



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE  
FACULTAD DE LETRAS

**EL ORDEN SUJETO-VERBO EN ENUNCIADOS INTERROGATIVOS TOTALES DEL ESPAÑOL**

Tesis presentada como requisito parcial para obtener el grado de Doctora en Lingüística

**DANIA RAMOS MARTÍN**

Profesor director: Dr. Carlos González

Profesor codirector: Dr. Xose Padilla

Profesora Informante Interna: Dra. Verónica Orqueda

Profesor Informante Externo: Dr. Matías Jaque

Profesor Informante Externo: Dr. Xavier Villalba

Octubre de 2021

## Resumen

El orden de palabras en español ha merecido la atención de múltiples investigadores y teóricos a lo largo de décadas. Desde los estudios tipológicos a los variacionistas, se ha intentado poner en claro cuáles son los factores que intervienen en la secuenciación de constituyentes oracionales. Mientras que la declaración ha estado en el epicentro de los esfuerzos investigativos, la interrogación, por su parte, ha exhibido un panorama menos ventajoso. Tanto si se revisa la tradición gramatical, como la producción científica de los últimos años, es posible observar que las referencias sobre esta temática se concentran en torno a la interrogación parcial o con pronombre o adverbio interrogativo. No abundan, sin embargo, las investigaciones basadas en corpus de habla real, que midan cuantitativamente cómo se conjugan factores de diversa naturaleza en la manifestación de cierto orden sintáctico. Tampoco se conocen a la fecha estudios que indaguen sobre la variación dialectal que experimenta el fenómeno.

La presente investigación, de carácter descriptivo y con un enfoque funcional, aborda el tema de orden de constituyentes oracionales en enunciados interrogativos totales del español. El objetivo principal es comprender el fenómeno de la alternancia sujeto-verbo en enunciados provenientes de textos escritos contenidos en el Corpus de Referencia del Español del Siglo XXI (CORPES). Mediante análisis estadístico descriptivo e inferencial se observó la relación entre las variables que intervinieron en el estudio (de carácter lingüístico y extralingüístico) y se determinó la influencia de estas, de manera individual y conjunta, en el fenómeno estudiado.

Los resultados muestran que el orden VS es mayoritario, pero no categórico; y que el fenómeno de la posposición del sujeto exhibe una variación dialectal, que tiende a predominar en las zonas dialectales de *España*, y *México* y *Centroamérica*. Los factores lingüísticos que explican la posposición del sujeto son los mismos que influyen, al menos de manera teórica, en la secuenciación VS de la declaración. Sin embargo, dada su frecuencia, puede plantearse que la fuerza ilocutiva, efectivamente, actúa en favor de este orden, pues es el más frecuente. El análisis inferencial develó que, cuando actúan todas las variables al unísono, son la transitividad verbal, la función pragmática del enunciado, la definitud del sujeto y la cantidad de argumentos del verbo (en ese orden) las que influyen de manera significativa en la secuenciación de constituyentes. Dicho de otra manera, la posposición del sujeto se puede predecir, en gran medida, a partir de la coaparición de estas variables.

Esta Tesis Doctoral fue financiada por la Comisión Nacional de Investigación Científica y Tecnológica, FONDECYT, actual Agencia Nacional de Investigación y Desarrollo de Chile, ANID, en el marco del Programa de Formación de Capital Humano Avanzado, Beca Doctorado Nacional 2015.

*A mi hijo, el fruto más fecundo de estos años de doctorado*

*A mis padres, por su sacrificio ejemplar*

## Índice de contenidos

|  |            |
|--|------------|
| <b>ÍNDICE DE TABLAS</b>  | <b>8</b>   |
| <b>ÍNDICE DE GRÁFICOS</b>  | <b>10</b>  |
| <b>ÍNDICE DE FIGURAS</b>   | <b>13</b>  |
| <b>INTRODUCCIÓN</b>  | <b>16</b>  |
| <b>CAPÍTULO 1. MARCO TEÓRICO Y CONCEPTUAL</b>  | <b>19</b>  |
| <b>1.1. LAS INTERROGATIVAS PARCIALES Y TOTALES</b>   | <b>19</b>  |
| <b>1.2. EL ORDEN DE PALABRAS</b>   | <b>28</b>  |
| 1.2.1. SOBRE LOS CONCEPTOS DE ORDEN BÁSICO Y NO MARCADO  | 28         |
| 1.2.2. EL ENFOQUE VARIACIONISTA  | 33         |
| 1.2.3. EL ORDEN DE PALABRAS EN ENUNCIADOS INTERROGATIVOS TOTALES EN LA TRADICIÓN GRAMATICAL      | 35         |
| 1.2.4. EL ORDEN DE PALABRAS EN ENUNCIADOS INTERROGATIVOS TOTALES EN LA INVESTIGACIÓN LINGÜÍSTICA | 39         |
| <b>1.3. HIPÓTESIS DE TRABAJO, PREGUNTAS DE INVESTIGACIÓN Y OBJETIVOS DE INVESTIGACIÓN</b>        | <b>51</b>  |
| <b>1.4. VARIABLES INCLUIDAS EN EL ESTUDIO</b>  | <b>53</b>  |
| 1.4.1. ZONA GEOGRÁFICA   | 53         |
| 1.4.2. SEMÁNTICA VERBAL  | 55         |
| 1.4.3. TIPO DE VERBO SEGÚN REALIZACIÓN VALENCIAL   | 61         |
| 1.4.3.1. Los verbos pseudoimpersonales   | 65         |
| 1.4.3.2. Instancias de descripción de la variable  | 65         |
| 1.4.4. TEMPORALIDAD VERBAL   | 68         |
| 1.4.5. MORFOLOGÍA DEL SUJETO   | 70         |
| 1.4.6. REFERENCIALIDAD DEL SUJETO  | 72         |
| 1.4.6.1. La animacidad   | 73         |
| 1.4.6.2. La definitud  | 78         |
| 1.4.7. COMPLEJIDAD DEL SUJETO  | 81         |
| 1.4.8. FUNCIÓN PRAGMÁTICA DEL ENUNCIADO INTERROGATIVO  | 84         |
| 1.4.8.1. Categorías analíticas del presente estudio  | 97         |
| <b>CAPÍTULO 2. METODOLOGÍA</b>   | <b>107</b> |
| <b>2.1. EL CORPUS</b>  | <b>107</b> |
| 2.1.1. HERRAMIENTA DE SELECCIÓN DEL CORPUS. EL CORPES XXI  | 107        |
| 2.1.2. CRITERIOS DE EXCLUSIÓN  | 107        |
| 2.1.2.1. Filtrado 1  | 107        |
| 2.1.2.2. Filtrado 2  | 108        |
| 2.1.2.3. Otras decisiones metodológicas  | 113        |
| 2.1.3. PRINCIPALES DIFICULTADES EN EL MANEJO DE LA HERRAMIENTA DIGITAL EN LÍNEA                  | 114        |
| 2.1.4. LIMITACIONES EN EL MANEJO DEL CORPUS  | 116        |

|  |            |
|--|------------|
| 2.1.5. DESCRIPCIÓN DE CASOS CONTENIDOS EN EL CORPUS  | 117        |
| 2.1.6. DECISIONES METODOLÓGICAS REFERIDAS A LA DESCRIPCIÓN                                   | 119        |
| 2.1.6.1. Sobre la estimación y uso de los conceptos de transitividad y estructura argumental | 119        |
| 2.1.6.2. Sobre la determinación del lugar ocupado por el circunstante                        | 120        |
| 2.1.6.3. Sobre la determinación de características referenciales del sujeto                  | 121        |
| 2.1.6.4. Sobre la variable complejidad del sujeto  | 122        |
| 2.1.6.5. Sobre la determinación de la función del enunciado interrogativo                    | 123        |
| <b>2.2. EL ANÁLISIS ESTADÍSTICO</b>  | <b>124</b> |
| 2.2.1. Naturaleza de las variables   | 125        |
| 2.2.2. Las pruebas estadísticas  | 125        |

---

## **CAPÍTULO 3. RESULTADOS Y DISCUSIÓN** **127**

|  |            |
|--|------------|
| <b>3.1. ANÁLISIS DE FRECUENCIA Y ASOCIACIÓN DE VARIABLES</b>                                     | <b>127</b> |
| 3.1.1. EL ORDEN PREDOMINANTE EN LAS INTERROGATIVAS TOTALES DEL CORPUS                            | 128        |
| 3.1.2. VARIABLE EXTRALINGÜÍSTICA   | 130        |
| 3.1.2.1. El orden del sujeto respecto del verbo según zona geográfica                            | 130        |
| 3.1.3. VARIABLES REFERIDAS AL PREDICADO  | 132        |
| 3.1.3.1. El orden del sujeto respecto del verbo según semántica verbal                           | 132        |
| 3.1.3.2. El orden del sujeto respecto del verbo según tiempo verbal                              | 137        |
| 3.1.3.3. El orden del sujeto respecto del verbo según realización valencial                      | 140        |
| 3.1.3.4. El orden del sujeto respecto del verbo según transitividad                              | 145        |
| 3.1.4. VARIABLES REFERIDAS AL SUJETO   | 151        |
| 3.1.4.1. El orden del sujeto respecto del verbo según referencialidad del sujeto                 | 151        |
| 3.1.4.2. El orden del sujeto respecto del verbo según morfología del sujeto                      | 161        |
| 3.1.4.3. El orden del sujeto respecto del verbo según complejidad del sujeto                     | 169        |
| 3.1.5. VARIABLE SEMÁNTICO-PRAGMÁTICA   | 174        |
| 3.1.5.1. El orden del sujeto respecto del verbo según la función pragmática del enunciado        | 174        |
| 3.1.6. RECAPITULACIÓN DE LA ESTADÍSTICA DESCRIPTIVA Y DEL ANÁLISIS DE ASOCIACIÓN ENTRE VARIABLES | 184        |
| <b>3.2. ANÁLISIS DE REGRESIÓN LOGÍSTICA</b>  | <b>185</b> |
| 3.2.1. OBJETIVOS   | 185        |
| 3.2.2. HERRAMIENTA DE ANÁLISIS   | 185        |
| 3.2.3. VARIABLES Y CATEGORÍAS DE ANÁLISIS  | 186        |
| 3.2.4. SOBRE LA INTERPRETACIÓN DE LOS DATOS  | 187        |
| <b>3.3. RESULTADOS REGRESIÓN LOGÍSTICA SIMPLE</b>  | <b>188</b> |
| 3.3.1. VARIABLE ZONA GEOGRÁFICA  | 189        |
| 3.3.2. VARIABLE SEMÁNTICA VERBAL   | 190        |
| 3.3.3. VARIABLE REALIZACIÓN VALENCIAL  | 191        |
| 3.3.4. VARIABLE TRANSITIVIDAD VERBAL   | 192        |
| 3.3.5. VARIABLE REFERENCIALIDAD DEL SUJETO   | 193        |
| 3.3.5.1. Animacidad  | 193        |
| 3.3.5.2. Definitud   | 194        |
| 3.3.6. VARIABLE MORFOLOGÍA DEL SUJETO  | 195        |
| 3.3.7. VARIABLE COMPLEJIDAD DEL SUJETO   | 196        |
| 3.3.8. VARIABLE FUNCIÓN PRAGMÁTICA DEL ENUNCIADO INTERROGATIVO                                   | 197        |
| 3.3.9. RECAPITULACIÓN ANÁLISIS DE REGRESIÓN  | 199        |
| <b>3.4. RESULTADOS REGRESIÓN LOGÍSTICA MÚLTIPLE</b>  | <b>200</b> |

|                                 |            |
|---------------------------------|------------|
| <b>CAPÍTULO 4. CONCLUSIONES</b> | <b>202</b> |
| <b>BIBLIOGRAFÍA</b>             | <b>210</b> |
| <b>ANEXOS</b>                   | <b>222</b> |

## Índice de tablas

|  |     |
|--|-----|
| Tabla 1. Frecuencias absolutas (F) y relativas (%) de interrogativas parciales (IntP) e interrogativas totales con sujeto implícito (SujI), respecto del total de casos revisado por zona.....   | 31  |
| Tabla 2. Separación en zonas dialectales y países que las integran. Fuente: adaptación de CORPES XXI.....  | 54  |
| Tabla 3. Propuesta de clasificación de Freed (1994), tomando como eje el contenido literal de la interrogativa y su contexto de uso .....  | 89  |
| Tabla 4. Clasificación de interrogativas según Escandell (1999), tomando como eje central la interpretación del enunciado según la intención del hablante .....  | 95  |
| Tabla 5. Resumen de tipos de interrogativa propuestos por Freed (1994) y Escandell (1999) .....  | 96  |
| Tabla 6. Cantidad total de casos por zona y cantidad de casos filtrados para obtener 100 datos analizables en cada un de ellas .....   | 112 |
| Tabla 7. Frecuencias absolutas (F) y relativas (%) de los criterios de selección y exclusión de las interrogativas filtradas por zona geográfica .....   | 112 |
| Tabla 8. Frecuencias totales y absolutas de los cruces entre verbos mentales y animacidad del sujeto. Los datos dentro de la categoría verbos mentales se han agrupado por orden (SV y VS) .....   | 135 |
| Tabla 9. Tabla cruzada tiempo verbal y orden del sujeto. Los niveles de la primera variable han sido reducidos a dos: tiempos simples y tiempos compuestos .....   | 140 |
| Tabla 10. Tabla cruzada de las variables realización valencial y orden del sujeto. Los niveles originales –bivalentes y trivalentes– se han fundido en uno solo: más de un argumento. ....   | 144 |
| Tabla 11. Residuos corregidos tipificados para los cruces de las variables realización valencial y orden del sujeto. Los niveles originales –bivalentes y trivalentes– se han fundido en uno solo: más de un argumento .....   | 144 |
| Tabla 12. Tabla cruzada de las variables transitividad y animacidad .....  | 147 |
| Tabla 13. Frecuencias absolutas y relativas de sujetos humanos en predicados transitivos, agrupados por orden del sujeto (SV y VS) .....   | 147 |
| Tabla 14. Tabla cruzada de las variables animacidad y orden del sujeto en las interrogativas .....   | 148 |
| Tabla 15. Residuos corregidos tipificados de la prueba de independencia (chi-cuadrado) entre las variables animacidad del sujeto y transitividad verbal .....  | 148 |
| Tabla 16. Tabla cruzada de las variables presencia de complemento circunstancial y orden del sujeto.....   | 149 |
| Tabla 17. Residuos corregidos tipificados para la asociación de las variables referidas a la referencialidad del sujeto (animacidad y definitud) y orden de palabras. Los residuos por encima de +2 o por debajo de -2 muestran cruces estadísticamente significativos ..... | 160 |

|  |     |
|--|-----|
| Tabla 18. Tabla cruzada de las variables Morfología del sujeto y Orden del sujeto. La primera de ellas se ha simplificado, dejando solo tres niveles: sujetos léxicos, infinitivos y sujetos clausulares, y sujetos pronominales ..... | 165 |
| Tabla 19. Frecuencias absolutas de pronombres personales en las interrogativas del corpus, según zona geográfica .....   | 168 |
| Tabla 20. Tabla cruzada de categorías funcionales de Freed (1994) y Orden del sujeto ...   | 179 |
| Tabla 21. Proporción de casos con orden VS según zona geográfica .....   | 182 |
| Tabla 22. Análisis de regresión logística entre orden del sujeto y zona geográfica .....   | 190 |
| Tabla 23. Análisis de regresión logística entre orden del sujeto y semántica verbal .....  | 191 |
| Tabla 24. Análisis de regresión logística entre orden del sujeto y estructura argumental .   | 192 |
| Tabla 25. Análisis de regresión logística entre orden del sujeto y transitividad verbal .....  | 193 |
| Tabla 26. Análisis de regresión logística entre orden del sujeto y su animacidad .....   | 194 |
| Tabla 27. Análisis de regresión logística entre orden del sujeto y definitud del sujeto .....  | 195 |
| Tabla 28. Análisis de regresión logística entre orden del sujeto y morfología del sujeto ..  | 196 |
| Tabla 29. Análisis de regresión logística entre orden del sujeto y complejidad del sujeto  | 197 |
| Tabla 30. Análisis de regresión logística entre orden del sujeto y complejidad del sujeto  | 198 |
| Tabla 31. Análisis de regresión logística entre orden del sujeto y morfología del sujeto ..  | 199 |
| Tabla 32. Resultados de la regresión logística múltiple. Variables estadísticamente significativas .....   | 200 |

## Índice de gráficos

|  |     |
|--|-----|
| Gráfico 1. Distribución de datos en las macrozonas España y América (izquierda) y en las zonas dialectales (derecha). Elaboración: Mercedes Sánchez, coordinadora del CORPES XXI. Fuente: Internet.....  | 55  |
| Gráfico 2. Proporciones de interrogativas con anteposición y posposición del sujeto en el corpus.....  | 129 |
| Gráfico 3. Casos con orden SV y VS, agrupados por zona geográfica. Se presentan las zonas en orden decreciente tomando en cuenta la proporción de casos con orden VS .....   | 131 |
| Gráfico 4. Representación de mosaicos de las variables Orden del sujeto y Zona geográfica .....  | 132 |
| Gráfico 5. Representación de las proporciones de las categorías pertenecientes a la variable Semántica verbal en el corpus .....   | 133 |
| Gráfico 6. Casos con orden SV y VS, agrupados por semántica verbal. Se presentan las categorías en orden decreciente según su número total de casos. Las barras expresan la proporción de casos para cada orden .....  | 134 |
| Gráfico 7. Representación de mosaicos de las variables Orden del sujeto y Semántica verbal .....   | 136 |
| Gráfico 8. Proporción de casos por categoría de la variable tiempo verbal .....  | 137 |
| Gráfico 9. Casos con orden SV y VS, agrupados por tiempo verbal. Se presentan las categorías en orden decreciente según su número total de casos. Las barras expresan la proporción de casos para cada orden. Las categorías marginales (menos de 10 casos) se han agrupado en la etiqueta “Otros” ..... | 138 |
| Gráfico 10. Representación de mosaicos de las variables Orden del sujeto y Tiempo verbal .....   | 138 |
| Gráfico 11. Proporción de casos por categoría dentro de la variable Realización valencial .....  | 141 |
| Gráfico 12. Casos con orden SV y VS, agrupados por realización valencial. El eje Y muestra las frecuencias absolutas. Se presentan las categorías en orden decreciente según su número total de casos. Las barras expresan la proporción de casos para cada orden .....                                  | 142 |
| Gráfico 13. Representación de mosaicos de las variables Orden del sujeto y Realización valencial.....  | 143 |
| Gráfico 14. Proporción de casos por categoría dentro de la variable Transitividad.....   | 145 |
| Gráfico 15. Casos con orden SV y VS, agrupados por transitividad verbal. Se presentan las categorías en orden decreciente según su número total de casos. Las barras expresan la proporción de casos para cada orden .....   | 146 |
| Gráfico 16. Representación de mosaicos de las variables Orden del sujeto y Transitividad verbal .....  | 150 |
| Gráfico 17. Proporción de casos por categoría dentro de la variable Animacidad del sujeto .....  | 151 |

|   |     |
|---|-----|
| Gráfico 18. Casos con orden SV y VS, agrupados por animacidad. El eje Y muestra las frecuencias absolutas. Se presentan las categorías en orden decreciente según su número total de casos. Las barras expresan la proporción de casos para cada orden.....   | 152 |
| Gráfico 19. Representación de mosaicos de las variables Orden del sujeto y Animacidad del sujeto.....   | 154 |
| Gráfico 20. Representación de mosaicos de las variables Orden del sujeto y Animacidad del sujeto (con solo dos niveles: animado e inanimado) .....  | 155 |
| Gráfico 21. Proporción de casos por categoría dentro de la variable Animacidad del sujeto .....   | 156 |
| Gráfico 22. Casos con orden SV y VS, agrupados por definitud. El eje Y muestra las frecuencias absolutas. Se presentan las categorías en orden decreciente según su número total de casos. Las barras expresan la proporción de casos para cada orden.....  | 157 |
| Gráfico 23. Representación de mosaicos de las variables Orden del sujeto y Definitud ...  | 157 |
| Gráfico 24. Representación de mosaicos de las variables Orden del sujeto y Definitud (con dos niveles: definido e indefinido) .....   | 159 |
| Gráfico 25. Tendencias de las combinaciones entre los niveles de las variables definitud, animacidad y el orden del sujeto. Los números del eje vertical expresan las proporciones de casos en cada una de las agrupaciones representadas en el eje horizontal .....  | 161 |
| Gráfico 26. Proporción de casos por categoría dentro de la variable Morfología del sujeto .....   | 162 |
| Gráfico 27. Casos con orden SV y VS, agrupados por morfología del sujeto. Se presentan las categorías en orden decreciente según su número total de casos. Las barras expresan la proporción de casos para cada orden. Las categorías con casos marginales (menos de 10) se han agrupado en la etiqueta “Otros” .....   | 163 |
| Gráfico 28. Representación de mosaicos de la relación entre las variables Orden del sujeto y Morfología del sujeto. En el eje vertical se muestran (de abajo hacia arriba) los mosaicos correspondientes a: sustantivo, solo o con artículo (Sust.), sustantivo con singularizador (Sust.+Sing.), nombre propio o título (Nombre), Usted, pronombre posesivo o demostrativo (Dem.), infinitivo (Inf.), indefinido (Indef.), Tú, Ustedes, sujeto clausular (SC), pronombre numeral (Num.), nosotros (Nos.), Él o Ella, y Yo..... | 164 |
| Gráfico 29. Representación de mosaicos de las variables Orden del sujeto y Morfología del sujeto (se han reducido los niveles a tres: sujeto pronominal, infinitivo y sujeto clausular y sujeto léxico).....  | 166 |
| Gráfico 30. Representación de la distribución de frecuencias absolutas de sujetos con pronombre personal, según zona geográfica .....   | 169 |
| Gráfico 31. Proporción de casos por categoría dentro de la variable Complejidad del sujeto .....  | 170 |
| Gráfico 32. Casos con orden SV y VS, agrupados por complejidad del sujeto. El eje Y muestra las frecuencias absolutas. Se presentan las categorías en orden decreciente según su número total de casos. Las barras expresan la proporción de casos para cada  |     |

|   |     |
|---|-----|
| orden. Las categorías con casos marginales (menos de 10) se han agrupado en la etiqueta “Otros” .....   | 171 |
| Gráfico 33. Representación de mosaicos de las variables Orden del sujeto y Complejidad del sujeto.....  | 172 |
| Gráfico 34. Representación de mosaicos de las variables Orden del sujeto y Complejidad del sujeto (se han reducido los niveles a 3, según su grado de complejidad).....   | 174 |
| Gráfico 35. Frecuencias relativas de las 14 categorías de función pragmática incluidas en el análisis .....   | 175 |
| Gráfico 36. Frecuencias relativas para anteposición y posposición del sujeto, según tipos funcionales de Freed (1994) .....   | 176 |
| Gráfico 37. Frecuencias relativas para anteposición y posposición del sujeto, según tipos generales propuestos por Escandell (1999). Se ha mantenido en el gráfico el tipo Interrogativa relacional de Freed (1994) por no existir en la propuesta de la autora . | 180 |
| Gráfico 38. Gráfico de mosaico para función pragmática del enunciado interrogativo y Orden del sujeto. Se han agrupado las 14 categorías originales según los tipos más frecuentes en el corpus .....   | 183 |

## Índice de figuras

|   |     |
|---|-----|
| Figura 1. Representación del mecanismo de interacción de factores lingüísticos y extralingüísticos en la determinación del orden del sujeto en la interrogación total.... | 35  |
| Figura 2. Jerarquía de definitud en la determinación del tópico oracional.....  | 80  |
| Figura 3. Variables más significativas en la determinación del orden sujeto-verbo. Escala de influencia .....   | 201 |
| Figura 4. Representación de la influencia de factores lingüísticos en el orden sujeto-verbo .....   | 201 |

## **Agradecimientos**

Aquel marzo de 2015 llegué a Chile llena de ilusiones. Sabía que se abría para mí, con el doctorado que empezaría, el camino de crecimiento intelectual, profesional y personal que tanto estaba ansiando. Con el punto final de esta tesis, ese camino se cierra o, mejor, se convierte en la posibilidad de abrir otros caminos. Durante el “viaje” que ha constituido esta modesta investigación, han sido muchas las personas con quienes he podido contar y las instituciones que me han albergado. Quiero agradecer profundamente:

Al gobierno de Chile y a la PUC, que me facilitaron los recursos materiales necesarios para llevar a cabo esta investigación.

A mi tutor Carlos González, quien asesoró esta tesis desde su sabiduría como lingüista y gramático. Mil gracias por sus repetidas lecturas y su disposición para responder mis dudas, aún a distancia y en horarios encontrados.

A mi codirector Xose Padilla, de la Universidad de Alicante, España, con quien hice mi primera pasantía, un momento crucial en el primer arranque de mi investigación. Gracias por su atinada lectura y por el apoyo brindado en estos años.

A la profesora Teresa Oteíza, antigua jefa de doctorado de la Facultad de Letras de la UC, quien me acogió cariñosamente como alumna novata y en cierto sentido indefensa. Su talento y nobleza inspiraron mi generación.

A mis profesores del Doctorado en Lingüística de la Facultad de Letras, cuyo entusiasmo e ímpetu para estudiar y descubrir los “misterios” del lenguaje es envidiable.

A la profesora Nicole Delbecque, de la Universidad Católica de Lovaina, Bélgica, con quien tuve el honor de hacer mi segunda pasantía doctoral. Su humildad y su rigor académico enaltecen las ciencias del lenguaje. También a Bert Cornillie, quien facilitó mi paso por la KU Leuven.

A mis compañeros y colegas de generación, por formar parte de ese necesario apoyo y contención que uno requiere cuando emigra, por ayudarme a enfrentar el enorme desafío que me impuso el doctorado. También a los amigos que dejé en la isla y a los que fui haciendo en mis viajes académicos, porque con su cariño y confianza pude sortear más de un bajón emocional.

A Camilo Quezada y Omar Sabaj, por compartir generosamente conmigo sus tiempos y sus conocimientos, por ayudarme a hacer una tesis mejor que la que había concebido.

A Germán, compañero de doctorado y de vida, por ayudarme tantísimo a resolver pragmáticamente los problemas más críticos de mi investigación; por ser una inspiración constante, por tantos pretextos para reír cada día.

A mi hijo Renzo que, sin saberlo, dio el impulso final a este informe de tesis, en medio de la pandemia y de todas las crisis que vinieron con ella. Sin su alegría, habría sido todo más difícil y aburrido.

A mis familias chilena y cubana, en especial a mis padres, por las constantes de mis avances, gracias por la compañía y el amor incondicional.

Al tiempo y su indulgencia para conmigo, a la vida por permitirme llegar hasta aquí con cordura y salud razonables.

A las ciencias del lenguaje, por fin, por regalarme un motivo más para que mis días tengan sentido.

## Introducción

El orden de palabras en la oración es un tema fértil que ha ocupado un espacio importante en la producción científica desde hace décadas. El aporte teórico pionero fue el establecimiento de una tipología de lenguas, de acuerdo con este criterio. Se ha afirmado que el español es una lengua de tipo SVO, es decir, la secuencia más natural o común es la que antepone el sujeto al verbo en la oración declarativa, y la mayoría de las investigaciones más recientes con respecto a la interrogación se ha dedicado a explicar las condicionantes de un orden diferente (VSO), establecido como reglamentario (RAE y ASALE, 2009; Bosque y Gutiérrez Rexach, 2009; Toribio, 2000), en aquellas estructuras con pronombre o adverbio interrogativo, denominadas interrogativas parciales (*¿Qué plato ha cocinado tu abuela?*). Elementos interesantes en el debate científico son la existencia de un patrón de orden diferente en el área caribeña, que destaca por anteponer el sujeto (*¿Qué tú dices?*), así como la alta incidencia de la expresión de sujetos pronominales.

Con respecto a la secuenciación de constituyentes en la interrogación total se propone el principio teórico que establece un orden no marcado VSO (*¿Viajaron ellos a Chile?*). Por otra parte, se carece de estudios cuantitativos de corpus diversificados de habla real que den cuenta de la existencia del fenómeno y de la influencia de factores lingüísticos y extralingüísticos, así como de estudios comparativos entre la lengua hablada y la escrita. No obstante, una investigación previa realizada por la misma autora de esta tesis, sobre el orden en interrogativas totales de Cuba y España (Ramos, 2013), esbozó algunos resultados que permiten esperar variación regional y dialectal tanto en el español peninsular como en el latinoamericano.

La presente investigación se propone indagar sobre el orden más frecuente adoptado por el sujeto respecto del verbo en la interrogación total en los principales países de habla hispana, a partir de una recopilación de casos provenientes de textos escritos, recogidos en el Corpus del Español del Siglo XXI<sup>1</sup>. Así, se toman como precedente teórico tanto las gramáticas tradicionales como las investigaciones empíricas con que se cuenta hasta la fecha,

---

<sup>1</sup> <https://www.rae.es/banco-de-datos/corpes-xxi>

a partir de las cuales se ha identificado un grupo de factores lingüísticos potencialmente influyentes en la selección de una determinada secuencia de constituyentes.

Un número importante de las obras y artículos consultados muestra consenso en cuanto a la tendencia del español a colocar el sujeto después del verbo (orden verbo-sujeto) en las oraciones interrogativas totales o absolutas (Gili Gaya, 1958 [1943]; NGLÉ, 2009; Aijón & Serrano, 2012). Uno de los principales referentes teóricos es Escandell (1999), quien afirma que este orden es el único posible para enunciados interrogativos en los que el emisor pretende despejar una incógnita –sobre la que puede tener un grado variable de conocimiento– dado su carácter de juicio tético.

Sin embargo, como se ha mencionado, se ha constatado un orden diferente en el área caribeña. Además, nuestras observaciones del habla de variantes de lengua que no han sido exhaustivamente estudiadas con respecto al tema que nos ocupa, como es el caso de la chilena, nos han permitido identificar la existencia del mismo fenómeno de la anteposición del sujeto, o de una ordenación lineal del enunciado, de manera más o menos generalizada, tal como sucede en el Caribe. Con motivo de ello, nos ha parecido interesante investigar qué sucede en zonas que permanecen virtualmente inexploradas, e intentar esbozar generalidades con respecto a todo el entorno hispánico.

En contraposición con el panorama que exhibe el área de la fonética y la fonología, donde existen proyectos integradores, como es el Atlas Multimedia de la Prosodia del Espacio Románico (AMPER)<sup>2</sup>, no se conocen propuestas investigativas integradoras que brinden descripción de la interrogación total desde el punto de vista de la sintaxis, la semántica y la pragmática. Hasta donde tenemos conocimiento, no hay, tampoco, un estudio integral de corpus que dé cuenta del fenómeno de la anteposición o posposición del sujeto en la interrogación total.

El estudio que proponemos pretende llenar este vacío investigativo. Por tanto, se centrará en el orden de constituyentes, con especial énfasis en el orden que adopta el sujeto respecto del verbo, en las oraciones interrogativas denominadas totales, generales o absolutas

---

<sup>2</sup> Consúltese información sobre el proyecto en [http://stel.ub.edu/labfon/amper/cast/amperinternacional\\_presentacion.html](http://stel.ub.edu/labfon/amper/cast/amperinternacional_presentacion.html)

en la bibliografía. Estas estructuras son empleadas, fundamentalmente, para corroborar la correspondencia del contenido proposicional expresado en la oración con un estado de cosas del mundo, por lo que requieren típicamente una respuesta de “sí” o “no”.

Establecidos estos principios teóricos, el problema de investigación puede formularse de la siguiente manera:

¿Cuál es el orden que muestran los constituyentes sujeto-verbo en las interrogativas totales o absolutas del español?

En consonancia con Flores-Ferrán (2007), estimamos que, de cara a las variaciones dialectales que puede experimentar cualquier fenómeno lingüístico (el orden de constituyentes, en el caso que nos ocupa), los factores que condicionan el uso de una u otra forma plantean en sí mismos un problema lingüístico que es necesario investigar.

Partiendo de esta necesidad de exploración, el trabajo que se perfila es de carácter descriptivo y responde a un enfoque teórico funcional. Los resultados obtenidos serán novedosos en el área de los estudios gramaticales, pues permitirán conocer los patrones de orden de la interrogación total existentes en el entorno hispánico, cuáles factores lingüísticos favorecen la elección de los hablantes y cómo se distribuyen estos patrones a nivel dialectal. Además, el estudio será el primero en combinar gran parte de las variables lingüísticas presentes en las investigaciones conocidas hasta la actualidad, relacionadas mayormente con la interrogación parcial, y analizar cómo se comportan en la interrogación absoluta.

El informe de tesis doctoral está organizado en tres capítulos. En el Capítulo 1 se recogen los antecedentes teóricos y conceptuales que sirven de base para la justificación del estudio y para la selección de las variables; asimismo, se plantean las hipótesis y los objetivos de la investigación. En el Capítulo 2 se describe la metodología, tanto en lo referido a la conformación del corpus, como a la selección de herramientas de análisis cualitativo y cuantitativo (estadístico). En el Capítulo 3 se describen y discuten los resultados obtenidos. Por último, en las conclusiones, se retoman los objetivos y se resumen los principales resultados, las limitaciones y las proyecciones de la investigación.

