

EL CORPUS PRESEEA APLICADO A LA ENSEÑANZA DE LA LENGUA: LA EXPRESIÓN DEL SUJETO PRONOMINAL

THE PRESEEA CORPUS APPLIED TO LANGUAGE TEACHING: THE EXPRESSION OF THE PRONOMINAL SUBJECT

Jorge Roselló Verdeguer

Universitat de València

Resumen

La determinación de los factores que condicionan la expresión del sujeto pronominal en español y cómo debe abordarse su enseñanza a estudiantes extranjeros es un tema recurrente en los estudios lingüísticos. En especial, se han observado las dificultades que tienen muchos aprendientes, sobre todo anglófonos, para adquirir un parámetro ajeno a ellos, como es el sujeto nulo. En el presente trabajo, tras un breve repaso sobre algunos estudios sobre la adquisición de los sujetos pronominales en español, proponemos un nuevo enfoque basado en la lingüística variacionista, más concretamente en los estudios realizados por los equipos del grupo PRESEEA en distintos lugares del mundo hispánico. Explicamos los factores que dicho grupo, utilizando técnicas estadísticas, considera más significativos en la presencia del sujeto pronominal y, de acuerdo con ellos, realizamos una propuesta didáctica que pueda ser de utilidad a los estudiantes de español como lengua extranjera.

Abstract

The consideration of the conditioning factors in the expression of the pronominal subject in Spanish and how to teach this issue to foreign students is a recurring topic in linguistics studies. In particular, the difficulties that many learners, especially English speakers, have when acquiring a foreign parameter as is the case with null subjects. In this paper, after a brief review of some studies on the acquisition of pronominal subjects in Spanish, we propose a new approach based on variationist linguistics, more specifically on the studies carried out by the PRESEEA group teams in different places in the Hispanic world. We explain the factors that this group, using statistical techniques, considers most significant in the presence of the pronominal subject and, according to them, we make a didactic proposal that may be used by students of Spanish as a foreign language.

PALABRAS CLAVE: *sujeto nulo, sociolingüística, adquisición, español L2, PRESEEA*

KEY WORDS: *null pronominal subject, sociolinguistics, acquisition, Spanish L2, PRESEEA*

1 INTRODUCCIÓN: ALGUNOS ESTUDIOS TEÓRICOS SOBRE EL TEMA

La enseñanza de la presencia o ausencia del sujeto gramatical con verbos conjugados en forma personal en español es un tema controvertido porque para su correcto uso se deben tener en cuenta, además de factores gramaticales, otros fenómenos pragmáticos y discursivos. La opción de transmitir al estudiante de español como lengua extranjera la idea de que los sujetos, por regla general, se omiten, puede ser la opción más generalizada debido a que los morfemas verbales ya aportan información suficiente acerca de la persona sujeto, frente a otras lenguas como el inglés o el francés. No obstante, la decisión final no resulta tan fácil, ya que —como tendremos ocasión de ver— son muchos los factores que entran en juego en la decisión final.

Jiménez Fernández y Gómez Marzo (2019) se refieren a varios trabajos teóricos sobre el tema y piensan que la mayoría de los estudios realizados se ciñen a una lista de principios formales que resultan demasiado polisémicos y ambiguos, por lo que los aprendientes de ELE hablantes de una L1 sin posibilidad de omitir sujetos, como es el caso del inglés, no encontrarán una explicación adecuada con la que paliar los errores que cometen.

En realidad, repasando diversos estudios experimentales que se han realizado sobre estudiantes de español como lengua extranjera, se observa que los aprendientes no hacen un uso agramatical del pronombre sujeto, aunque sí inapropiado, ya que lo utilizan en contextos donde un nativo, normalmente, no haría uso de él.

El problema, a nuestro juicio, es que los listados de propuestas que se utilizan para abordar el tema, al utilizar ejemplos muy concretos, se ciñen exclusivamente a cuestiones gramaticales y se dejan fuera elementos contextuales y pragmáticos que también serían útiles para explicar determinados usos.

En este sentido, es muy interesante la aportación realizada por López Rueda (2015), que incluye en su estudio sobre la adquisición del pronombre personal sujeto en aprendientes anglófonos una vertiente más pragmática y discursiva. En efecto, esta autora cita trabajos provenientes de la sociolingüística, como el realizado por Bentivoglio (1987) sobre la presencia del pronombre personal sujeto en el habla de Caracas. En dicho estudio se muestran una serie de factores que contribuyen a la presencia del pronombre, y que no habían sido citados anteriormente. Es el caso de determinados tipos de verbo que favorecen la aparición del pronombre, como son los verbos cognitivos (*yo creo que no es así*), verbos perceptivos (*yo no lo vi*), de dicción (*yo te digo que fue él*) o de volición (*yo quiero seguir contigo*).

También se fija López Rueda (2015) en estudios que ponen el foco en la conversación coloquial cuando analizan la presencia del pronombre personal sujeto, como el realizado por Brenes Peña (2011), que, tras analizar conversaciones coloquiales del corpus Val.Es.Co., llega a conclusiones similares, sobre todo en el factor de la ambigüedad, ya que, señala esta autora, el pronombre personal sujeto apenas se utiliza para desambiguar la persona verbal, puesto que esta información se suele deducir por el contexto. Como se ve en (1), son la propia progresión de la conversación junto con la historia conversacional

previa las que hacen que el sujeto de la forma verbal *tenía* no pueda ser otro que el emisor mismo.

