico y de uso de la lengua. Estos dos grupos de exposición se mostraron distintos en la influencia del inglés, del género y de la edad de llegada. Los recién llegados son diferentes únicamente por el género del hablante. Son las mujeres recién llegadas las que presentan una tasa pronominal alta. Este resultado sitúa a las mujeres caribeñas como las promotoras en la variación pronominal en Nueva York. En cambio, los inmigrantes son distintos por la edad a la que llegaron y el nivel de competencia en inglés. Los inmigrantes que llegaron antes de los 12 años a Estados Unidos y que tienen un inglés excelente utilizan más el pronombre. La hipótesis del contacto de lenguas sólo toma fuerza con los inmigrantes, porque son sólo estos los que muestran cambios en la realización del pronombre, en correlación con la competencia excelente en inglés que poseen y con la edad tan temprana de llegada a Estados Unidos.

En los dos grupos de exposición, recién llegados e inmigrantes, los análisis mostraron que la variable externa que tuvo mayor fuerza predictiva en la tasa pronominal fue distinta en cada grupo regional. Los análisis también señalaron que la fuerza del cambio por el contacto de lenguas fue mínima en tres de los grupos. La hipótesis del contacto con el inglés adquirió sentido únicamente en los inmigrantes que son caribeños.

Los recién llegados caribeños se diferencian por el nivel de educación que tienen; utilizan más el pronombre cuando tienen un nivel más alto de educación. Los recién llegados continentales son distintos por el género del informante; las mujeres son las que mayor tasa

pronominal presentan. Los inmigrantes caribeños se distinguen por la edad a la que vinieron a Estados Unidos y el nivel de competencia en inglés. Los que tienen menos de 12 años y saben un inglés excelente son más propensos a usar el pronombre. Los inmigrantes continentales sólo son diferentes por los años que tienen. Tener más edad se correlaciona con el aumento del pronombre.

En definitiva, hay continuidad regional en la tasa pronominal de los recién llegados y los inmigrantes en Nueva York. El constante flujo de inmigrantes latinoamericanos y caribeños contribuye al mantenimiento de las normas de habla de sus países. También, los datos corroboraron que la hipótesis del contacto de lenguas sólo es válida para los informantes que llevan más tiempo en la ciudad y han estado más expuestos a la lengua de contacto. Es decir, los que tienen más conocimiento del inglés son los inmigrantes y, específicamente, los que vienen del Caribe. Los recién llegados presentan una cierta resistencia a la influencia del inglés en el uso pronominal probablemente debido a la poca exposición con esta lengua y a que su nivel de conocimiento es menor.

La variación pronominal de los recién llegados y los inmigrantes es una combinación de las fuerzas de la continuidad con el español de Latinoamérica y las presiones del cambio por el contacto con el inglés.

En el capítulo 5, documentamos qué quedaba igual en las gramáticas que guían el uso de los pronombres personales en el español en Nueva York de los grupos de exposición y qué cambiaba bajo el impacto del contacto con el inglés y otras variedades de español. El resultado fue que los recién llegados habían transmitido partes intactas de su gramática a los inmigrantes, pero también los inmigrantes habían modificado su gramática como consecuencia del contacto con otras variedades hispanas y el inglés.

Observamos que todas las variables internas fueron muy significativas en los grupos de exposición y en los grupos regionales. El cambio se detecta en el orden diferente de las variables y de los factores. Los inmigrantes siguen representando la mayor cantidad de variación en la ocurrencia del pronombre dentro de la primera generación.

Descubrimos que la fuerza de la continuidad del español en Latinoamérica en la gramática interna en el español de los recién llegados y los inmigrantes es más fuerte que la fuerza de cambio. Los recién llegados y los inmigrantes comparten características comunes. Todas las variables internas del estudio fueron muy significativas en el impacto del uso del pronombre. Pero al comparar las jerarquías de variables y de factores el cambio se empieza a hacer visible en los inmigrantes que presentan órdenes diferentes al de los recién llegados.