## Capítulo 1. Marco teórico y conceptual

El presente capítulo recoge los principales elementos teóricos y conceptuales en torno al tema que nos ocupa. Este tiene la siguiente estructura interna: primeramente, se hace un breve barrido bibliográfico sobre las interrogativas parciales y totales; en segundo lugar, se recogen algunos elementos teóricos relacionados con el orden de palabras en general, y sobre la secuencia de constituyentes oracionales en interrogativas totales; en tercer lugar, se declaran la hipótesis de trabajo y las preguntas y objetivos de investigación. Por último, se incluye una reflexión teórica sobre las variables contenidas en el estudio.

### 1.1. Las interrogativas parciales y totales

La primera conceptualización de los enunciados interrogativos en la tradición gramatical hispánica, fuera del ámbito de la retórica, pertenece a la *Nueva Edición Reformada* de la *Gramática de la Real Academia Española* (1931), obra en la que se clasifica las oraciones interrogativas siguiendo la tradición de Andrés Bello (1984 [1847]). En ella se plantea que las estructuras usadas para interrogar pueden ser directas o indirectas<sup>3</sup>, según estén constituidas por una sola oración, o formen parte de otra, como sujetos o complementos. En el texto se consigna además que la oración interrogativa directa no es más que un estado mental intermedio entre la aseveración y la negación. El hablante expresa este estado mental mediante una duda que constituye la incógnita de la oración, que a su vez puede recaer sobre el predicado verbal —en cuyo caso no se precisa de pronombres interrogativos— o sobre el sujeto, el predicado nominal, y los complementos directo, indirecto y circunstancial.

La pregunta directa se usa para informarnos de algo que ignoramos (*¿Qué hora es?*), para expresar una duda (*¿Qué le habrán dicho, que tan enojada está con nosotros?*), para sugerir el signo contrario a lo que se propone en la pregunta que estamos haciendo (*¿De la pasada edad, qué me ha quedado?*), o bien para dar a conocer otros significados como

---

<sup>3</sup> Si bien sería interesante rastrear el comportamiento del fenómeno de orden del sujeto en interrogativas indirectas, no nos es posible incluirlas en la investigación, dado que el CORPES XXI no cuenta con esta categoría, o sea, no existe una manera de recuperar de este tipo de estructuras. Las interrogativas directas, en cambio, pueden ser recuperadas introduciendo los signos de puntuación (*¿?*)

extrañeza, admiración, repugnancia, horror; la interrogación es en este caso una “figura oratoria” (p. 325).

Desde una perspectiva psicológica, Lenz (1944 [1920]) sugiere que nos comunicamos por medio de la oración interrogativa cuando queremos formular una declaración, pero tenemos una duda; en ese caso “formulamos la pregunta para completar lo que nos falta en la declaración que proyectamos hacer” (p. 65). Este autor clasifica las oraciones interrogativas según se quiera “saber si todos los elementos expresados corresponden o no al análisis de la representación total” (p. 65), en cuyo caso la interrogación sería dubitativa o general, con respuestas de “sí” o “no”; o si se inquiere por el elemento que falta para formular la declaración, con lo cual estaríamos en presencia de una interrogativa parcial o determinativa, que expresa el elemento que falta por medio de una partícula interrogativa, usualmente un pronombre.

Por su parte, Gili Gaya (1958) [1943], señala que las oraciones interrogativas son usadas para dirigirnos a uno o varios oyentes con ánimo de resolver una duda. La marca fundamental es la inflexión final ascendente de la voz para dar a conocer que se espera una respuesta. El autor separa este tipo de oraciones en generales, que preguntan por la verdad o falsedad de un juicio, como en *¿Ha llegado tu padre?*; y parciales en las que “la duda no recae sobre el predicado de la oración, sino sobre el sujeto o sus cualidades, o sobre cualquiera de los demás elementos de la oración”, como en *¿Quién ha venido?* (p. 38). Estas últimas estructuras se caracterizan por presentar adverbios o pronombres interrogativos.

A medio camino entre la fonología y la pragmática, pero privilegiando la primera de las disciplinas, Navarro Tomás (1966) propone una clasificación de las oraciones interrogativas que combina elementos acústicos (entonación) con el propósito comunicativo del hablante. El autor acuñó el término de “absoluto” para hacer referencia al enunciado interrogativo que expresa una duda con respecto a la realidad, sobre la cual no se tiene ningún grado de conocimiento o presuposición y cuya incógnita abarca todo el contenido de la frase. Muchas de las investigaciones en el ámbito de la fonética utilizan esta etiqueta para hacer referencia a las interrogativas totales o generales de la tradición gramatical.

Otros tipos de interrogativas, según Navarro Tomás son:

-la relativa: se formula para cerciorarse de algo sobre lo que se tiene alguna idea.

Presenta, por tanto, algún grado de suposición, pero la incógnita afecta a toda la oración:  
*¿Podré yo elegir la ruta que quiera?*

-la restrictiva: la duda recae principalmente sobre uno de los elementos de la frase: *¿Es mañana cuando se marchan?*

-la aseverativa: la interrogación se inclina a la aseveración, de modo que la oración es interrogativa solo en su forma: *¿Verdad que fue muy bueno contigo?*

-la pronominal: la estructura interrogativa está encabezada por un pronombre interrogativo que concentra la duda del emisor.

-la reiterativa: con ella se repite algo que ha sido mencionado anteriormente, ya sea para aclarar algún elemento o para realzar la importancia de lo dicho, como en la secuencia *¿Están decididos a seguir adelante?, ¿Que si están decididos?*

-la exclamativa: en este tipo de estructuras existe un grado de suposición con respecto al estado de cosas sobre el cual se interroga, y se emplean para expresar sorpresa, admiración, entusiasmo o temor. Este tipo de estructura se distingue de las oraciones exclamativas propiamente dichas porque conservan algún grado de incertidumbre con respecto a lo dicho: *¿Pero aún están en la puerta esos muchachos?*

-la hipotética: la interrogación parte de una suposición, que espera ser confirmada por la respuesta: *¿Si será verdad lo del testamento?*

-la alternativa: como indica su nombre, existe una disyunción que condiciona la respuesta: *¿Quieres té o helado?*

Volviendo a las obras normativas del español, el *Esbozo de una nueva gramática de la lengua española* (RAE, 1973) hace alusión a los enunciados interrogativos como aquellos utilizados “cuando nos dirigimos a uno o varios oyentes para que resuelvan una duda o nos digan algo que ignoramos” (p. 359). En conformidad con un criterio semántico-lógico, esta obra clasifica las interrogativas en dos grupos: generales, cuando se pregunta sobre la verdad o falsedad de un juicio (*¿Recibió usted mi carta?*); y parciales, con las que el hablante no pregunta por el predicado, sino por el sujeto o cualquier otro elemento de la oración.

Una perspectiva de clasificación que se aparta de la tradición gramatical es la de García (1987). La autora parte del presupuesto de que la oración interrogativa plantea una duda o

incógnita del emisor que debe ser despejada por el receptor. Tomando como eje el análisis comunicativo, propone una clasificación de las oraciones interrogativas a partir del grado de desconocimiento de la incógnita y de la estructura de esta.

La investigadora diferencia los elementos de la estructura de la incógnita que están unidos por relaciones paradigmáticas de tipo temático, de los que forman parte de una oposición que se excluye mutuamente, pero se agrupa en relaciones antonímicas. Obsérvese los ejemplos siguientes: *¿Quién viene a la reunión de hoy?* y *¿Max viene a la reunión?* En el primer caso la incógnita puede ser despejada por elementos de un conjunto con iguales posibilidades de aparición: Juan, Pedro, Vivian; en el segundo, se pregunta por la realización o no de una acción: las respuestas posibles son “sí” o “no”. En consecuencia, la autora propone distinguir en el análisis el plano horizontal y el vertical de los ejes sintagmático y paradigmático, respectivamente. La diferencia entre ambos planos radica en que el vertical está integrado por los elementos desconocidos que forman la estructura de la incógnita, mientras que el plano horizontal está formado por elementos conocidos, más o menos compartidos por los interlocutores en el momento en que se formula la pregunta. Ambos planos no pueden divorciarse de la situación y el contexto (García 1987, p. 18).

Teniendo esto en cuenta, la autora propone para el español, básicamente, tres grupos de preguntas: el primero abarca las estructuras en las que todos los elementos son desconocidos en igual grado (*¿Qué hora es?*); en el segundo, uno de los miembros de la serie es más probable que otro (*¿Cuándo fue que llegó Pablo?*); en el tercero se trata de dilucidar si se ha realizado o no la acción, o si existe o no en la realidad objetiva, una cualidad o un estado; los miembros de la oposición de la incógnita son desconocidos en igual grado (*¿Paso por aquí mi marido?*).

Otra propuesta que difiere de la tradicional con respecto a la nomenclatura y el análisis efectuado en la clasificación es la de Rabanales (1996), quien sugiere dividir las preguntas en reales e irreales, según la intención comunicativa del emisor. En las preguntas reales, se espera recibir información en torno a una duda auténtica. También llamadas por el autor como indagativas o inquisitivas, estas estructuras funcionan como “auténticas preguntas”, como es el caso de *¿Habló contigo mi cuñado?* y *¿Quién te llamó por teléfono?* Las preguntas irreales conservan la forma de la interrogación, pero en ellas “no se indaga, sino

que se busca satisfacer otros fines de la comunicación (expresar un sentimiento, hacer una declaración, manifestar una duda, sorpresa, enojo, indignación; exhortar, ordenar, etc.), que no implican por lo mismo otro enunciado que funciona como respuesta” (p. 237). Es el caso de las preguntas retóricas o formuladas a uno mismo, como *¿Me estaré volviendo loco?*

Los elementos que permiten reconocer una interrogación indagativa, según el autor, son la entonación (en anticadencia de manera general), la presencia de formas interrogativas como pronombres o adverbios, y el contexto o la situación (en el caso de la comunicación oral) que permite esclarecer la “actitud indagativa” del emisor.

Desde el punto de vista gramatical, Rabanales establece una clasificación de acuerdo con la estructura oracional: a) la pregunta directa (que nos interesa especialmente en esta investigación), constituye ella misma una oración; b) la pregunta semidirecta depende de un verbo *dicendi*, como *preguntar*, *decir*, *contestar* (*Dime, ¿encontraría trabajo?*); y c) la pregunta indirecta, que es también subordinada, no contempla una pausa entre el verbo y la cláusula subordinada (*Me gustaría saber si vendrá esta noche*<sup>4</sup>).

Entre las preguntas indagativas el autor distingue las totales de las parciales. La diferencia fundamental entre ellas es el ámbito de acción de la incógnita que pretende despejarse. En las primeras abarca todo el enunciado interrogativo, “porque todo lo que en él se dice es ignorado por el emisor” (p. 246), en tanto en las segundas lo ignorado se restringe solo a una parte del enunciado.

A su vez, las totales y parciales pueden ser presuntivas o no presuntivas. Las totales no presuntivas son aquellas con las cuales el emisor demanda información de la cual no presume nada (*¿Recibió usted mi carta?*). Las presuntivas, por su parte, contienen algún grado de suposición con respecto a lo que se pregunta, y poseen marcas como las expresiones “acaso” o “es que” (*¿Es que no viven tus padres?*). Las parciales no presuntivas concentran la duda o el desconocimiento en la partícula interrogativa, y representan un desconocimiento real, es decir, no se presume nada sobre el elemento desconocido (*¿Cuándo estarás listo?*). En cambio, en las presuntivas la duda recae en una sola “parte” del enunciado, que se distingue por algún tipo de énfasis y que coincide normalmente un argumento que el verbo selecciona

---

<sup>4</sup> Ejemplo tomado de Serra (s/f).

de manera más lejana, como es el caso de *¿Estarás lista EN UNA HORA?* o *¿El perro entró POR LA COCINA?* Según aclara el autor, “lo que distingue básicamente las totales de las parciales es el grado de ignorancia que ponen de manifiesto de parte del que pregunta con respecto a lo que se desea saber (...) y a las presuntivas de las no presuntivas, el revelar si esta misma persona no tiene, respectivamente, alguna presunción con respecto a lo mismo” (p. 248).

Tanto en lo que respecta a la sintaxis, como a la semántica y la pragmática, Escandell (1999) hace una propuesta descriptiva de los enunciados interrogativos que nos parece completa, si bien creemos que algunas de sus afirmaciones pueden complementarse a partir de la implementación de investigaciones empíricas o con análisis de muestras de habla real. La autora establece, primeramente, una distinción entre los conceptos de *pregunta* e *interrogación*, y afirma que debe establecerse una diferenciación en la caracterización de una estructura oracional a partir de lo formal y lo semántico, y no teniendo en cuenta los objetivos comunicativos del hablante (p. 3931). “Por lo tanto, dentro del terreno de la reflexión gramatical no es conveniente identificar ambas nociones ya que, al hacerlo, se están identificando equivocadamente estructura gramatical y propósito discursivo” (1999, p. 3932). Así, reserva el término *pregunta* para aquellos enunciados interrogativos que pretenden obtener información y el de *oración interrogativa* para referirse a aspectos gramaticales y semánticos.

En consecuencia, la autora precisa que la propiedad significativa fundamental de una oración interrogativa es la de contener una incógnita o una variable. En este sentido, estos enunciados son estructuras abiertas, que requieren de un elemento para cerrar el contenido proposicional: la respuesta. En los enunciados interrogativos parciales este elemento responde a la incógnita introducida por el pronombre interrogativo; en las totales es “sí” o “no”.

Según la autora, el operador interrogativo o foco se concentra en los pronombres interrogativos en las interrogativas parciales, elementos que reciben además una prominencia prosódica; y abarca toda la oración en las interrogativas totales. Compárense estos ejemplos (el foco se marca en versalitas):

- a)     ¿*QUIÉN* ha roto la ventana?
- b)     ¿*HA LLEGADO ANDRÉS*?

Escandell propone distinguir, desde el punto de vista gramatical, las “interrogativas neutras” de las “orientadas”. La diferencia entre unas y otras es que las segundas están afectadas por alguna de las marcas formales que, según la autora, restringen la interpretación, a saber: poseen esquemas entonativos que indican atribución del contenido proposicional transmitido, presentan un orden de palabras marcado (SV en el caso de las interrogativas totales), y cuentan con términos de polaridad negativa, así como con partículas introductoras.

El primer tipo de interrogativa neutra lo constituyen las preguntas directas (están formadas por una sola oración) mediante las cuales se pide información. Ellas, a su vez, se subdividen en preguntas reales (la presencia de la variable interrogativa responde directa y efectivamente a un desconocimiento real por parte del emisor, que no sabe cuál es el elemento que debe remplazar a la incógnita que plantea); preguntas de examen (el desconocimiento del emisor con respecto a la incógnita es bajo, y formula la pregunta para saber si el interlocutor conoce o no la solución), como en *¿En qué año llegó el hombre a la Luna?* (el profesor al alumno); interrogativas problemáticas (se parte del supuesto de que nadie puede dar una respuesta concreta a la interrogación que el emisor plantea), como en *¿Cómo acabar con la guerra en el mundo?*; y deliberativas (son aquellas en las que el emisor se dirige a sí mismo), como en *¿Dónde he puesto las llaves?*

Otra clase de enunciado interrogativo neutro son las interrogativas directivas, utilizadas para formular peticiones, principalmente asociadas al grado de cortesía del enunciado (*¿Me trae la cuenta?*, *¿Podrías abrir la ventana?*).

Las interrogativas marcadas<sup>5</sup> u orientadas, que contienen al menos uno de los indicadores gramaticales de orientación interpretativa, se diferencian además por las condiciones pragmáticas en que son emitidas, es decir, han de ser tomados en cuenta la posición discursiva, las intenciones y los objetivos del emisor. Así, se distinguen:

a) las atribuidas: interrogativas por medio de las cuales el emisor hace oír palabras, efectiva o supuestamente pronunciadas, por otro hablante como en *¿Conque María se empeña en ir?*

---

<sup>5</sup> Para consultar los ejemplos por cada uno de los subtipos, consúltese el epígrafe 61.5 del capítulo pp. 3978-3987.

Dentro de este subtipo se encuentran las repetitivas o interrogativas de eco, estructuras que son usadas para repetir en su totalidad o en parte palabras que acaban de ser emitidas por otro hablante. Ellas, a su vez, se dividen en recapitulativas (su finalidad es expresar la actitud del emisor ante las palabras que acaba de emitir su interlocutor); especificativas (emitidas con la intención de obtener precisiones sobre una parte del enunciado que ha sido vaga o inespecífica); explicativas (solicitan, bien la explicación de una unidad cuyo significado se desconoce, bien la repetición de un término que no se ha escuchado correctamente); y las de copia (el hablante, a instancias de su interlocutor, repite un enunciado que él mismo había emitido con anterioridad).

También dentro de las atribuidas se encuentran las anticipativas, con las cuales el emisor se adelanta, de manera real o ficticia, a una posible intervención de su interlocutor, como en *¿Que qué hago yo aquí vestido de esta manera y a estas horas de la noche?* Son parte de esta clasificación las hipotéticas o exploratorias, con las que el emisor avanza una suposición sobre un estado de cosas; las interpretativas, que se producen como resultado de una interrupción, por parte del emisor, del discurso del otro; y las exclamativas, que hacen referencia a un hecho que es manifiesto para ambos interlocutores.

b) las retóricas: interrogativas que contienen algunos de los marcadores relacionados con la negación (negación externa, términos de polaridad o partículas de inversión argumentativa) por medio de las cuales el emisor comunica que no es totalmente neutral con respecto al contenido proposicional de su enunciado, sino que favorece explícitamente una determinada opción, a saber, aquella que presenta el signo contrario al que aparece en su enunciado.

Como última referencia citaremos la *Nueva gramática de la lengua española*, la más reciente de la Real Academia Española y la Asociación de Academias de la Lengua Española (RAE y ASALE, 2009), obra que dedica el capítulo 42 a “La modalidad. Los actos de habla. Construcciones imperativas, interrogativas y exclamativas” (pp. 3113-3222).

En la obra, se dividen las oraciones interrogativas en dos grupos: las directas y las indirectas. De acuerdo con la clasificación tradicional, las interrogativas directas se clasifican en totales o disyuntivas y parciales o pronominales. Las totales tienen a su vez dos subclases, según las opciones entre las que el oyente debe elegir a la hora de brindar una respuesta: de

“sí” o “no” (o polares), como en *¿Tienes frío?*; y alternativas, en las cuales se elige entre dos o más opciones, como en *¿Prefieres este libro o el otro?* En las interrogativas parciales los pronombres, determinantes, y adverbios interrogativos proporcionan la incógnita que debe ser satisfecha en la respuesta.

Las interrogativas totales y parciales son también llamadas cerradas y abiertas respectivamente, teniendo en cuenta que ponen de manifiesto las segundas (en oposición a las primeras) la existencia de una serie de opciones no expresas, restringidas por el tipo de interrogativo empleado.

Como variantes de oraciones interrogativas totales, la NGLE cita las estructuras introducidas por *si* (registradas en la lengua antigua), las formadas por un solo grupo adjetival o nominal (*¿Cansado?*) y las constituidas por “elementos infraoracionales” (*¿Verdad o fantasía?*).

A modo de síntesis, las gramáticas y autores que se ajustan a la tradición coinciden en afirmar que la oración interrogativa parte de la existencia de una duda que puede ser expresada por medio de dos tipos fundamentales de construcción: la interrogativa total y la parcial, cuyas características difieren a nivel semántico, sintáctico y prosódico. Además, los autores de estas obras hacen referencia a la presunción o presuposición por parte del hablante a la hora de formular el enunciado interrogativo, lo cual ofrece un elemento importante para la caracterización de estas estructuras a nivel comunicativo. A la par, se reconoce el empleo de la interrogación con otros propósitos, diferentes al de preguntar, como es el caso de la expresión de la sorpresa, la crítica o reproche, el avance de una hipótesis, etc.

A partir de los referentes conceptuales anteriormente reseñados, entenderemos como enunciado interrogativo total o absoluto aquel cuya marca gramatical fundamental es la ausencia pronombre o adverbio interrogativo. Este tipo de enunciado busca, básicamente, corroborar la correspondencia del contenido proposicional expresado en la oración con un estado de cosas del mundo, por lo que requiere típicamente una respuesta de “sí” o “no”. En la lengua oral, este tipo de interrogación puede distinguirse por poseer una entonación característica (ascendente en el español estándar, circunfleja en algunas variedades

caribeñas<sup>6</sup>); en la escritura es marcada ortográficamente por signos de interrogación (¿?). Asimismo, no existe consenso con respecto al orden de palabras por el cual se formaliza la interrogación total española: mientras la tradición gramatical establece una secuencia no marcada de verbo-sujeto (con excepción de algunos autores que admiten un orden más flexible), en investigaciones recientes se ha establecido que para este tipo de enunciados la inversión del orden de constituyentes es opcional, es decir, existiría libertad de orden.

## 1.2. El orden de palabras

### 1.2.1. Sobre los conceptos de orden básico y no marcado

Es importante para una investigación como la que nos ocupa partir del concepto de orden o patrón básico. Como se sabe, los estudios tipológicos acuñaron este término para referirse a las cláusulas declarativas, afirmativas, independientes, con dos constituyentes nominales, con sujeto definido, agentivo y humano y objeto paciente y definido; y con un verbo que designa una acción (Siewierska, 1988; Suñer, 1982; López Meirama, 1997; Rivas, 2008).

Así, Greenberg (1963), pionero en este tema, propone que el esquema básico para el español es el de sujeto-verbo-objeto (SVO)<sup>7</sup>. La investigación lingüística ha corroborado que este es, efectivamente, el orden estadísticamente favorecido (Bentivoglio & Weber, 1986; López Meirama, 1997; Delbecque, 2011). Se habla, entonces, de orden no marcado, entendido como “el que aparece por defecto, es decir, el más neutral, en el sentido de no estar sujeto a condiciones contextuales específicas y de ser compatible con el mayor número de contextos de uso posibles” (Leonetti, 2014, p. 3). Partiendo de esta base, otro tipo de secuenciación constituiría un fenómeno de “inversión” o “dislocación” (Leal, 2005; Padilla,

---

<sup>6</sup> No es objeto de la presente investigación estudiar las variaciones sufridas en la entonación de este tipo de enunciado. Existen múltiples investigaciones sobre el tema. Para profundizar, consúltense Escandell (1999) y Navarro Tomás (1966).

<sup>7</sup> Existen, sin embargo, contrapropuestas. Tal es el caso de Contreras (1991, citado por Liceras, 1994), quien propone que el orden básico para el español es el de VOS. “La llamada inversión libre no es sino la consecuencia de que el sujeto pueda recibir caso *in situ* (a la derecha o a la izquierda del predicado)” (p. 232).

2001, 2005; Villalba, 2019), y los sujetos concernidos por este movimiento han sido clasificados como ‘sujetos no canónicos’ (Leonetti, 2018). Es por ello que, tanto en obras normativas como en la literatura científica, es común reconocer la expresión de ‘inversión del sujeto’ cuando se hace referencia a la posposición de este constituyente respecto del verbo en enunciados interrogativos. Sobre esta base, y asumiendo que existe un orden generalizado diferente en la oración interrogativa, las estructuras que son el foco de análisis en esta investigación ostentarían, por oposición al orden básico, lo que se ha llamado un ‘orden marcado’. Hay, sin embargo, suficientes razones para pensar que no todas las inversiones o posposiciones del sujeto pertenecen a un orden marcado, y es el caso de las estructuras gramaticales de interrogativas parciales o la focalización contrastiva (Leonetti, 2018), donde el orden reglamentario o no marcado es justamente el que tiene el sujeto pospuesto (VSO).

Entre las lenguas romances, es la hispánica la que mayor flexibilidad posee con respecto al orden de palabras. En la escritura, existe una tendencia a la ordenación lógica, mientras que en la oralidad se tiende a la tematización o topicalización (Hernando, 2005). Asimismo, la suficiencia de las desinencias verbales para la determinación del sujeto de una oración trae como consecuencia la abundante presencia de sujetos implícitos. Entonces, los conceptos de orden básico y no marcado adquieren otro matiz cuando se pasa del nivel teórico a la realidad lingüística, pues puede haber una discrepancia con el orden más frecuente<sup>8</sup>. Esta discrepancia es muy evidente en lenguas como el español.

Un análisis de la variable presencia/ausencia de sujeto gramatical probablemente llevará a corroborar la hipótesis de que el orden más frecuente en español es el de (S)VO<sup>9</sup>. Por ejemplo, Padilla (2005) demostró que, si bien esta secuencia es la más común en la conversación, donde el contexto situacional condiciona especialmente el uso de sujetos tácitos, también lo es para otros tipos de intercambio lingüístico, vale decir textos escritos, pues esta tendencia se conserva también en el caso de sujetos referidos a personas no participantes en el diálogo y a entidades carentes del rasgo +humano.

---

<sup>8</sup> De hecho, varios autores cuestionan la existencia de un orden básico determinado por la sintaxis. López Meirama (2006) menciona la prevalencia de factores funcionales (como la ordenación tema-remata) en la secuenciación de constituyentes.

<sup>9</sup> Se ha ofrecido evidencia tanto empírica como teórica (Bolinger, 1954).

En el caso de nuestra investigación, el corpus con el que trabajamos y los recursos tecnológicos que teníamos a mano impidieron un análisis exhaustivo de la frecuencia de oraciones con sujeto implícito. En primer lugar, el CORPES XXI<sup>10</sup> es una herramienta diseñada fundamentalmente para investigaciones de carácter lexicológico; por tanto, no existe un recurso para el etiquetado de funciones sintácticas, con lo cual es virtualmente imposible hacer un rastreo de estructuras sintácticas. A pesar de este inconveniente, mediante un análisis manual, pudimos observar, de manera muy impresionista, el comportamiento de la variable presencia/ausencia del sujeto en el corpus escrito haciendo un sondeo del fenómeno en el filtrado de los casos objeto de análisis (100 casos por zona, según quedará explicitado en el capítulo metodológico). La Tabla 1 muestra la proporción de oraciones con sujeto implícito y con pronombre interrogativo, presentes en el grupo de casos filtrados para alcanzar el centenar de casos útiles por zona. Si se asume que en la interrogación parcial el sujeto implícito es la regla (ver Barbosa, 2001 y Rosemeyer, 2018), puede comprobarse que, efectivamente, el fenómeno de la expresión lexical del sujeto es un fenómeno minoritario, pues el porcentaje de sujetos tácitos puede alcanzar, al menos teóricamente, el 50% de la muestra, como sucede en las zonas Andina y España.

---

<sup>10</sup> Las razones que nos llevaron a seleccionar este corpus están expresadas en el capítulo metodológico.

| <i>Zona dialectal</i>   | <i>Número total de casos*</i> | <i>Nro. Casos filtrados**</i> | <i>IntP</i> |    | <i>SujI</i> |    | <i>Total %</i> |
|-------------------------|-------------------------------|-------------------------------|-------------|----|-------------|----|----------------|
|                         |                               |                               | F           | %  | F           | %  |                |
| <b>Antillas</b>         | 4083                          | 335                           | 52          | 16 | 94          | 28 | 44             |
| <b>Andina</b>           | 4612                          | 373                           | 51          | 14 | 137         | 37 | 50             |
| <b>Caribe cont.</b>     | 7673                          | 331                           | 59          | 18 | 81          | 24 | 42             |
| <b>Chilena</b>          | 4220                          | 407                           | 60          | 15 | 135         | 33 | 48             |
| <b>España</b>           | 24651                         | 415                           | 64          | 15 | 141         | 34 | 49             |
| <b>Méx. y Centroam.</b> | 11078                         | 370                           | 58          | 16 | 106         | 29 | 44             |
| <b>Río de la Plata</b>  | 12172                         | 364                           | 44          | 12 | 125         | 34 | 46             |

Tabla 1. Frecuencias absolutas (F) y relativas (%) de interrogativas parciales (IntP) e interrogativas totales con sujeto implícito (SujI), respecto del total de casos revisados por zona

\*Total de casos obtenidos después de un primer filtrado automático con criterios definidos en capítulo metodológico, y a partir de los cuales se realizó la búsqueda de casos útiles.

\*\*Se trata de la cantidad de casos del total por zona que fue necesario revisar para obtener 100 casos susceptibles de ser descritos para alcanzar los objetivos de la investigación.

El estudio de los factores que condicionan la aparición del sujeto en las interrogativas que integran el corpus recogido para esta investigación no fue contemplado como objetivo del proyecto. Se sabe que la expresión de este constituyente es un fenómeno que puede estar condicionado por elementos de diversa naturaleza: establecimiento del tópico oracional o de una unidad discursiva; expresión de información focal; clarificación del referente del sujeto; ambigüedad de la forma verbal; correferencialidad con el sujeto precedente; e incluso naturaleza lexical del verbo (Silva-Corvalán, 2001, p. 173). Por otra parte, “la expresión del sujeto (...) tiende a resaltar su agentividad, y con ella toda una variada gama de significados más particulares que se ponen de manifiesto en la interacción” (Aijón y Serrano, 2012, p. 135), de los cuales no daremos cuenta aquí.

Por el momento, haremos abstracción de estos factores para concentrarnos en lo que nos interesa: conocer cómo se manifiesta el orden del sujeto respecto del verbo en las interrogativas totales, dadas ciertas condicionantes lingüísticas y contextos sociales de uso (zona geográfica).

Volviendo al tema de la inversión de constituyentes, si se revisan las referencias históricas al orden de palabras en el español, se verá que la tradición gramatical contemplaba el fenómeno de la inversión del sujeto como una cuestión estilística, perteneciente al ámbito

de la retórica (Leal, 2005; Liceras, 1994; Hernando, 2005). Las gramáticas normativas más modernas comenzaron a ofrecer explicaciones de naturaleza lingüística al fenómeno, pero se observa una multitud de propuestas que otorgan relativa importancia a los factores morfosintácticos, semánticos y pragmáticos. Otros análisis han venido desde diversas teorías lingüísticas, la funcionalista y la generativa entre las más destacadas.

Entre los elementos que se han destacado como influyentes en la determinación del orden de constituyentes oracionales, desde el punto de vista gramatical, están el tipo de construcción, las propiedades estructurales del sujeto, el tipo de verbo, la longitud de los sintagmas, la estructura informativa, entre otros (Fernández, 1993; Liceras, 1994; Leonetti, 2014).

Es conocida la propuesta de la perspectiva funcional de la oración, preconizada por los autores del Círculo Lingüístico de Praga, según la cual la secuenciación de elementos oracionales se debe a necesidades informativas, situándose en primer lugar los elementos menos informativos y más prominentes en el discurso, y relegando a la parte final del enunciado los elementos más informativos o nuevos. También el orden de palabras se ha explicado desde la pragmática. Así, se ha llamado la atención sobre la frecuencia con que los usuarios de la lengua alteran el orden básico antes mencionado para lograr ciertos efectos, sobre todo en el coloquio (Padilla, 2005). En estos casos “el orden de las palabras responde a la función pragmática de la topicalización y al realce informativo, a veces meramente subjetivo, de los elementos” (Hernando, 2005, p. 168). A propósito de este particular, se ha afirmado que son los elementos pragmáticos los que determinan, en primera instancia, el orden de un enunciado.

La presente investigación se concentrará en los factores sintácticos, semánticos, pragmáticos y dialectales que condicionan el orden de palabras en la interrogación total española. Es importante aclarar que, si bien hacemos referencia a otros componentes oracionales, nuestro foco es el orden que adopta el sujeto respecto del verbo, y en esta variación se concentrarán nuestros esfuerzos investigativos. Partimos de la base de que el sistema de la lengua impone ciertas restricciones al hablante, y que también proporciona un margen de libertad dentro del cual las construcciones emitidas por ellos presentan tendencias

claras, susceptibles de ser descritas; determinadas por una combinación de factores que pueden ser más o menos preponderantes.

Como resulta intuitivo, y según será visto en su momento, el orden que adopta el sujeto respecto del verbo en la interrogación total en español exhibe una variación, de la cual intentaremos dar cuenta, aun cuando reconocemos, al igual que Fernández (1993), que el estudio del orden en una lengua muy flexible como es el español “implica aventurarse en la tarea de desenmarañar un intrincado complejo de factores y ello, además, una vez que los tengamos localizados” (p.113). En este sentido, debe ser aclarado que no es nuestra intención agotar todos los factores que pudieran estar convergiendo en el fenómeno de ordenación de constituyentes en la interrogación –tarea que sería extenuante y en alguna medida imposible, por su complejidad–; más bien, se trata de establecer un diálogo con la literatura previa, donde se han identificado algunos elementos, tanto lingüísticos como extralingüísticos, que son potencialmente influyentes en el orden (en muchos casos, la investigación empírica así lo ha probado).

Por último, retomando las nociones de orden marcado y no marcado, intentaremos arrojar luz sobre la afirmación de que el orden no marcado para la interrogación total es el de VS. En consecuencia, partiremos por evitar usar la expresión “inversión del sujeto”, pues responde a una visión reduccionista del orden de palabras. Preferimos hablar en términos de alternancia sujeto-verbo. Coincidimos con Delbecque (1988), cuando afirma que “en el caso de la posición del sujeto, la variación es demasiado masiva para que uno se limite a llamarla libre sin más, y al mismo tiempo demasiado recalcitrante para poder ajustarla a reglas que la condicionarían al cien por ciento” (p. 49)

### **1.2.2. El enfoque variacionista**

El variacionismo, como afirma Martín Butragueño (1994) “es la teoría o perspectiva que defiende que el argumento cuantitativo tiene alguna importancia en el estudio del lenguaje” (p. 33). Los estudios de variación sintáctica encuentran su fundamentación en que “el comportamiento lingüístico, como otros comportamientos humanos, puede estar sujeto a variación estadística, de forma que la sintaxis podría estudiarse de acuerdo con patrones de

uso subyacentes que se computen probabilísticamente” (Sankoff, 1973, citado por Serrano, 2005, p. 156).