- (1) A: § me hacía mis espaguetis↑ mi tortilla de patatas↑
B: tus tortitas ↑ [(RISAS)]
C: [(RISAS)] §
A: § tortitas no llegué a hacer/// tenía yo allí mi cafetera↑ /// mi cama↑ que no logré→ /// ¡qué pena

El factor clave que permite explicar la totalidad de los casos es el contraste o contraposición de sujetos, pero no referida a «la existencia de una oposición o enfrentamiento entre los interlocutores, sino al deseo del hablante de destacar su actitud, sus ideas, su opinión, sus costumbres o preferencias frente al resto de los participantes que se relacionan con él» (Brenes Peña, 2011: 126). Así, en (2) aparecen de manera reiterada un gran número de pronombres personales sujeto que podrían haber sido omitidos sin ningún problema. El elemento que justifica su expresión en el discurso, y que hace que esta no resulte agramatical, es el contraste existente entre las posiciones adoptadas por los dos participantes:

- (2) B: pero si yo no te pido más tiempo ↓ yo lo que te pido es que estés SEGURO/ porque si no estás seguro pues/ oye↓ si quieres lo dejamos ¿me entiendes? yo no lo quiero dejar/ yo por mí ya sabes que/ yo te quiero mucho y yo↑ o sea- si ya sabes que por mí NO/ pero si tú ves que necesitas un tiempo/ o-yo que sé o que no estás seguro de que me quieras→/ o/ es que no lo SÉ/ entonces ya eso es lo que tú pienses// yo por mí [ya]

En realidad, el factor que incide más en la utilización del pronombre no es la utilización de un verbo de opinión, sino la pretensión del hablante de realzar su actitud por encima del resto. No obstante, esa enfatización se asocia a las creencias y opiniones, de ahí que la frecuencia del pronombre personal sujeto sea mayor cuando aparecen verbos cognitivos (pensar, creer, asegurar o confirmar).

2 LA COMPETENCIA CONVERSACIONAL

Por tanto, nos encontramos ante un nuevo enfoque del fenómeno que estamos analizando, que toma como centro las conversaciones, ya sean estas coloquiales o de carácter formal. La conversación, como han puesto de relieve muchos especialistas, es la forma prototípica de usar el lenguaje. Por ello, pensamos, cualquier fenómeno lingüístico debe analizarse en el marco de la conversación porque a través de esta interacción lingüística se adquieren la mayor parte de las estructuras sintácticas de la lengua, los elementos discursivos y los actos comunicativos lingüístico-pragmáticos, como es el caso de la presencia del sujeto pronominal.

La conversación se concibe no como una unidad lingüística superior a la oración, sino como una actividad comunicativa, un tipo de interacción social que posee una estructura y unas unidades propias e independientes. Su estructura es jerárquica, de manera que las unidades principales se van uniendo y combinando hasta formar la actividad comunicativa completa (Cestero, 2000, 2005). La unidad básica principal es el turno de habla, que podemos definir como «el periodo de tiempo que comienza cuando una persona empieza

a hablar y concluye cuando dicha persona deja de hacerlo» (Cestero, 2000: 20). Otro concepto fundamental es el de alternancia, esto es, cuando un hablante finaliza su mensaje y otro hablante toma la palabra para emitir el suyo, y que da lugar al cambio de turno (Cestero, 2000, 2005). Todos estos términos van a ser muy relevantes para nuestro estudio, ya que uno de los factores que la sociolingüística tiene en cuenta en la aparición del pronombre personal sujeto es el cambio de turno del hablante en la conversación.

Por otra parte, la competencia conversacional se ha llevado también a la enseñanza de lenguas extranjeras, ya que es uno de los pilares básicos en la adquisición y la enseñanza de estas. Esto ha supuesto también concienciar a los profesionales de la necesidad de trabajar con textos reales, tanto en el aula como en la elaboración de materiales y manuales, lo que ha propiciado que se utilicen los corpus lingüísticos, «que posibilitan y facilitan la utilización de materiales discursivos auténticos» (Cestero, 2005: 72). En esta línea, nosotros vamos a referirnos al corpus PRESEEA (Proyecto para el Estudio del Español de España y América) y su posible utilización en las clases de español como lengua extranjera en relación con el tema que estamos tratando, la presencia o ausencia del sujeto pronominal en situaciones comunicativas habituales.

3 LA SOCIOLINGÜÍSTICA VARIACIONISTA: LA METODOLOGÍA DEL GRUPO PRESEEA

La sociolingüística ha estudiado los condicionantes lingüísticos y sociales que afectan a la aparición del pronombre personal sujeto en las distintas variedades del español hablado. El grupo PRESEEA ha analizado el fenómeno en distintas zonas del mundo hispánico utilizando una metodología similar (la «Guía PRESEEA de estudio de la expresión del sujeto pronominal», de Guerrero y Manjón-Cabeza 2021¹) y basándose siempre en corpus elaborados según las directrices del citado grupo.

La metodología empleada, siguiendo a Moreno Fernández (2005 y 2021), se basa en entrevistas realizadas a hablantes residentes en núcleos urbanos monolingües o bilingües con una población hispanohablante bien asentada. La selección de informantes para las entrevistas se hace siguiendo los criterios fijados en un muestreo por cuotas, creados a partir de tres variables: sexo/género, edad y nivel de instrucción o nivel sociocultural. En algunas comunidades bilingües también se incluye la lengua habitual o materna, que sirven para analizar las posibles interferencias². Con respecto al tamaño de la muestra, los núcleos con un número de habitantes inferior a 500 000 suelen tener una muestra de 54 informantes, mientras que aumentaría a 72 en aquellas ciudades más pobladas. Los materiales se recogen en conversaciones semidirigidas grabadas con algún dispositivo a la vista, en lugares de fácil acceso para los informantes y con buenas condiciones para su registro (es decir, ausencia de ruidos que dificulten después la interpretación de las palabras). Las entrevistas se realizan proponiendo a los informantes módulos temáticos (saludos, el tiempo, lugar donde vive, familia y amistad, costumbres, anécdotas de su vida, etc.), con una duración aproximada de 45 minutos. Los criterios y normas propuestos son

¹ Esta guía se basa en la elaborada por Bentivoglio, Ortiz y Silva-Corvalán en el año 2011. Para más información sobre las guías se puede consultar la siguiente dirección: <https://preseea.uah.es/documentos-preseea-de-investigacion>.