Los informantes de esta muestra, a pesar de las diferencias en género, edad, edad de llegada, clase social, y competencia en inglés, utilizan los pronombres en modos significativos que reafirman en parte su identidad como miembros de la misma comunidad de habla hispana en Nueva York. Pero la sensibilidad de los inmigrantes mostrada en el orden diferente de algunas variables y algunos factores refleja que se está produciendo una adaptación lingüística como consecuencia del contacto que tienen con otros grupos de hablantes (por ejemplo, de segunda generación), y por tanto, de la exposición con el inglés y otras variedades del español.

En los inmigrantes, la evidencia del impacto con el inglés comienza a emerger en la debilidad de algunos de los modelos que comparten con los recién llegados. Esto confirma que la continuidad y el cambio son fuerzas muy significativas dentro de la primera generación.

Los diferentes análisis han revelado la importancia de las variables internas y de las fuerzas regionales y de contacto que definen el español de los recién llegados y los inmigrantes. Hemos analizado cuál de las variables y de los factores produce mayor variación y hemos

observado qué variables internas y qué jerarquías de factores son las que tienen mayor efecto predictivo sobre la ocurrencia pronominal, examinando que los recién llegados y los inmigrantes se distinguen.

Bibliografía

- Alba, Orlando. 1988. "Estudio sociolingüístico de la variación de las líquidas finales de palabra en el español cibaño". In R. Hammond & M. Resnick (Eds.), *Studies in Caribbean Spanish Dialectology* (pp. 1–12). Washington DC: Georgetown University Press.
- Álvarez Nazario, Manuel. 1978. "El español de Puertorriqueños en Nueva York". *Philológica Hispaniensa: In Honorem Manuel Alvar: I Dialectología*, ed. por J. Fernández Sevilla, et al., 69–80. Madrid: Gredos.
- Appel, René and Pieter Muysken. 1987. *Language Contact and Bilingualism*. London: Edward Arnold.
- Barrenechea, Ana María y Alicia Alonso. 1969. "Cuantificación del uso de los pronombres personales sujetos en español". *Atlas do II congresso internacional da associaçao de lingüística e filología da América Latina (ALFAL)*. Sao Paulo: Universidad de Sao Paulo.
- Barrenechea, Ana María y Alicia Alonso. 1977. "Los pronombres personales sujetos en el español hablado en Buenos Aires". *Estudios sobre el español hablado en las ciudades principales de América*, ed. por Juan Lope Blanch, 333–49. México: Universidad Nacional Autónoma de México.
- Baumgardner, Robert. J. 1982. *Sociolinguistic aspects of personal pronouns: A Macro/Micro Análisis*. Tesis Doctoral, Universidad del Sur de California.
- Bayley Robert y Lucinda Pease-Álvarez. 1996. "Null and Expressed Pronoun Variation in Mexican-Descent Children's Spanish". In *Sociolinguistic Variation: Data, Theory, and Analysis*, ed. by Jennifer Arnold, Renée Blake, and Brad Davidson, 85-99. Stanford, Calif., CSLI.
- Bayley Robert y Pease-Álvarez, Lucinda. 1997. "Null Pronoun Variation in Mexican-Descent Children's Narrative Discourse". *Language Variation and Change* 9.349-371.
- Bayona, Patricia. 2006. "Sociolinguistic competence in the use of Colombian pronouns of address". *Proceedings of the 2006 Canadian Linguistics Association Annual Conference*, ed. por Claire Gurski y M. Radisic, disponible en: <http://ling.uwo.ca/publications/CLA2006/Bayona.pdf>
- Bentivoglio, Paola. 1987. *Los sujetos pronominales de primera persona en el habla de Caracas*. Caracas: Universidad Central de Venezuela.
- Bookhamer, Kevin. 2012. *The variable grammars of the Spanish subjunctive in second generation bilinguals in New York*. CUNY Graduate Center Tesis Doctoral.