A partir de las aportaciones de Lavandera (1975), el método variacionista, originalmente concebido para dar cuenta de la variación fonológica, fue aplicado a fenómenos sintácticos. Después de superar varios escollos metodológicos (ver Serrano (2007) para más detalle), los estudios de variación sintáctica experimentaron un impulso considerable. Respecto del orden de palabras en español, destacan los trabajos sobre la posición del sujeto en la declaración de Silva-Corvalán (1982), Bentivoglio (1986 y 2003) y Delbecque (1988).

En el caso de la presente investigación, se estima que el orden del sujeto es un fenómeno variable, que coocurre con otros factores de naturaleza lingüística y extralingüística. Si bien en el ámbito de la morfosintaxis no es preciso, como se ha apuntado en repetidas ocasiones, hablar de alternancia de orden como dos formas para expresar lo mismo (Silva Corvalán, 2001; Serrano 2007 y 2009), sí es factible -o por lo menos razonable- hablar de variantes que pueden estar en distribución complementaria débil, es decir, que ellas pueden ser empleadas por los hablantes sin que se establezca una distinción funcional.

Hemos dejado fuera de nuestro análisis la influencia de factores sociodemográficos, pues ha sido demostrado en varias ocasiones que la variación gramatical se debe en gran medida a factores internos (lingüísticos) y no externos, como la edad, el sexo, la etnia o el nivel socioeconómico (Silva-Corvalán, 2001; Lizárraga & Mora-Bustos, 2010). Nuestro estudio incluye, sin embargo, una variable extralingüística, zona geográfica, partiendo de las referencias encontradas respecto de la variación dialectal que experimenta el orden de palabras en el ámbito de habla hispana.

El enfoque variacionista parece el óptimo para el estudio del fenómeno del orden de palabras en un corpus de habla real (escrita en este caso), como el que se propone. El empleo de esta perspectiva responde a una concepción funcional de la lengua, según la cual el cambio y la variación son fenómenos siempre presentes, las categorías poseen márgenes en ocasiones difusos y las estructuras sintácticas son flexibles (Givón, 1995). Los hechos lingüísticos, por demás, son el resultado de las relaciones entre los elementos sintácticos, semánticos y

pragmáticos (Pinuer, 2007), que se ponen de manifiesto en el uso. Asumiendo estos principios, el fenómeno que nos ocupa podría ser representado de la siguiente manera:

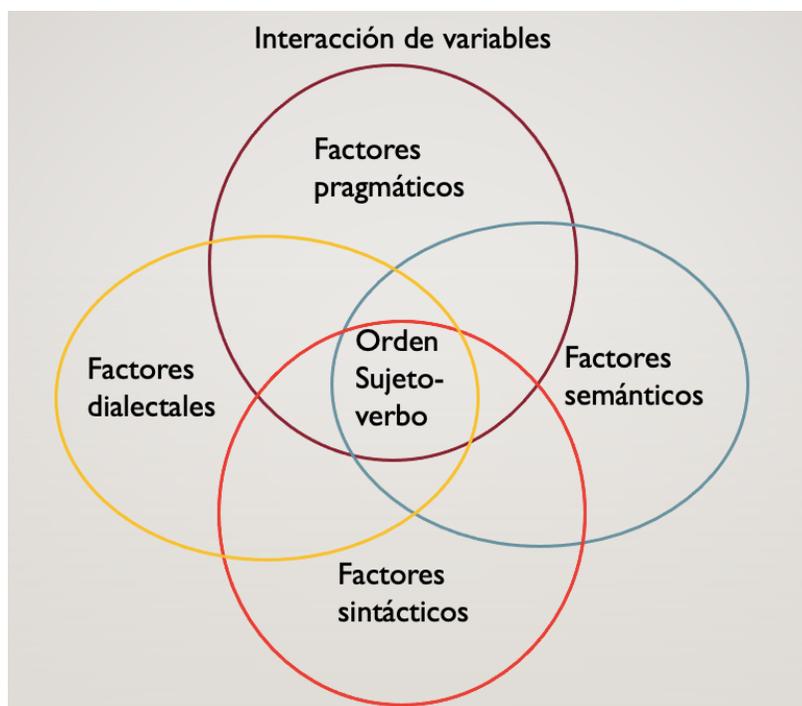


Figura 1. Representación del mecanismo de interacción de factores lingüísticos y extralingüísticos en la determinación del orden del sujeto en la interrogación total.

### 1.2.3. El orden de palabras en enunciados interrogativos totales en la tradición gramatical

El orden de palabras en oraciones interrogativas totales es un tema al que no se le ha dedicado mucho espacio en las obras normativas del hispanismo. En los referentes más antiguos, como es el caso de la *Gramática española* de Nebrija (1492), la *Gramática de la lengua castellana destinada al uso de los americanos*, de Andrés Bello (1847) y las obras de la Real Academia Española anteriores al *Esbozo de una nueva gramática de la lengua española*, de 1973, nada se dice al respecto, si bien se conceptualiza la oración interrogativa desde diferentes puntos de vista.

Seco (1978 [1930]), en su *Manual de gramática española*, hace una referencia breve al orden de palabras en oraciones interrogativas, al afirmar que esta “no se diferencia de la afirmativa, sino en signos interrogantes en la escritura y en la entonación especial del que habla, aunque por lo general la interrogativa lleva delante el predicado, lo cual no ocurre muchas veces en la afirmativa” (p. 212).

Rodolfo Lenz, en *La oración y sus partes* (1944 [1920]), afirma que uno de los recursos de la lengua española para indicar la presencia de una interrogativa es “poner la palabra acentuada al comienzo, y como tal se considera el verbo, pues con su afirmación o negación se afirma o se niega toda la oración” (p. 67). Según el autor, las declarativas pronunciadas como interrogativas (es decir, con la entonación propia de estas últimas) contribuyen a atenuar lo que llama “carácter perentorio de la pregunta dubitativa”. Es así que una oración como *¿Tu hermano vendrá?* no sería expresión de una duda real.

Por su parte Gili Gaya, en *Curso superior de sintaxis española* (1958 [1943]), hace referencia a la anteposición del verbo, como recurso para indicar el interés del hablante o emisor en ese elemento en particular. El autor no descarta el mantenimiento del orden lineal, pero solo en casos en que el sujeto se desgaja en un grupo fónico independiente, como ocurre en *El criado, ¿ha traído la carta?* Esta tendencia es más visible, según el gramático, cuando el sujeto tiene determinantes o modificadores y se torna largo. Por ejemplo: *Las personas más cultas de la ciudad, ¿podrán soportar tanta chabacanería?*

El autor menciona la relativa libertad que tiene el español para establecer el orden de los elementos oracionales, si se le compara con otras lenguas, como el francés, pero añade que “es indudable que [la inversión] da a la oración el carácter sintético que corresponde a la pregunta general” (p. 46).

En el *Esbozo de una nueva gramática de la lengua española* (RAE, 1977[1973]) los autores toman partido por la libertad de orden: el verbo de las interrogativas generales o totales puede anteponerse o posponerse al sujeto y a los demás elementos, con la misma libertad de construcción que las enunciativas. Interviene en el orden dado a los elementos oracionales la intención del hablante de destacar alguna información que se estima más importante o “expresiva” (p. 359). Se afirma que “cualquier oración enunciativa puede convertirse en interrogativa general sin necesidad de alterar su estructura sintáctica (...) Esta

libertad de construcción se debe a que la curva melódica interrogativa adquiere en español su carácter peculiar desde el comienzo de la pregunta, y no solo al final, como sucede en otras lenguas” (p. 396).

El capítulo 61 de la *Gramática descriptiva de la lengua española* (1999) contiene un análisis semántico y pragmático de los enunciados interrogativos, a cargo de María Victoria Escandell. La autora descarta la existencia un orden libre para las estructuras interrogativas gracias a la intervención de la entonación como elemento suficiente para establecer el contraste con la declaración: “ni las estructuras sin inversión son intercambiables en todos los contextos, ni las propiedades sintácticas de unas y otras son totalmente coincidentes” (Escandell 1999, p. 3952).

En consecuencia, la investigadora propone para la interrogación española un orden no marcado verbo-sujeto, en oposición a un orden marcado de sujeto-verbo. La razón de esta propuesta proviene desde la semántica lógica: la anteposición del verbo en oraciones como *¿Ha decidido ya Juan cuándo va a convocar la reunión?* permite interpretar el contenido preposicional como juicio tético, es decir, no habría ningún contenido presupuesto, y toda la información se plantearía como nueva.

El orden marcado o de sujeto-verbo favorece, según la autora, la interpretación de la interrogación como juicio categórico (Casado Velarde, 1993). En otras palabras, la oración pasa a tener una estructura bimembre y el sujeto se tematiza, con lo cual es contenido presupuesto; el predicado en estos casos es remático y constituye la información nueva. Escandell asume una posición radical, afirmando que “esta organización informativa, que asevera la relación entre sujeto y predicado, es por definición incompatible con la modalidad interrogativa, que se caracteriza precisamente por dejar en suspenso el signo correspondiente a dicha relación” (p. 3952).

Además, la autora propone que estructuras como *¿El ministro ha dimitido?* no tienen una variable, ya que preexiste una proposición completa. Por tanto, la interrogación actúa sobre una proposición cerrada, desacorde a lo que, desde la lógica, se entiende como interrogación, a saber, una proposición abierta. Las estructuras con orden sujeto-verbo, según esta propuesta, son un recurso discursivo para indicar la atribución del contenido

proposicional a otra persona o la existencia de una declaración previa<sup>11</sup>. Así, las interrogativas sin inversión no serían neutras.

Otro elemento que la autora utiliza para dar evidencia de la inequivalencia de las estructuras con y sin inversión es la presencia de una locución de polaridad negativa como ‘el más mínimo esfuerzo’ en los siguientes ejemplos:

a. *¿Ha hecho Juan el más mínimo esfuerzo por ayudarme?*

b. *\*¿Juan ha hecho el más mínimo esfuerzo por ayudarme?*

Según la propuesta de Escandell, la segunda secuencia sin inversión es agramatical porque, en tanto parte de una metaproposición, rechaza en principio la presencia de elementos de polaridad negativa. Solamente la interrogación legitima este tipo de elementos. Siguiendo este presupuesto, la secuencia *b* sería impropia en español.

Por último, en la *Nueva gramática de la lengua española* (RAE y ASALE, 2009) no encontramos referencias al orden de palabras en interrogativas totales, pero sí en interrogativas parciales o con pronombre interrogativo. Según los autores, en este tipo de estructuras también es reglamentaria la secuencia verbo-sujeto.

No obstante, en la obra se alude a la existencia de un patrón diferente, existente sobre todo en el Caribe. En esta zona pueden anteponerse al verbo pronombres monosilábicos (*¿Qué tú dices?*) o de dos o más sílabas (*¿Cómo usted sabe que yo tengo que ver con una muchacha llamada Maritza?*). Este patrón de orden diferente parece tener varias causas: suplección de la pérdida del valor distintivo de las desinencias verbales por aspiración de la -s (p.3169), o herencia del español hablado en Islas Canarias (lo cual no proporciona una explicación, sino una evidencia del desplazamiento del fenómeno). La primera de las explicaciones al fenómeno, según enuncian los propios autores está sujeta a verificación, pues en otras zonas del continente americano, como Chile, donde también se da el fenómeno de aspiración de la -s final, no suelen abundar los sujetos preverbiales (p. 3170).

---

<sup>11</sup> La autora complementa su argumentación con la referencia a esquemas melódicos específicos para interrogativas sin inversión. El orden sujeto/verbo, afirma, requiere la existencia de un patrón circunflejo, propio de enunciados que son atribuidos a otros.

A modo de resumen, puede decirse que el estudio sistemático del tema de orden de palabras en la oración interrogativa es relativamente reciente en las obras normativas y descriptivas del hispanismo. Si bien se da cabida a la hipótesis de que la diferencia entre la aseveración y la interrogación absoluta puede ser establecida, en última instancia, por la entonación (Seco, 1978; Gili Gaya, 1958 [1943]; RAE, 1977[1973]), existe una fuerte tendencia a plantear como orden no marcado el de verbo-sujeto (Lenz, 1944 [1920]; Escandell, 1999; NGLE, 2009). Creemos que la revisión de las obras más icónicas nos ha permitido ver que el estudio del orden de constituyentes ha pasado de considerar criterios puramente formales y centrados en el habla peninsular a contemplar elementos de orden semántico y pragmático, como es el caso de la intención del hablante. Asimismo, se ha comenzado a atender la variación existente en el ámbito hispanohablante. Con todo, como es propio de propuestas más apegadas a la tradición, las afirmaciones se articulan a partir la intuición lingüística de los autores, en calidad de hablantes nativos de la lengua, provenientes de un análisis particular de la realidad. Si bien son válidas estas prácticas en el ámbito teórico, se necesita un estudio de la realidad comunicativa de cada comunidad de habla, a fin de acercarse más los análisis al uso cotidiano de la lengua.

#### **1.2.4. El orden de palabras en enunciados interrogativos totales en la investigación lingüística**

La revisión bibliográfica exhaustiva realizada para la confección del proyecto arrojó que las investigaciones sobre el orden de palabras en enunciados interrogativos están, en su mayoría, referidas a la interrogación parcial o con pronombre o adverbio interrogativo. Abundan, además, los trabajos sobre el orden sujeto-verbo predominante en el Caribe, así como sobre la expresión de sujetos nulos, fenómeno que es común en esta zona lingüística.

Sobre el orden de palabras en interrogativas absolutas se ha dicho poco. Como pudo verse en las referencias hechas a las obras normativas y descriptivas en la sección anterior, existe una tendencia a considerar el orden verbo-sujeto (en adelante VS) como el no marcado en la interrogación española. Si bien se alude a la facultad que tiene el usuario o hablante de la lengua de anteponer el sujeto en interrogativas totales directas, esta posibilidad se ve descartada en las parciales, las cuales exigen un orden invertido con respecto al patrón básico

(Bosque y Gutiérrez Rexach, 2009; Toribio, 2000; RAE, 2009; Leonetti, 2018)<sup>12</sup>. Otro tipo de secuenciación, sobre todo la de sujeto-verbo (en adelante, SV) tanto en interrogativas totales como parciales, ha tenido diferentes justificaciones, que reseñaremos a continuación. Podrá notarse una relación temporal dispar entre los artículos, debida a la poca producción literaria que existe en torno al tema que nos ocupa.

En su artículo “Pérdida de transposición del sujeto en interrogativas pronominales del español del Caribe” (1983), Núñez Cedeño explica que el español estándar puede convertir una oración enunciativa en una interrogativa sin alterar su orden sintáctico en el caso de las interrogativas dubitativas (o interrogativas que no solicitan información, sino que expresan sorpresa, dan énfasis o evitan la ambigüedad<sup>13</sup>), y añade que en este tipo de construcción, la interrogación cuestiona el hecho que se plantea a través de la duda, como en la siguiente secuencia:

- a) *José va*
- b) *¿José va?*

Lo que diferencia a dos enunciados como *¿El hombre estudia?* y *¿Estudia el hombre?* es la “trasposición” del sujeto, que convierte la primera oración en dubitativa y a la segunda en interrogante, o estructura que demanda información.

El autor propone que posposición del sujeto no es obligatoria. Afirma que una construcción con orden SV, puede ser identificada como interrogante o pregunta si tiene una inflexión prosódica característica de este tipo de estructura. La ausencia de trasposición puede deberse, propone Núñez, a la intención del hablante de hacerlo foco de atención. En el primero de los ejemplos anteriormente citados, el sintagma *el hombre* es el elemento focal o temático; en tanto en el segundo es el verbo *estudia*. Sentencia el autor que “la trasposición no llega a afectar radicalmente lo que se solicita. Sólo que de un elemento que era

---

<sup>12</sup> Bosque y Gutiérrez Rexach (2009) precisan que cuando el pronombre interrogativo es un argumento del verbo, la inversión es obligatoria, como en *¿Qué dijo Pepe?* En cambio, cuando no es un argumento, como en *¿Cuándo llamó María?*, la inversión es general, pero no siempre obligatoria.

<sup>13</sup> Consúltese Escandell (2000) para especificaciones de tipo sintáctico y semántico de este tipo de enunciados interrogativos, que la autora clasifica como metarrepresentacionales.

temáticamente central se pasa ahora a un plano de menos tematicidad (...) El hablante puede dejar el sujeto en su lugar, o lo puede trasladar; si hace esto, la trasposición de sujeto es facultativa” (1983, p. 41).

Un referente más reciente es “Sintaxis y uso interpretativo”, de María Victoria Escandell (2000). En el artículo, la autora vuelve sobre su propuesta teórica (Escandell, 1999) en torno al orden de palabras y la entonación como indicadores que establecen una inequivalencia entre enunciados aparentemente idénticos e intercambiables en todos los contextos, como es el caso de los siguientes:

a) *¿Ha aprobado Juan?*

b) *¿Juan ha aprobado?*

Para Escandell, la segunda de las estructuras puede utilizarse para manifestar sorpresa, con lo cual admitiría una paráfrasis del tipo *¿Es cierto que Juan ha aprobado?*; o para inferir información nueva a partir de otra ya conocida, y validarla. Tanto el orden de palabras como la entonación imponen “restricciones” en la interpretación, al descartar la posibilidad de ser la segunda de las estructuras una pregunta. A este tipo de construcciones, o sea, a las que tienen el orden sujeto-verbo, la autora las denomina “interrogativas metarrepresentacionales”.

Con respecto a las condiciones de naturaleza sintáctica que obligan a la anteposición del verbo en interrogativas totales, la autora recurre a explicaciones brindadas por el generativismo: un movimiento sintáctico permitiría que el verbo suba a una posición más alta que el nodo Inflexión, motivada por la presencia de un operador interrogativo en el Sintagma Complementante. Partiendo de la semejanza semántica entre estos enunciados y los parciales, podría decirse que en ambos existe una variable, representada por un constituyente (pronombre interrogativo) en las parciales; y por una categoría que concentra el valor afirmativo o negativo de la proposición en las totales. “Cabría suponer que también esta categoría puede contener un rasgo interrogativo que explicaría que su comportamiento sintáctico sea idéntico al de cualquier otro constituyente interrogado. Será, por tanto, la anteposición de esta categoría la que impediría la subida del sujeto” (2000, p. 232).

En Aijón y Serrano (2012) puede encontrarse un análisis de la variación posicional de los sujetos de primera y segunda persona en cláusulas interrogativas e imperativas. Los autores también se oponen al principio de que la variación libre a que hacen referencia algunos teóricos no tenga repercusiones significativas. A su vez, apuntan a que el fenómeno no puede explicarse solo a partir de las condicionantes pragmáticas o de estructura informativa (p. 132). Así, proponen una concomitancia de factores intralingüísticos y extralingüísticos en la ordenación de elementos en las construcciones oracionales. Estos factores “se ven conectados en último término por los valores cognitivos que fundamentan la estructura y el uso de la gramática” (p. 133).

La propuesta teórica de los autores defiende que “son factores abstractos y generales de índole discursivo-cognitiva los que permiten explicar más adecuadamente la colocación del sujeto en estos tipos de cláusulas, pero también las múltiples repercusiones significativas de dicha variación en contextos concretos” (p.160). A su vez, destacan que un análisis completo, que establezca las relaciones entre ellos, solo puede hacerse a partir de textos interaccionales reales y diversificados. Este elemento es destacable, dado que se sabe que muchas de las propuestas de descripción (y explicación) de fenómenos de índole gramatical parten de la intuición de los autores, y de modelos de abstracción preconcebidos, sin base ajustada a la realidad comunicativa.

La secuenciación VS, tanto en interrogativas parciales como absolutas, es vista por estos autores como un recurso para oponer construcciones que tienen prototípicamente funciones diferentes: proporcionar información (enunciados declarativos, cuyo orden no marcado es SV) y pedir información (enunciados interrogativos). La posposición del sujeto con respecto al verbo en cláusulas interrogativas totales se plantea como no marcada, al reducir su agentividad y aumentar su informatividad: en ellas “se hace focal al sujeto, subrayando su responsabilidad en el marco de la pregunta; el efecto pragmático es especialmente notorio en aquellos casos en que el referente del sujeto es la segunda persona y si se trata de preguntas que propiamente solicitan una información, esto es, no retóricas ni gramaticalizadas”, como en *¿Estabas tú al corriente del asunto?* (p.142).

Con respecto al orden SV en cláusulas interrogativas parciales, señalan que abunda este tipo de estructuras en las variedades caribeñas y en textos poéticos. No explican, sin embargo,

qué causas puede tener esta variación. No obstante, los autores mencionan que preguntas introducidas por pronombres interrogativos de valor causal o modal, como es el caso de *por qué*, tienden a favorecer el orden SV.

En lo que concierne a las interrogativas absolutas, los autores reconocen que “poseen una elasticidad mayor que las parciales en lo que respecta a la colocación del sujeto expreso” (p.143). Sin embargo, apoyan la idea de un orden prototípico VS. En consecuencia, las de orden SV son pragmáticamente “menos interrogativas” y están asociadas icónicamente a las oraciones declarativas. En este sentido, una estructura como *¿Tú estabas al corriente del asunto?* se ajustaría más a usos como expresión de sorpresa ante una información previa, compartida por los interlocutores, reforzada por una entonación que señala incredulidad o contrariedad. Concluyen los autores que “la anteposición del sujeto es adecuada para toda estructura interrogativa (sea parcial o total) que en realidad no solicita una información, sino que modaliza una toma de postura” (p. 145).

Como ejemplo de entorno propicio para la emisión de preguntas con orden SV, los autores citan el debate político, donde la función primaria de la interrogación no suele ser la de obtener información o una respuesta concreta sobre lo preguntado. Reconocen, sin embargo, que existen factores discursivos y socioestilísticos que favorecen la posposición, incluso en estos contextos, como puede ser la comunicación escrita y planificada.

En el ámbito de la cortesía verbal parece óptimo, según Aijón y Serrano, hablar de un orden conveniente, en términos de la imagen que quiere proyectar el hablante ante el oyente. Así, cuando la imagen propia se ve amenazada, como sucede en las peticiones, se tiende a la anteposición del sujeto. En coherencia con la propuesta de mayor agentividad del sujeto en las estructuras SV, este orden sugiere que se espera una respuesta afirmativa o se pretende reducir la posibilidad de negación. De modo inverso, el orden VS pudiera generar un matiz interpretativo descortés, al sugerir mayor incredulidad hacia el contenido de la interrogación (p.148).

Ej: *¿Y dice usted que está abierto al diálogo?*

Puede decirse que, con la excepción de las referencias directas al enunciado interrogativo total anteriormente citadas, la mayor producción científica con respecto al orden de palabras en interrogativas se encuentra relacionada con la interrogación parcial. Esta

realidad parece estar motivada por el abundante tratamiento recibido por obras descriptivas y normativas. Además, la existencia comprobada de un patrón diferente para la zona caribeña ha provocado una profusión de publicaciones y un amplio abanico de propuestas de análisis.

Un fenómeno que se ha tratado ampliamente en los estudios de orden de palabras en español es el de la expresión/omisión del sujeto lexical. Este es un tema que hay que mencionar necesariamente, por ser clasificada esta lengua romance como *pro-drop* o de sujeto nulo. Esta condición complica el análisis del fenómeno que nos ocupa, porque en ausencia de sujeto (o de su expresión fonética) no se puede dilucidar cuál es el orden de elementos, con referencia a este constituyente oracional.

Si bien nuestro centro de análisis no es la declaración, las investigaciones a que haremos referencia nos proporcionan pistas sobre lo que puede suceder con la interrogación respecto de la presencia/omisión el sujeto y los tipos más comunes de este último. Puede deducirse que, por analogía, los sujetos expresos más frecuentes en enunciados declarativos podrían serlo también en enunciados interrogativos, además de que podrían repetirse los patrones de selección de estos, según variables lingüísticas de diversa naturaleza.

Morales (1997) hace un estudio de la aparición del sujeto pronominal<sup>14</sup> en enunciados declarativos de tres dialectos del español: Madrid (España), San Juan (Puerto Rico) y Buenos Aires (Argentina). Los resultados muestran que, de manera general, son los pronombres de las dos primeras personas del discurso los que más aparecen de forma expresa. Con respecto a la anteposición del sujeto, la variedad puertorriqueña tiene mayor porcentaje (58%), seguida por las variedades porteña (30%) y madrileña (28%). El tipo de verbo se muestra como la variable más significativa para la anteposición del sujeto, siendo los de actividad mental (como *pensar, creer, asumir*), conjugados en primera persona, los que más favorecen el orden SV. Estos verbos “parecen crear contextos propicios para el refuerzo del sujeto; en ese sentido construyen cierto tipo de contexto contrastivo que impide la elisión o la posposición del mismo sujeto” (p. 157).

Otra de las variables significativas es el tipo de construcción: las construcciones de

---

<sup>14</sup> Solo se incluyeron los sujetos de primera persona (*yo*), segunda persona (*tú*), y las dos terceras personas, singular y plural (*él-ella, ellos-ellas*).

relativo de objeto y pregunta indirecta favorecen el orden SV. Respecto a estas últimas, el orden verbo-sujeto alcanzó solo el 2%. En la variante española, los pronombres más utilizados en posición antepuesta son los de segunda persona. Por otra parte, tanto los sujetos elididos (variante madrileña), como los expresados (variante puertorriqueña), se utilizan para darle continuidad al tópico. La autora propone que para el habla caribeña (San Juan de Puerto Rico, en el caso de esta investigación), podría estarse forjando un proceso automático de refuerzo pronominal y que los factores que más influyen en la utilización de los sujetos pronominales son los sintácticos y pragmáticos.

Por su parte, Bentivoglio (2003) estudia el sujeto explícito en el habla venezolana (caraqueña), comparando los resultados con los obtenidos por Delbecque (1991, citado por la autora en el trabajo referenciado) a partir del análisis de textos escritos por trece autores españoles y seis hispanoamericanos. Las variables que intervinieron en el estudio fueron la morfología del sujeto, la extensión del sintagma nominal, los rasgos semánticos del referente, la función gramatical, y la presencia/ausencia del sujeto en el contexto precedente (p. 237). Los resultados mostraron que, en la mayoría de los casos, el sujeto aparece antepuesto. Favorecen la posposición del sujeto en ambos *corpus* la presencia de sujetos conformados por infinitivos o con complemento preposicional u oración subordinada. Lo mismo sucede con los sujetos que representan una acción o proceso. Por último, la ausencia del referente del sujeto bajo análisis en el contexto precedente parece favorecer el orden VS.

Volviendo al ámbito de la interrogación, en un trabajo referido a la posición del sujeto español bajo el lente generativista, Liceras (1994) hace referencia al orden de palabras tanto en interrogativas totales como parciales. Con respecto a las parciales, la autora explica la obligatoriedad de “inversión” sujeto-verbo en el caso de sujetos léxicos a partir de una regla de movimiento que sitúa al foco interrogativo en la posición de Complementante. Tal es el caso de estructuras como *¿Qué han hecho Juan y María?*

En interrogativas totales se propone como opcional la inversión del orden lineal, pues no existe –pensando en interrogativas no marcadas– un sintagma focalizado, sino que el foco abarca toda la oración, como en *¿Ha llamado Juan?*, *¿Juan ha llamado?* (p. 246).

Por otra parte, la posición del sujeto en las interrogativas parciales caribeñas ha llamado especialmente la atención de los investigadores, al constatarse un orden diferente al del

español estándar. Lipski (1977) se refiere al orden SV con pronombres de las tres personas del singular en países como Cuba, Puerto Rico y República Dominicana. En especial para los hablantes de la variante cubana, el orden lineal es preferido como normal en el habla espontánea, frente al orden VS, más poético o afectado. El investigador hace referencia a los tiempos que parecen favorecer la anteposición del sujeto (el presente, seguido por el pretérito e imperfecto), si bien aclara que no existe suficiente evidencia para postular un orden para los demás tiempos. Asimismo, reseña algunas de las explicaciones dadas al fenómeno de la inversión sintáctica, muchas de ellas superadas posteriormente con el aporte de nuevas investigaciones<sup>15</sup>: la preponderancia de la segunda persona, dado que el pronombre tónico se convertía en átono mediante un proceso de desplazamiento del acento hacia el pronombre interrogativo, en lugar de hacia el sujeto; la influencia de las lenguas africanas o la compensación morfológica.

Lantolf (1980) presenta los resultados de una encuesta aplicada a puertorriqueños de diferentes edades, sexos y nivel sociocultural, radicados en New York. Existe, según la investigación, una preferencia por la interpolación del *tú* en todas las edades y niveles socioculturales. Esa preferencia es seguida por el pronombre *yo*, los sujetos nominales, *usted*, *ustedes*, *nosotros* y *él*, en orden decreciente. Los más jóvenes prefieren la anteposición del sujeto, excepto en el caso de *él*, en tanto los más viejos optan por la posposición, salvo en el caso del pronombre *tú*.

Lapesa (1992) hace un repaso de las teorías explicativas en torno a la anteposición del sujeto en las interrogativas parciales, y da cuenta de la existencia del fenómeno, de forma corroborada, en el Caribe, Río de la Plata y Canarias. Sin embargo, el autor reporta estructuras con orden SV en hablantes de Palencia y del nordeste leonés (España), si bien de manera muy impresionista, al ser parte de su experiencia como usuario de la lengua. El filólogo reconoce la generalización del fenómeno en el Caribe, reportada desde 1929. No obstante, “eso no excluye la posible antigüedad de su germen: ha podido permanecer siglos y siglos en estado latente y haber cundido más tarde, incluso hace poco, por acción de nuevas circunstancias favorables” (p. 548).

---

<sup>15</sup> Ver Núñez (1993).

Junto con la mayoría de los teóricos, Lapesa propone la secuencia VS como la no marcada para la interrogación parcial directa. Los casos en los que antepone el sujeto son usados, según el autor, para transmitir un interés inquisitivo, sorpresa, admiración o actitud recriminatoria (1992, p. 550), hacia un interlocutor o hacia sí mismo. Tal es el caso de enunciados como los siguientes:

- a) *¿Por qué tú no me has dicho eso antes?*
- b) *¿Cómo él ha podido sobreponerse a tantas dificultades?*
- c) *¿Para qué yo me habré esforzado tanto?*

El autor apuesta por el factor de uso heredado. Estima que interrogativas directas e indirectas tienen semejanzas estructurales que se deben a la “pervivencia del uso latino correspondiente” (p. 552). El teórico propone valorar la probabilidad de la influencia de la interrogación indirecta sobre la directa como una suerte de continuidad o “contagio” de una estructura a otra.

Ordóñez y Olarrea (2006) proponen explicar la diferencia de patrones de orden existente en las variedades caribeñas (especialmente del español dominicano<sup>16</sup>) y el español estándar a partir de criterios morfológicos: básicamente, se trata de factores tipo de sujeto y la complejidad del pronombre interrogativo.

Dadas las evidencias recogidas en el área del Caribe, todo apunta a que los sujetos pronominales en singular son más aceptados en la anteposición que los plurales (la segunda persona más que la tercera y la primera). Por otra parte, los sujetos constituidos por frases nominales tienden a estar antepuestos, y son usados frecuentemente con frases interrogativas complejas como *por qué, en qué lugar, cuál de...* en lugar de con pronombres interrogativos.

En una investigación llevada a cabo por los autores con sujetos dominicanos, mediante el uso de la técnica de cuestionario, se supo que el 100% de los hablantes acepta el orden lineal con el pronombre singular de segunda persona en interrogativas parciales, en contraste con el número de informantes que no acepta la anteposición con frases nominales (15-18%).

---

<sup>16</sup> Véase Toribio (2000) para más información sobre la hipótesis de que el español dominicano estaría pasando por un proceso de cambio de una variedad *pro-drop* (de sujetos omitidos y posverbiales, como es el caso del español estándar) a una variedad *non-pro-drop* (de sujetos expresos y preverbiales).

El número decrece significativamente cuando las frases se hacen más complejas, por ejemplo, con una cláusula de relativo (12%). Los hablantes mostraron preferencia por el orden lineal con la segunda persona, en detrimento de la primera y la tercera. Además, las frases interrogativas complejas permiten la no inversión con un mayor porcentaje que las que son simples.

La propuesta teórica de los autores radica en que los pronombres en el español caribeño, contrario al resto de las variedades, se comportan como elementos débiles o deficientes (*weak pronouns*), tal como sucede con los clíticos. Es por ello que son más difíciles de invertir, incluso si ellos no son clíticos. En consecuencia, estructuras como *¿Qué José quiere?* serían agramaticales, partiendo del principio de que solo pueden anteponerse pronombres personales, que funcionan como elementos débiles.

En un análisis sobre forma y función pragmática de interrogativas parciales puertorriqueñas, Brown y Rivas (2011a) encontraron que las estructuras con sujeto omitido y orden VS son las más comunes en el corpus estudiado (28% cada una). Las interrogativas SV alcanzan un 10%<sup>17</sup>.

Estas estructuras podrían identificarse por su función: alrededor del 50% de interrogativas con sujeto omitido y con posposición del sujeto son usadas para solicitar información, lo que los autores llaman “función externa”. Las que tienen orden SV son usadas con un valor citativo o una “función expresiva” en el 62% de los casos. Esta es la razón, afirman los autores, que podría explicar la presencia de una forma que “viola” la regla de “inversión” sujeto-verbo (p. 310).