² También es posible atender a otras variables, como profesión, ingresos económicos, condiciones de alojamiento o 'modos de vida', pero se tratan como variables de post-estratificación.

un mínimo metodológico que busca la comparabilidad de los materiales reunidos en los distintos centros de investigación asociados al proyecto, aunque cada uno de ellos es libre de realizar otras investigaciones complementarias (actitudes lingüistas, ampliación de grupos etarios, etc.).

Antes de entrar en el análisis de los factores determinantes de la presencia o ausencia de sujeto pronominal desde parámetros sociolingüísticos, hemos de realizar algunas precisiones terminológicas. La sociolingüística que se ha encargado de analizar este fenómeno ha sido la llamada sociolingüística variacionista o laboviana (Labov, 1972), basada en la cuantificación y en el análisis probabilístico, lo que la aleja de los procedimientos empíricos intuitivos tradicionales. Labov considera que la lengua es variable y que la variación no es simplemente una mera manifestación de la diversidad consustancial a las actuaciones lingüísticas, sino que constituye una pieza esencial de su aprendizaje, adquisición y uso. Se crea la noción de *variable*, definida como la posibilidad de optar por varias maneras diferentes de expresar un mismo contenido («[different] ways of saying the same thing», Labov 1972: 323), de modo que las alternativas o 'variantes' adquieren un significado social. De una manera más específica, se puede decir que una variable sociolingüística es un elemento lingüístico que covaría con otros factores lingüísticos y con un número determinado de variables extralingüísticas independientes, como la clase social, la edad, el sexo, el grupo étnico o el estilo contextual, entre otras.

Los primeros estudios variacionistas se dieron en el nivel fonológico y su extensión al resto de niveles constituyó un problema complejo y un motivo de controversia, lo que llevó a algunos autores, como Lavandera (1978), a declarar inadecuada la extensión del concepto de variable sociolingüística más allá del nivel fonológico, ya que la mayor o menor presencia de la variable sintáctica puede deberse a las diferencias de significado que estas formas conllevan y que la hacen más o menos compatibles con un contexto comunicativo determinado. Por tanto, la covariación sociosintáctica tiene que ser interpretada de manera especial. Aun estando de acuerdo en líneas generales con esta postura, Silva-Corvalán (2004: 62) hace notar que los estudios de variación sintáctica no tienen necesariamente que examinar o establecer posibles correlaciones sociolingüísticas para justificar su importancia. El análisis de numerosas variables sintácticas ha demostrado que, en muchos casos, el efecto de factores sociales, como el sexo, la educación o la edad, sobre la elección de una variante es mínimo o inexistente, mientras que los factores lingüísticos que condicionan la variación han resultado mucho más interesantes, por lo que hay que atender también a factores pragmáticos³. Sin embargo, es también cierto, como ha señalado la sociolingüística, que es necesaria la cuantificación para realizar estas afirmaciones, y así lo señalan ciertos estudios dialectales, que han demostrado el distinto comportamiento de los pronombres sujeto en distintas zonas hispanicas.

En efecto, la cuantificación es un elemento muy importante en el variacionismo, que ha tenido que desarrollar una metodología compleja y rigurosa (Moreno Fernández, 1994),

³ En este sentido, Serrano (2006: 127) tampoco ve obstáculo en extender el término 'variación sintáctica' a los casos que interactúan con el discurso y la pragmática, porque «la sintaxis se halla enmarcada siempre en un discurso particular». Por lo tanto, en el caso que nos ocupa, la presencia del pronombre debe entenderse «como una variante de su ausencia con alguna motivación comunicativa, dado que, si formalmente no es necesario el sujeto, funcionalmente da indicios de serlo».

diseñada específicamente para abordar el análisis de su objeto de estudio. Así pues, los principales atractivos de la cuantificación en el estudio de la regla variable son, por un lado, la posibilidad de obtención de medidas absolutas y relativas en la descripción de los fenómenos atendiendo a la influencia que ejercen una serie de factores contextuales lingüísticos y extralingüísticos sobre una determinada variable lingüística y, por otro, la elaboración de modelos estadísticos inferenciales predictivos que contribuyen a una rigurosa estimación del estado de un determinado fenómeno lingüístico variable y cuyos resultados pueden ser extrapolados a la población global de la que proceden las muestras de habla analizadas.

4 FACTORES DETERMINANTES EN EL USO DEL PRONOMBRE PERSONAL SUJETO

El primer dato reseñable al intentar explicar el fenómeno es la variabilidad en el uso de los pronombres personales sujeto dentro del mundo hispánico. Como se puede observar en la Tabla 1, su presencia oscila entre un 45,6 % en Barranquilla (Colombia) y un 16,2 % en Huncayo (Perú).