- Cabré-Sans, Yolanda, y Anna Gavarró. 2006. "Subject distribution and verb classes in Chile Catalán". *Cascadilla Proceedings Project*, disponible en: Seneca.uab.es/ggt/Reports/GGT-06-23.pdf
- Cameron, Richard. 1992. *Pronominal and null subject variation in Spanish: constraints, dialects, and functional compensation*, Tesis Doctoral, Universidad de Pennsylvania.
- Cameron, Richard. 1993. "Ambiguous agreement, functional compensation, and non-specific tú in the Spanish of San Juan, Puerto Rico, and Madrid, Spain". *Language Variation and Change* 5.305-34.
- Cameron, Richard. 1994. "Switch reference, verb class, and priming in a variable syntax". *Papers from the 30th Regional Meeting of the Chicago Linguistic Society: The Parasession on Variation in Linguistic Theory*. ed. por Beals, K. et. al. Chicago Linguistic Society 2.27-45.
- Cameron, Richard. 1995. "The scope and limits of switch reference as a constraint on pronominal subject expression". *Hispanic Linguistics* 6/7.1-27.
- Cameron, Richard. 1996. "A Community-Based Test of a Linguistic Hypothesis". *Language in Society* 25.61-111.
- Cantero Sandoval, Gustavo. 1976. "Peculiaridades en el empleo del pronombre personal, YO, en el habla culta de la ciudad de México". *Anuario de letras: Revista de la facultad de Filosofía y Letras*. Universidad de México XIV.233-7.
- Carvalho, Ana M. y Michael Child. 2011. "Subject Pronoun Expression in a Variety of Spanish in Contact with Portuguese". In *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics*, ed. por Jim Michnowicz y Robin Dodsworth, 14-25. Somerville: Cascadilla Proceedings Project.
- Carvalho, Ana M., Naomi Shin, y Rafael Orozco. 2015. *Subject pronoun expression in Spanish: A cross-dialectal perspective*. Georgetown University Press.
- Cedergren, Henrietta J. 1978. "En torno a la variación de la /s/ final de sílaba en Panamá: análisis cuantitativo". In H. López Morales (Ed.), *Corrientes actuales en la dialectología del Caribe hispánico* (pp. 80-103). Río Piedras: Universidad de Puerto Rico.
- Cedergren, Henrietta J., and David Sankoff. 1974. "Variable Rules: Performance as a Statistical Reflection of Competence". *Language* 50.333-355.
- Clark, Brady. 2005. "On Stochastic Grammar". *Language* 81.207-217.
- Davidson, Brad. 1996. "'Pragmatic Weight' and Spanish subject pronouns: The pragmatic and discourse uses of tú and yo in spoken Madrid Spanish". *Journal of Pragmatics* 26.543-65.

- Enríquez, Emilia V. 1984. *El pronombre personal sujeto en la lengua española hablada en Madrid*. Madrid: Consejo Superior de Investigaciones Científicas.
- Erker, Danny. 2010. "A Subsegmental Approach to Coda /s/ Weakening in Dominican Spanish". *International Journal of the Sociology of Language*: 9–26.
- Field, Andy. 2009. *Discovering statistics using SPSS* (3rd ed.). London: Sage Publications.
- Flores-Ferrán, Nydia. 2002. *A sociolinguistic perspective on the use of subject personal pronouns in Spanish narratives of Puerto Ricans in New York City*. Munich: Lincom-Europa.
- Flores-Ferrán, Nydia. 2004. "Spanish subject personal pronoun use in New York City Puerto Ricans: Can we rest the case of English contact?" *Language Variation and Change* 16.49–73.
- Flores-Ferrán, Nydia. 2005. "La expresión del pronombre personal sujeto en narrativas orales de puertorriqueños en Nueva York". *Contactos y contextos lingüísticos: El español en los Estados Unidos y en contacto con otras lenguas*, ed. por Luis A. Ortiz y Manel Lacorte López, 119–29. Madrid: Vervuert.
- Flores, Nydia., and Jacqueline Toro. 2000. "The persistence of dialect features under conditions of contact and leveling". *The Southwest Journal of Linguistics* 19.31–42.
- García, Erica. 1975. *The Role of Theory in Linguistic Analysis: The Spanish Pronoun System*. Amsterdam: North-Holland.
- Goodall, Grant. 2004. "On the Syntax and Processing of Wh-questions in Spanish". *Proceedings of the 23rd West Coast Conference on Formal Linguistics*, Benjamin Schmeiser, Vineeta Chand, Ann Kelleher and Angelo Rodriguez (eds.). Somerville, MA: Cascadilla Press.
- Gutiérrez, Manuel. J. 1995. "On the future of the future tense in the Spanish of the southwest". *Spanish in four continents: studies in language contact and bilingualism*, ed. por Carmen Silva-Corvalán, 214-223. Washington, DC: Georgetown University Press.
- Guy, Gregory R. 2005. "Grammar and Usage: A Variationist Response". *Language* 81.561-663.
- Guy, Gregory R. 2007. "Letter to the Editor". *Language* 83.2-4.
- Hochberg, Judith. 1986. "Functional Compensation for /s/ Deletion in Puerto Rican Spanish". *Language*, 62:609–621.
- Hurtado, Luz Marcela. 2005. "La expresión del pronombre personal sujeto en narrativas orales de puertorriqueños en Nueva York". *Contactos y contextos lingüísticos: El español en los*