El análisis estadístico realizado por los autores tuvo como variables la función pragmática, la temporalidad verbal, la función sintáctica, el tipo de verbo y la animacidad del pronombre interrogativo. Los resultados muestran que los dos factores más influyentes son tipo de verbo y función sintáctica del pronombre interrogativo.

---

<sup>17</sup> A pesar de esta baja cifra, los autores aclaran que “si solo tenemos en cuenta aquellos tipos de preguntas en las que es posible la variación en cuanto a la posición sujeto-verbo (N=464), el porcentaje de interrogativas SV aumenta al 27%. Esto quiere decir que más del 25% de las construcciones interrogativas con sujeto explícito se construyen sin inversión del sujeto-verbo en esta variedad del español” (2011a, p. 307).

Las funciones de sujeto, complemento directo, complemento indirecto, o complemento circunstancial de los pronombres interrogativos favorecen el uso del orden SV, mientras la función de complemento predicativo propicia el orden inverso. Por otro lado, los verbos como *creer, querer, decir, ir, hacer, tener* y *ver* favorecen este orden. Los verbos *ser, estar* y *llamarse*, en cambio, muestran un índice probabilístico bajo.

En “Subject-verb word order in Spanish interrogatives”, los mismos autores se refieren a propuestas teóricas (Toribio (2000) y Morales (1989), citados por Brown y Rivas, 2011), según las cuales el español del Caribe estaría perdiendo flexibilidad y habría comenzado un proceso tendiente al orden rígido de SVO, al igual que sucede con lenguas romances como el francés. Otros elementos que podrían estar formando parte de esta tendencia son la expresión de sujetos pronominales y el uso de pronombres con verbos en infinitivo (Ej: *para yo ir*).

Brown y Rivas (2011) analizaron 882 preguntas parciales emitidas en 27 horas de conversación con hablantes de la variante puertorriqueña. En el análisis intervinieron variables como: orden de palabras (SV, VS, o sujeto nulo), tipo de sujeto (primera, segunda, tercera personas; sujeto nulo o lexical), tipo de pronombre interrogativo, tipo de verbo y tipo de enunciado interrogativo (retórico, citativo o pregunta). Contrario a lo que otras investigaciones mostraban con respecto al orden de elementos oracionales en interrogativas parciales, solo el 14% de las construcciones con sujeto expreso tenía un orden SV. “Estos resultados sugieren que la no inversión no es tan prevalente en el español de Puerto Rico como generalmente se asume, y que la asociación de las interrogativas sin inversión con el español de este país y los dialectos caribeños puede constituir un estereotipo sociolingüístico” (p. 35). Por otra parte, cuando el orden es SV, los pronombres son, sobre todo, *tú* o *usted* (73%).

Con respecto a los factores que más favorecen el uso del sujeto expreso de manera general, el tipo de verbo es la variable más significativa, si bien las únicas formas con peso estadístico son *ser, estar* y *llamarse*. Asimismo, las preguntas retóricas y citativas tienden a tener sujeto expreso, contrario a lo que sucede con las interrogativas en el sentido clásico (estas ostentan un alto índice de sujetos nulos). La presencia del sujeto en las interrogativas citativas puede estar en consonancia, según los autores, con la intención del emisor de darle

mayor peso pragmático al enunciado. Así, con este recurso “el hablante asegura una narración más dramática de su historia” (2011, p. 38); a su vez, el enunciado se torna más personalmente relevante.

Entre los factores que más favorecen el orden SV están el tipo de sujeto, el tipo de interrogativa y la categoría del verbo. Con los pronombres *yo*, *tú* y *usted*, se ve favorecido el orden SV. Con *él* y *ella*, la tendencia a este ordenamiento es bastante menor. Los sujetos léxicos desfavorecen este orden. Igualmente, los verbos como *pensar*, *creer* y *querer* influyen en esta ordenación, mientras que las interrogativas citativas y retóricas tienen en el 55% de los casos el sujeto antepuesto.

En síntesis, puede decirse que el tema que nos ocupa (orden de palabras en enunciados interrogativos absolutos) ha sido poco tratado, y que los referentes teóricos son muy variados. Los mayores aportes han venido desde el cognitivismo (Aijón y Serrano, 2012) y el generativismo (Escandell, 2000 y Ordóñez y Olarrea, 2006), aun cuando abundan investigaciones de corte experimental que han brindado datos interesantes que tomaremos como antecedentes. No existe consenso con respecto al orden sistemático, pero gran parte de las propuestas teóricas apuntan a un orden VS para enunciados que buscan información, frente a un orden SV para enunciados que tienen otras funciones comunicativas.

El mayor caudal de investigaciones en cuanto a orden de palabras en interrogación española se encuentra relacionado con la interrogación parcial, pero existe igualmente una profusión de estudios con metodologías y objetivos dispares, que hace que los resultados sean poco sistematizables. La expresión/omisión del sujeto en algunas variedades del español es un tópico que toca tangencialmente nuestro objeto de estudio, pero que ofrece información que se debe tener en cuenta para la determinación de posibles variables. Con respecto al orden de palabras en interrogativas parciales, se ha estudiado la influencia de factores como el tipo de verbo, tipo de sujeto, extensión de este, función pragmática del enunciado, tipo de interrogativa, tipo de pronombre interrogativo, etc. Si bien existe abundante evidencia de que es el habla caribeña la que tiene mayor tendencia al orden SV, investigaciones recientes (Brown y Rivas, 2011 y 2011a) apuntan a que estas afirmaciones pudieran corresponder más a un “estereotipo sociocultural” que a datos empíricamente comprobables. El orden VS, en consecuencia, está asociado sistemáticamente con la interrogación neutral o de demanda de información.

### 1.3. Hipótesis de trabajo, preguntas de investigación y objetivos de investigación

Este estudio aborda el tema del comportamiento del orden de constituyentes en enunciados interrogativos absolutos en el entorno hispánico. Se trata, pues, de una investigación de carácter descriptivo, que pretende profundizar en el fenómeno de la alternancia sujeto-verbo y ampliar los referentes investigativos ya existentes. Partiendo de los datos reportados en el español del Caribe, según los cuales habría, en principio, un orden diferente en la interrogación con respecto al español peninsular, y de la observación del comportamiento lingüístico de hablantes de variedades dialectales que no han sido estudiadas en profundidad respecto del tema que nos ocupa, esbozamos la siguiente **hipótesis**:

*El orden establecido como prototípico para la interrogación total del español (VS) puede exhibir una variación a partir de la influencia de factores lingüísticos (sintácticos, semánticos y pragmáticos) y extralingüísticos (zona geográfica).*

Partiendo de esta hipótesis, la investigación buscará responder a la siguiente pregunta general:

¿Qué caracteriza el fenómeno de la alternancia sujeto-verbo en las interrogativas totales?

Y las siguientes preguntas específicas:

1. ¿Cuál es el orden sintáctico predominante en interrogativas absolutas en lengua española, en especial referencia a la relación entre sujeto y verbo?
2. ¿Qué factores sintácticos, semánticos y pragmáticos favorecen ese orden?
3. ¿Qué zonas o regiones con patrones particulares pueden identificarse?

En consecuencia, la presente investigación tiene los siguientes objetivos:

#### **Objetivo general:**

Comprender el fenómeno de la alternancia sujeto-verbo en las interrogativas totales del español.

#### **Objetivos específicos:**

1. Determinar el patrón de orden sintáctico predominante en interrogativas totales españolas.

2. Caracterizar la influencia de factores sintácticos, semánticos y pragmáticos que intervienen en la secuenciación sujeto-verbo de las interrogativas totales del español.
3. Describir la distribución dialectal del fenómeno de la alternancia sujeto-verbo en la interrogativa total, partiendo de las zonas geográficas tradicionalmente identificadas para el español.

## **1.4. Variables incluidas en el estudio**

### **1.4.1. Zona geográfica**

La variable zona geográfica es la única variable de carácter extralingüístico contenida en el análisis. La motivación principal de su inclusión es la comprobación, mediante la abundante evidencia lingüística contenida en la literatura, de que existe un patrón de orden diferente en el Caribe. Creemos que vale la pena estudiar si la llamada tendencia a un orden de elementos rígido en estas variantes de lengua (es decir, la ordenación lineal de los enunciados con secuencia SV(O)) es solo típica (si se corrobora la hipótesis) del Caribe, o si es una tendencia existente en alguna otra zona lingüística.

Por otra parte, existen autores que creen que las condicionantes del orden de palabras en español no ofrecen variación dialectal, sino que corresponden a características generales de la lengua española (Silva-Corvalán, 1982).

La zonificación dialectal a la que responde esta investigación es la que está establecida en el CORPES XXI. La herramienta de la Real Academia recoge 10 categorías: España, Andina, Antillas, Caribe Continental, Chile, México y Centroamérica, Río de la Plata, Estados Unidos, Guinea Ecuatorial y Filipinas. Las razones de exclusión de las últimas 3 zonas están expresadas en el acápite metodológico. En la Tabla 2 aparece el detalle de zonas dialectales y países que las integran.

| <b>Gran área</b>         | <b>Zona geográfica</b>              | <b>Países</b>   |
|--------------------------|-------------------------------------|---|
| <b>España</b>            |                                     |   |
| <b>América</b>           | Andina (And)                        | Bolivia, Ecuador, Perú  |
|                          | Antillas (caribeña) (Ant)           | Cuba, Puerto Rico, República Dominicana                       |
|                          | Caribe continental (Car)            | Colombia, Venezuela   |
|                          | Chilena (Chi)                       | Chile   |
|                          | México y Centroamérica (Mex)        | Costa Rica, El Salvador, Guatemala, México, Nicaragua, Panamá |
|                          | Río de la Plata (Río) <sup>18</sup> | Argentina, Paraguay, Uruguay                                  |
|                          | Estados Unidos                      |   |
| <b>Filipinas</b>         |                                     |   |
| <b>Guinea Ecuatorial</b> |                                     |   |

Tabla 2. Separación en zonas dialectales y países que las integran. Fuente: adaptación de CORPES XXI

Como elemento adicional, el Gráfico 1 ilustra la distribución de datos en las macrozonas del corpus y en las zonas dialectales tradicionales propuestas para el español de América<sup>19</sup>. Esta distribución correspondiente a las grandes áreas lingüísticas del mundo hispánico se establece a partir del “cruce de criterios diferentes, entre los que figuran la población, el volumen de publicaciones, la cantidad de ediciones digitales de periódicos y revistas” (RAE, 2013, p.7). Como puede verse, hay un predominio de información proveniente de España, mientras las zonas Chile, Caribe Continental y Andina tienen las proporciones más bajas.

<sup>18</sup> Las abreviaturas colocadas en paréntesis han sido usadas para codificar los ejemplos del corpus que conformamos para la investigación, seguidas del número correlativo para cada uno de ellos.

<sup>19</sup> Han sido varias las propuestas de variación dialectal, pero fue la propuesta de Henríquez Ureña la que abrió el camino de la caracterización dialectal moderna.

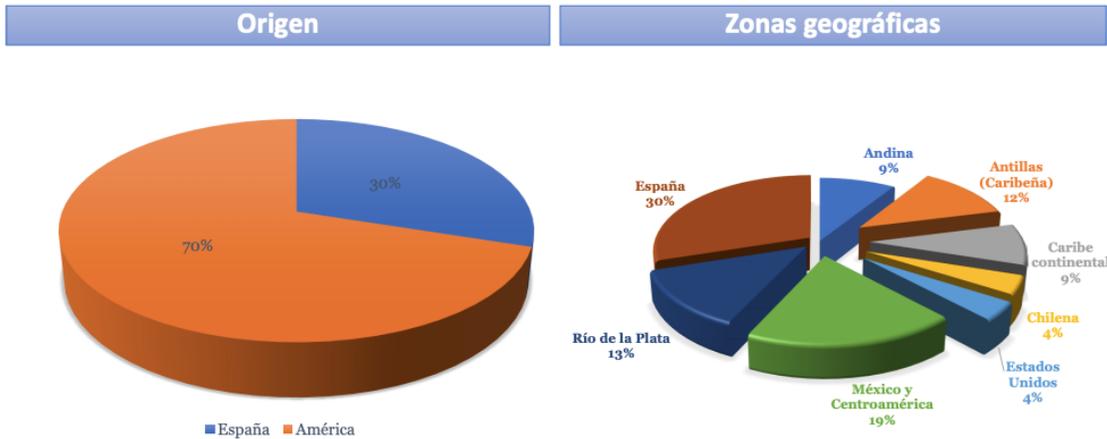


Gráfico 1. Distribución de datos en las macrozonas España y América (izquierda) y en las zonas dialectales (derecha).  
Elaboración: Mercedes Sánchez, coordinadora del CORPES XXI. Fuente: Internet

Debe ser apuntado que somos conscientes de que el complejo panorama de la diversidad dialectal del español no puede ser captado con una división general en zonas como la que se propone, especialmente porque existe una riqueza al interior de ellas que es virtualmente ignorada. Solo en España, país y zona dialectal a la vez, se ha constatado la existencia de tres variedades: la andaluza, la canaria y la castellana (Moreno, 2018). La asunción de esta separación es justificada solo por razones metodológicas: se trata de obtener una mirada general de la variación diatópica del fenómeno que nos ocupa.

#### 1.4.2. Semántica verbal

Existe evidencia a favor de que la categorización de unidades léxicas está en estrecha relación con la selección de estructuras sintácticas en una lengua (Crous & Gràcia, 2015; García-Miguel, 2005). Una de las clasificaciones de verbos en español es la que se refiere a su tipo semántico o, lo que es lo mismo, su modo de significar (Cano-Aguilar, 1981). Establecer una clasificación en torno a este criterio es una tarea difícil, dada la complejidad del fenómeno de la significación, lo cual implica un desafío para las labores de discretización. En el caso de los verbos es aún más complejo el establecimiento de fronteras significativas, pues en ellos hay una “variedad de componentes o aspectos experienciales que pueden ser codificados o evocados simultáneamente” (Albertuz, 2007, p. 7).

La tradición gramatical ofrece una amplia gama de clasificaciones de esta categoría sintáctica, bajo criterios de diversa índole y con diverso grado de sistematicidad. También en la investigación lingüística se observan clasificaciones cuyo origen no es aclarado en un número alto de casos. En consecuencia, haremos referencia en este apartado a algunos de los autores que particularizan en una clasificación referida al significado de los verbos y esclareceremos cuál se ajusta más a nuestros objetivos.

Lenz (1944 [1920]) propone establecer una separación según el sujeto al que haga referencia el verbo. Puede hablarse, en consecuencia, de verbos meteorológicos (verbos impersonales que predicen sobre fenómenos atmosféricos, como *hacer calor, llover, tronar*); verbos del reino animal y vegetal (*brotar, florecer, bramir, gruñir*); verbos que hacen referencia a fenómenos psíquicos del hombre tales como fenómenos sentimentales, morales, estéticos, volitivos e intelectuales; o a fenómenos fisiológicos de carácter interno, como *tener hambre, estar soñoliento, digerir* o externo: *comer, andar, beber*; por último, están los verbos referidos a actividades culturales.

Cano-Aguilar (1981), por su parte, establece una clasificación de verbos transitivos teniendo en cuenta diferentes factores como la agentividad, el tipo de objeto que tiene (exige) el verbo y el significado de este. De este modo, propone distinguir entre:

a) Verbos de acción resultativa: son verbos cuyo objeto denota el resultado de una acción o estado designado por el verbo. El ejemplo prototípico es *hacer*, con sus formas lexicalizadas, formadas a partir de la unión del verbo con un sustantivo o un adjetivo. Es el caso de *hacer entrega (de algo)*, que equivale a la unidad léxica *entregar (algo)*. También se incluyen en este grupo los grupos verbales que no se corresponden con una unidad léxica, como *hacer caso, hacer el favor*, etc. Otros verbos como *fabricar, producir, realizar, inventar, pintar, escribir*, etc, tienen carácter resultativo.

b) Verbos de carácter causal: se refieren a una acción que da lugar a un determinado acontecimiento, como es el caso de *causar, provocar, suscitar*. Estas formas verbales cuentan normalmente con sujetos abstractos o nombres de acontecimientos.

c) Verbos de modificación: indican una acción ejercida sobre algo que, como resultado de esa acción, es modificado o alterado. Entre los más comunes, encontramos *arreglar, componer, corregir, modificar, preparar*.

d) Verbos de objeto afectado: el objeto de esta clase de verbos no es creado ni modificado, sino que “es algo sobre lo cual se hace algo” (p. 65): *alegrar* (a alguien), *enfadar* (a alguien), *traicionar*, *felicitar*, *engañar*, *usar*, *manejar*, etc.

e) Verbos de movimiento: como su nombre indica, son verbos en el desplazamiento o cambio de lugar de lo indicado por el objeto. En este grupo se encuentran *colocar*, *poner*, *llevar*, entre otros.

f) Verbos con objeto de lugar: implican también un movimiento, realizado por el sujeto, pero sin desplazar o mover el objeto. Ejemplos de estos verbos son *recorrer*, *cruzar*, *atravesar*, *rodear*.

g) Verbos de posesión: son los referidos al acto de tener (*poseer*) o llegar a tener algo (*tomar*), así como dejar de tener (*perder*), pasar a tener (*dar*, *quitar*).

h) Verbos de actitud: designan la actuación de alguien o algo, con el objetivo de que una acción o hecho pueda tener lugar. Es el caso de verbos como *impedir*, *obligar*, *empujar* (a hacer algo), *impulsar*, *evitar* (en su sentido de impedir), etc.

i) Verbos de percepción: están incluidos en este grupo los verbos de percepción física como *ver*, *mirar*, *sentir*, *percibir*, *contemplar*, *escuchar*; y los de percepción intelectual, como *saber*, *conocer*, *entender*, *comprender*, entre otros.

j) Verbos de voluntad, emoción y sentimiento: estos verbos tienen en común con los actitudinales el hecho de que expresan una postura del alguien frente a algo. Sin embargo, en el caso de este grupo, está presente la marca volitiva. “Designan estados o situaciones que podríamos caracterizar como emocionales” (p. 197). Ejemplos de este tipo son *querer*, *desear*, *temer*, *amar*, *odiar*, etc.

k) Verbos de comunicación verbal: estos verbos contienen en última instancia el sentido de ‘decir’, o sea, indican la expresión de un pensamiento o idea. Tradicionalmente también se los ha llamado *verba dicendi*. En este grupo se encuentran *decir* (como verbo más general), *contar*, *referir*, *anunciar*, *declarar*, *afirmar*, etc.

Las clasificaciones antes aludidas presentan como principales falencias su relativa generalización y su extensión. Ello trae como consecuencia que se dejen fuera predicados (en el caso de Cano-Aguilar no se incluyen los verbos intransitivos), o que estos puedan

pertenecer a más de un tipo según las descripciones que se ofrecen; por tanto, la asignación del verbo que se analice a un tipo de predicado se convierte en una tarea demasiado engorrosa<sup>20</sup>. En consecuencia, se ha decidido asumir la clasificación de Albertuz (2007), que es la que opera en el proyecto Alternancias de Diátesis y Esquemas Sintáctico-Semánticos del Español (ADESSE), dirigido por García Miguel<sup>21</sup>. El autor establece “clases y subclases verbales identificando las áreas de experiencia o dominios cognitivos conceptualizados primariamente por los verbos, lo que se revela tanto en el núcleo semántico de sus designaciones como en el tipo de relaciones semánticas que mantienen con el verbo y/o entre sí las diferentes entidades actanciales evocadas” (p. 3). Este autor se basa en la propuesta de Halliday y Matthiessen (2013 [2004]) quienes, por demás, establecen una extensa subclasificación que no abordaremos en este trabajo, por parecernos innecesaria. Basten las macroclases de verbos establecidas en el proyecto ADESSE para adjudicarle una esfera de la experiencia a cada verbo, con el objetivo de saber si esto puede o no influir en el orden de palabras. Cabe aclarar que las clasificaciones semánticas, en la mayoría de los casos, son discutibles e inexactas por la complejidad asociada al significado verbal. Esta propuesta nos parece, sin embargo, coherente y robusta, por ser eminentemente funcionalista.

Son seis las macroclases que propone Albertuz (*op. cit.*) y que usaremos para nuestra clasificación<sup>22</sup>:

a) material: relacionado con lo que sucede de manera exterior al sujeto; referido a acciones y eventos que se encuentran en el mundo que nos circunda (son procesos relacionados con ‘hacer’ y ‘suceder’, como *crear, romper, llevar, poner, golpear*, etc).

b) mental: relativo al mundo interno del hablante, a los procesos de conciencia (*sentir*,

---

<sup>20</sup> En Levin (1993) se puede encontrar una exhaustiva clasificación de los verbos en inglés que ha sido retomada en varias ocasiones para la clasificación de los verbos españoles.

<sup>21</sup> Disponible para consultas en <http://adesse.uvigo.es/>

<sup>22</sup> La clasificación acá recogida carece de una de las clases establecidas por Halliday y Matthiessen (2013) y que estaba originalmente contenida en el proyecto ADESSE. Se trata de los procesos conductuales, conceptualizados en el texto que hemos usado como: la manifestación externa de actividades internas, la exteriorización de procesos de conciencia y de estados fisiológicos (*reír, dormir, respirar*). Estos procesos están a medio camino entre los materiales y mentales. En una posterior categorización de los procesos verbales, el proyecto asignó los verbos pertenecientes a los procesos conductuales a las categorías material o mental según el significado expresado por el verbo en una expresión concreta; y añadió la categoría de verbos de modulación, que se recoge en este trabajo.

*desear, pensar*).

c) relacional: se refiere a la relación que puede establecerse entre un fragmento de la experiencia y otro, con lo cual se caracterizan e identifican fenómenos (procesos de *ser* y *tener*).

d) verbal: referido a las relaciones simbólicas construidas en la conciencia y que se representan en la forma lingüística (*decir, hablar, responder*). Estos procesos comparten rasgos con los mentales y relacionales.

e) existencial: relativo a la existencia, mediante la cual los fenómenos pueden ser reconocidos como tal (ellos existen (son), suceden). Los verbos más icónicos son *haber, ocurrir, nacer, existir*, entre otros. Esta categoría está a medio camino entre los procesos relacionales y materiales<sup>23</sup>.

f) modulación: agrupa a verbos que no pertenecen a ningún dominio de la experiencia, sino que suelen combinarse con otro típicamente en infinitivo (*hacer ver, ayudar a salir, tratar de comprender*). Es este último el que se adscribe a un dominio de experiencia. Estos verbos actúan como semiauxiliares<sup>24</sup>.

Tal como propone el autor, la manera más eficiente de asignar un verbo a una macroclase es determinar el significado básico expresado por el verbo, tratando de aislarlo de su contexto lingüístico (vale decir, tipo de construcción sintáctica en la que se encuentra y tipo de argumentos que presenta) en la medida de lo posible.

Con respecto a la investigación lingüística en torno a la influencia de la semántica verbal en la determinación de la estructura lineal del enunciado, Ocampo (1991), por ejemplo, postula que el significado verbal influye de manera determinante en el orden de constituyentes. Con respecto a los verbos de una valencia en español (más adelante trataremos la estructura argumental), el autor propone establecer una distinción entre verbos

---

<sup>23</sup> De manera similar, Breñes (2011) utiliza una división en cuatro grupos: verbos enunciativos, cognitivos, de estado, y de actividad.

<sup>24</sup> Los miembros del proyecto ADESSE no cuentan con publicaciones posteriores a la reseñada en este trabajo (2007) donde se ofrezca una conceptualización de este tipo de proceso. La descripción acá recogida se ha conformado sobre la base de la comunicación personal de la autora con el coordinador del dicho proyecto, José M. García-Miguel.

de estado que expresan significados existenciales (*haber, existir*) y otros tipos de verbo. Esta clasificación es útil al autor para proponer un orden preestablecido: el primer tipo de predicado condiciona que el sujeto se posponga al verbo (VS), en tanto el segundo favorece la secuencia SV<sup>25</sup>. Esta evidencia es complementada por Brown y Rivas (2011), quienes afirman que los verbos copulativos como *ser, estar y llamarse* favorecen la posposición del sujeto.

Por su parte, Morales (1997), en su estudio sobre la declaración en tres dialectos (Madrid, San Juan y Buenos Aires) sugiere que los verbos de actividad mental (como *pensar, creer y asumir*) favorecen la anteposición del sujeto en el 47% de las oraciones con pronombres de primera y segunda persona del singular, y en un 41% de las que tienen pronombre de tercera persona del singular. Según la autora, estos verbos, en tanto expresan acciones propias de los seres humanos, tienden a favorecer el refuerzo del sujeto y su anteposición. “Los hablantes tienden a expresar más el sujeto pronominal cuando este lleva a cabo una acción mental relacionada con la manifestación de una opinión o comentario, tanto sea del propio hablante como de otra persona” (p. 157). Brown y Rivas (2011) refuerzan esta postura en su análisis de interrogativas parciales puertorriqueñas, al encontrar que son los verbos *creer* y *querer* los que más influyen en la anteposición del sujeto (100% y 80%, respectivamente), seguidos por los verbos de dicción (54%) y, además, *ir* y *hacer*, con 43 y 42% de casos con este orden, respectivamente. Estos resultados son respaldados por Rosemeyer (2016), quien también encontró, en un estudio sobre interrogativas parciales caribeñas, que los verbos de creencia y *dicendi* favorecen la anteposición del sujeto, mientras que los de estado aumentan las probabilidades de la posición.

Con respecto al orden que asume el sujeto cuando se enfoca el fenómeno desde el punto de vista pragmático, abundante bibliografía hace referencia, además, a que los verbos de carácter presentativo y existencial favorecen la posposición del sujeto en la declarativa (Contreras, 1978; Fernández Soriano, 1993), tanto si este es portador de una información

---

<sup>25</sup> Esta mirada del orden de palabras desde la semántica verbal está en estrecha relación con el comportamiento sintáctico de los verbos. Los verbos de existencia y presentación, por ejemplo, en tanto verbos intransitivos, seleccionan un solo argumento, que suele carecer de agentividad. El rol semántico de paciente, como se verá más adelante, determina en un número importante de casos la posición de este argumento, con función sintáctica de sujeto.

conocida como nueva, como en el siguiente ejemplo, tomado de Rivas (2008, p. 901).

V: iba a la facultad, estudiaba: - Antropología, iba con su motito una: - tipo, qué se yo, una Siambre:ta, una de esas motos.

F: sí

V: bueno, un día se la roban, de la puerta de la universidad. Sale, no (es)tá *la mo:to*.

Por otra parte, los verbos de movimiento como *llegar* y *venir*, que forman parte del grupo de presentativos en español (Hatcher, 1956), también aparecen delante del sujeto de manera regular (Bentivoglio & Weber, 1986), pues típicamente sirven para introducir participantes humanos en el discurso (p. 30).

### 1.4.3. Tipo de verbo según realización valencial

Las gramáticas tradicionales clasifican los verbos según lo que se denomina “valencia verbal”. La propuesta original, planteada por Tesnière (1959), establece que cada verbo posee una serie de actantes (participantes, argumentos, según la terminología) a los cuales es susceptible de regir. Esta valencia, también llamada estructura argumental, tiene influencia en el orden de palabras. Por ejemplo, puede establecerse un comportamiento diferente, sintácticamente hablando, entre predicados transitivos (de más de un argumento) e intransitivos (de un solo argumento). Así, se plantea que el primer tipo de predicado favorece el orden SV, mientras el segundo favorece la secuencia VS (Fernández Ramírez, 1986; Delbecque, 1987 y 1988; Rodríguez, 2008; Gattei *et al*, 2015, Rivas, 2013). Sin embargo, a pesar de los órdenes marcado/no marcado antes mencionados, se ha llamado la atención sobre la poca sistematicidad con la que los verbos intransitivos posponen el sujeto. La investigación lingüística ha arrojado que en estos casos existe una tendencia hacia la anteposición del sujeto (Alfonso y Melis 2010; Rivas 2008, Bogard, 2019<sup>26</sup>), pero la posposición sigue siendo un fenómeno frecuente en este tipo de estructura (Silva-Corvalán, 1982). Podría hablarse, en consecuencia, de una alternancia con el orden SV. En Sánchez (2008), por ejemplo, se encuentra evidencia a favor de la anteposición del sujeto en oraciones de predicado

---

<sup>26</sup> El estudio de este autor con textos de los siglos XIII, XVI y XIX mostró que los sujetos de verbos intransitivos pasaron de un orden mayoritariamente pospuesto (dos primeros siglos) a la anteposición (XIX).

intransitivo. La posposición, afirma la autora, es típica de estructuras pasivas pronominales (*Se reabrió la universidad y fui profesor*<sup>27</sup>) y las biactanciales con verbos pseudoimpersonales, como el caso del verbo gustar (*Me gusta el fútbol*). Recalca la autora que “las cifras tan cercanas de sujetos preverbiales y postverbiales revelan la ausencia del predominio numérico de un orden determinado” (p. 95) y que, en el caso de los verbos inacusativos (cuyo concepto discutimos a continuación), puede estar ocurriendo una pérdida de la posposición como marca distintiva de este tipo de predicados. La autora establece una amplia clasificación semántica de predicados intransitivos y concluye que los rasgos semánticos de ciertas clases verbales como las que expresan existencia (‘existir’ o ‘faltar’) y experimentación emotiva (‘agradar’ o ‘importar’) tienen una fuerte influencia en la determinación del orden (en estos casos la secuencia más común es la de VS). Esas clases verbales están asociadas pragmáticamente al acto de presentar una entidad en el discurso (p. 102); en este sentido, sería la variable semántico-pragmática la que adquiriría preponderancia en la linearización del enunciado. También Suárez (2016) comparte esta postura, al afirmar que los predicados de movimiento de dirección inherente muestran en castellano medieval una tendencia a presentar el sujeto pospuesto precisamente por su carácter presentativo. Son los factores de naturaleza semántica y pragmático-discursiva, en consecuencia, los que determinan la posición del sujeto.

A la par, existe bastante consenso entre los gramáticos sobre la existencia de dos grupos con características semánticas y sintácticas diferentes dentro de los verbos intransitivos: se trata de los inergativos y los inacusativos<sup>28</sup>. Perlmutter (1978) propuso, en el marco de la gramática generativa, que los verbos inacusativos cuentan con un sujeto que es realmente un objeto profundo, lo cual condiciona su papel de tema o paciente, y su orden pospuesto al verbo, tal como lo haría típicamente un objeto, o argumento interno del verbo. En cambio,

---

<sup>27</sup> Ejemplo tomado de la autora p. 89.

<sup>28</sup> En lenguas como el italiano y el francés existen indicadores morfosintácticos para diferenciar ambos tipos de verbos (ver Mendikoetxea, 1999 y Suárez, 2016). Estos recursos son inexistentes en español. Para nuestra lengua, se utilizan algunos diagnósticos de inacusatividad que, no obstante, en muchos casos no son conclusivos (Ver Suárez, 2016), tales como la participación en construcciones de participio absoluto (*Terminada la tarea*) y de sujetos desnudos (sin determinantes), como en *¿Quedan personas?* Van Valin (1990) incluye la aspectualidad como un aspecto importante de la inacusatividad. Para más información sobre este particular, ver Batsiukova (2004).

los verbos inergativos cuentan con un sujeto agente, que cumple la misma función que el sujeto de los verbos transitivos; por tanto, suele aparecer antepuesto al verbo (Beletti, 1987; Fernández Soriano, 1993; López Meirama, 2006; Leonetti, 2018). Los cuestionamientos o adscripciones a estos conceptos, que son muchos<sup>29</sup>, quedan fuera del alcance de la presente investigación, por lo que se usarán las nociones de inergatividad e inacusatividad de manera operativa, y según las cualidades semánticas referidas a la agentividad del referente que denota el sujeto en cada construcción.

Con respecto a la literatura científica, en Alfonso y Melis (2010) se pone en evidencia mediante un análisis de lengua en uso que, si bien los verbos inergativos muestran una preferencia por la anteposición del sujeto (74% de la muestra), los verbos inacusativos también presentan sujeto preverbal en el 54% de los casos, lo que habla en contra de una preferencia por el orden VS. A un resultado similar llega Ortiz-López (2009) en una investigación sobre las secuencias SV y VS en el español del Caribe mediante la aplicación de encuestas de elección de orden y juicio de aceptabilidad. Los sujetos participantes prefieren el orden SV para los verbos inergativos en un 83% de los casos; para los inacusativos prefieren la misma secuencia en un 71% de los casos. Volviendo al trabajo de Alfonso y Melis (2010), estos autores postulan que el orden de palabras en estos predicados es predecible a partir de la presencia mayoritaria de modificadores circunstanciales (88% de las cláusulas). Así, cuando el circunstante está pospuesto al verbo, el sujeto aparece a la izquierda de este (95% de las oraciones); cuando está en posición inicial, el sujeto se pospone al verbo (80% de las oraciones)<sup>30</sup>.

Otras investigaciones (Kahane, 1950; Soriano, 1993) habían ya apuntado que la anteposición de complementos adverbiales condiciona la posposición del sujeto, como es el caso del siguiente ejemplo: *En Navidad te dará Juan el regalo* (Soriano, 1993, p. 122). También Fernández (1986) postula que el complemento directo y los complementos adverbiales y preposicionales guardan una relación muy estrecha con el verbo, lo cual causaría que, si se anteponen a este por razones expresivas, se posponga el sujeto: *Todo lo*

---

<sup>29</sup> Por solo poner un par de ejemplos, en López Meirama (2006) y Suárez (2016) puede encontrarse una discusión en torno a la inacusatividad de los verbos de movimiento.