Localidad	Tasa pronominal
Barranquilla, Colombia (Pérez Córdoba y Gomes Camacho, 2019)	45,6 %
Cartagena, Colombia (Pérez Córdoba y Gomes Camacho, 2019)	45,3 %
San Juan, Puerto Rico (Cameron, 1993)	44,8 %
Valledupar, Colombia ((Pérez Córdoba y Gomes Camacho, 2019)	41,8 %
Barranquilla, Colombia (Orozco y Guy, 2008)	35,7 %
Nueva York, EEUU (Otheguy y Zentella, 2012)	34,0 %
Málaga, España (Ávila y Segura, 2022). Nivel bajo	29,0 %
Los Ángeles, EEUU (Silva-Corvalán, 1994)	28,9 %
Santiago de Chile (Martínez-Lara, Guerrero y González, 2020)	25,3 %
Sevilla, España (Repede, 2019). Nivel alto.	24,8 %
New Jersey, EEUU (Flores Ferrán, 2007)	24,0 %
Valencia, España (Roselló, 2021)	21,8 %
México DF (Lastra y Martín, 2015)	21,7 %
Madrid, España (Cameron, 1993)	20,9 %
Granada, España (Manjón Cabeza, Pose y Sánchez, 2017)	17,5 %
Huncayo, Perú (Cerrón-Palomino, 2017)	16,2 %

Tabla 1. Porcentaje de presencia de sujeto pronominal en distintas zonas de habla hispana

Hay que advertir que no todos los estudios se realizaron bajo el mismo paraguas metodológico, por lo que los anteriores datos solo nos sirven para dar cuenta de la heterogeneidad del fenómeno en los distintos territorios donde está presente la lengua española. A pesar de ello, es un dato importante en la enseñanza del idioma, porque, en términos generales, parece que en los territorios americanos (Colombia, Puerto Rico, zonas hispanas de Estados Unidos) se hace un uso más intenso del pronombre que en el español europeo (Valencia, Madrid, Granada).

Vamos a centrarnos, pues, en aquellos estudios que han utilizado una misma metodología (la del grupo PRESEEA) para observar cómo han operado, cuáles han sido sus resultados

y en qué medida podemos mostrarlos para ofrecer un modelo del uso de los pronombres personales sujeto con verbos conjugados aplicable a la enseñanza del español como lengua extranjera.

En primer lugar, los grupos que han trabajado con la metodología PRESEEA son, hasta el momento, los de México DF (Lastra y Martín, 2015), Granada (Manjón-Cabeza, Pose y Sánchez, 2016), Sevilla (Repede, 2019), Santiago de Chile (Martínez-Lara, Guerrero y González, 2020), Valencia (Roselló, 2021) y Málaga (Ávila y Segura Lores, 2022)⁴. También lo han hecho Pérez Córdoba y Gómez Camacho (2019) en el español hablado del Caribe colombiano (Barranquilla, Cartagena y Valledupar), si bien solo han estudiado unas pocas variables⁵.

En todos ellos se ha utilizado la guía propuesta por Bentivoglio, Ortiz y Silva-Corvalán (2011), en la que se establece una variable dependiente (la expresión o no del sujeto pronominal) y once independientes:

- Persona del sujeto (primera, segunda, tercera...).
- Especificidad del sujeto (específico o no específico).
- Modo (indicativo o subjuntivo).
- Tiempo verbal (presente, pretérito, imperfecto, futuro...).
- Progresividad (progresivo y no progresivo).
- Perfectividad (tiempo compuesto y tiempo simple).
- Ambigüedad (forma morfológicamente ambigua y forma verbal no ambigua).
- Clase semántica del verbo (procesos mentales, verbos de estado, dicendi, de actividad...).
- Tipo de cláusula (afirmativa, negativa, interrogativa, preguntas sin palabra interrogativa).
- Correferencialidad.
- Turno de habla

A estas variables se añaden la de tipo social: edad, sexo, nivel sociocultural o nivel de estudios y, en las comunidades bilingües, lengua habitual. Algunos grupos han añadido otras variables, como la secuencia textual, las formas perifrásticas, la ambigüedad fonética potencial o la longitud de la forma verbal que, en algún caso, han resultado ser significativas.

La propia guía excluye toda una serie de supuestos en los que la presencia o la ausencia del pronombre es obligatoria. En un enfoque destinado a la enseñanza habría que incluir estos supuestos para facilitar la tarea al estudiante, ya que —en términos generales— su aprendizaje no reviste mucha dificultad. Se utiliza siempre el sujeto en los siguientes supuestos:

- a) Frases idiomáticas: *Yo qué sé, Tú sabes...*, si no presentan variación.
- b) Usos enfáticos con *mismo / misma*: *yo mismo lo hice; lo hiciste tú misma*.

⁴ El grupo de Sevilla ha publicado los resultados del sociolecto alto y el de Málaga, el nivel de instrucción bajo.

⁵ Para una visión global puede verse también el trabajo de Martín-Butragueño (2020).

c) Sujetos focales, bien porque el sujeto es información nueva que responde a una pregunta (*¿Quién lo trajo? Lo traje yo*), bien porque el sujeto es indiscutiblemente necesario por ser foco de contraste (*Pepe estudia y yo trabajo*).

Y no se utiliza, frente a otros idiomas que sí lo harían como el inglés, en los siguientes:

d) Verbos impersonales, como haber o hacer (*Ø hay una persona en la calle; Hace frío*).

e) Cláusulas relativas con sujeto relativizado (*un estudiante que Ø era de Alcalá*).

f) Cláusulas con se impersonal (*Se come bien en Alcalá*).