- Estados Unidos y en contacto con otras lenguas*, ed. por Luis A. Ortiz López y Manel Lacorte, 187–97. Madrid/Frankfurt: Iberoamericana/Vervuert.
- Klee, Carol y Lynch, Andrew. (2009). *El español en contacto con otras lenguas*. Washington DC: Georgetown University Press.
- Labov, William. 1972. “The isolation of contextual styles”. In *Sociolinguistic patterns* (pp. 70–109). Philadelphia, Pennsylvania: University of Pennsylvania Press.
- Labov, William. 2001. *Principles of Linguistic Changes: Social Factors*. Oxford: Blackwell.
- Lafford, Barbara. 1986. “Valor diagnóstico-social del uso de ciertas variantes de /s/ en el español de Cartagena, Colombia”. In R. Nuñez-Cedeño, I. Paéz, and J. Guitart (Eds.), *Estudios sobre la fonología del español del Caribe* (pp. 53–75). Caracas: La Casa de Bello.
- Lapidus, Naomi, y Ricardo Otheguy. 2005. “Overt nonspecific ellos in the Spanish of New York”. *Spanish in Context* 2.157–74.
- Lipski, John. 1994. *Latin American Spanish*. Londres: Longman.
- López Morales, Humberto. 1992. *El español del Caribe*. Madrid, España: MAPFRE.
- Matos Amaral, P., y Scott Schwenter. 2005. “Contrast and the (non-) occurrence of subject Pronouns”. *Selected Proceedings of the 7th Hispanic Linguistics Symposium*, ed. por D. Eddington, 116–27. Somerville: Cascadilla Press.
- Meillet, Antoine. 1921. *Linguistique historique et linguistique générale*. Paris: Champion.
- Mendard, Scott. 2002. *Applied Logistic Regression Analysis*. 2nd ed. Thousand Oaks, Calif.:Sage.
- Montes Miró, Rosa G. 1986. “Factores discursivos en el análisis de los pronombres personales sujeto en español”. *Morphé* 1.45–71.
- Montrul, Silvina. 2002. “Incomplete acquisition and attrition of Spanish tense/aspect distinctions in adult bilinguals”. *Bilingualism: Language and Cognition* 5.1: 39-68.
- Montrul, Silvina. 2004. “Subject and object expression in Spanish heritage speakers: A case of morpho-syntactic convergence”. *Bilingualism: Language and Cognition* 7.2: 125–142.
- Montrul, Silvina. 2009. “Incomplete Acquisition of Tense-Aspect and Mood in Spanish Heritage Speakers”. *The International Journal of Bilingualism* 13.3: 239-269.
- Morales, Amparo. 1986. “La expresión de sujeto pronominal en el español de Puerto Rico”. *Gramáticas en contacto: análisis sintácticos sobre el español de Puerto Rico*, 89-100. Madrid: Editorial Playor.