<sup>30</sup> En López Meirama (2006) se puede encontrar una interesante reflexión en torno a la relación entre la inacusatividad de los verbos y lo que se ha denominado como ‘Inversión locativa’ (*locative inversion*).

*absorben entonces las empresas imperiales de España.* Este fenómeno, según afirma el autor, es frecuente tanto en la lengua literaria como en la hablada. Más recientemente, y desde el punto de vista histórico, Bogard (2019) demuestra que la posposición del sujeto en construcciones transitivas está en estrecha relación con la existencia de un constituyente diferente del sujeto en posición preverbal (p. 23).

Silva-Corvalán (1982) atribuye este fenómeno de la posposición del sujeto respecto del verbo cuando hay elementos adverbiales en el primer lugar de la oración a que estos usualmente introducen verbos de carácter presentativo, cuyo sujeto es información nueva y, por tanto, pospuesta: *Atrás venían los médicos.* En esta misma cuerda pragmática, Mendikoetxea (1999) se refiere al fenómeno de “inversión locativa” que tiene lugar frecuentemente con verbos inacusativos que denotan existencia y aparición. Se trata de estructuras en las que se antepone el complemento locativo y, como consecuencia, el sujeto se coloca después del verbo de manera reglamentaria: *En el bosque existen hadas y enanitos;* con lo cual posponer el elemento circunstancial generaría una secuencia agramatical: *\*Hadas y enanitos existen en el bosque.* Parece, sin embargo, que la secuencia agramatical se debe más, en este caso específico, a la anteposición de sustantivos escuetos que a la posposición del complemento locativo. En efecto, la secuencia *Existen hadas y enanitos en el bosque* es perfectamente gramatical.

En cualquier caso, según propone la autora, la posposición del sujeto en oraciones declarativas que tienen el complemento locativo antepuesto tiene su fundamento en la función discursiva de este tipo de construcción, vale decir: “la de introducir información conocida (la que aporta el sintagma preposicional locativo), sin que la información que aporta el argumento tema tenga que ser necesariamente nueva en el discurso” (p. 1613). Es así que la autora relaciona la anteposición de los complementos circunstanciales y la posposición del sujeto solo con determinados verbos, semánticamente más ‘ligeros’, es decir, verbos que aportan la existencia de una entidad (están incluidos acá también verbos inergativos, con este matiz semántico), en cuyo significado está incluida la noción de locación, pudiendo aparecer esta de manera implícita o explícita.

### 1.4.3.1. Los verbos pseudoimpersonales

Mención aparte merece una clase verbal que muestra un comportamiento semántico y sintáctico bastante homogéneo. Se trata de los verbos pseudoimpersonales, descritos por Alcina y Blecua (1975)<sup>31</sup>, cuyas construcciones tienen las siguientes particularidades (Melis y Flores, 2007, p. 8):

- i) seleccionan preferentemente sujetos inanimados, no agentivos, que se colocan después del verbo en la posición del objeto, y que no se analizarían como sujetos de no ser por las marcas de concordancia con el verbo;
- ii) presentan con mucha frecuencia una estructura oracional en función de sujeto;
- iii) se combinan con un objeto, al parecer regido por el verbo, de referente por lo general humano, que, contrariamente a lo que hace esperar el modelo de la oración biactancial en español, se codifica como objeto indirecto y no como objeto directo;
- iv) permiten que este objeto indirecto ocupe la posición inicial de tópico y, en virtud de su naturaleza humana y su prominencia, cobre apariencia de sujeto.

Estos verbos suelen tener, salvo casos marginales, un orden OI-V-S, dada la topicalidad del objeto indirecto antes mencionada, que caracteriza a la clase verbal. Los verbos de afección psíquica como gustar son los más icónicos. Existen, no obstante, otros dos tipos, según estiman Melis y Flores (2007): los existenciales (*bastar, suceder*) y los modales (*parecer, convenir*).

### 1.4.3.2. Instancias de descripción de la variable

Teniendo en cuenta los referentes investigativos anteriormente mencionados, y con el objetivo de emplear una clasificación simple de los predicados que permita agilizar el proceso de descripción y posterior análisis estadístico, al tiempo que pueda recoger la mayor cantidad de información diferenciada posible, decidimos diferenciar tres instancias descriptivas. La

---

<sup>31</sup> Entre los más comunes encontramos abrumar, afligir, agobiar, agradar, alarmar, alegrar, aliviar, angustiar, apenar, apetecer, bastar, competir, complacer, conmover, convenir, costar, deprimir, divertir, disgustar, doler, exasperar, fastidiar, gustar, honrar, importar, impresionar, incumbir, interesar, molestar, mortificar, pesar, preocupar, repugnar, sorprender, urgir, etc. (Hernanz, 1999).

primera de ellas tuvo como propósito conocer la estructura argumental del verbo en cuestión. Para ello, no se tuvo en cuenta la naturaleza sintáctica del argumento, es decir, no se hicieron distinciones entre argumentos externos, internos u oblicuos.

Siguiendo la propuesta de García Miguel (2007), en la determinación de la valencia verbal se contemplaron no solo los argumentos semánticamente obligatorios, sino cualquiera que estuviera relacionado con el sentido del verbo en el uso que se le dé al mismo. Partiendo de referentes teóricos previos, este autor establece una diferenciación entre el “potencial valencial”, propio de un elemento léxico, que recoge los argumentos potenciales o posibles para la clase léxica del verbo, aunque estos estén expresados o no; y la ‘realización valencial’ que consiste en la expresión de determinados argumentos, en una “interacción entre el potencial valencial del verbo, la construcción sintáctica de la cláusula y factores discursivos” (p. 24). Esta diferenciación parte de la convicción de que los verbos o clases verbales no están necesariamente asociados con determinados argumentos de manera inherente, sino que puede existir más de un esquema semántico-sintáctico para un predicado, el cual emerge de la variabilidad presente en las experiencias lingüísticas comunicadas por los hablantes (Alfonso & Melis, 2011). Según señala García Miguel (*op. cit.*) la estructura argumental de un verbo habla de la frecuencia con que se expresan ciertos participantes en una escena denotada por el verbo. Es por ello que, en rigor, los verbos pueden contar con una o más listas de roles, según la frecuencia con que estos se presenten.

Planteado este modo de proceder, la tipología que usamos en la clasificación es la empleada por la RAE (2009) y los estudios gramaticales tradicionales, vale decir, diferenciamos entre:

a) predicados monovalentes o de un lugar, o sea, que presentan un solo argumento. Ej:

*¿Volverá Papá?* (Ant28<sup>32</sup>)

b) predicados bivalentes o predicados que presentan dos argumentos. Ej: *¿Alguno de los factores siguientes producen la salida de orina?* (And2)

c) predicados trivalentes, que expresan tres argumentos. Ej: *¿Me da usted la venia?* (Esp14)

---

<sup>32</sup> Esta nomenclatura corresponde a la abreviatura de la zona a la que pertenece el ejemplo, seguida del número de enunciado dentro del corpus.

Un segundo nivel de análisis tuvo en cuenta la transitividad del verbo, entendida como la capacidad que tienen algunos verbos de requerir un complemento directo (como *comprar*, *hacer*, *poner*). La definición de transitividad de Campos (1999) añade que el complemento directo de los verbos transitivos es de carácter nominal y es regido por el verbo sintácticamente y semánticamente. Contrariamente, los verbos que no llevan complemento directo (como *florecer*, *trabajar*) se denominan intransitivos (RAE, 1973). Dentro de ellos, se hizo una distinción entre verbos inacusativos e inergativos, dado que estos se conocen como dos clases sintácticas dentro de los predicados intransitivos (Mendikoetxea, 1999).

Quedaron excluidos de esta clasificación de la estructura argumental, desde luego, los predicados atributivos, por no ser susceptibles de un análisis de estructura argumental<sup>33</sup>, es decir, siguiendo la propuesta de Campos (1999), los verbos copulativos no rigen semánticamente a su complemento, por lo que no pueden ser verbos transitivos. Para ellos, se estableció una etiqueta que permitiera su diferenciación con respecto a las estructuras predicativas.

En un tercer nivel, añadimos a la descripción el factor presencia/ausencia de modificadores circunstanciales (CC) y su posición en la cláusula (antepuestos, pospuestos, e intercalados).

A continuación, se muestra cómo quedó conformada esta variable y sus niveles de análisis:

---

<sup>33</sup> Estas estructuras carecen de un verbo que sea núcleo predicativo. La relación entre sujeto y predicado es de carácter ecuativo o identificativo (RAE y ASALE, 1999). Por tanto, no puede hablarse de rección verbal.

| <b>Instancias de descripción</b>                    | <b>Niveles de descripción</b> |
|---|-------------------------------|
| <b>Realización valencial del predicado</b>          | Monovalente                   |
|   | Bivalente                     |
|   | Trivalente                    |
| <b>Transitividad verbal</b>                         | Transitivo                    |
|   | (Intransitivo) Inergativo     |
|   | (Intransitivo) Inacusativo    |
| <b>Presencia de modificador circunstancial (CC)</b> | Sin CC                        |
|   | Con CC antepuesto             |
|   | Con CC pospuesto              |
|   | Con CC intercalado            |

#### **1.4.4. Temporalidad verbal**

Esta variable ha sido poco estudiada con respecto al tema que nos ocupa. En la bibliografía revisada encontramos escasas referencias y dos de ellas ofrecen evidencia dispar. Lipski (1977) hace referencia a que los tiempos presente, pretérito e imperfecto favorecen, en ese orden, la anteposición del sujeto. Sin embargo, el propio autor reconoce que no encontró en su investigación evidencia suficiente para postular un orden para los demás tiempos.

Rodríguez (2008) plantea que las oraciones con verbo compuesto favorecen la anteposición, lo que podría asociarse con el uso de determinados tiempos verbales, en tanto Brown y Rivas (2011) tuvieron en cuenta la variable temporal sin encontrar una influencia significativa en lo que respecta al orden sujeto-verbo.

Seguimos en nuestra investigación la clasificación temporal propuesta por las gramáticas tradicionales. En consecuencia, agrupamos los verbos en las siguientes categorías temporales, tomadas de la Nueva Gramática de la Lengua Española, RAE y ASALE (2009):

i. Tiempos simples:

a) Modo indicativo:

- (1) Presente (canto)
- (2) Pretérito perfecto simple (canté)
- (3) Pretérito imperfecto (cantaba)
- (4) Futuro simple (cantaré)
- (5) Condicional simple (cantaría)

b) Modo subjuntivo:

- (1) Presente (cante)
- (2) Pretérito imperfecto (cantara o cantase)
- (3) Futuro simple (cantaré)

ii. Tiempos compuestos:

a) Modo indicativo:

- (1) Pretérito perfecto compuesto (he cantado)
- (2) Pretérito pluscuamperfecto (había cantado)
- (3) Pretérito anterior (hube cantado)
- (4) Futuro compuesto (habré cantado)
- (5) Condicional compuesto (habría cantado)

b) Modo subjuntivo:

- (1) Pretérito perfecto compuesto (haya cantado)
- (2) Pretérito pluscuamperfecto (hubiera o hubiese cantado)
- (3) Futuro compuesto (hubiere cantado)

Una complicación adicional sobreviene cuando pensamos en las perífrasis y locuciones verbales, pues a menudo en estas estructuras se mezclan valores modales, temporales y aspectuales, contrario a lo que sucede con los verbos sintéticos<sup>34</sup> (Genta, 2008; Real Academia, 2009). En consecuencia, estimamos que una solución para evitar la ambigüedad en la clasificación, y facilitar el posterior análisis estadístico, es establecer una categoría independiente para estas construcciones verbales. Entendemos por perífrasis verbal aquella “estructura compleja formada por un verbo auxiliar, con mayor o menor grado de gramaticalización, y una forma auxiliada formada por una estructura nominal (infinitivo, gerundio, participio)” (Genta, 2008, p. 43), que no da lugar a dos predicados diferentes<sup>35</sup> (RAE y ASALE, 2009), como es el caso de *ir a leer*, *estar estudiando*, *llegar a ser*. Por su parte, se ha entendido por locuciones verbales aquellas expresiones que tienen como núcleo un verbo y “que forman parte del léxico como fórmulas fijas: *querer decir* (significar), *hacer saber* (comunicar)” (RAE y ASALE, 2009, p.2131). Se le asignó a cada caso el tiempo morfológico del verbo principal.

#### 1.4.5. Morfología del sujeto

Esta variable tiene importancia tanto si hablamos de presencia/omisión del sujeto gramatical como del orden que adquiere este último con respecto al verbo. Con respecto al primer elemento, Morales (1997) plantea que son los pronombres de primera y segunda persona los que menos se omiten en oraciones declarativas. Por su parte, Fernández Ramírez (1986) menciona que *usted* es el pronombre que menos se omite por razones de cortesía y equívocidad. En un análisis hecho por el autor a partir de un corpus literario, este pronombre aparece más frecuentemente detrás del verbo (82 ejemplos vs 17, en un análisis estadístico aplicado a un corpus literario). El autor hace alusión a determinadas estructuras, fijadas por el uso, como es el caso de *Como usted guste* y *Como usted comprenderá*. Cuando este

---

<sup>34</sup> No sucede así, desde luego, con todos los tipos de perífrasis verbales. En el caso de las perífrasis *ir a+infinitivo* está bastante claro el valor temporal, salvo en los usos correspondientes a las interrogaciones orientadas y las expresiones exclamativas y lexicalizadas (RAE y ASALE, 2009).

<sup>35</sup> Del mismo modo que con las perífrasis verbales, las locuciones conforman un solo grupo verbal (aunque con características sintácticas diferentes). Hemos marcado las existentes en nuestro corpus (aunque muy escasas) como parte del grupo de perífrasis verbales.

pronombre se pospone al verbo, suele formar parte de perífrasis verbales, como en *Tienen ustedes que encontrarse*, *Puede usted pasar*, etc.

En lo referente a la posición que adquiere el sujeto con respecto al verbo, en un estudio sobre interrogativas parciales en Puerto Rico, Lantolf (1980) cita los aportes de otros autores en torno a la influencia del factor tipo de sujeto en la anteposición o posposición de este. Así, Bergen (1976) establece que la anteposición es generalizada con los pronombres *tú*, *yo*, *vos*, *usted* y *ustedes*, en tanto la posposición es obligatoria con *nosotros* o con un nombre propio. Sobre este presupuesto, el primer autor mencionado afirma que es el pronombre *yo* el que mayor ocurrencia tiene en posición antepuesta. *Nosotros*, *ustedes* y *él*, en ese orden, tienden a favorecer la posposición. El único sustantivo propio testeado (*Juan*) obtuvo 52% de anteposición versus 48% de posposición, por lo que parece haber al menos en este corpus una alternancia en el orden. Otras investigaciones sobre la influencia que tiene este último tipo de sujeto en la linealidad de las preguntas arrojan que puede, en efecto, ser un factor importante. Lipski (1977) había propuesto que el sujeto lexical es un elemento determinante en la anteposición del sujeto, en tanto Rodríguez (2008) encuentra una tendencia semejante. Kahane (1950), por su parte, había afirmado que tanto los pronombres personales como los demostrativos tienden a anteponerse al verbo.

En estudio sobre la posición del sujeto respecto del verbo con un corpus de lengua escrita, Delbecque (1988) encontró que los sujetos clausulares, los infinitivos y los sujetos constituidos por sustantivos sin determinantes son los más propensos a la posposición y los pronombres personales, demostrativos e indefinidos muestran una tendencia contraria<sup>36</sup>. Los sustantivos introducidos por artículo o singularizador, los más frecuentes en el corpus analizado, muestran una clara preferencia por la anteposición.

En Bentivoglio (2003) puede encontrarse un estudio de la influencia que tienen los sujetos pronominales, nominales, verbales y de oración subordinada en la determinación del orden en oraciones declarativas. Solo los sujetos conformados por verbos en infinitivo tienen un alto índice de posposición (89% de los casos). Con respecto a las formas pronominales,

---

<sup>36</sup> La autora aclara, sin embargo, que los resultados de probabilidades alcanzados en su estudio para los factores mencionados, si bien muestran una tendencia clara hacia la anteposición o la posposición del sujeto, alcanzan no más del 20% de la muestra, por lo cual tienen un valor relativo (p. 163).

se observa en esta investigación una diferencia entre el corpus oral y escrito: el pronombre se pospone más en la muestra de escritura que en la de habla.

Para la presente investigación, se agruparon los sujetos en las siguientes categorías:

- pronombre de primera persona del singular.
- pronombre tú (segunda del singular).
- pronombre usted (segunda persona del singular).
- pronombre de tercera persona del singular.
- pronombre de primera persona del plural.
- pronombre de segunda persona del plural.
- pronombre de tercera persona del plural.
- pronombre posesivo.
- pronombre demostrativo.
- pronombre indefinido.
- pronombre numeral.
- sustantivo común, solo o con artículo definido.
- sustantivo con singularizador (posesivo, demostrativo, indefinido, cuantificador o relativo)
- nombre propio o título.
- forma verbal de infinitivo (con o sin complementos).
- sujeto clausular.

#### **1.4.6. Referencialidad del sujeto**

Las categorías referenciales (o parámetros referenciales) están asociadas a la naturaleza semántica de una expresión y su forma de referir (Von Heusinger y Kaiser, 2003). Existen varias subcategorías, como animacidad, definitud, genericidad, especificidad, género, número, entre otras. Por la naturaleza de nuestro corpus, en esta tesis nos referiremos

específicamente a la primera y la última de ellas.

#### **1.4.6.1. La animacidad**

La **animacidad** es una categoría léxica que describe una propiedad del referente designado (Von Heusinger & Kaiser, 2003, p. 43). Está relacionada, más concretamente, con las características prototípicamente humanas de ese referente (Ji & Liang, 2018). Esta categoría parece tener alguna influencia en fenómenos sintácticos de diversa naturaleza. Por ejemplo, se sabe que la interacción entre la animacidad, la definitud y la especificidad puede tener consecuencias sintácticas: puede mediar en la selección del modo verbal, en la marcadez del objeto directo, entre otros fenómenos (ver al respecto Von Heusinger & Kaiser, 2003; Aissen, 2003; Lizárraga & Mora-Bustos, 2010). Algunos autores han estudiado la relación entre animacidad y agentividad, entre ellos Fox (1995, citado por Rivas 2008), quien revela que existe una correlación entre ambos rasgos semánticos: el 97% de los sujetos agentivos contenidos en su estudio tienen también el rasgo +humano. López Meirama (1997, citado por Rivas 2008, p. 900) menciona que el sujeto de los verbos con participantes agentivos (*blasfemar, rechistar, trabajar*) aparecen regularmente delante del sujeto, mientras que los que no exigen un participante agentivo (*brillar, refulgir o resplandecer*), tienden a posponerse. A partir de esta evidencia, cabría suponer que el rasgo +humano puede influir de manera determinante en la linealidad del enunciado, dada su asociación con la agentividad de los sujetos en un número importante de casos.

También con respecto a la relación entre la agentividad/pasividad y el rasgo de animacidad, es conocida la generalización acerca de que los sujetos inanimados tienden a favorecer el orden VS (Hatcher, 1956), dado su carácter de paciente; mientras que los sujetos animados se anteponen al verbo, pues tienden a asumir un rol agentivo que los lleva a convertirse en tópico (De Miguel, 1989; Alonso Cortés, 2001).

De lo anterior da cuenta Bentivoglio (2003), cuya investigación arrojó que los sujetos animados tienen índices bajos de posposición con respecto al verbo (30% en el corpus escrito, 13% en el oral). Si bien la autora no establece una categoría de animacidad, puede deducirse que en la medida en que los sujetos pierden este rasgo, aparecen pospuestos en mayor proporción. La mayor proporción la exhiben los sujetos que denotan procesos o acciones,

con un 93% en el corpus hablado y un 55% en el corpus escrito. También Rodríguez (2008), en su estudio de corpus diacrónico, da cuenta de un comportamiento semejante. El autor observó que en las oraciones con sujetos inanimados y objetos humanos se produce la posposición del sujeto en la totalidad de los casos (100% de ocurrencia). Sin embargo, la confiabilidad de los resultados podría cuestionarse, pues las oraciones con este tipo de constituyente conforman solo el 0,5% de su muestra. Rosemeyer (2018) realizó un estudio de orden del sujeto en interrogativas parciales del español Caribe y el portugués brasileño; sus hallazgos confirman que la animacidad es un factor que favorece la anteposición del sujeto.

Dialogando con estas afirmaciones, aparecen las investigaciones de Rivas (2008), quien no observa una preferencia por sujetos animados o inanimados en posición posverbal, salvo en los datos provenientes del corpus oral que formó parte de su análisis. En estos casos, los sujetos inanimados triplican en número a los animados. Por tanto, el autor estima que la animacidad no es un criterio fiable para predecir el comportamiento del orden de constituyentes. A ello se añaden las aportaciones de Alfonso y Melis (2010), quienes encontraron que el 64% de los sujetos animados aparece antepuesto con respecto del verbo, pero que también el 54% de los inanimados se antepone, lo que estaría hablando de una ligera tendencia a la anteposición en este último tipo.

En esta investigación se siguió la escala propuesta de Aissen (2003), y se dividió la categoría animacidad en tres subcategorías: **humano**, **animado** e **inanimado**. A su vez, esta escala cubre una serie de categorías léxicas que exhiben mayor o menor animacidad. Usamos la conocida escala de Dixon (1979) que va de lo más humano, como prototipo de animacidad (siendo el centro el emisor del discurso) a lo menos animado (o inanimado). Así, la mayoría de las lenguas exhibe el orden de animacidad reproducido a continuación:

### **Escala de animacidad**

Pronombre de primera persona > pronombre de segunda persona > pronombre de tercera persona y pronombres demostrativos > sustantivos propios > sustantivos comunes humanos > sustantivos comunes animados<sup>37</sup>> sustantivos comunes inanimados<sup>38</sup>.

A tono con Rivas (2008), cuando se trataba de pronombres personales, se atendió a la animacidad de sus referentes. Hay que aclarar también que la pertenencia de una entidad a una de estas categorías no es algo que pueda hacerse a partir de criterios definidos, pues en muchos casos las concepciones en torno a ciertos objetos pueden variar de persona a persona (Von Heusinger & Kaiser, 2003), por lo que es fundamental tener en cuenta el contexto lingüístico para colocar cada sujeto gramatical dentro de una determinada clasificación. En esta misma línea, Ji y Liang (2018) afirman que la determinación de la animacidad del referente puede ser en ocasiones una tarea difícil, dado que es una noción que no depende solo de los atributos biológicos del referente denotado, sino del tratamiento que haga el hablante de este referente, es decir, que el modo de ‘ver’ la entidad es definitorio en la determinación de sus cualidades más o menos animadas.

De acuerdo con los estudios cognitivistas, la animacidad puede verse como una jerarquía en la que hay miembros más o menos prototípicos: “el ser humano puede considerarse como el prototipo de la categoría de animación, mientras que el ‘yo’ puede ser el prototipo dentro de la categoría humana. Los atributos del ser humano y el yo constituyen las características prototipo de la animación, y la ‘tipicalidad’ de otros miembros de la categoría depende de su similitud con el prototipo o de cuántas características prototípicas posean” (Ji y Liang, 2018, p.73). De acuerdo con lo anteriormente expuesto, las características del referente son asignadas por el hablante según ciertos niveles de empatía. Puede, en consecuencia, haber un tratamiento animado de una entidad que no lo es en un estado real de cosas. Además, la animacidad varía según diferentes contextos: por ejemplo, cuando nos referimos a un pato, podemos estar haciendo referencia al animal, a un tipo de

---

<sup>37</sup> Acá se incluyen los animales.

<sup>38</sup> Estas entidades nunca funcionan como ‘agentes’ controladores de los procesos, si bien en algunos casos, como se detallará más adelante, pueden contener diversos grados de animacidad en función de la empatía o interpretación de los hablantes.

carne comestible, incluso a un juguete<sup>39</sup>. Por último, es bastante frecuente encontrar en el habla ejemplos del paso de lo inanimado o abstracto a lo animado mediante un proceso de metáfora o metonimia. En nuestro corpus abundan casos de este tipo: *¿Puede el periodismo decir la verdad?* (Ant 7).

Ahora bien, la escala propuesta anteriormente no da cuenta de lo que sucede al interior de las entidades inanimadas, que suelen ofrecer un abanico de posibilidades respecto de la identificación con un referente más o menos animado. En la presente investigación se han asumido los aportes de Ji y Liang (2018) respecto del tratamiento de la animación en entidades inanimadas. De acuerdo con la propuesta de los autores, tanto la animación como la inanimación no son categorías discretas, sino espacios categoriales con gradientes, donde la interpretación y la concepción de los hablantes desempeñan un papel importante. Partiendo de las jerarquías tradicionalmente propuestas para las entidades animadas, los autores proponen la siguiente gradación, que va de mayor a menor animación dentro de los sustantivos inanimados:

#### **Escala de animación en sustantivos inanimados**

sustantivos colectivos > sustantivos espaciales y temporales > sustantivos concretos > sustantivos abstractos psicológicos > sustantivos abstractos no psicológicos.

A continuación, ahondaremos un poco en cómo se comporta el gradiente anteriormente mencionado al interior de cada una de las categorías:

-sustantivos colectivos: pueden ser entidades físicas o psicológicas en función de la concepción que el hablante tenga de ellos. Por ejemplo, pueden hacer referencia a una comunidad u organización, en tanto grupo de personas como en *¿Y el Estado está apuntando en esa dirección?* (Car84), en cuyo caso son capaces de ejercer acciones humanas; o pueden hacer alusión a entidades inanimadas como ‘la escuela’ en la expresión ‘ir a la escuela’.

-sustantivos espaciales y temporales: hacen referencia a orientaciones (norte, sur), localizaciones generales (área, zona) y específicas (ciudad, país). Son consideradas entidades físicas, salvo en casos en los que se usen metonímicamente, pasando a aludir a las personas

---

<sup>39</sup> Ejemplo tomado de (Ji y Liang, 2018).

que se reúnen (o que existen) en estos espacios, como en *¿El Sur existe para el Norte?* (Chi73), o en *¿Negociará Israel con dicho Ejecutivo?* (Esp60).

-sustantivos concretos: hacen referencia a objetos abstractos o sustancias (sol, agua), artefactos (auto, casa), comidas (pan, leche), y a fenómenos naturales (tormenta, lluvia). Denotan entidades físicas, aunque en algunas culturas son concebidas como auto-propulsadas, o sea, capaces de movimiento autónomo, como sucede con los seres humanos.

-sustantivos abstractos psicológicos: si bien se ubican en la parte más baja de la escala, dado que los sustantivos abstractos no poseen características físicas ni psicológicas, en ocasiones pueden portar rasgos de cognición humana, sensibilidad, comunicación o actividad psicológica o social, en cuyo caso serán contemplados como psicológicos. Pueden en ocasiones designar por metonimia a las personas que instigan esos eventos psicológicos, como en la frase ‘los estudios muestran’.

-sustantivos abstractos no psicológicos: son considerados en su mayoría como portadores de atributos inanimados (precio, tarifa), estados (condición), eventos (explosión, consecuencia), procesos (inflación, reducción), relaciones (contrariedad, correlación) y formas (arco, espiral). Estos últimos son los que ocupan el último lugar en la escala de animacidad a la que hemos hecho referencia.

Además de la clasificación antes propuesta, puede usarse una prueba operativa en casos confusos: se trata de la marca diferenciada de objeto para identificar los sujetos animados. En pocas palabras, los sujetos que tienen el rasgo de animacidad son susceptibles de llevar la preposición *a* si cambian su función a la de objeto directo<sup>40</sup>. Así se pone de manifiesto en los siguientes ejemplos, tomados de la RAE (2009), donde existe una posibilidad de alternancia en el uso de la preposición *a*: *Es difícil parar (a) un tren en marcha; atacar (a) los aviones o bombardear (a) las ciudades costeras*. Tal como se comenta en esa obra, “la elección de la preposición acentúa el proceso de personificación —y, por tanto, el grado de animacidad que

---

<sup>40</sup> Si bien se ha demostrado que no es una condición taxativa, la presencia de la animacidad y la definitud tienen gran influencia en la marcación diferenciada de objeto (para más detalles, consultar Lizárraga y Mora-Bustos, 2010).

se atribuye a esos seres—, mientras que, si se evita, se entiende que se habla de cosas materiales sin capacidad de acción” (p.4145)<sup>41</sup>.

Por tanto, en casos como *¿El Estado está apuntando en esa dirección?*, puede observarse el grado de animacidad del sustantivo colectivo mediante su transposición en una estructura eventualmente transitiva, susceptible de recibir marca de objeto diferencial. Por ejemplo, nos resulta más natural la estructura “Tenemos que observar al Estado” que “Tenemos que observar el Estado”, lo que ofrece una pista inconfundible de la interpretación de esta entidad como animada por parte del sujeto emisor.

#### 1.4.6.2. La definitud

El segundo índice de referencialidad que tendremos en cuenta en la presente investigación es la **definitud**, noción de naturaleza gramatical que es útil para la construcción de una interpretación por parte del receptor (Leonetti, 1999). Así, los determinantes, entre los que se encuentran en español los artículos, los determinantes demostrativos, posesivos y cuantificadores y elementos léxicos que denotan identidad o cantidad, como *diversos*, *diferentes*, *numerosos*, etc (Laca, 1999), “permiten conectar la información léxica contenida en el sintagma nominal con la información contextual que los hablantes emplean para construir la interpretación de los enunciados” (Leonetti, 1999, p. 789).

Como afirman Von Heusinger y Kaiser (2003, p. 44), “la definitud (...) indica que el referente discursivo asociado con la expresión definida puede ser identificado con un elemento o entidad ya introducido en el discurso”. De este modo, las personas u objetos suelen ser introducidas en el discurso mediante un artículo o determinante indefinido, en tanto entidad no identificable (referente no conocido para el receptor), mientras que las subsecuentes menciones suelen estar anteceditas por determinantes definidos, como parte de una identificación esperable, a partir de un conocimiento previo (Du Bois 1980, p. 206-207).

---

<sup>41</sup> Para profundizar en este tema de la marca de objeto diferencial (DOM, por sus siglas en inglés), consúltense a Von Heusinger y Kaiser (2003) y Aissen (2003).

Leonetti (1999) apunta que esta identificación puede ser mediante percepción directa, mención previa en el mismo discurso, o mediante conocimiento extralingüístico compartido.

Hay que apuntar que esta relación entre la familiaridad de los hablantes con el referente a que hace alusión el sujeto y la definitud es la más manejada por la tradición gramatical. Sin embargo, como afirma el autor antes citado, en ocasiones el artículo definido no aporta información conocida o dada. Es el caso de sintagmas nominales ricos en descripciones, altamente informativos como sucede con *la derrota de los sarracenos en la batalla de Poitiers*, cuyo empleo sirve para introducir información nueva. Del mismo modo, el artículo definido puede introducir información no mencionada anteriormente o integrada en el universo discursivo; es lo que ocurre cuando el emisor supone que el receptor podrá identificar el referente, sin que sea familiar. Ej: *Cuidado con el escalón*.

Si bien la lista de determinantes y lexemas que indican definitud (o indefinitud) es extensa y heterogénea en español, citaremos aquí los casos más icónicos, que servirán a la investigadora para reconocer de manera más expedita a qué categoría se ajusta determinado sujeto. Están incluidos en la categoría formal de definitud los pronombres definidos tanto personales como demostrativos (*yo, tú, aquel*); los nombres propios (*John, Londres*); y las frases nominales con modificadores de la categoría de artículo definido (*el niño*), posesivos (*su bicicleta*) y cuantificadores universales (*todos los estudiantes, cada día*).

Por su parte, se clasifican como indefinidas aquellas construcciones nominales precedidas por un artículo indefinido (*un chico*); los sustantivos sin artículo que denotan masa (*había agua en el jarrón*); los sustantivos plurales contables (*piñas*); sustantivos plurales precedidos por cuantificadores numerales (*tres chicos*), o indefinidos (*alguien, algo, varios, ningún*). De acuerdo con la tradición gramatical, se establecerán tres categorías: sintagma nominal **definido**, sintagma nominal **indefinido** y sintagma nominal sin determinante (ver Laca, 1999) o **escueto**<sup>42</sup>.

En lo referente al orden de palabras, se ha planteado la existencia de una jerarquía de

---

<sup>42</sup> Herrera (2017) propone insertar los sintagmas nominales en construcciones existenciales con haber, como prueba para diferenciar definitud e indefinitud de los determinantes. Así, los determinantes definidos no pueden aparecer en este tipo de construcción, contrario a lo que sucede con los indefinidos.

‘definitud’ en la determinación de la posición que ocupa un argumento, según el valor referencial del sujeto. Así, los elementos definidos suelen ser más topicales (y por tanto tienden a anteponerse) que los indefinidos, y estos a su vez, son más topicales que los no referenciales (o sea, los que carecen de elemento determinante). La siguiente clina ha sido tomada de López Meirama (2006).