Para la elaboración de una guía al estudiante, se tendrán en cuenta los casos en los que es posible la alternancia, como en (3):

- (3) o sea **[Ø] empiezan** por ejemplo este año **[Ø] empiezan** ya un poco a rebelarse pero como **yo** los **conozco** del año pasa(d)o pues **[Ø]** noo **me la suelen pegar** demasia(d)o/ demasia(d)o/ alguna vez sí/ yy **[Ø] les cojo** bastante el tranquillo ¿no?/ yy **[Ø]estoy** a gustísimo con ellos/ y **yo** no sé si **ellos están** a gusto conmigo pero me parece que vamos todos un pocoo bien ¿no?/ ee me lo **paso yo** bomba con ellos VALE_M331_10⁶

En la Tabla 2 podemos observar las variables independientes que han resultado ser estadísticamente significativas en las diversas comunidades en las que se ha trabajado con metodologías similares. Todas obtienen una significación asintótica o p valor por debajo de .005, a excepción de la ambigüedad en el corpus de Málaga, que comentaremos más tarde.

Variable	Málaga (Ávila y Segura, 2022)	Valencia (Roselló, 2021)	Sevilla (Repede, 2019)	Granada (Manjón, Pose, Sánchez, 2016)
Persona del sujeto	.000	.000	.000	.000
Especificidad	.000	.022	.000	.000
Ambigüedad	.950	.000	.000	.000
Clase semántica del verbo	.000	.000	.000	.000
Correferencialidad	.000	.000	.000	.000
Turno de habla	.005	.000	.000	.008
Perífrasis verbal	.002	.001	.002	.002

⁶ Los ejemplos están extraídos de los corpus de Valencia (VALE) y Granada (GRAN). Seguimos el sistema de codificación indicado en el grupo PRESEEA: la ciudad: VALE (Valencia) o GRAN (Granada), seguido del código sociolingüístico del informante: H321. Sexo: H(ombre), M(ujer); grupo de edad: 1 (20-34), 2 (35-55), 3 (> 55); nivel sociocultural o de instrucción: 1 (bajo), 2 (medio), 3 (alto); en el caso de Valencia, lengua habitual: 1 (valenciano), 2 (castellano). Los tres dígitos finales indican el número del informante.

Tabla 2. Comparativa de resultados de variables significativas en diferentes corpus de PRESEEA España⁷

Pero el objetivo de los trabajos del grupo PRESEEA va más allá de analizar la relación entre dos variables, y para ello recurre a un análisis estadístico multivariante. El programa Goldvarb X (Sankoff, Tagliamonte y Smith, 2005) permite averiguar cuál es la probabilidad de que se seleccione una de las dos variantes de la variable dependiente cuando actúan de forma combinada todas las variables independientes. Para ello compara todos los modelos posibles de combinación de las variables o grupos de factores y selecciona el que mejor se ajusta a los datos: determina si existe o no relación entre una o más variables independientes y de dependiente; a continuación, mide la magnitud de esta relación, y, por último, estima la probabilidad de que se produzca un hecho similar en función de los valores asignados a las variables independientes. Los pesos probabilísticos por encima de 0.500 favorecen la aparición del pronombre, mientras que los que se hallan por debajo no favorecen dicha aparición. Los resultados se pueden ver en la Tabla 3.

Variable	Málaga (Ávila y Segura, 2022)	Valencia (Roselló, 2021)	Santiago de Chile (Martínez, 2020)	Granada (Manjón, Pose, Sánchez, 2016)
1.ª persona del singular (yo)	.777	.648	.417	.646
Sujeto específico	.622	.563		.555
Ambigüedad			.587	.701
Verbos de proceso mental	.642	.572	.638	.648
No correferencial con el sujeto precedente	.578	.644		.520
Cambio de turno de habla	.546		.523	.581
Perífrasis verbal modal	.609	.356		.467

Tabla 3. Comparativa de resultados del análisis de regresión logística binomial para la presencia de sujeto

Con estos datos podríamos transmitir al estudiante extranjero unas directrices claras sobre cuándo se deben utilizar los pronombres personales sujeto en español. En primer lugar, con respecto a la persona gramatical, es la primera persona del singular la más utilizada en prácticamente todas las zonas hispánicas estudiadas, lo que se ha vinculado a la naturaleza egocéntrica del discurso (Silva-Corvalán, 1997, 2001), que muestra la necesidad que tiene el hablante de dejar huella de su presencia, como podemos ver en (4).

- (4) es totalmente distinto y no me importaría/ para NADA/irme a vivir HOY a la TORREVIEJA que **yo** conocí// no a la Torre vieja que un día/ nos llevamos a mi madre/ mi marido y yo/ recordando/ yo tenía en la retina aquella Torre vieja/ de calle ANCHÍSIMA/ ee las casas preciosas/ y mi decepción fue que la calle era muy estrechita// aquello era normal// pero claro/ **yoo** la veía de otra forma// ee en mi retinaa estaba otra

⁷ Adaptada de Ávila y Segura (2022).

cosa/// y era el pueblo/ eran lo mismo/ con la única diferencia que uno era de mar y otro de montaña//
laa- la única diferencia es que Sarrión se ha mantenido igual/ porque es un pueblo en la montaña// yy- y
Torreveja es- es un pueblo totalmente EXPLOTA(D)O y es lo que dicen que le pasó a Benidorm/// **yo**
cuando fui a Torreveja de hecho [Ø] no lo conocí ¿eh?/ [Ø] conocí mi calle/// pero [Ø] ya no me
acordaba// y me ha costa(d)o mucho trabajo recordar dónde estaba el horno// dónde estaba el colegio
que **yo** fui// [Ø] no lo recordaba para nada/ dónde estaba la modista que me hi- me hizo el traje de
comuni3n/ VALE_M322_03

Con respecto a la especificidad, un sujeto es específico cuando hace referencia a un ser, real o imaginario, que se considera identificable, al menos para el emisor (5). En caso contrario, se dice que es inespecífico, como en (5). Es interesante explicar al estudiante extranjero las diferencias entre ambos pronombres:

- (5) A mi madre le salen las comidas/ de aquí a Lima más ricas que nada/ digo “mamá ¿cómo espesas Ø estas lentejas?” [...] Es me salen a mí flotando las lentejillas (risas) // “a saber cómo haces **tú** las lentejas” GRAN_M12_024
- (6) entonces/ es una cosa que me encanta porquee/ mm [Ø] **te das cuenta**/ al cabo del tiempo de lo que/
TÚ un día plantastes en esos alumnos// [Ø] llevo MUY pocos años dando clase pero// mm yaa/ pues//
de los primeros que [Ø] tuve/ pues [Ø] **te vas dando cuenta**/ que [Ø] se acuerdan de ti/ que lo que [Ø]
les **dijistes** en clase/ les ha servido para algo/// entonces/ ess/ muy- una gran satisfacción para mí
VALE_H132_02

En líneas generales, la presencia del pronombre es mayor cuando el referente es específico, mientras que con un referente inespecífico se tiende a no utilizar el pronombre personal sujeto. Manj3n, Pose y S3nchez (2017) demuestran que, en Granada, las generaciones m3s j3venes son las que promueven la inespecificidad en el discurso, como un modo de refuerzo «de la solidaridad y familiaridad» de la que hablaban Brown y Gillman (1960). A pesar del inter3s del llamado «tú impersonal», los estudios del PRESEEA no entran a analizarlo con profundidad.

Las formas verbales ambiguas (sobre todo el pret3rito imperfecto de indicativo) favorecen el empleo del pronombre sujeto con mayor frecuencia que cuando se utilizan formas verbales no ambiguas, como podemos ver en (7):

- (7) Pero/ ahora est3/ un poquito agrandado porque compraron el piso/ mm de... del vecino// y lo agrandaron// y **yo** ten3a dos pisos// yo ten3a/ bueno el/ el local// y/ y otro piso donde **yo** viv3a// y donde en fin/ los bajos de mi dormitorio y eso// que lo agrand3// GRAN_H32_033

Si nos centramos en la clase sem3ntica del verbo, son los verbos asociados a procesos mentales (*pensar, opinar, creer...*) los que m3s favorecen la presencia del sujeto. Normalmente este tipo de verbos aparecen en di3logos en los que los informantes tienen que argumentar sobre alguna cuesti3n que se les ha planteado (yo pienso que..., yo creo que...), como podemos ver en (8):

- (8) a m3 me da de comer (risas)/ a m3 me sirve pa(ra) comer (risas)/ a los dem3s no lo s3/ ¡hombre!// **yo creo** que sirve mucho para ir por la vida/ lo que pasa es que la gente no la aplica/ los fil3sofos los primeros ¿eh?/ [...] **yo creo** que la gente es poco coherente/// entonces/ la coherencia te la da el sentido com3n y la l3gica [...] **yoo veo** que muchas veces la gente hace co- cosas totalmente incoherentes y dices ¡pero bueno!/// a lo mejor yo soy muyy- lo analizo mucho desde el punto de vista l3gico// no lo s3// pero **yo**

pienso que hay partes de la Filosofía que es bueno utilizar muchísimo ¿eh?// y bueno la Ética yaa/ no te quiero ni contar// pero bueno// eso ... VALE M232_01

La correferencialidad con el sujeto precedente es, según Silva-Corvalán (2003: 849), el factor estadísticamente más significativo en todos los estudios realizados. Por tanto, si el sujeto no es mencionado en la oración precedente, tiende a explicitarse en la siguiente intervención (9), lo cual no ocurre con la misma frecuencia si el sujeto es el mismo (10):

- (9) siempre había excepciones/ siempre había/ oficiales y suboficiales/ que mantenían el espíritu militar/ pero eran los menos// la **mayoría** procuraba/ escurrir el bulto// con lo cual/ **yoo** pasé desdee entrar con un cierto espíritu de que aquello tendría UN VALOR/ VALE_H332_12
- (10) Llegó un momento en/ que vino un señor// de fuera/ de la Línea de la Concepción/// y **Ø** trajo a sus jugadores [GRAN_H31_051]

El cambio de turno en la conversación suele llevar aparejada una mayor frecuencia en el uso del pronombre. Ya hemos puesto de relieve la importancia de la conversación en la enseñanza de idiomas, y este factor va unido a las intervenciones que se producen en los intercambios lingüísticos (11):

- (11) E: Y conectando sus estudios con/ su profesión actual ¿qué puede decir de ella?
I: Pues/ curiosamente que cuando yo estudié la carrera que hice Farmacia nunca pensé que me iba a dedicar a investigación científica como me he dedica[d]o. GRAN_M33_018

Aunque es una variable añadida por algunos grupos a la guía propuesta por Bentivoglio, Ortiz y Silva-Corvalán (2011), las perífrasis verbales son una variable significativa y hay una tendencia a que las formas perifrásticas patrocinen la ausencia de sujeto, frente a las no perifrásticas. Manjón-Cabeza, Pose y Sánchez (2016: 198-200) consideran que esta variable, por sí misma, es poco fiable y que las muestras tendrían que ser interrelacionadas con otros factores como la semántica verbal, aunque los resultados que ellos extraen después de un análisis más minucioso no son concluyentes. Sí parece tener más relación con la aparición del pronombre sujeto la longitud de la forma verbal (es obvio que las perífrasis son más extensas), y, en este sentido, se ha observado que las formas largas, es decir, con un mayor número de sílabas, tienden a una menor presencia del sujeto pronominal. Además, se constata que esa tendencia es más acentuada en las perífrasis aspectuales (12) que en las modales (13):

- (12) Y yo he dicho/ «pues **vamos a montar** algo»/ (risas) con las barras de labios nos forramos (risas). GRAN_H12_021
- (13) O sea **yo me puedo fumar** un/ un porro/ o meterme lo que quiera siempre y cuando/ a los demás los respete GRAN_H21_044

En cuanto a las variables sociales han resultado ser, en líneas generales, poco importantes en relación con la presencia o ausencia del pronombre personal. No obstante, el factor edad sí es significativo en algunos lugares (Granada, Santiago de Chile, Sevilla, Barranquilla, Ciudad de México), y la tendencia es, casi siempre la misma, esto es, que la frecuencia de uso es mayor en el grupo etario de 55 años en adelante y que, por el contrario, son los jóvenes los que menor explicitación hacen de él.