- Morales, Amparo. 1982. "La posición de sujeto en el español de Puerto Rico a la luz de la clase semántica verbal, la oposición tema-remata y tópico oracional". *Lingüística Española Actual* 4.23–37.
- Morales, Amparo. 1986. *Gramáticas en contacto: análisis sintácticos sobre el Español de Puerto Rico*. Puerto Rico: Editorial Playor.
- Morales, Amparo. 1997. "La hipótesis funcional y la aparición de sujeto pronominal: el español de Puerto Rico", *Hispania*, 80, 1, 153-165.
- Moser, Karolyn. 2002. "En torno a la reciente situación de las formas de tratamiento de la 2da persona singular en el español de Costa Rica". *Memoria del VIII Congreso de Filología, Lingüística, y Literatura Carmen Naranjo*, San José.
- Müller, Max. 1871. *Lectures on the science of language*. Nueva York: Scribner.
- Navas Sánchez-Elez, María Victoria. 1997. "Factores lingüísticos y extralingüísticos que determinan la alternancia de las variantes de -"s" en un dialecto luso-español, el barranqueño". *Revista de filología románica*, 14(1), 391–410.
- Newmeyer, Frederick J. 2006. "Letter to the Editor". *Language* 82.705-706.
- Newton, Rae R., and Kjell Eric Rudestam. 1999. *Your statistical consultant: Answers to your data analysis questions*. London: Sage Publications.
- Ordóñez, Francisco and Antxon Olarrea. 2001. "Weak Subject Pronouns in Caribbean Spanish and XP Pied-Piping". In Julia Herschensohn, Enrique Mallen, & Karen Zagana (eds.), *Features and Interfaces in Romance: Essays in Honor of Heles Contreras.*, Amsterdam: John Benjamins, pp 223-238.
- Otheguy, Ricardo. 2014. "Variationist sociolinguistics and linguistic theory in the context of pronominal perseveration". En Torres Cacoullos, Dion y Lapierre, *Linguistic Variation: Confronting Fact and Theory*. Rotledge.
- Otheguy, Ricardo, Ana Celia Zentella, and David Livert. 2007. "Language and dialect contact in Spanish in New York: Towards the formation of a speech community". *Language* 83.1-33.
- Otheguy, Ricardo y Ana Celia Zentella. 2007. "Apuntes preliminares sobre el contacto lingüístico y dialectal en el uso pronominal del español en Nueva York". *Spanish in contact: Policy, social and linguistic inquiries*, ed. by Kim Potowski & Richard Cameron. Amsterdam & Philadelphia: John Benjamins Publishing Co. pp. 275-296.
- Otheguy, Ricardo y Ana Celia Zentella. 2012. *Spanish in New York: language contact, dialectal leveling, and structural continuity*. Nueva York: Oxford University Press.

- Pampel, Fred. 2002. *Logistic Regression: A Primer*. Thousand Oaks, Calif.:Sage.
- Pérez Sala, Paulino. 1973. *Interferencia lingüística del inglés en el español hablado en Puerto Rico*. Hato Rey, Puerto Rico: Inter American University Press.
- Poplack, Shana (ed.). 2000. *The English History of African American English*. Oxford: Blackwell.
- Poplack, Shana, and David Sankoff. 1987. "The Philadelphia story in the Spanish Caribbean". *American Speech* 62.291-314.
- Poplack, Shana, and Sali Tagliamonte. 1999. "The grammaticalization of going to in (African American) English". *Language Variation and Change* 11.315-42
- Poplack, Shana, and Stephen Levey. 2010. "Contact induced grammatical change: A cautionary tale". En Peter Auer y Jürgen Erich Schmidt (eds.), *Language and space: An international handbook of linguistic variation*, vol. 1, 391–418. Berlin: Mouton de Gruyter.
- Posio, Pekka. 2011. "Spanish subject pronoun usage and verb semantics revisited: First and second person singular subject pronouns and focusing of attention in spoken Peninsular Spanish". *Journal of Pragmatics*, 43(3), 777-798.
- Raña-Risso, Rocío. 2013. *A corpus-based sociolinguistic study of subject pronoun placement in Spanish in New York*. Ph.D. dissertation. The Graduate Center of the City University of New York, New York, NY.
- Rissel, Dorothy. 1989. "Sex, attitudes, and the assibilation of /r/ among young people in San Luis Potosí", Mexico. *Language variation and change*, 1(3), 269–283.
- Samper Padilla, José Antonio. 1990. *Estudio sociolingüístico del español de Las Palmas de Gran Canaria*. Las Palmas: La Caja de Canarias.
- Sapir, Edward. 1921. *Language*. Nueva York: Harcourt, Brace & Company.
- Shin, Naomi y Cecilia Montes-Alcalá. 2014. "El uso contextual del pronombre sujeto como factor predictivo de la influencia del inglés en el español de Nueva York". *Sociolinguistics Studies*, 8(1). 85-110.
- Shin, Naomi, y Ricardo Otheguy. 2013. "Social class and gender impacting change in bilingual settings: Spanish subject pronoun use in New York". *Language in Society*, 42 429-452.
- Shin, Naomi, y Ricardo Otheguy. 2009. "Shifting sensitivity to Continuity of Reference". *Español en Estados Unidos y otros contextos de contacto: Sociolingüística, ideología y*