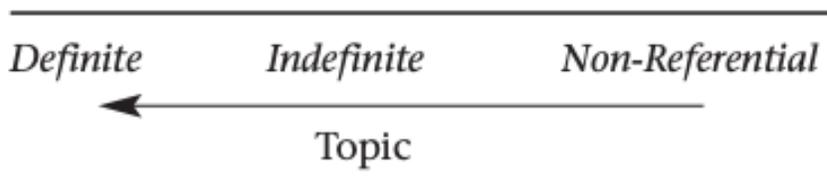


Figura 2. Jerarquía de definitud en la determinación del tópico oracional

Abundante bibliografía da cuenta de la validez de esta jerarquía. Fernández Soriano (1993) afirma que el mero hecho de que un sustantivo aparezca sin determinante condiciona que el orden reglamentario sea VS. Laca (1999) explica que el orden anteriormente descrito se debe a que los sujetos de este tipo de estructura son no topicales, es decir, se trata de construcciones téticas, en las que no se predica de un sujeto o el sujeto es foco contrastivo. Los sujetos sin determinante en posición preverbal, por su parte, sirven a funciones discursivas presentativas y son más comunes en la prosa narrativa. Rivas (2008) ofrece evidencia de lo anterior en un análisis de un corpus de oraciones intransitivas; el autor encontró que el 93% de los sujetos en posición preverbal son identificables. “Se puede entonces llegar a la conclusión de que en español existe una tendencia acusada a evitar sujetos no identificables en posición preverbal” (p. 906). Esta misma investigación arrojó que la categoría identificación es más predictiva de la posición del sujeto que la de animacidad, pues sólo el 63,5% de los sujetos preverbales de su corpus posee este rasgo, frente a 93% de sujetos indetificables. Por otra parte, los sujetos no identificables suelen colocarse en posición posverbal (en el corpus oral, por ejemplo, no aparece ningún caso de sujeto preverbal; y en el corpus escrito el número de sujetos posverbales se duplica con respecto a los preverbales). Sin embargo, los sujetos identificables siguen siendo más frecuentes en esta posición que los no identificables. Esto se explica por la rareza de los sujetos no identificables en español. Tendencias semejantes fueron encontradas por el mismo autor en una investigación sobre el orden del sujeto en oraciones principales y subordinadas (Rivas, 2013). Delbecque (1988) explica que ejemplos de este tipo ocupan solo el 3,6% del corpus de su

investigación. Sin embargo, la investigadora identifica 109 casos de un total de 221 en los que el sujeto se antepone. Dada la diversidad de hallazgos en esta cuestión, es importante contar con una mayor cantidad de datos para poder probar o rechazar la hipótesis de que los sujetos no identificables se posponen como regla general.

#### 1.4.7. Complejidad del sujeto

La investigación seguirá, con algunas modificaciones, la propuesta de Delbecque (1998) y Bentivoglio (2003). De este modo, agruparemos los sujetos en los siguientes tipos de estructura sintáctica:

a) unidad nominal mínima (pronombre, nombre propio, sustantivo o infinitivo escuetos, títulos) (UM). Ej: *¿Podrá **Calibán** prescindir alguna vez de ese lenguaje e imponer el suyo, el auténtico?* (Ant66)

b) sintagma nominal compuesto por un elemento nominal y exclusivamente un determinante nominal (det.+SN). Ej: *¿Puede **el periodismo** decir la verdad?* (Ant7)

c) elementos nominales sin determinante, yuxtapuestos o coordinados; también elementos pronominales modificados por adverbios (SNcoord). Ej: *¿Serán **Manny Ramírez y Greg Maddux** las piezas que faltaban para que Joe Torre lograra engranar el motor de los Dodgers de Los Ángeles?* (Car43)

d) elementos nominales sin determinante, acompañados de uno o varios complementos de nombre (SN+compl). Ej: *¿**Alguno de los factores siguientes** producen la salida de orina?* (And2)

e) sintagma nominal con determinante nominal, modificado por un adjetivo o un adverbio (det+SN+mod). *¿**Las razones culturales** verdaderamente primaron sobre las estructurales?* (Mex94)

f) sintagmas nominales introducidos por un determinante, modificados por una frase preposicional o una frase nominal en aposición; también dos sintagmas nominales coordinados, estos últimos pueden tener modificadores simples (det+SN+fprep). Ejs: *¿**Son el medio y nuestra biografía** las fuerzas que moldean el carácter?* (Ant88); *¿**Algunos tipos de liderazgo** son más oportunos que otros en determinadas circunstancias?* (Esp54)

g) sintagmas nominales con determinante nominal, coordinados, que contengan complementos preposicionales o subordinadas adjetivas (det+SNcoord). Ej: *¿No es suficiente el resultado de la autopsia y el tiempo que transcurrió desde que el niño presentó dificultad respiratoria y el momento en que se llamó a los paramédicos?* (Ant56)

h) sintagma nominal compuesto por sustantivo más determinante nominal y una cláusula de relativo (det.+SN+CR) o al menos dos complementos preposicionales. Ej: *¿Este mundo en el que nosotros vivimos es el mejor de todos los mundos posibles?* (Ant16)

i) sujeto constituido por una cláusula sustantiva (CS). Ej: *¿Acaso lo que pasa con el mundo vegetal o animal no se va replicando hasta aplicarse en el ser humano?* (Chi66)

j) sintagma nominal con estructura más compleja, no descrita en las categorías anteriores (SNCompl). Ej: *¿No asoma aquí también la ardua misión humana de trabajar la tierra y dominarla para que produzca frutos perennes de amor, justicia, vida y alegría?* (Esp71)

Bentivoglio (2003) observó que, en el caso de que el sujeto esté constituido solo por un pronombre, un sustantivo con un determinante y un nombre propio, la tendencia es a la anteposición. En cambio, cuando el sujeto está constituido por un verbo, por cláusulas completivas o de relativo, o tiene una estructura compleja (más de un modificador, por ejemplo) aparece pospuesto. A resultados similares había llegado Delbecque (1988): los sujetos clausulares son los que más condicionan la posposición. A este tipo de estructura le siguen, en orden decreciente, los elementos nominales modificados por frases preposicionales, los sustantivos introducidos por un determinante y modificados por adjetivos o elemento coordinado; y los elementos nominales coordinados o yuxtapuestos.

Estos resultados pueden estar relacionados con lo que se ha dado en llamar principio de pesantez sintáctica<sup>43</sup>: “los constituyentes de mayor longitud prefieren ocupar la posición final” (Fernández Soriano (1993), p. 130), esto debido a la dificultad de procesamiento que implica la posición inicial de un constituyente extenso. Obsérvese la mayor naturalidad del

---

<sup>43</sup> Promulgado por Hawking (1983).

segundo de los ejemplos brindados por ella:

(1) Todos los partidarios de la abolición de cualquier tipo de medida coercitiva para con los disidentes de las opiniones gubernamentales manifestaron su intención de firmar el manifiesto.

(2) Manifestaron su intención de firmar el manifiesto todos los partidarios de la abolición de cualquier tipo de medida coercitiva para con los disidentes de las opiniones gubernamentales.

Bogard (2017) también hace alusión a este fenómeno, cuando afirma que las oraciones complementarias o completivas se sitúan detrás del verbo debido a su complejidad estructural, o sea, que “las estructuras sintácticamente más complejas tienden a seguir a las más ligeras, o a ubicarse en el margen derecho de la construcción completa (p. 14). Esto es especialmente evidente cuando se trata de oraciones relativas de sujeto con verbos presentativos, pues este tipo de estructura tiene un alto valor remático (introduce información nueva en el discurso). Sin embargo, el autor plantea que las oraciones de relativo no adjetivas pueden anteponerse al verbo principal: *El que iba junto a ella trató de sacarla del automóvil.*

A propósito del tema que nos ocupa es conocido el trabajo de Bolinger (1954). El autor propone que, efectivamente, los elementos más largos van en posición final porque portan información nueva. Es decir, que hay una concomitancia de factores, uno de naturaleza sintáctica, y otro de naturaleza discursiva, siendo este último el que prima. El autor propone, en otras palabras, que la información nueva suele ser más dilatada, extensa, y que por eso la estructura que le sirve de forma va en posición final. Semejante afirmación podría ser puesta en duda, pues no siempre la información nueva suele ser más extensa o dilatada. Se impone, entonces, un análisis de habla real. La discusión de los datos arrojados por esta tesis podrá dar cuenta de la concomitancia de los factores discursivo y sintáctico o la prevalencia de uno de ellos.

#### **1.4.8. Función pragmática del enunciado interrogativo**

El análisis del rol pragmático de la interrogación en el discurso se relaciona con el orden de constituyentes de la cláusula interrogativa. Es un presupuesto ya aceptado que la oración interrogativa tiene características semánticas y sintácticas que determinan su significado (presencia de palabras interrogativas como pronombres o adverbios, uso de determinados patrones entonativos, presencia de una variable que puede o no ser especificada por la respuesta), que son independientes de su función pragmática, y deben, por tanto, ser analizadas por separado (Sinclair & Van Gessel, 1990; Freed, 1994; Escandell 1988 y 1999). A su vez, se ha estudiado la correlación que existe entre la función comunicativa de la interrogación y su estructura sintáctica (Escandell, 1999; Freed, 1994; Brown & Rivas, 2011a; Aijón & Serrano, 2012). La evidencia empírica muestra que puede haber una relación más o menos directa entre estas variables (Sinclair & Van Gessel, 1990; Freed, 1994, Brown & Rivas, 2011a). Se ha asociado, por ejemplo, el orden VS a la interrogación canónica (es decir, a lo que conocemos como pregunta) y el orden SV a la interrogación ecoica y a la retórica (Brown & Rivas, 2011a; Aijón & Serrano, 2012).

Una parte importante del análisis lingüístico de estas oraciones es conocer cuáles son los elementos que ellas tienen en común en la conversación o en el discurso (en el uso), es decir, qué funciones pueden desempeñar o, lo que es lo mismo, cuál es el significado pragmático de la interrogación. Considerando la naturaleza de nuestro corpus, es pertinente hablar de enunciado, entendido como “la realización de una determinada oración por parte de un determinado hablante en unas determinadas condiciones y circunstancias” (Escandell, 1988, p. 275).

La revisión bibliográfica efectuada para la conformación del marco teórico de esta tesis arrojó que existen tres clasificaciones que pueden servir a nuestros objetivos de investigación. A continuación, las presentamos y comentamos.

Freed (1994) ofrece una taxonomía de la función discursiva de las interrogativas en un corpus conversacional de lengua inglesa. Los 16 tipos de interrogativas establecidos por la autora tienen en cuenta el significado literal de la pregunta (o la información buscada por la interrogación, aun cuando esto no implique la existencia de una respuesta) y se agrupan en

cuatro categorías funcionales, según el propósito del enunciado en el intercambio dialógico (p. 625), que se describirán a continuación. La autora postula la existencia de un *continuum* que va de la información buscada (*information sought*) a la información transmitida por el enunciado (*information conveyed*). Así, el enunciado interrogativo puede situarse entre los extremos de búsqueda de información y el de expresión estilística (relacionado con la información personal que determinados enunciados interrogativos transmiten, a partir del presupuesto de que hay un alto conocimiento de la variable que encierra la interrogación), lo cual implica un movimiento desde el dominio público de información hasta el dominio retórico (situado en el ámbito personal). Resumiremos las cuatro categorías funcionales propuestas por la autora<sup>44</sup>:

I. **Interrogativas de información externa.** Las interrogativas que pertenecen a este grupo demandan información externa al discurso. Se trata de las preguntas como manifestación prototípica de la interrogación, pues el objetivo del emisor es obtener información nueva sobre la vida de las personas, hechos que les son familiares, entorno social o físico, etc. Los tipos establecidos por la autora son los siguientes: (a) información pública, (b) información social, (c) invitación social, (d) información deíctica.

(a) A: Hay una tienda en el centro de la ciudad que vende excelentes productos de belleza...

B: *¿Dónde es eso?*

A: Cerca del teatro municipal.

(b) A: *¿Vienes a la fiesta esta noche?*

B: No sé, depende de si termino temprano mis quehaceres domésticos.

(c) A: Iré a tomarme una cerveza, *¿quieres acompañarme?*

B: Sí.

(d) A: Está llegando una música muy interesante, *¿la escuchas?, ¿de allá afuera?*

---

<sup>44</sup> Los ejemplos incluidos en cada categoría son traducciones libres hechas por la investigadora, dada la necesidad de contar con ejemplos más independientes de contextos demasiado específicos en la lengua original en la que está escrita la propuesta (inglés).

B: Ah sí, es buena.

Ha de destacarse que estas categorías han sido determinadas a partir del corpus que tomó la autora como referencia. Si se tiene en cuenta una práctica comunicativa más general, podría decirse que el primero y segundo tipos forman parte de lo que se conoce como pregunta real o inquisitiva, que busca recabar información sobre algo que el emisor desconoce. Para los efectos de esta investigación, convendría fusionarlos.

II. **Interrogativas de habla.** Como su nombre lo indica, se relacionan directamente con la conversación misma. Esta categoría abarca los enunciados que son utilizados para clarificar, repetir o confirmar información precedentemente brindada por el interlocutor. Al igual que las que pertenecen a grupo anterior, este tipo de interrogativa espera una respuesta. La autora deslinda los siguientes tipos: (a) interrogativas para clarificar información (clarificadoras), que piden información adicional relacionada con el contenido de un enunciado previo, (b) interrogativas repetitivas, que repiten literalmente un enunciado previo (c) interrogativas confirmativas, que verifican, a partir de una repetición, una información previa (dada).

(a) A: Tiene un nieto, pero no es de su hijo, sino del hijo de su segunda esposa.

B: Oh, *¿se ha vuelto a casar?*

A: Sí.

(b) A: *¿Pasaste de largo anoche?*

B: *¿Cómo?*

A: *¿Pasaste de largo anoche?*, quiero decir, no dormiste nada...

(c) A: Dejaron esto en el trabajo y nadie lo quiso, me dije: "Me lo llevo". Creo que está bien.

B: Oh, *¿alguien dejó esto en el trabajo?*

A: Sí.

III. **Interrogativas relacionales.** Se trata de interrogativas que buscan abrir, mantener o cerrar el vínculo conversacional entre hablante y oyente. Están incluidas en este grupo (a) las de foco conversacional, las cuales están orientadas a introducir un elemento que luego se ampliará en la conversación, (b) las de información compartida, que se encargan de introducir

temas conocidos por ambos interlocutores, (c) las fáticas, destinadas a comprobar que el interlocutor sigue el hilo de la conversación y (d) las de elaboración, que promueven la conversación a partir de una especificación en la información contenida en el intercambio comunicativo. Este tipo de preguntas frecuentemente no requiere respuesta.

(a) A: Estuve conversando con él por un rato y empecé a molestarme, *¿sabes lo que me molesta?* Que comiencen a hablar más de los demás...

(b) A: *¿Recuerdas cuando fuimos al cine con Manuel?*

B: Sí

A: Ese día estaba radiante

(c) Imagínate que ella no puede hacer algo por sí misma, *¿sabes a lo que me refiero?* Ni siquiera es capaz que sacrificarse para lograr algo para sí misma, menos lo haría por un amigo.

(d) A: Me parece que últimamente hay problemas de relación en este grupo, *¿no te parece?*

B: No lo sé...

IV. **Interrogativas expresivas.** Agrupa las preguntas cuya respuesta conoce el hablante. El objetivo de la emisión de la pregunta es expresar un punto de vista o transmitir información. Están incluidas en este grupo las preguntas (a) las didácticas, típicas de entornos de enseñanza donde el emisor sabe la respuesta y usa la pregunta para controlar el conocimiento de los alumnos; (b) las retóricas, que transmiten una información u orientan a el punto de vista del emisor, por lo que suponen una declaración; (c) las humorísticas que expresan información desde un punto de vista inesperado; (d) las autodirigidas y (e) las referidas, que hacen alusión a episodios pasados o eventos imaginados (estas últimas se refieren al tipo de pregunta que se usa, por ejemplo, en recuentos, donde se remeda algo que ha dicho alguien antes).

(a) A: *¿Sabes lo que significa IoT?* La sigla, del inglés, designan un fenómeno que se denomina el internet de las cosas (Internet of Things), *¿qué es exactamente?*

(b) A: Salgo 45 minutos antes de una clase. Solo por estar segura de llegar a tiempo.

B: ¿Tanto tiempo?

A: Sí. *¿No será demasiado?* Entonces llego temprano.

B: Yo no llego tan temprano.

(c) A: Olvidé preguntarte algo. Qué vamos a... eh...

B: ¿Qué vamos a hacer? Vamos a tener que dar muestras de sangre. *¿No te conté esa parte?*

A: Ja ja. Qué gracioso.

(d) A: Piénsalo bien, yo creo... ¿qué creo yo? Yo creo que es igual...

B: ¿Tú crees?

(e) A: Estábamos hablando sobre mi lactancia después de los dos años. Me preguntó por qué lo hacía, le dije “*¿Tienes algún problema con eso?*” Ella dijo no...

B: Tu creíste que estaba criticándote...

Esta propuesta ha seguido una metodología coherente para la clasificación, separando el sentido o contenido del enunciado del significado pragmático. El criterio de clasificación no está articulado en función de los objetivos perseguidos por el emisor, sino del contenido de las oraciones. Nos parece que la tipología ofrecida por la autora es bastante exhaustiva y podrá aportarnos significativamente al análisis de la presente investigación, en tanto la muestra proviene de diversas fuentes y realidades comunicativas. Respecto de las categorías funcionales propuestas por la autora, creemos que, por ejemplo, en la primera de ellas (las interrogativas de información externa) no es productivo hacer una separación entre los dos primeros subtipos, pues ambos buscan finalmente el mismo objetivo, que no es otro que pedir información sobre algo que el emisor desconoce. La Tabla 3, continuación, resume la propuesta de clasificación antes referenciada. Se han fusionado los dos primeros tipos de la primera categoría funcional:

| <b>Categoría funcional</b>  | <b>Subcategorías</b>  |
|---|---|
| <b>Interrogativas externas (buscan información/acción)</b>  | -Interrogativas sobre información pública o social<br>-Invitaciones<br>-Interrogativas deícticas  |
| <b>Interrogativas de habla (relacionadas con la conversación misma)</b>                           | -Interrogativas clarificadoras<br>-Interrogativas repetitivas<br>-Interrogativas confirmativas  |
| <b>Interrogativas relacionales (orientadas a la relación verbal/social de los interlocutores)</b> | - Interrogativas de foco conversacional<br>- Interrogativas de información compartida<br>- Interrogativas fáticas<br>-Interrogativas de elaboración   |
| <b>Interrogativas expresivas (transmiten información)</b>   | -Interrogativas didácticas<br>-Interrogativas retóricas<br>-Interrogativas humorísticas<br>-Interrogativas autodirigidas<br>-Interrogativas referidas |

Tabla 3. Propuesta de clasificación de Freed (1994), tomando como eje el contenido literal de la interrogativa y su contexto de uso

Otra de las propuestas de clasificación es la de Rabanales (1996), cuyo sello es la importancia que le concede el autor a la situación comunicativa como elemento fundamental para determinar el significado de un enunciado<sup>45</sup>. La primera distinción que se establece es la de **preguntas reales e irreales**, que tiene en cuenta intención comunicativa del emisor. En las preguntas reales se espera recibir información en torno a una duda auténtica. También llamadas **indagativas o inquisitivas**, estas estructuras constituyen “auténticas preguntas”, como es el caso de *¿Habló contigo mi cuñado?* y *¿Quién te llamó por teléfono?* Las

<sup>45</sup> Sin embargo, puede notarse que sus criterios de clasificación abarcan diferentes ámbitos de manifestación de la lengua. Es así que, si bien la pragmática es el eje fundamental, se siguen también los criterios sintácticos, semánticos y prosódicos. Esta mezcla es, desde nuestro punto de vista, uno de los elementos que debilita la propuesta del autor.

preguntas irreales son, por su parte, una interrogación solo en su forma, pues en ellas “no se indaga, sino que se busca satisfacer otros fines de la comunicación (expresar un sentimiento, hacer una declaración, manifestar duda, sorpresa, enojo, indignación; exhortar, ordenar, disuadir, etc.), que no implican por lo mismo otro enunciado que funciona como respuesta” (p. 237). Es el caso de las preguntas retóricas o formuladas a uno mismo, como *¿Me estaré volviendo loco?*

Dentro de las preguntas indagatorias, el autor establece una subclasificación que responde a criterios diversos:

a) Grado de desconocimiento de la incógnita contenida en el enunciado por parte del emisor. Según este criterio, las interrogativas pueden ser:

**-no-presuntivas:** aquellas en las cuales se indaga por uno o varios elementos sobre los cuales el emisor no tiene ninguna hipótesis (*¿Recibió usted mi carta?*).

**-presuntivas:** como contraparte de las preguntas anteriores, son aquellas suponen un determinado grado de presunción con respecto a la variable del enunciado, deducible en muchos casos de la situación comunicativa. Poseen a menudo marcas léxicas como “acaso” o “es que” en interrogativas totales (*¿Es que no viven tus padres?*). En el caso de las totales presuntivas, aparece un constituyente con la misma función sintáctica y valor semántico que un pronombre o adverbio interrogativo, como es el caso de *¿Estarás lista en una hora?*, donde ‘en una hora’ es equivalente a *¿cuándo?* Estas interrogativas coinciden parcialmente con las orientadas de Escandell (1999), como veremos más adelante.

b) Posicionamiento dentro de la conversación. El autor distingue las interrogativas según ellas cumplan determinados roles dentro del diálogo:

**-inceptivas:** son las interrogativas que comienzan el diálogo.

**-reactivas:** constituyen una réplica o reacción a lo dicho anteriormente por el interlocutor. Estas, a su vez, pueden ser: confirmativas (solicitan información para comprobar que se ha entendido bien; ello supone repetir total o parcialmente el enunciado anterior, por lo que funcionan como preguntas ecoicas), aclarativas (como las anteriores, se basan en el enunciado previo, pero el emisor tiene la intención de aclarar algo que no oyó bien o que no

entendió) o completivas (adelantan parte del enunciado que el interlocutor está emitiendo, por suponer cómo terminará).

Esta subclasificación es muy semejante a la propuesta de interrogativas relacionales y de habla de Freed (1994), anteriormente reseñada.

c) Grado de condicionamiento de la respuesta por parte del emisor. Bajo esta perspectiva, pueden distinguirse:

-**preguntas no-inductoras**: el emisor no induce una respuesta determinada, es decir, no hay en el enunciado ningún elemento que oriente la interpretación. Siempre son preguntas totales: *¿Pagaste la cuenta?*

-**semi-inductoras**: se trata de interrogativas que contienen un pronombre o adverbio interrogativo, que condiciona la naturaleza de la respuesta:

A – *¿Qué buscas?*

B – Mi reloj.

-**inductoras**: la interrogativa contiene un elemento que condiciona la respuesta, al partir de una hipótesis previa. Comúnmente, se trata de partículas confirmativas: *Está mejor que nunca, ¿no es cierto?*

Como puede deducirse, esta clasificación guarda una estrecha relación con la presunción antes mencionada, es decir, el grado de condicionamiento de la respuesta deriva del conocimiento previo que tenga el emisor con respecto a la incógnita. Desde nuestro punto de vista, tal diferenciación en la categorización no afecta, en esencia, el tipo de interrogativa planteado, por lo que es redundante.

Asimismo, creemos que no debe confundirse el condicionamiento de la respuesta con el ámbito de acción de la incógnita. En consecuencia, en el caso de la propuesta de Rabanales, se incurre en una doble clasificación semántica, por demás innecesaria, al hablar de interrogativas totales y parciales, por un lado, y de no inductoras y semi-inductoras, por otro. En ambos casos se trata del mismo fenómeno. Tal como demuestra Escandell (1999) en su reflexión sobre interrogación y foco, no es correcto identificar tipo de interrogativa según la respuesta que completa la variable contenida en el enunciado, con grado de suposición con

respecto a la incógnita. Es la delimitación del foco la que hará posible conocer en qué medida el hablante tiene o no alguna presuposición con respecto a la incógnita.

Por último, resumiremos la propuesta de Escandell (1999). Una acotación importante es que la clasificación pragmática de enunciados interrogativos que propone la autora establece una primera distinción de naturaleza gramatical: las oraciones interrogativas pueden dividirse en neutras (aquellas cuyas características gramaticales impiden que haya algún tipo de orientación en su interpretación); y orientadas (aquellas que contienen algún elemento que orienta la interpretación, como es el caso de algunos esquemas entonativos y de las partículas introductorias).

Entre los factores que condicionan la interpretación la autora incluye el orden de palabras: propone que la secuencia SV es típica de estructuras cerradas, con lo cual su uso en la interrogación condiciona la preexistencia de una proposición e inhabilita al enunciado para contener una variable. A pesar de ello, y basándonos en la evidencia científica de la existencia de un orden cuando menos alternante en la interrogativas caribeñas (Morales, 1997; Brown y Rivas, 2011), nosotros pensamos que el acto de petición de información puede realizarse también con enunciados cuya estructura sintáctica carezca de la posposición del sujeto. Por tanto, utilizaremos la clasificación que propone la autora solo en lo referido al significado pragmático, es decir, en lo que atañe a las interpretaciones discursivas de los enunciados, en el supuesto de que las estructuras interrogativas con orden SV pueden estar presentes en cualquiera de los usos que la autora reseña para el tipo de enunciado del que nos ocupamos.

Por otra parte, marcaremos oportunamente los casos en los que está presente alguna de las partículas introductoras que la autora menciona (citativas, consecutivas, de orientación invertida y de orientación constante), pues estas marcas formales servirán de apoyo a la clasificación de enunciados. Respecto de la negación y la polaridad negativa, se registrarán los casos en los que aparece para estudiar su concomitancia con el orden de constituyentes.

A continuación, presentamos la taxonomía que propone la autora, y complementaremos con algunos ejemplos propios:

1. **Interrogación epistémica:** se trata de aquellas estructuras que contienen una incógnita que representa un valor desconocido. Se encuentran en esta categoría varios subtipos:

1.1. **Pregunta real:** el hablante solicita información mediante una oración interrogativa directa. En consecuencia, se espera una respuesta verbal, que otorga un valor a la incógnita del enunciado: *¿Compró Amaya la bicicleta?* Constituye el ejemplo prototípico de pregunta, pues el emisor desconoce cuál es el elemento que debe remplaza la incógnita.

1.2. **Pregunta de examen:** el emisor, de común, conoce la respuesta y formula la pregunta para saber si el interlocutor puede otorgarle un valor a la variable que contiene el enunciado. Este tipo de pregunta está marcada por el espacio institucional en que se usa: el entorno escolar o de instrucción formal: *¿Cuáles son los tipos de reacciones químicas?* (dicho por el profesor al alumno).

1.3. **Pregunta problemática:** también supone un conocimiento nulo de la incógnita que contiene el enunciado por parte del emisor. Sin embargo, no se espera una respuesta concreta: *¿Existe solución para la pobreza?*

1.4. **Pregunta deliberativa:** mediante este tipo de enunciado el emisor se dirige a sí mismo. Se trata de una ficción dialógica, en la que el propio emisor es quien tiene (aunque no la dé) la respuesta a la pregunta: *¿Dónde tendré la cabeza?*

2. **Interrogación directiva:** es la que busca “dirigir la actuación del interlocutor” (Escandell, 1988, p. 555). En esta categoría se incluyen los siguientes subtipos:

2.1. **Petición:** su objetivo es interrogar “sobre el cumplimiento o la realización de una acción por parte del interlocutor” (p. 554): *¿Me sirves jugo?* Tanto este tipo como el siguiente forman parte de la cortesía del enunciado. En este caso, la solicitud reporta un beneficio para el emisor.

2.2. **Ofrecimiento:** como sucede con el tipo anterior, la interrogación sirve para solicitar una acción que debe ser ejecutada por el interlocutor: *¿Quieres ir al cine?* También perteneciente a la cortesía, este tipo de interrogativas se sitúa en el extremo opuesto al tipo anterior, pues la acción que se espera beneficia al propio interlocutor: *¿Te apetece un café?*

3. **Interrogativas atribuidas:** son aquellas mediante las cuales el emisor reproduce de manera real o supuesta parte del discurso de su interlocutor (ya sea anterior o por producirse, según veremos más adelante), o sus propias palabras. En este grupo se distinguen:

3.1. **Interrogativas repetitivas o ‘eco’:** repiten total o parcialmente palabras efectivamente emitidas por un participante en el intercambio, por lo que tienen un carácter retrospectivo. El emisor puede tener varios objetivos: expresar una actitud ante lo recientemente dicho mediante interrogativas ‘recapitulativas’ (a), obtener una explicación o especificación de algo que no ha quedado claro o se desconoce (b), o repetir algo que él mismo ha dicho antes, ante una solicitud del interlocutor (c):

(a) –¡Sal de mi casa!

–¿*Que salga de tu casa?*, ¿seguro?

(b) –En nuestro viaje conocimos Chigualoco.

–¿*Conocieron qué?*

–Chigualoco, una playa de la IV Región.

(c) –¿Tú ya conoces Madrid?

–¿Qué?

–¿*Tú ya conoces Madrid?*

3.2. **Interrogativas anticipativas:** mediante ellas el emisor se adelanta a lo que va a decir su interlocutor. Ellas pueden avanzar una hipótesis (a), adelantar una presunción sobre lo que interlocutor está diciendo (b), o expresar sorpresa (equivalente a una exclamación, como en (c)). Funcionan, en consecuencia, como cita prospectiva.

(a) ¿Qué hay que hacer? ¿*salir a comprar?*

(b) –Cuando hablé con ella me sentí rara...

–¿*Te molestaste?*

(c) ¿*Pero ya estás aquí?* (expresión que se usa en el recibimiento de alguien que llega sorpresivamente).

4. **Interrogativas retóricas:** con ellas el emisor sugiere una respuesta que se corresponde con el signo contrario al que aparece en el enunciado. Parten de la presuposición de que el interlocutor conoce también la respuesta, pues existe un conocimiento previo compartido: ¿*No es bueno este libro?*, ¿*Has hecho algo acaso?*

A continuación, la Tabla 4 resume la propuesta de clasificación de la autora:

| <b>Tipo</b>                      | <b>Categorías</b>      | <b>Elementos distintivos</b>   |
|----------------------------------|------------------------|--|
| <b>Pregunta epistémica</b>       | Pregunta real          | Contienen una incógnita de valor desconocido, pero varían en el grado de conocimiento de esta por parte del emisor |
|                                  | Pregunta problemática  |  |
|                                  | Pregunta de examen     |  |
|                                  | Pregunta deliberativa  |  |
| <b>Interrogación directiva</b>   | Petición               | Motivan una acción en el interlocutor  |
|                                  | Ofrecimiento           |  |
| <b>Interrogativas atribuidas</b> | Repetitivas o ‘de eco’ | Interrogativas de carácter retrospectivo   |
|                                  | Anticipativas          | Interrogativas de carácter prospectivo   |
| <b>Interrogativas retóricas</b>  | Pregunta retórica      | Conocimiento previo compartido (interrogativa que parte de una presuposición)                                      |

*Tabla 4. Clasificación de interrogativas según Escandell (1999), tomando como eje central la interpretación del enunciado según la intención del hablante*

La Tabla 5 recoge una puesta en común de las propuestas de Escandell (1999) y Freed (1994). Partiendo de la base de que son las clasificaciones más completas y coherentes encontradas en la literatura, una comparación entre ambas permite destacar las siguientes diferencias, que serán utilizadas en la conformación de las categorías analíticas del presente estudio.

| <b>Freed (1994)</b>   | <b>Escandell (1999)</b>                              |
|---|--|
| -Interrogativas sobre información pública o social<br>-Interrogativas deícticas   | -Pregunta real                                       |
| -Invitaciones   | -Petición y ofrecimiento                             |
|   | -Pregunta problemática                               |
| -Interrogativas didácticas  | -Pregunta de examen                                  |
| -Interrogativas autodirigidas   | -Pregunta deliberativa                               |
| -Interrogativas retóricas   | -Interrogativa retórica                              |
| -Interrogativas humorísticas  |  |
| -Interrogativas clarificadoras<br>-Interrogativas repetitivas<br>-Interrogativas confirmativas<br>-Interrogativas referidas                         | -Interrogativas eco<br>-Interrogativas anticipativas |
| - Interrogativas de foco conversacional<br>- Interrogativas de información compartida<br>- Interrogativas fáticas<br>-Interrogativas de elaboración |  |

*Tabla 5. Resumen de tipos de interrogativa propuestos por Freed (1994) y Escandell (1999)*

A continuación, procedemos a comentar las principales diferencias organizadas en la Tabla 5:

- a) Freed (1994) incluye un tipo de interrogación que no existe en la propuesta de Escandell: las interrogativas relacionales, que creemos que son fundamentales en intercambios comunicativos reales.
- b) La propuesta de Escandell (1999) es más específica respecto del tipo de interrogación que busca información social/pública o que quiere provocar una acción en el interlocutor, por lo que sus categorías parecen más prácticas para describir un corpus con una amplia gama de temas/intercambios posibles, como el nuestro.
- c) Las preguntas para clarificar información de Freed (1994), pertenecientes a la interrogación del habla, no parecen atribuibles ni a las interrogativas eco, ni a las anticipativas

que Escandell (1999) agrupa dentro de las interrogativas atribuidas, por tanto, merecen ser contempladas como un tipo aparte. Por otra parte, las preguntas eco y las repetitivas de ambas clasificaciones hacen referencia a la misma noción.

d) Las preguntas anticipativas de Escandell (1999) no parecen pertenecer a ninguna de las clasificaciones de Freed (1994), si bien es claro que se corresponden con el tipo funcional de las interrogativas relacionadas con la conversación de esta última autora, por lo que deben ser mantenidas como tipo independiente.

e) La propuesta de clasificación de interrogativas expresivas de Freed (1994), que encierran un grado bajo de desconocimiento de la incógnita, es más amplia que la de Escandell, pero los principales tipos –que creemos los más productivos para la presente investigación– están en ambas propuestas: interrogativas didácticas (o de examen), interrogativa deliberativa (o autodirigida) e interrogativa retórica.

f) La separación por categorías funcionales de Freed permite establecer un *continuum* que va de información elicitada a información transmitida, lo que es lo mismo que decir, de bajo a alto grado de desconocimiento de la incógnita. Creemos que esta tipología puede ser útil para poder extraer conclusiones respecto del orden de palabras.