5 CONCLUSIONES

La expresión del sujeto pronominal es uno de los fenómenos sintácticos más analizados por la sociolingüística hispánica. Tampoco ha pasado inadvertido a los numerosos especialistas que lo han estudiado con fines didácticos a la enseñanza del español como lengua extranjera. Así, por ejemplo, Gili Gaya (1989: 228) consideraba indispensable en la enseñanza de nuestra lengua a extranjeros «corregir su tendencia a enunciar todos los verbos con sujeto pronominal, a fin de evitar la redundante pesadez que esto comunica al estilo». Similar opinión tenía el Esbozo (1973: 421), que consideraba conveniente «llamar la atención de los traductores que, por contaminación del texto traducido o por apresuramiento, no advierten a veces la machacona pesadez que comunica al estilo la repetición insistente de los sujetos pronominales». Solamente consideraba el texto académico un empleo correcto de los pronombres sujeto en los casos de énfasis expresivo o para evitar alguna ambigüedad posible.

Así pues, vistos los estudios realizados por la sociolingüística, más concretamente la sociolingüística de tipo variacionista del grupo PRESEEA, sobre la presencia del sujeto pronominal con verbos conjugados, hemos establecido una serie de directrices aplicables a las clases de ELE que, creemos, resultan claras y, a la vez, operativas.

6 REFERENCIAS BIBLIOGRÁFICAS

- Ávila, Antonio Manuel y Alba Segura Lores (2022): «Estudio de las variables predictoras de la expresión del sujeto pronominal en el corpus PRESEEA Málaga. Nivel de instrucción bajo», *Anuario de Letras*, 10 (2), 57-93.
- Bentivoglio, Paola (1987): Los sujetos pronominales de primera persona en el habla de Caracas, Caracas, Universidad Central de Venezuela.
- Bentivoglio, Paola, Luis A. Ortiz y Carmen Silva-Corvalán (2011): «La variable expresión del sujeto pronominal. Guía de codificación», [en línea]: <http://preseea.linguas.net/Portals/0/Metodologia/guia_codificacion_sujetos_julio_2011.pdf>. [Consulta: 02/11/2023].
- Brenes Peña, M. Ester (2011): «La personalización de enunciado. Análisis pragmalingüístico del pronombre personal sujeto en la oralidad», en Bustos, José Jesús et alii, eds., *Sintaxis y análisis del discurso hablado en español. Homenaje a Antonio Narbona*, Sevilla, Publicaciones de la Universidad de Sevilla, vol. 1, 121-136.
- Brown, Roger y Albert Gilman (1960): «The pronouns of power and solidarity», en Fishman, Joshua, ed., *Reading in the Sociology of Language*, La Haya, Mouton, 252-275.
- 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, 305-334.
- Cerrón-Palomino, Álvaro (2018): «Variable Subject Pronoun Expression in Andean Spanish: a Drift from the Acrolect», *Onomázein*, 42, 53-73.
- Cestero, Ana María (2000): *El intercambio de turnos de habla en la conversación. Análisis sociolingüístico*, Alcalá de Henares, Servicio de Publicaciones de la Universidad de Alcalá.
- Cestero, Ana María (2005): *Conversación y enseñanza de lenguas extranjeras*, Madrid, Arco/Libros.
- Flores Ferrá, Nydia (2007): «Los Mexicanos in New Jersey: Pronominal Expression and Ethnolinguistic Aspects», en Holmquist, Jonathan, Augusto Lorenzino y Lotfi Sayahi, eds., *Selected Proceedings of the Third Workshop on Spanish Sociolinguistics*, Somerville, Cascadilla Proceedings Project, 85-91.
- Gili Gaya, Samuel (1989): *Curso superior de sintaxis española*, Barcelona, Vox, 15.ª ed.
- Guerrero, Silvana y Antonio Manjón-Cabeza (2022): *Guía PRESEEA de estudio de la expresión del sujeto pronominal*, Universidad de Alcalá de Henares, [en línea]: <<https://preseea.uah.es/documentos-preseea-de-investigacion>>. [Consulta: 02/11/2023].