- pedagogía*. En Manel Lacorte y Jennifer Leeman. Madrid / Frankfurt: Iberoamericana/Vervuert, pp. 111-136.
- Silva-Corvalán, Carmen. 1982. "Subject expression and placement in Mexican-American Spanish". En J. Amaste y E. Olivares (eds.), *Spanish in the United States: sociolinguistic aspects*, 93-120. Nueva York. Cambridge U.P.
- Silva-Corvalán, Carmen. 1988. "Oral narrative along the Spanish-English bilingual continuum". *On Spanish, Portuguese and Catalan linguistics*, ed. por John Staczek, 172-184. Washington, DC: Georgetown University Press.
- Silva-Corvalán, Carmen. 1994. *Language contact and change: Spanish in Los Angeles*. Nueva York: Oxford University Press.
- Silva-Corvalán, Carmen. (ed.) 1995a. *Spanish in four continents: studies in language contact and bilingualism*. Washington, DC: Georgetown University Press.
- Silva-Corvalán, Carmen. 1995b. "The gradual loss of mood distinctions in Los Angeles Spanish". *Language variation and change* 6:255-272.
- Silva-Corvalán, Carmen. 2001. *Sociolingüística y pragmática del español*. Georgetown: Georgetown University Press.
- Silva-Corvalán, Carmen, y Tracy Terrell. 1989. "Notas sobre la expresión de futuridad en el español del Caribe". *Hispanic Linguistics* 2:191-208.
- Sorace, Antonella. 2005. "Syntactic optionality at interfaces". In L. Cornips and K. Corrigan (eds). *Syntax and Variation: Reconciling the Biological and the Social*, 46-111. Amsterdam: John Benjamins.
- Stewart, Miranda. 2003. "'Pragmatic weight' and face: pronominal presence and the case of the Spanish second person singular pronoun tú". *Journal of Pragmatics* 23.191-206.
- Tagliamonte, Sali. 2002. "Comparative sociolinguistics". *The handbook of language variation and change*, ed. by J. K. Chambers, Peter Trudgill, and Natalie Schilling-Estes, 729-63. Oxford: Blackwell.
- Tagliamonte, Sali. 2006. *Analysing sociolinguistic variation*. Cambridge University Press.
- Thomason, Sarah, G. 2001. *Language Contact: An Introduction*. Washington, DC: Georgetown University Press.
- Toribio, Almeida Jacqueline. 2004. "Convergence as an optimization strategy in bilingual speech: Evidence from code-switching". *Bilingualism: Language and Cognition*, 7, 165-173. Vallduví, E. (1991), *The Informational Component*, New York, NY: Garland.

- Travis, Catherine. 2007. "Genre effects on subject expression in Spanish: Priming in narrative and conversation". *Language Variation and Change* 19:101–35.
- Valdivieso, Humberto, y Juanita Magaña. 1988. "Variación lingüística: La /s/ implosiva en Concepción". *Revista de Lingüística Teórica y Aplicada*, 26, 91–103.
- Weinreich, Ulrich. 1953. *Languages in contact*. The Hague: Mouton.
- Zagona, Karen. 2002. *The syntax of Spanish*. Cambridge University Press.
- Zubizarreta, María Luisa. 1998. *Prosody, Focus, and Word Order*. Cambridge, Massachusetts: The MIT Press.
- Zapata, Gabriela, Liliana Sánchez, and Jacqueline Toribio. 2005. "Contact and contracting Spanish." *International Journal of Bilingualism* 3-4, 377-395.
- Zentella, Ana Celia. 1997. *Growing Up Bilingual: Puerto Rican Children in New York*. Cambridge, MA: Blackwell Publishers.