Teniendo en cuenta lo anteriormente comentado, se decidió tomar en cuenta para un estudio piloto las categorías recogidas en la Tabla 5, con el objetivo de saber si son productivas y si la definición aportada por las autoras es útil para describir los datos de nuestro corpus. Se han agrupado en filas los tipos de interrogativa que son equivalentes en las tipologías de ambas autoras. El contenido del epígrafe siguiente recoge la sistematización del estudio piloto.

#### **1.4.8.1. Categorías analíticas del presente estudio**

La productividad de las categorías originalmente propuestas se puso a prueba mediante la realización de un estudio exploratorio en el que intervinieron tres expertos. Se procedió a la distribución entre ellos de 400 casos del corpus, de manera que cada uno de los expertos tuviera en común 100 casos de análisis. Una vez realizada la descripción, la puesta en común arrojó que hubo un acuerdo del 67% como promedio (73% entre los expertos 1 y 2; y 60% entre los expertos 2 y 3). Con ello, se procedió a la discusión de casos donde no hubo concordancia. Como resultado de este proceso, se pudo ajustar los conceptos a los 14 tipos

de interrogativa presentes en el corpus. Así, se realizó una redefinición que tuviera en cuenta tanto los usos discursivos más comunes como las marcas textuales típicas de cada uno de ellos. El último paso del proceso de descripción fue la unificación de los criterios analíticos y su aplicación a todo el corpus.

A continuación, se presentan las categorías definitivas de interrogativas totales o absolutas con las que se trabajó en la clasificación. En todos los casos, los ejemplos han sido extraídos del corpus.

**Pregunta real:** el emisor solicita una información que desconoce mediante una pregunta directa (Escandell, 1999). Típicamente aparecen como parte de diálogos presentes en obras literarias o en entrevistas, por lo que requieren respuesta.

Ejemplo:

Como lo explica Ricardo Blanco, encargado de las Comunicaciones Globales y Asuntos Públicos de Productos de Google en América Latina, han ido avanzando en distintos países, captando los diferentes modismos como por ejemplo España y Brasil. Y esta vez captaron un cierto número de modismos en el habla del continente americano con el idioma español. "Hemos reconocido acentos de personas hispanas que llevan mucho tiempo en Estados Unidos, y que pueden tener similares acentos a los argentinos o chilenos", explicó.

*-¿Esta información recopilada les puede servir para otras aplicaciones?*

-Por supuesto, imagínate que en Estados Unidos ya se avanzó con una versión que les permite captar el acento personalizado. También esto ha permitido desarrollar un sistema en que en algunos videos en inglés en Youtube se pueda colocar automáticamente los subtítulos. (Chi 23)

En el corpus aparecen también las preguntas que forman parte de las secciones informativas de páginas web que ofrecen servicios o bienes de consumo, o bien de páginas institucionales. Estas preguntas también suponen un emisor (en el ejemplo a continuación sería el productor/comercializador de cierto producto de belleza), y un interlocutor (clientes potenciales):

A continuación algunas preguntas frecuentes que el potencial consumidor del producto puede tener antes de decidirse a usar el mismo.

*¿Este es un producto aprobado para su consumo humano?*

Es un producto completamente testeado dermatológicamente en diversas pruebas de laboratorio, y ha demostrado cumplir con todas las normas de bioseguridad para ser aprobado como producto cosmético de venta libre. (Ant50)

**Interrogativa *subjectio*:** debe su nombre a la figura retórica clásica (Escandell, 1999). A caballo entre la interrogativa didáctica y la anticipativa, sirve para organizar el discurso, introduciendo o expandiendo un argumento. Este tipo de preguntas es frecuente en textos periodísticos, en donde abundan las ficciones dialógicas. Al igual que en las interrogativas didácticas, el emisor se erige como una instancia de saber (Bürki, 2010). Tal como sucedería en un intercambio real de habla, estas preguntas suponen un interlocutor, “y se convierte en el pivote sobre el cual descansa la respuesta del enunciador” (p. 360). Son seguidas típicamente de una respuesta, o son retomadas en el contexto posterior.

Ejemplo:

También son consideradas como acciones de ese tipo ciertas medidas de discriminación inversa o positiva, que se toman: 1) en consideración de aspectos como el sexo o la raza, que son criterios sospechosos o potencialmente prohibidos; o 2) frente a situaciones de especial escasez de bienes deseados, como suele ocurrir en puestos de trabajo o cupos universitarios, lo que conlleva como contrapartida una restricción a derechos de otras personas.

*¿Pero, son válidas estas acciones a pesar de la existencia de la premisa de igualdad de todas las personas ante la ley?* Pues sí, porque la igualdad formal, reconocida por el artículo 13 de la Constitución, supone que todos los individuos, como sujetos de derechos, deben ser tratados con la misma consideración y reconocimiento ante la ley. (Car35)

**Interrogativa *problemática*:** expresa una duda auténtica por parte del emisor que no puede ser respondida de una manera cabal. Normalmente constituye la coda de un argumento, por lo que suele cerrar el discurso o marcar un cambio de tópico. Típicamente se plantea la pregunta, pero no se responde ni se desarrolla. Se corresponde con la interrogativa problemática de Escandell (1999), de donde hemos tomado el nombre.

Ejemplo:

Y si analizamos el balance de petrolíferos, incluyendo gasolina, diesel y combustóleo, llegamos a la conclusión de que las exportaciones netas de hidrocarburos líquidos comenzaron a descender desde 2003. La tendencia parece irreversible.

*¿Conviene mantener exportaciones de crudo a su máximo nivel posible para empezar a importarlo en unos cuantos años?* El debate sobre nuestra plataforma de producción y de exportación de petróleo resulta inevitable.

Resulta urgente llevar a cabo una evaluación detallada del potencial de recursos petroleros por descubrir y hacerlo con el mayor rigor técnico. (Méx.80)

**Interrogativa anticipativa:** tal como plantea Escandell (1999) este tipo de interrogativa es utilizada para adelantar una hipótesis sobre lo que cree el propio emisor o sobre lo que espera que va a decir el interlocutor. Ella parte, por tanto, de una presuposición. Este tipo de enunciado puede aparecer en instancias dialógicas, a continuación de una pregunta real, como se muestra en el ejemplo (1). En estos casos, se requiere una respuesta que constituye el rechazo o afirmación de la hipótesis contenida en el enunciado interrogativo.

En otras ocasiones, aunque el enunciado adelanta una hipótesis del emisor, el objetivo es criticar o reprochar cierto comportamiento del interlocutor. En el ejemplo (2) el emisor nos informa sobre la condición de trabajador no oficial de su tío y de sus prácticas poco probas. El enunciado interrogativo, en consecuencia, tiene la intención de marcar una actitud de desaprobación en el emisor. Es equivalente, en consecuencia, a una aseveración. Podría parafrasearse de la siguiente manera: “En este país se hizo una revolución y lo que estás haciendo no está bien”.

Desde el punto de vista formal, estos enunciados suelen tener un marcador que orienta su interpretación, como es el caso de partículas introductorias *es que* y *acaso* o términos de polaridad negativa (adverbio *no*).

Ejemplos:

(1) Turing es el apodo de un personaje que trabaja en el Estado interceptando mensajes de la oposición. Y otro de los personajes centrales es Kandinsky, un hacker que envía virus informáticos a las multinacionales y a las computadoras del Estado.

— ¿Por qué el pseudónimo de Kandinsky?, *¿el pintor ruso tiene algo que ver en la novela?*

— Cuando estaba escribiendo la novela tenía en mi escritorio un calendario con los cuadros de Kandinsky. Nada más que eso, pero estoy seguro de que los críticos le encontrarán algún significado interesante. (And47)

(2) En su condición de albañil y plomero, mi tío debe de ser uno de los únicos cubanos que nunca han trabajado para el Estado. Con materiales robados, o en contubernio con gente bien conectada, siempre se las arregló para vivir de sus trabajos por cuenta propia. Un día le dije:

Tío, *¿tú no sabes que en este país se hizo una revolución?*

¿Ah, sí? -se sorprendió y contestó-: ¿Y a mí qué me importa? (Ant4)

**Interrogativa de foco:** de manera semejante a la interrogativa de foco conversacional descrita por Freed (1994), estos enunciados introducen una temática o elemento nuevo que luego se ampliará en el discurso. Por consiguiente, esta interrogativa no busca respuesta, y esta no es proporcionada en el texto. A diferencia de las interrogativas *subjectio*, estos enunciados no tienen el objetivo de introducir un argumento, sino que sirven para construir un relato, que puede articular voces propias o externas. Del mismo modo, se diferencian de las anticipativas antes aludidas en no que parten de una presunción o hipótesis previa.

Ejemplo:

Será difícil "porque la gimnasia ha cambiado mucho". Tanto, que casi no la reconoce. El nuevo código de puntuación ha variado el valor de los ejercicios y, sobre todo, la forma en que éstos son valorados por los jueces.

*¿Le gustan al campeón olímpico las nuevas normas?* "El código no es malo para el tipo de gimnasia que yo hago, pero no me gusta", reflexiona el gimnasta, que llegó a la élite casi por sorpresa cuando se proclamó campeón del mundo absoluto en 2003 ante miles de enfervorizados seguidores en Anaheim (California). Su palmarés se completa con tres medallas olímpicas: la de la polémica y dos platas, en la prueba por equipos y en barra fija. (Esp36)

**Interrogativa repetitiva:** repite total o parcialmente lo que se ha dicho previamente en la conversación, por lo que se encuentran típicamente en entrevistas, relatos de ficción o transcripciones de intercambios de habla real. Tienen un carácter retrospectivo, pues dependen de lo que se ha dicho inmediatamente antes. Se corresponde con las interrogativas recapitulativas de Escandell (1999).

Ejemplo:

-Ay Tola... Yo tengo pensión que de pronto a Uribe le empiece a patinar el coco. Me han contaó que ya no tiene los bríos de antes, que lo notan achicopalao, que se le ve por encima el desaliento... Que se parece a Bolívar tuberculoso en su quinta de Santa Marta dececionao de haber arao en el mar...

-*¿Bolívar aró en el mar?* ¿Y no se le hundió el recatón?

-Es que Colombia es un toro duro de lidiar Maruja...

En ocasiones la función de la interrogativa surge de una interpretación del contexto, y de la presuposición de la existencia de una interrogativa previa que se retoma o repite. Las interrogativas repetitivas tienen en ocasiones marcadores que orientan su interpretación, como es el caso de las conjunciones *si* o *que*, que funcionan como partículas introductoras.

Ejemplo:

El resultado preocupa más por el futuro que por el presente: Venezuela será el rival directo de la Argentina en el Preolímpico sudamericano, que dará sólo un lugar para Pekín. Será en los primeros días de enero de 2008, en nuestro país. Pero con este antecedente, la situación se vuelve alarmante. **“¿Si sirve el partido de hoy? Mucho. ¿Si me preocupa el resultado? Mucho”**. Verborrágico como es, el DT Jon Uriarte no usó muchas palabras para describir su sensación. (Río 84)

**Interrogativa confirmativa:** buscan verificar algo que se deduce de lo anteriormente dicho. De manera semejante a las repetitivas, las confirmativas tienen un carácter retrospectivo, pues se retoma parte del contenido previo de la conversación. Desde el punto de vista formal, suelen hacer uso de elementos correferenciales (pronombres, sinónimos, etc). Van seguidas típicamente de una respuesta. Se corresponden con las interrogativas explicativas de Escandell (1999) y las confirmativas de Freed (1994), de donde hemos tomado el nombre.

Ejemplo:

(1) -Usted señaló como otro problema el de los caminos.

-El problema se presenta como siempre, en las áreas rurales, ya que a los caminos no se les ha dado mantenimiento, porque prácticamente en La Concha las autoridades edilicias actuales donde se han concentrado es en el área urbana, pero las comunidades rurales las han abandonado.

**-¿Esto quiere decir que la infraestructura vial se centra en el casco urbano?**

-Así es, hay un abandono, porque en las visitas que hemos realizado casa a casa hemos visto que hay una escuela en la comunidad de Los Mercados, la cual está socavada en la parte de abajo y a esto, el alcalde ha hecho caso omiso (...) (Méx97)

(2) -Vale, y ¿por qué te ha dado por el hígado ahora?

-Porque, realmente, es el único órgano capaz de regenerarse en un cuerpo humano adulto. Un donante de hígado puede servir para varios receptores, solo hay que cortarlo en trozos y, una vez dentro, recupera el tamaño necesario en función del cuerpo del receptor.

-¿Eso es regeneración?

-Sí, trasplante y regeneración. Una forma actual de medicina regenerativa, disciplina que ambiciona los mismos resultados obtenidos con el hígado, pero a partir de cualquier otro órgano o tejido humano. (Esp41)

**Interrogativa referida:** reproduce de manera literal un enunciado emitido en un intercambio previo, ya sea imaginado o real (1). También pueden reproducir interrogantes que preexisten a la argumentación o exposición de que se trate, en el caso de los textos escritos no dialógicos (2). En ambos casos la interrogativa es usada como una referencia que luego será usada con un determinado propósito (en el caso del primer ejemplo a continuación es contribuir a la argumentación; en el caso del segundo ejemplo tiene fines meramente informativos). Desde el punto de vista formal, típicamente aparecen entre comillas, o son anunciadas con dos puntos. Si bien implican una respuesta, esta puede no aparecer. Se corresponden con las interrogativas eco de Escandell (1999) y las interrogativas referidas de Freed (1994), de donde hemos tomado el nombre.

(1) Ante el complaciente panorama de exculpación a los violadores, la abogada Luz Almeida, en un lúcido artículo titulado “Resistirse poco”, se preguntaba: “*¿No será la libertad sexual del agresor la que se está protegiendo?*”. Y es que, en muchas ocasiones, las sentencias son auténticas chapuzas. ¡No en balde, los jueces rematan sus faenas de aliño con un martillazo! (Esp2)

(2) Las preguntas claves que se habrán de responder son: “**¿trabaja el paciente de manera regular?**”, “¿le gusta su trabajo?”, “¿puede estar a la altura de lo que tiene que hacer?”, etc. En el caso de un estudiante o de una ama de casa, adecuar las preguntas que deberían hacerse. (Chi87)

Se diferencian de las repetitivas en que estas últimas repiten total o parcialmente palabras previamente emitidas por un participante en la conversación, por lo que tienen un carácter retrospectivo al interior del discurso.

Como es normal en el manejo de categorías que no tienen límites precisos, en ocasiones puede haber solapamientos que complejizan el proceso de descripción. A continuación, puede encontrarse un ejemplo de pregunta problemática que puede también considerada como interrogativa referida.

Ejemplo:

Cabe preguntarse si los postulados socialdemócratas han perdido atractivo en una sociedad que parece haber modificado sus paradigmas. Se da la paradoja de que, aceptado generalmente que la peor crisis que hemos sufrido se debe al triunfo de los principios conservadores de la desregulación, ésta se haya llevado por delante a la izquierda que preconiza el predominio del Estado.

*¿Ha muerto la socialdemocracia de éxito tras generalizar en Europa el Estado de bienestar?* ¿Se han producido excesos de solidaridad provocando una lógica reacción pendular de tinte individualista? ¿Qué papel ha desempeñado la caída del imperio soviético en el hundimiento de los socialismos democráticos que ya no parecen útiles como coartada frente al comunismo? Son muchos los interrogantes que se plantean en un debate sobre lo que nos está pasando. En ellos mojan sus plumas los gurús de la izquierda. (Esp24)

El hecho de que el emisor declara que la pregunta forma parte de las interrogantes que existen en un debate actual hace pensar que podría estar reproduciendo preguntas preexistentes al momento de la emisión.

**Interrogativa fáctica:** sirven para comprobar el vínculo entre emisor y receptor, preguntando por un elemento relacionado con el contenido de la conversación, pero sin requerir respuesta de manera obligatoria. Forman parte de las interrogativas de habla Freed (1994), de donde hemos tomado el nombre.

Ejemplo:

Habíamos ido al Museo de La Habana, en la plaza de la catedral, a ver no tanto la colección de muebles y pinturas antiguas que albergaba el edificio colonial, sino los medios puntos de vidrio que a Boris le causaban alegría. En cada puerta y ventana rematada por un semicírculo de vidrio emplomado se detuvo a explicarme cómo habían ido evolucionando a lo largo de la colonia estos adornos transparentes que, al filtrar la luz del trópico, teñían los pisos de cantera con colores de caramelo y al mismo tiempo ayudaban a matizar su calor.

*“¿Ves tú, Alma, cómo somos cubanos los cubanos?”*, exclamaba Boris entusiasmado. "Esta casa del siglo diecisiete no es El Escorial, estas ventanas de colores no son la catedral de Salamanca, por más que la piedra es piedra y el vidrio emplomado es vidrio (...)". (Méx41)

**Interrogativa de información compartida:** de acuerdo con las interrogativas homónimas descritas por Freed (1994), estos enunciados se encargan de introducir temas conocidos por ambos interlocutores. Desde el punto de vista formal suelen presentar pronombres personales o indefinidos (alguien, ustedes) y normalmente no buscan una respuesta.

Ejemplo:

*-¿Recuerda usted el auge de la biotecnología en los setenta y posterior declinación por expectativas no satisfechas? ¿Qué diferencia esencial aprecia ahora?*

-Hay una respuesta muy simple y directa para esto, a comienzos de los setenta el crecimiento de la biotecnología fue dirigido por las técnicas. La gente era capaz de entender las cosas pero desde un prisma reactivo, es decir, podían entender qué era lo malo que estaba ocurriendo y luego reparar esas cosas. En cambio hoy se pueden anticipar los efectos y ofrecer productos. (Chi48)

Como puede verse en el ejemplo, el interlocutor responde a la segunda pregunta, relacionada con la diferencia entre el auge biotecnológico de los 60 (punto referencial que establece la interrogativa de información compartida) y la realidad actual. El emisor parte de la presuposición de que el interlocutor conoce de qué le están hablando.

**Interrogativa clarificadora:** piden información adicional relacionada con el contenido de la conversación (Freed, 1994). Dependen del discurso inmediatamente anterior, por lo que son retrospectivas. En tanto duda real gatillada por el contenido del discurso previo, requieren una respuesta.

Ejemplo:

- Oites Tola ¿vos me podés acompañar a que me fileteen este estraple que se va estrenar hoy mi nieto en el desfile del orgullo gay?

- *¿Vos tenés un nieto gay?*

- El menor de Graciela, Carlos Adonis. (Car97)

**Interrogativa retórica:** preguntas que parten de una presuposición y cuya respuesta, sugerida por el emisor y típicamente ausente, es contraria al signo contenido en la pregunta (Escandell, 1999). Típicamente son introducidas por *acaso*.

Ejemplo:

Si eso es así, ¿cuáles serán esos partidos? La respuesta ya la sabemos. Los mismos que hoy dominan el binominal. Entonces, ¿en qué consiste el cambio de mecánica electoral? ¿Qué va a aportar para la oxigenación política un proporcional corregido con cifra repartidora?

Acaso, en este modelo ¿no se genera también exclusión y distorsión entre votos y escaños? O acaso, *¿no sigue existiendo la supremacía de la lista?*

A mi entender un "proporcional puro" es el que mejor asegura la relación entre votos y asientos parlamentarios. En ese caso, la exclusión de las fuerzas menores es natural y no el resultado de la ingeniería electoral ni el cálculo político. Al mismo tiempo, es un modelo que no le tiene miedo a la incertidumbre. (Chi24)

**Interrogativa deliberativa:** Forma parte de una ficción dialógica en la que el emisor se formula a sí mismo una pregunta, como parte de una disquisición personal (Escandell, 1999). En tanto es parte de un diálogo con uno mismo, no se requiere una respuesta, si bien el emisor la conoce. Como marca formal, este tipo de interrogativa suele estar integrada en un recuento o una exposición en la que prima la primera persona del singular. Se corresponde con el tipo de interrogativa autodirigida, de Freed (1994).

Ejemplos:

(1) No he visto maqueta alguna para el nuevo proyecto, pero ya se oyen voces lejanas. ¡Puerto deportivo! ¡Canchas de deporte! ¡Playas artificiales! ¿Hoteles? Y yo me pregunto, *¿a nadie le interesa la paz, lo diferente, la sensación que uno puede escaparse dentro de esta ciudad?*

Como humilde ciudadano, le pediría al señor alcalde y a todo su séquito de lumbreras que antes de esas apetitosas fotos, una tarde a eso de las nueve de la noche se fueran a Los Baños del Carmen y se sentaran en una de sus mesas. Con la tranquilidad que da esa hora, se pidieran algo fresco, abrieran bien los ojos y se dedicaran a eso tan poco popular, que es mirar el mar al atardecer y soñar. (Esp99)

(2) Me asaltan a ratos, como ahora, inquietudes sobre el escribir, pasión que llevo en mí sin que siquiera intente detenerla. Algunas de ellas son: *¿es esta infame dedicación lo que la escritura se merece? ¿Es ésta la forma de sacar de mí lo mejor que puedo dar?* ¿Son, desde mis manos, éstos los mejores poemas que haya podido escribir? Y ante todas aparece un rotundo no. (Car78)

**Petición u ofrecimiento:** buscan provocar una acción en el interlocutor (Escandell, 1988). Este tipo integra las invitaciones de Freed (1994). Nótese que en el ejemplo que se cita a continuación la interrogativa es equivalente a una orden. Desde el punto de vista formal, este tipo de interrogativa suele tener algún verbo modal, como es el caso de 'podría' en el siguiente ejemplo:

- En efecto, señor Pérez, y pase el retintín. Cambiemos pues, de tercio: *¿podría usted mencionarme algunas muestras representativas de los orígenes de la literatura española?*

- Sí puedo, gracias a su tenaz magisterio: las jarchas mozárabes, el Cantar de Mio Cid (Mió, señor Mairena, no Mío, note cómo le hago caso), la Disputa del alma y el cuerpo... (Esp73)

## **Capítulo 2. Metodología**

### **2.1. El corpus**

#### **2.1.1. Herramienta de selección del corpus. El CORPES XXI**

Para responder a las preguntas de investigación (ver 1.3), empleamos el Corpus del Español del Siglo XXI (CORPES). La decisión de trabajar con esta herramienta respondió a la necesidad de contar con datos de habla real, provenientes de todo el ámbito de habla hispana. El CORPES, disponible on-line, ha sido elaborado por la Real Academia de la Lengua Española. Está formado por textos escritos (90%) y orales (10%), procedentes de España, América, Filipinas y Guinea Ecuatorial. Al momento de la consulta del corpus digital, en su versión beta (0.83), este poseía casi 225 millones de formas, recopiladas entre los años 2001 y 2012. El CORPES posee un 70% de datos de origen americano y un 30 % precedente de España. De los textos escritos contenidos en el corpus, 40% corresponden a libros, 40% a publicaciones periódicas, 7,5% a material de internet y 2,4% a miscelánea.

El recurso —actualmente en su versión 0.94— permite una recuperación selectiva, es decir, los casos pueden ser obtenidos filtrando por origen, zona lingüística, país, medio (i.e., texto escrito u oral), tipo de texto (e.g., novela, entrevista, etc.) y área temática. En el caso de la presente investigación, se trabajó solamente con datos provenientes de textos escritos y se procedió a extraer todos los casos del corpus que contaran con el signo de apertura de la interrogación (¿). Posteriormente se realizó la exportación de los más de 134.000 casos arrojados por la búsqueda. El proceso de filtrado de los datos se realizó en dos partes que se detallan a continuación.

#### **2.1.2. Criterios de exclusión**

##### **2.1.2.1. Filtrado 1**

En una primera fase se diseñó una herramienta digital que permitió el filtrado automático de casos inútiles a partir de criterios simples como los siguientes:

a) oraciones de una sola palabra, pues debe haber, al menos, dos elementos oracionales para poder, en principio, hablar de algún orden.

b) oraciones que inician con pronombre interrogativo.

c) oraciones que comienzan con la estructura preposición+pronombre interrogativo (es importante aclarar que hubo que describir una a una el tipo de estructura, al no contar un sistema de *parseo* asequible).

d) oraciones que constituyen expresiones fáticas o muletillas (Rabanales, 1996) como: *¿es posible?*, *¿es cierto?*, *¿me entiendes?*, *¿tú crees?*, *¿se puede?*, *¿sabía usted que...?*

e) datos provenientes de Filipinas y Guinea Ecuatorial, pues en estos países existe un porcentaje considerable de hablantes que no domina el español.

Mediante este procedimiento quedaron excluidas 66.245 oraciones y un total de 68.489 quedó por clasificar.

#### **2.1.2.2. Filtrado 2**

Con el total de interrogativas que quedó sin filtrar, procedimos a dar un orden aleatorio a los datos. Posteriormente, se procedió a seleccionar un total de 100 casos útiles por cada una de las siete zonas incluidas en el estudio. Este número fue definido arbitrariamente, en el supuesto de que la cantidad total de casos obtenida (700) sería razonable para un estudio como el que se propone. En el proceso de selección, los casos fueron etiquetados según los criterios de inclusión o exclusión definidos, hasta cubrir el número de casos útiles necesarios por zona (100).

Así, en un documento Excel, registramos todos los casos que contemplamos en el proceso de selección y los etiquetamos uno a uno de manera manual, siguiendo la siguiente nomenclatura:

a) Casos incluidos, es decir, aquellos en los que hay un sujeto y un verbo de manera explícita y que no cumplen con los criterios de exclusión (declarados más abajo): **Inc**.

Ej.: *¿La idea del yo puede desaparecer del cerebro?* Punset, Eduardo: *El alma está en el cerebro. Radiografía de la máquina de pensar*. Madrid: Aguilar, 2006.

b) Casos excluidos por constituir oraciones interrogativas parciales: **IntP**.

Ej.: ¿Qué opinas de guitarristas como Neil Young o Lou Reed considerados guitarristas limitados pero con encanto? *Elmundo.es. Encuentro digital con Ariel Rot*. [www.elmundo.es](http://www.elmundo.es): elmundo.es, 2001-05-10.

c) Casos excluidos por contar con sujeto implícito (o tácito): **SujI**.

Ej.: ¿Podemos llegar a simular las sensaciones, emociones, olores, sabores y hasta experiencias totales como las sexuales? Piscitelli, Alejandro: *Ciberculturas 2.0. En la era de las máquinas inteligentes*. Buenos Aires: Paidós, 2002.

d) Casos excluidos por constituir oraciones impersonales en las cuales no es posible deslindar un sujeto<sup>46</sup>: **Imp**.

Ej.: ¿Hay ideas comunes en las obras que se exhibirán? «Dos mujeres ganaron en el Salón EL COMERCIO». *Elcomercio.com*. Quito: [elcomercio.com](http://elcomercio.com), 2005-12-01.

e) Casos excluidos por no constituir estructuras oracionales: **NOrac**.

Ej.: ¿Preguntas e hipótesis impertinentes? Elizalde, Antonio: «Parque Pumalín: aporte al desarrollo regional». *Programa Chile Sustentable: Parque Pumalín. Obstáculo u oportunidad para el desarrollo*. Santiago de Chile: LOM Ediciones, 2002.

f) Casos excluidos por no ser posible una clasificación, debido a omisiones que afectan la gramaticalidad de la estructura oracional o su significado; forman parte también de este criterio de exclusión los casos de transcripciones truncas, así como de interrogativas indirectas transcritas como interrogativas directas: **S/C**.

Ej.: ¿Ya que hace un momento tocamos el tema, háblame cómo surgió la idea de hacer un disco con Silvio Rodríguez? Henríquez Lagarde, Manuel: «Maremoto cubano». *La Jiribilla*. La Habana: [Lajiribilla.cu](http://Lajiribilla.cu), 2002-11-23.

---

<sup>46</sup> En consonancia con Real Academia Española y Asociación de Academias de la Lengua Española (ASALE) (2009), estimamos que las oraciones impersonales son aquellas en las que no existe sujeto expreso ni aparece sobreentendido. Así, se excluyeron las oraciones con verbos meteorológicos, con *haber*, con *se* impersonal, etc.

g) Casos excluidos por constituir oraciones que forman parte de conjuntos mayores, cuyo análisis aislado es imposible por no tener independencia de sentido. En este caso están las interrogativas disyuntivas que aparecen como oraciones independientes<sup>47</sup>: **IntD**.

Ej.: ¿O es que su oposición a la repostulación del jefe del Estado es también parte de la estrategia blanca en su conjunto? Arache, Patricia: «Al PRD hay que dejarlo ir». [A]hora.com.do. Santo Domingo: ahora.com.do, 2003-08-25.

h) Interrogativas cuyo sujeto es una oración subordinada sustantiva relativa sin determinante (o factiva)<sup>48</sup>, pues el sujeto es raramente antepuesto en la enunciación, y no podría, en consecuencia, saberse si la posposición se debe a la influencia de la interrogación: **SujSS**.

Ej.: ¿Es una regla que la novela policial escrita en Latinoamérica esté marcada por la pobreza? Planas, Enrique: «Vigencia del policía romántico». *Elcomercio.com.pe*. Luces. Lima: elcomercio.com.pe, 2009-02-01.

Aún cuando existen contextos en los que es posible la anteposición del sujeto, asumimos que su representatividad en el corpus sería tan pequeña que es preferible la exclusión de casos, con el fin de evitar un desbalance de categorías. El análisis posterior a la clasificación de las oraciones arrojó que, efectivamente, los 49 casos encontrados tenían el sujeto pospuesto, o sea, que la proporción de orden VS es del 100%. La alternancia, en consecuencia, queda descartada.

i) Construcciones con ‘se’, medias e impersonales: Se trata de construcciones que se construyen con el pronombre átono ‘se’, frecuentes con todo tipo de verbos, que semánticamente “se interpretan como oraciones con un sujeto implícito indeterminado, que se desconoce, se sobreentiende o no interesa mencionar” (Mendikoetxea, 1999a, p. 1635).

---

<sup>47</sup> La selección de la muestra se hizo haciendo una aleatorización de los casos extraídos. En consecuencia, las interrogativas que en el corpus original aparecen ordenadas siguiendo determinados criterios, adquieren un orden diferente en nuestro corpus, producto del mecanismo antes mencionado. Por tanto, en el caso de que hubiere en el CORPES dos interrogativas consecutivas, estas tendrán un nuevo orden. Es el caso de las disyuntivas que –normalmente– dependen de una interrogativa previa. Aparecen, por tanto, aisladas, razón por la cual se decidió excluirlas, al no tener independencia semántica.

<sup>48</sup> Para abundar en las características de estos sujetos, el orden marcado y no marcado según la clase semántica del verbo, véase Delbecque y Lamiroy (1999).

Las oraciones pasivas y medias se asemejan en que los sujetos son inanimados de tercera persona y corresponden al objeto de la oración activa correspondiente. Además, en ambas se deduce la presencia de un agente. Sin embargo, las medias son “proposiciones estativas”, que requieren la presencia de un modificador adverbial (Mendikoetxea, 1999a): *Esta camisa se lava muy bien con lejía*. Dentro de las estructuras medias con ‘se’ existe una subclase, llamada medias-impersonales: *A estos niños se les asusta fácilmente*. Este tipo de estructura tiende a tener el sujeto pospuesto, por lo que no se podría decir si la posposición se debe a la interrogación o a un orden natural de elementos en la declaración. En efecto, de las 44 instancias encontradas en el corpus, 9 ostentaron el orden SV (20% de los casos), por lo que puede afirmarse que la posposición del sujeto es mayoritaria, aunque no categórica. Por tanto, parece este ser un criterio válido de exclusión.

j) Oraciones con topicalizaciones, pues constituyen construcciones marcadas, tanto desde el punto de vista sintáctico como pragmático (**Top**):

Ej.: ¿Es así como funciona la política? Méndez, Lucía: «Cristina Cifuentes: indignidad, crueldad, desastre». *El Mundo*. Madrid: elmundo.es, 2018-04-26.

k) Oraciones con vocativo (**Voc**):

Ej.: ¿Vos te imaginás Maruja vos y yo de areperas, peliando por el rincón de la cama, alegando cuál de las dos desocupa la bacinilla? Gallego, Carlos Mario: *La era Uribe contada por las dos lengüilargas de Colombia*. Buenos Aires: Aguilar, 2007.