- Jiménez Fernández, Ángel y Lucía Gómez Marzo (2019): «El sujeto nulo en español para aprendientes anglófonos: tratamiento en manuales de ELE y reflexiones didácticas», *MarcoELE: revista de didáctica ELE*, 28, 1-25, [en línea]: <<https://marcoele.com/sujeto-nulo-para-anglofonos/>>. [Consulta: 02/11/2023].
- Labov, William (1972: *Sociolinguistic Patterns*, Philadelphia, University of Philadelphia.
- Lastra, Yolanda y Pedro Martín-Butragueño (2015): «Subject Pronoun Expression in Oral Mexican Spanish», en Carvalho, Ana Maria, Rafael Orozco y Naomi Lapidus SHIN, eds., *Subject Pronoun Expression in Spanish: A Cross-dialectal perspective*, Washington, Georgetown University Press, 39-57.
- Lavandera, Beatriz (1978): «Where Does the Sociolinguistic Variable Stop?», *Language in Society*, 7, 171-178.
- López Rueda, Susana (2015): «Adquisición de los sujetos pronominales en español por aprendientes anglófonos», *MarcoELE: Revista de didáctica ELE*, 20, [en línea]: <<https://marcoele.com/adquisicion-de-los-sujetos-pronominales/>>. [Consulta: 02/11/2023].
- Manjón-Cabeza, Antonio, Francisca Pose y Francisco José Sánchez (2016): «Factores determinantes en la expresión del sujeto pronominal en el corpus PRESEEA Granada», *Boletín de Filología*, 51(2), 181-207.
- Manjón-Cabeza, Antonio, Francisca Pose y Francisco José Sánchez (2017): «El factor social edad y la expresión del sujeto pronominal en el español hablado en Granada», *LEA*, 39, 31-51.
- Martín-Butragueño, Pedro (2020): «An approach to subject pronoun expression patterns in data from the “Project for the Sociolinguistic Study of Spanish in Spain and America”», *Spanish in Context*, 17(2), 294-316.
- Martínez-Lara, José Alejandro, Silvana Guerrero y Carlos González (2020): «(Yo) opino: la variación del sujeto pronominal en el habla de Santiago de Chile», en Rogers, Brandon y Mauricio Figueroa Candía, eds., *Lingüística del castellano chileno: estudios sobre variación, innovación, contacto e identidad*, Málaga, Vernon, 286-308.
- Moreno Fernández, Francisco (1994): «Status quaestionis: sociolingüística, estadística e informática», *Lingüística*, 6, 95-154.
- Moreno Fernández, Francisco (2005): «Corpus para el estudio del español en su variación geográfica y social. El corpus PRESEEA», *Oralia*, 8, 123-139.
- Moreno Fernández, Francisco (2021): «Metodología del “Proyecto para el estudio sociolingüístico del español de España y América” (PRESEEA)», *Documentos PRESEEA de investigación*, *Documentos PRESEEA de trabajo*, 1». Alcalá de Henares, Editorial Universidad de Alcalá, [en línea]: <<https://doi.org/10.37536/preseea.2021.doc1>>. [Consulta: 02/11/2023].
- Orozco, Rafael y Gregory R. Guy (2008): «El uso variable de los pronombres sujetos: ¿qué pasa en la costa Caribe colombiana?», en Westmoreland, Maurice y Juan Antonio Thomas, eds., *Selected Proceedings of the 4th Workshop on Spanish Sociolinguistics*, Somerville, MA, Cascadilla Proceedings Project, 70-80.
- Otheguy, Ricardo y Ana Celia Zentella (2012): *Spanish in New York: Language contact, dialectal leveling, and structural continuity*, Oxford, Oxford University Press.
- Pérez Córdoba, Alder Luis y Roberto Gomes Camacho (2019): «A expressão do pronome pessoal sujeito no espanhol falado no Caribe colombiano», *Estudos Linguísticos (São Paulo)*, 1978), 48, 1, 404-424, <https://doi.org/10.21165/el.v48i1.2148>.
- RAE (1973): *Esbozo de una nueva gramática de la lengua española*, Madrid, Espasa-Calpe.
- Repede, Doina (2019): «Condicionantes sociolingüísticos de la expresión del sujeto pronominal en el sociolecto alto del corpus PRESEEA-Sevilla», *Revista de Investigación Lingüística*, 22, 397-423.
- Roselló Verdeguer, Jorge (2021): «Factores determinantes en la expresión del sujeto pronominal en español hablado de Valencia», *LINRED*, 17, 1-24, [en línea]: <https://linred.web.uah.es/articulos_pdf/LR-articulo-16012021.pdf>. [Consulta: 02/11/2023].
- Sankoff, David, Sali A. Tagliamonte y Eric Smith (2005): *GoldVarb X. A variable rule application for Macintosh and Windows*, Toronto, University of Toronto.
- Serrano, M.^a José (2006): «Acción e interacción social en variación sintáctica y discursivo-pragmática», en Blas Arroyo, José Luis, Manuela Casanova Ávalos y Mónica Velando Casanova, eds., *Discurso y sociedad: contribuciones al estudio de la lengua en contexto social*, Castelló de la Plana, Publicacions de la Universitat Jaume I, 121-140.

Silva-Corvalán, Carmen (1994): *Language Contact and Change*. Spanish in Los Angeles, Oxford, Clarendon Press.

Silva-Corvalán, Carmen (1997): «Variación sintáctica en el discurso oral: problemas metodológicos», en Moreno Fernández, Francisco, ed., *Trabajos de Sociolingüística Hispánica*, Alcalá de Henares, Universidad de Alcalá, 115-135.

Silva-Corvalán, Carmen (2001): *Sociolingüística y pragmática del español*, Georgetown, Georgetown University Press.

Silva-Corvalán, Carmen (2003) «Otra mirada a la expresión del sujeto como variable sintáctica», en Moreno Fernández, Francisco et alii, coords., *Lengua, variación y contexto*. Estudios dedicados a Humberto López Morales, Madrid, Arco / Libros, 849-860.

Silva-Corvalán, Carmen (2004): «La expresión del sujeto desde una perspectiva variacionista», en Cabré, Teresa, Mercè Lorente y Elisabet Solé, eds, *Lingüística aplicada: cicle de conferències i seminaris 98-00*, Barcelona, Institut Universitari de Lingüística Aplicada de la Universitat Pompeu Fabra, 61-76.