Los casos totales revisados por zona quedaron distribuidos de la manera que se muestra en la Tabla 6. Es *Antillas* la zona que menos datos tiene en el corpus, por lo que el porcentaje de casos revisados es mayor que el promedio, mientras que *España* es la región más representada y la que menor proporción de casos revisados presenta. Por último, la zona *Chilena* no es la que menos casos tiene pero sí es la que mayor porcentajes de casos revisados presenta.

| <i>Zona dialectal</i>   | <i>Número total de casos</i> | <i>Número de Casos revisados</i> | <i>% de casos revisados</i> |
|-------------------------|------------------------------|----------------------------------|-----------------------------|
| <i>España</i>           | 24651                        | 415                              | 1,7                         |
| <i>Río de la Plata</i>  | 12172                        | 364                              | 3,0                         |
| <i>Méx. y Centroam.</i> | 11078                        | 370                              | 3,3                         |
| <i>Caribe cont.</i>     | 7673                         | 331                              | 4,3                         |
| <i>Andina</i>           | 4612                         | 373                              | 8,1                         |
| <i>Chilena</i>          | 4220                         | 407                              | 9,6                         |
| <i>Antillas</i>         | 4083                         | 335                              | 8,2                         |
| <b>Total</b>            | 68489                        | 2595                             | 3,7                         |

Tabla 6. Cantidad total de casos por zona y cantidad de casos filtrados para obtener 100 datos analizables en cada un de ellas

La Tabla 7 muestra el comportamiento de cada uno de los criterios de exclusión y exclusión en las siete zonas estudiadas:

| <i>Crit. Filt.</i> | <i>Ant.</i> |          | <i>Andina</i> |          | <i>Caribe</i> |          | <i>Chile</i> |          | <i>España</i> |          | <i>Méx.</i> |          | <i>Río</i> |          | <i>Total</i> |
|--------------------|-------------|----------|---------------|----------|---------------|----------|--------------|----------|---------------|----------|-------------|----------|------------|----------|--------------|
|                    | <b>F</b>    | <b>%</b> | <b>F</b>      | <b>%</b> | <b>F</b>      | <b>%</b> | <b>F</b>     | <b>%</b> | <b>F</b>      | <b>%</b> | <b>F</b>    | <b>%</b> | <b>F</b>   | <b>%</b> | <b>F</b>     |
| <i>IntP</i>        | 52          | 16       | 51            | 14       | 59            | 18       | 60           | 15       | 64            | 15       | 58          | 16       | 44         | 12       | 388          |
| <i>SujI</i>        | 94          | 28       | 137           | 37       | 81            | 24       | 135          | 33       | 141           | 34       | 106         | 29       | 125        | 34       | 819          |
| <i>Imp</i>         | 19          | 6        | 24            | 6        | 13            | 4        | 18           | 4        | 21            | 5        | 12          | 3        | 15         | 4        | 122          |
| <i>Norac</i>       | 56          | 17       | 44            | 12       | 58            | 18       | 71           | 17       | 64            | 15       | 73          | 20       | 55         | 15       | 421          |
| <i>S/C</i>         | 3           | 1        | 2             | 1        | 1             | 0        | 4            | 1        | 5             | 1        | 1           | 0        | 3          | 1        | 19           |
| <i>IntD</i>        | 2           | 1        | 1             | 0        | 4             | 1        | 10           | 2        | 1             | 0        | 0           | 0        | 6          | 2        | 24           |
| <i>Voc.</i>        | 0           | 0        | 1             | 0        | 1             | 0        | 0            | 0        | 2             | 0        | 0           | 0        | 2          | 1        | 6            |
| <i>Top.</i>        | 0           | 0        | 2             | 1        | 0             | 0        | 0            | 0        | 1             | 0        | 1           | 0        | 0          | 0        | 4            |
| <i>SujSS</i>       | 2           | 1        | 3             | 1        | 8             | 2        | 3            | 1        | 10            | 2        | 12          | 3        | 11         | 3        | 49           |
| <i>Pas.</i>        | 8           | 2        | 8             | 2        | 6             | 2        | 6            | 1        | 6             | 1        | 7           | 2        | 3          | 1        | 44           |
| <i>Inc.</i>        | 100         | 30       | 100           | 27       | 100           | 30       | 100          | 25       | 100           | 24       | 100         | 27       | 100        | 27       | 700          |
| <b>Total</b>       | 335         |          | 373           |          | 331           |          | 407          |          | 415           |          | 370         |          | 364        |          | 2595         |

Tabla 7. Frecuencias absolutas (F) y relativas (%) de los criterios de selección y exclusión de las interrogativas filtradas por zona geográfica

Los resultados exhibidos en la Tabla 7 ofrecen elementos interesantes. Lo primero que salta a la vista es el comportamiento homogéneo de las interrogativas parciales, cuyas proporciones oscilan entre 12% y 18%. El criterio de exclusión que mayor variedad exhibe es de el interrogativa con sujeto implícito (SujI). Si bien el promedio en el corpus es del 31%, las zonas de las Antillas y Caribe Continental son las que tienen menor proporción de casos de interrogativas con sujeto implícito (28% y 24%, respectivamente), mientras que España, Río de la Plata y Andina exhiben el comportamiento contrario (34% para las dos primeras zonas y 37% para la última). Ello respalda las investigaciones que han encontrado una alta expresión del sujeto en las variedades caribeñas, en comparación con el comportamiento de los sujetos explícitos, tanto nominales como pronominales, en las variedades de la Península. A su vez, es también España la zona con menor proporción de casos útiles; ello también se debe a la alta incidencia de sujetos tácitos en esta zona dialectal. La zona de Río de la Plata es la que contiene una menor proporción de interrogativas parciales (con pronombre interrogativo). Por último, las interrogativas útiles son, en promedio, el 27% de los casos filtrados.

### **2.1.2.3. Otras decisiones metodológicas**

A los criterios de exclusión/inclusión anteriormente mencionados se añadió una serie de parámetros para la consecución del análisis:

a) En el caso de oraciones interrogativas que ofrezcan una interpretación ambivalente con respecto a alguna de las variables (ver 2.1.5.), se recurrirá a la desambiguación contextual, con el objetivo de saber cuál es el valor específico que se deberá asignar a dicha variable.

Ejemplo: *¿Fue la presa cazada?* (Ant60).

En este caso, no queda claro si estamos en presencia de una estructura pasiva (con sujeto explícito: ‘la presa’) o de una cláusula con sujeto implícito, en cuyo caso ‘cazada’ sería modificador de ‘presa’ (fue *ell/ella* la presa cazada). Por lo tanto, debe recurrirse nuevamente a la herramienta digital del CORPES XXI para ampliar el contexto del caso y poder saber si la oración debe ser incluida (si es que cumple con la primera condición antes mencionada) o excluida (si cumple con la segunda).

b) En el caso de que exista una estructura que incluya dos interrogativas, solo se analizará la primera de ellas. En la mayoría de los casos, se mantiene el mismo orden sintáctico en ambas.

Ejemplo: *¿Todo el mundo es angelical, divino, nadie tiene contradicciones, desigualdades, irregularidades?* (Ant39)

Lo mismo aplica para los casos en los que la oración cuenta con más de un verbo:

Ejemplo: *¿El instrumento medirá, captará, la plenitud del fenómeno?* (Ant84)

c) En el caso de que existan perífrasis o locuciones verbales con el sujeto intercalado entre el auxiliar y el verbo auxiliado, se estimará igualmente esta posición del sujeto como una posposición, aunque sea de modo parcial:

Ejemplo: *¿Llegó la sociedad peruana alguna vez a pensarse en términos clasistas tal como éstos fueron formulados en el pensamiento socialista clásico?* (And80).

### **2.1.3. Principales dificultades en el manejo de la herramienta digital en línea**

El CORPES XXI no tiene una herramienta de recuperación de casos para estructuras gramaticales con determinadas funciones. Es un corpus diseñado para “obtener ejemplos de uso de los lemas y formas contenidos en él, ver los datos estadísticos más relevantes, comprobar las agrupaciones más frecuentes de una palabra (coapariciones), obtener la concordancia de palabras cuando aparecen en proximidad, así como clasificar y filtrar los resultados de la consulta mediante diversos criterios”, según se declara en el manual de consulta en línea<sup>49</sup>.

Ello implicó una serie de obstáculos a la hora de recuperar y filtrar los casos objeto de análisis:

1. La única forma de obtención de los casos como los que interesan en esta investigación es colocando en la casilla “forma” el signo de apertura de la interrogación (*¿*). Desde luego, la cantidad de casos que coinciden con ese criterio de búsqueda es gigantesca.

---

<sup>49</sup> Disponible en <https://apps2.rae.es/CORPES/org/publico/pages/ayuda/manual.view>

2. Uno de los elementos operativos que dificultan el trabajo con el corpus es que la ventana de trabajo caduca muy rápidamente, es decir, si el período de inactividad supera los 5 minutos, la página se reinicia y hay que colocar nuevamente todos los criterios de búsqueda de interés del investigador. Hay que llevar, entonces, una relación detallada de esos criterios y de la última página de consulta, por si es necesario retomar el trabajo. Estos elementos no son advertidos en el manual de uso, sino que el investigador los descubre sobre la marcha, con el consecuente retraso que ello implica en el plan de trabajo, pues hay que invertir horas en la familiarización con la herramienta.

3. Parte de esa familiarización es conocer cómo funciona la herramienta de exportación de casos. El CORPES XXI solo permite exportar mil (1000) casos a la vez. Cuando el caudal de casos coincidentes con el criterio de búsqueda es grande, como fue el caso de esta investigación (los casos superaban los 130.000), el proceso de exportación es especialmente demorado y no exento del riesgo de duplicación de documentos, pues corresponde al investigador controlar las páginas finales de exportación cada vez que decida cesar esta actividad (que necesita horas de trabajo). Por tanto, puede suceder que, ante un descuido, se obtengan exportaciones duplicadas. En nuestro caso, se obtuvo un total de 159 archivos con extensión txt.

4. El sistema de exportación está diseñado para el registro solo de los casos, sin contexto ampliado. Si el investigador, por alguna razón relativa a sus necesidades de análisis, necesita recuperar casos contextualizados, el trabajo deberá hacerlo manualmente, pulsando el botón (+) de ampliación de búsqueda, con las consecuentes dificultades que acarrea este tipo de trabajo. Como podrá suponerse, se necesita una gran cantidad de tiempo para hacer esta tarea. Por ejemplo, una vez extraídos los datos y seleccionados aleatoriamente en función de los criterios de selección de la muestra, no existe modo alguno de automatizar búsquedas referidas únicamente a unos ejemplos seleccionados en función de criterios que no están relacionados con los parámetros utilizados en la codificación de los datos. En consecuencia, la única manera de obtener el contexto ampliado es buscar caso a caso, poniendo cada uno de ellos en la casilla correspondiente a “forma”. Hay que cuidar, además, que haya un espacio entre el signo de interrogación, que es un criterio ortográfico, y la primera y última palabra. Sin ese conocimiento, no se encontrará ninguna coincidencia. Por ejemplo, para recuperar en un caso como *¿Existe el amor eterno?*, la

búsqueda sin espacios no arroja ninguna coincidencia porque el sistema de recuperación interpreta “*eterno?*” como una forma lingüística que no existe. Todo esto, nuevamente, implica días de familiarización y correos sistemáticos con el personal que diseñó el corpus, porque no se aclara en el manual de uso.

5. La sección de sugerencias rara vez funciona. Esto significa que cualquier duda que no esté comprendida en las preguntas frecuentes del manual de uso disponible *on-line*, debe ser despejada mediante la búsqueda de contactos fuera de la herramienta de trabajo con el corpus. En el caso de esta investigación, fue tremendamente útil el intercambio por correo electrónico con el coordinador del corpus, Guillermo Rojo, quien amablemente respondió cada una de las preguntas.

#### **2.1.4. Limitaciones en el manejo del corpus**

Como la mayoría de los *corpora* con muestras gigantescas, el CORPES XXI se sirve de la tecnología para la recogida de los datos. Ello, por supuesto, implica que están de antemano fijados los criterios para la clasificación: país, año, zona dialectal, medio, etc. Sin embargo, esta tecnología no permite distinguir entre, por ejemplo, un enunciado procedente de un hablante español que está efectivamente incluido en la rúbrica zona española y un chileno que ha escrito (o hablado) en un medio español, por lo que ha sido incluido en esta zona; o entre un chileno que es incluido en su zona dialectal o un español que ha sido, por ejemplo, entrevistado en Chile. Es el caso de los primeros 38 casos que devuelve el sistema de búsqueda, pertenecientes a una entrevista digital a Chayanne, cantante puertorriqueño, publicada en *Elmundo.es*. Los que preguntan son lectores de las más diversas partes del mundo, desde Europa hasta Asia. Controlar estas posibles fuentes de sesgo es prácticamente imposible, puesto que es información no codificada. Esperamos haber sorteado ese obstáculo mediante la aleatorización de los datos, herramienta que permite un nivel de representatividad equitativo para todos los casos, con lo cual se garantiza que los casos marginales no estén supra representados.

Además, el corpus ha sido concebido, mayormente, para investigaciones de corte lexicográfico, por lo que el motor de búsqueda de casos está especialmente pensado para ello. Al no contar con una herramienta para el *parseo* de los datos, es imposible filtrar casos según,

por ejemplo, determinadas funciones sintácticas como sujeto, verbo u objeto. Ello hizo que el proceso de selección y ordenación de los datos fuera dilatado y engorroso, lo cual, además, provocó que tuviéramos que reducir la cantidad original de casos objeto de análisis originalmente proyectada por no contar con tiempo suficiente para expandir la muestra.

### **2.1.5. Descripción de casos contenidos en el corpus**

Para cada una de las oraciones que quedaron finalmente incluidas en la muestra se procedió a registrar en una planilla Excel la siguiente información:

Variable Orden del sujeto: anteposición del sujeto (SV) y posposición del sujeto (VS).

Variable Zona geográfica: Andina, Antillas, Caribe Continental, Chile, España, México y Centroamérica y Río de la Plata.

Variable Semántica verbal: verbos materiales, mentales, relacionales, existenciales y de modulación.

Variable Realización valencial: predicados monovalentes, predicados bivalentes y predicados trivalentes.

Variable Transitividad verbal: transitivo, inergativo, inacusativo o predicado nominal.

Variable presencia de modificador circunstancial: sin CC, con CC antepuesto, con CC pospuesto y con CC intercalado.

Variable Tiempo verbal: Modo indicativo: Presente, Pretérito perfecto simple, Pretérito imperfecto, Futuro simple, Condicional simple, Pretérito perfecto compuesto, Pretérito pluscuamperfecto, Pretérito anterior, Futuro compuesto, Condicional compuesto; Modo subjuntivo: Presente, Pretérito imperfecto, Futuro simple, Pretérito perfecto compuesto, Pretérito pluscuamperfecto, Futuro compuesto.

Variable Morfología del sujeto: pronombre de primera persona del singular, pronombre tú (segunda del singular), pronombre usted, pronombre de tercera persona del singular, pronombre de primera persona del plural, pronombre de segunda persona del plural, pronombre de tercera persona del plural, pronombre posesivo, pronombre demostrativo, pronombre indefinido, pronombre numeral, sustantivo común, solo o con artículo definido,

sustantivo con singularizador (posesivo, demostrativo, indefinido, cuantificador o relativo), nombre propio o título, forma verbal de infinitivo (con o sin complementos), sujeto clausular.

Variable Animacidad del sujeto: humano, animado e inanimado.

Variable Definitud del sujeto: definido, indefinido y escueto.

Variable Complejidad del sujeto: UM (unidad nominal mínima: pronombre, nombre propio, sustantivo o infinitivo escuetos, títulos); det+SN (sintagma nominal compuesto por un elemento nominal y exclusivamente un determinante nominal); SNcoord (elementos nominales sin determinante, yuxtapuestos o coordinados; también elementos pronominales modificados por adverbios); SN+compl (elementos nominales sin determinante, acompañados de uno o varios complementos de nombre); det+SN+mod (sintagma nominal con determinante nominal, modificado por un adjetivo o un adverbio); det+SN+fprep (sintagmas nominales introducidos por un determinante, modificados por una frase preposicional o una frase nominal en aposición; también dos sintagmas nominales coordinados, estos últimos pueden tener modificadores simples); det+SNcoord (sintagmas nominales con determinante nominal, coordinados, que contengan complementos preposicionales o subordinadas adjetivas); det+SN+CR (sintagma nominal compuesto por sustantivo más determinante nominal y una cláusula de relativo o al menos dos complementos preposicionales); CS (sujeto constituido por una cláusula sustantiva); SNCompl (sintagma nominal con estructura más compleja).

Variable Función pragmática del enunciado: pregunta real, interrogativa *subjectio*, interrogativa problemática, interrogativa anticipativa, interrogativa de foco, interrogativa repetitiva, interrogativa confirmativa, interrogativa referida, interrogativa fática, interrogativa de información compartida, interrogativa clarificadora, interrogativa retórica, interrogativa deliberativa, petición u ofrecimiento, N/A<sup>50</sup>.

---

<sup>50</sup> La etiqueta N/A fue usada en casos en los que el contexto es insuficiente (o inexistente) para poder conocer la función pragmática del enunciado.

## 2.1.6. Decisiones metodológicas referidas a la descripción

### 2.1.6.1. Sobre la estimación y uso de los conceptos de transitividad y estructura argumental

Transitividad y estructura argumental son dos conceptos que distan de estar consensuados en la literatura científica. Respecto de la primera noción, es bien conocida la polémica existente entre los gramáticos. Algunas posturas teóricas establecen que existe transitividad siempre que haya complemento regido, otras estiman que es transitiva solo la cláusula cuyo verbo cuenta con un complemento directo (García Miguel, 1995). Con respecto a la estructura argumental, las posturas más extremas se debaten entre un origen léxico o construccionista de esta.

La descripción llevada a cabo en este trabajo sigue las propuestas más comunes en las gramáticas descriptivas del español respecto de lo que se entiende por transitividad, es decir, se estimará como transitivo el predicado donde exista un complemento directo exigido por el verbo. Esta decisión se basa, fundamentalmente, en una razón práctica: clasificar de este modo los predicados presentes en el corpus permitirá posteriormente una comparación con investigaciones previas, las cuales, en su mayoría, responden a este criterio de transitividad. Con respecto a la estructura argumental, se asumió una postura constructivista (Goldberg, 2002; García Miguel, 2007), que desde luego no ignora la existencia de una transitividad léxica: esta nos permite “hacer predecible cuál es el comportamiento más esperado [de un verbo] y (...) lo que efectivamente sucede en las oraciones que los hablantes usan en contextos comunicativos reales” (Ibáñez y Melis, 2015, p. 162).

Se trata, entonces, de construcciones, no de verbos, en las que se instancian ciertos argumentos. Podría hablarse, en consecuencia, de la ‘estructura argumental instanciada’ o ‘realización valencial’ en términos de García Miguel (2007). Ello implicó que no se considerara como transitiva una construcción que contenga un uso absoluto de un verbo lexicalmente transitivo, como es el caso de ‘comer’<sup>51</sup> cuando el objeto directo aparece tácito

---

<sup>51</sup> Sánchez (2008) denomina estas estructuras ‘intransitivizadas’, pues el verbo ha sido despojado de su complemento directo:

a) [La empleada] (S) [hirvió] (V) [el agua] (O)  
b) [El agua] (S) [hirvió] (V)

y es, por tanto, recuperable, como en el siguiente ejemplo: *Antonio come*. Siguiendo a Dryer, Sánchez (2008) estima que las estructuras de este tipo deben ser agrupadas por su comportamiento sintáctico junto a las intransitivas, dado que existe evidencia en otras lenguas -como es el caso del rumano- de que “el orden de los sujetos que acompañan a verbos transitivos sin objeto coincide con el de los sujetos intransitivos y no con el orden adoptado por los sujetos en cláusulas transitivas con un objeto nominal” (p. 89). Inversamente, se consideró transitiva la construcción cuyo verbo, a pesar de ser considerado como intransitivo, contenía un complemento directo. Tal es el caso de las estructuras con objetos directos cognados como *Julio murió una muerte penosa*.

También es importante mencionar que cuando aludimos a los conceptos de inacusatividad y inergatividad, estamos haciendo referencia no solo a tipos de verbos, sino a estructuras, tal como propone Mendikoetxea (1999). Así, se trata de identificar como ‘construcciones inacusativas’ las oraciones provenientes de predicados transitivos que se han inacusativizado, como es el caso de oraciones ‘se’, también llamadas pasivas medias (*La puerta se cerró*) y las construcciones con pasiva perifrástica (*Las puertas fueron cerradas*), cuyo agente queda implícito.

Respecto de la estructura argumental, siguiendo a Pavey (2010), se consideró que los predicados nominales contienen un solo argumento, es decir, el sujeto. Las estructuras que cuentan con una cópula no constituyen predicados verbales, y es por ello que no son susceptibles de un análisis actancial.

Los complementos de régimen preposicional, sobre los cuales se discute mucho (Alarcos, 1970; Seradilla, 1997; RAE, 2009, entre otros), fueron analizados también según su uso específico en la oración. Ello implica que, bien fueron clasificados como argumentos (y, por tanto, implicados semánticamente en una determinada construcción sintáctica), bien como modificadores circunstanciales (omisibles sin que se afecte el sentido expresado por el predicado).

#### **2.1.6.2. Sobre la determinación del lugar ocupado por el circunstante**

La anteposición o posposición del circunstante –elemento tenido en cuenta dentro de la variable *realización valencial* con el objetivo de contrastar nuestros datos con otras

investigaciones– puede ser confusa en algunos casos. Por ejemplo, las perífrasis verbales suelen contener adverbios intercalados entre el verbo auxiliar y el principal. Se estimó en estos casos que, si el sujeto y el verbo auxiliar están delante del CC, este último está en posición posterior aunque anteceda al verbo principal, como ocurre en *¿Los observadores árabes pueden verdaderamente tener acceso a la información en total libertad?*

### **2.1.6.3. Sobre la determinación de características referenciales del sujeto**

Además de los criterios clasificatorios reseñados en 1.4.6.1, correspondientes a los más comunes o abundantes en la expresión cotidiana, vale destacar otros casos de animacidad y definitud que pueden parecer confusos o difíciles de determinar. Se precisará a continuación, en consecuencia, qué decisión metodológica se tomó en cada caso.

a) Los sujetos constituidos por oraciones de infinitivo, en tanto alusivos a situaciones o cualidades abstractas, fueron considerados como inanimados y escuetos (1), salvo que contaran con un determinante, en cuyo caso serían inanimados y definidos (2).

(1) *¿Volverte músico fue un acierto o un desacierto?* (Rio78)

(2) *¿El ser médico te hace sentir superior?*

b) Los sujetos constituidos por demostrativos como *este*, *aquel*, *estos*, fueron considerados animados o inanimados en dependencia de la referencia. Por ejemplo, en (3) puede observarse que el demostrativo hace referencia a un elemento abstracto inanimado (estilo de gobierno); mientras en (4) hace alusión a una persona y es, por lo tanto, animado.

(3) *¿Este es el estilo del gobierno "con los pantalones bien puestos" proclamado por él en Rosario?* (Rio41)

(4) *¿Esta es Miss Colombia?* (And24)

c) En casos de uso en sentido figurado o personificaciones, los sujetos se consideraron animados, dado su contexto. Para confirmar la clasificación, se procedió a cambiar la función sintáctica del sintagma de sujeto a objeto directo usando ejemplos hipotéticos con verbos transitivos que no induzcan la presencia de un sujeto animado, con el objetivo de saber si eran susceptibles de recibir la preposición *a*, marca de animacidad (también llamada marca

diferencial de objeto). Así, sujeto de la oración (5), que entendemos que es animado, pudiera transformarse en (6), en cuyo caso se verifica la animacidad del referente (o al menos su capacidad de recibir el rasgo +animado, usando el mismo verbo).

(5) *¿Se mueven las galaxias?* (Mex88)

(6) *Observamos moverse a las galaxias.*

En el segundo de los ejemplos, las entidades a las que se hace referencia tienen un carácter agentivo, “en el sentido de entidades que desempeñan un comportamiento activo asimilable al de las personas o los animales” (RAE, 1999, p. 4146).

e) En casos cuyos referentes pueden ser ambiguos, se recurrió a una desambiguación ampliando el contexto en la concordancia que devuelve el CORPES XXI. Así, en el caso reproducido (7), se pudo comprobar que el emisor estaba haciendo referencia a unos personajes humanos, por lo que se clasificó el sujeto *alguno* como tal:

(7) *¿Alguno de ellos existió realmente o son hijos de la ficción?* (Car34)

#### **2.1.6.4. Sobre la variable complejidad del sujeto**

Las subcategorías establecidas para la investigación han sido determinadas siguiendo la bibliografía existente y las posibilidades teóricas de combinación de constituyentes para estructurar un sintagma nominal en español. Sin embargo, es evidente que el análisis de un corpus de habla real dista de tener compartimentos estancos en donde caben todos los ejemplos de una manera clara y tajante. En el proceso de descripción se encontraron, como era de esperarse, estructuras cuya composición no se correspondía con ninguna de las subcategorías establecidas. En consecuencia, se intentó aplicar un criterio lo más ajustado posible a estos casos específicos sin seguir ampliando la ya exhaustiva clasificación de subtipos con la que contábamos. Así, se mantuvo constante el criterio de complejidad que rige cada subcategoría y los casos poco claros se agruparon en las que más cercanas parecían. A continuación, se muestran algunos ejemplos y la decisión tomada en cada uno de ellos.

(1) *¿No le llamó la atención el abuelo paterno inglés de Salomón?* (Mex1)

(2) *¿Se acabaron los años de los concentrados vinos de garaje?* (Esp53)

(3) *¿Mi denuncia sobre el dengue que afectó a mi hija y mis dos sobrinos es cierto o mentira?* (Ant21)

(4) *¿Todo el mundo es angelical, divino, nadie tiene contradicciones, desigualdades, irregularidades?* (Ant39)

En (1) estamos en presencia de un sintagma nominal cuyo núcleo es modificado por varios adjetivos y un complemento preposicional. Entendemos que un ejemplo así debe ser agrupado en la subcategoría de ‘sintagmas nominales introducidos por un determinante, modificados por una frase preposicional’, atendiendo a que los sustantivos modificados por complementos preposicionales son más complejos que los modificados por adjetivos. Esta subcategoría, en consecuencia, englobaría ambos niveles de complejidad. Se incluyeron también en esta subcategoría los casos como (2), cuyo complemento preposicional contenía otra frase preposicional en su interior. Igual sucedió con ejemplos como (3), cuyo sujeto está constituido por un sustantivo que está modificado por un complemento preposicional con una subordinada adjetiva. Respecto de los casos cuyo sujeto contiene el cuantificador *todo*, como en (4), seguimos la postura de la RAE (2009). Este cuantificador universal modifica a grupos definidos y se considera un predeterminante. Las estructuras que lo contenían fueron clasificadas según la complejidad del resto de sus componentes, es decir, que no se creó una categoría especial para estos ejemplos.

Por último, queremos precisar que, en el caso de los sujetos compuestos por infinitivo, se estableció una diferenciación según apareciera como un constituyente escueto o tuviera más elementos.

#### **2.1.6.5. Sobre la determinación de la función del enunciado interrogativo**

Determinar la función que cumple un determinado enunciado en el contexto en el que ha sido emitido puede ser una tarea que algunos estimen imposible. La primera de las razones es el escaso consenso sobre qué se entiende por función, es decir, cuáles elementos se deben tomar en cuenta para constituir una etiqueta que resuma determinado uso en el discurso. Por otra parte, se discute cuál es la unidad lingüística que se debe analizar para determinar tales funciones (Sinclair & Van Gessel, 1990). Además, las clasificaciones de este tipo tienden a contener altos grados de subjetividad.

El eje central de nuestra clasificación lo constituyen las intenciones del hablante en el contexto en que se desarrolla el intercambio comunicativo. Esas intenciones han sido determinadas según la información contextual con la que contamos: se procedió a extraer el contexto anterior y posterior de cada uno de los casos del corpus, usando la herramienta de expansión del CORPES XXI. Dada la variabilidad en la extensión del contexto proporcionado por la herramienta, no establecimos un límite textual razonable para extraer conclusiones sobre la función pragmática del enunciado. Más bien, se ha tratado de aplicar la tipología definida en el marco teórico de esta investigación a partir de la información de que disponíamos. Cuando el contexto ha sido insuficiente o ambiguo, los casos se han marcado con la etiqueta N/A. Un análisis mucho más abarcador (más precisamente, que tenga en cuenta elementos prosódicos y también extralingüísticos) es deseable pero no posible en esta investigación, dada la naturaleza del corpus con que trabajamos.

Dado lo anterior, la nuestra es una propuesta preliminar susceptible de ser enriquecida con posteriores aportaciones; quiere ello decir que no pretende hacer (no puede, por las dimensiones del corpus) afirmaciones sobre todas las posibles funciones de los enunciados interrogativos en español. Más bien, se trata de registrar los usos más frecuentes presentes en los fragmentos analizados, siendo conscientes de que son un botón de muestra de una realidad mucho más heterogénea. Por último, vale destacar que, para paliar en alguna medida la falencia que implicaría la aplicación unilateral de criterios a los casos del corpus (y evitar que la subjetividad antes aludida sesgue la clasificación), se ha partido del consenso de tres expertos en la elaboración de la definición de las categorías y en la aplicación de estas últimas a la descripción de los casos que conforman el corpus.

## **2.2. El análisis estadístico**

La investigación que se propone es un estudio transversal de carácter descriptivo y sincrónico. Ello implica que su principal objetivo es observar lo que ocurre con dos o más variables (no existe manipulación en ellas por parte del investigador). Asimismo, no da cuenta de la evolución del fenómeno del orden de constituyentes en el tiempo, si bien los casos contenidos en el corpus ocupan un rango temporal de diez años (2002 y 2012). La

variable año no la hemos tenido en cuenta en el análisis por cuanto suponemos que los fenómenos lingüísticos de la naturaleza que investigamos requieren mucho tiempo para presentar cambios significativos, tal como se ha comprobado en investigaciones sobre orden del sujeto (Bogard, 2019).

Por tanto, el estudio estadístico que proponemos tiene como objetivo general estimar la relación y la influencia de factores lingüísticos y no lingüísticos (distribución geográfica) en el orden del sujeto de interrogativas totales del español.

Para ello, se aplicarán tres técnicas indispensables en este tipo de estudios:

- a) Selección de muestra.
- b) Descripción de las variables independientes (factores sintácticos, semánticos, pragmáticos y dialectales) y la variable dependiente (orden de palabras).
- c) Análisis estadístico de los datos.

### **2.2.1. Naturaleza de las variables**

Las variables comprendidas en el estudio son de carácter categórico, es decir, los valores contenidos en ellas no pueden ser sometidos a un criterio de orden. Cada variable o grupo de factores cuenta con niveles o factores en los que se discretiza, según el criterio fundamentado por el investigador, el fenómeno lingüístico de que se trate.

### **2.2.2. Las pruebas estadísticas**

La naturaleza categorial de las variables determina el tipo de análisis cuantitativo a realizar. Así, en términos de estadística descriptiva, se han utilizado tablas y gráficos de frecuencia para observar la distribución de los datos; el análisis inferencial abarca la prueba de chi-cuadrado para medir la relación entre las variables (su grado de dependencia) y el análisis de regresión logística simple y múltiple, que tiene la utilidad principal de que “permite averiguar cuál es el grado en que los grupos de factores explicativos determinan la variación de un elemento cuando todos ellos actúan conjuntamente; dicho de otro modo, el cálculo permite conocer la probabilidad general de que aparezca uno de los factores cuando actúan simultáneamente diversos grupos de factores lingüísticos y extralingüísticos”

(Moreno, 1994, p. 6). Básicamente, el cálculo de las probabilidades se realiza a partir de un procedimiento matemático que tiene en cuenta las frecuencias absolutas y relativas con que se manifiesta un factor o variable en relación con las otras, y cuando actúan todas al unísono. Tanto la estadística descriptiva como la inferencial se han realizado usando como herramienta principal el software R<sup>52</sup>.

En este punto debe ser aclarado que se ha evitado dar cuenta, en términos de determinación de tendencias, de los factores o niveles cuya cantidad total de casos es menor a diez, por considerar esta cifra como marginal<sup>53</sup>. Se ha procedido, siempre que ha sido teóricamente posible, a su fusión con otros niveles, con el objetivo de alcanzar una mayor precisión en las descripciones y posterior análisis inferencial de los datos. Además, es de esperar que los parámetros que constan de varios factores tengan una distribución desigual. Es por ello que deben ser eliminados, con el objetivo de evitar que la carencia de datos dificulte la identificación de regularidades en el corpus (Delbecque, 1987).

---

<sup>52</sup> RStudio Team (2021). RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA. URL <http://www.rstudio.com/>.

<sup>53</sup> Esta cifra se justifica en la existencia de dos niveles en la variable dependiente o respuesta. Estadísticamente, se espera contar con más de 5 ocurrencias en cada factor. En consecuencia, como la variable dependiente es binominal (anteposición y posposición del sujeto), se necesitan al menos 10 datos por cada factor.

## Capítulo 3. Resultados y discusión

### 3.1. Análisis de frecuencia y asociación de variables

El presente acápite recoge una descripción del corpus, sus tendencias generales en términos de ocurrencias y de relaciones entre las variables dependiente e independientes. Básicamente, cada epígrafe contiene dos partes: la primera, a modo de “fotografía”, recoge la distribución de casos (frecuencias absolutas y relativas) para cada una de las categorías de las variables que intervienen en el estudio, junto a las proporciones de anteposición y posposición del sujeto.

La segunda parte presenta el análisis de asociación entre variables o análisis bivariado, a partir de la confección de tablas de contingencia y de la realización de la prueba *chi-cuadrado*, que evalúa el grado de independencia entre las variables. Esta asociación se encuentra graficada en todos los casos mediante los diagramas de mosaico de R (Hornik, Zeileis & Meyer, 2006). Así, además de comentar los datos puramente descriptivos, se ofrece un panorama general de cómo se establecen las relaciones entre las variables.

Antes de pasar al contenido del apartado, precisaremos cómo se han construido los gráficos de mosaico antes mencionados y cuál es su utilidad para la investigación.

Los diagramas de mosaico de R permiten una visualización exploratoria óptima, a partir de la cual se puede observar de manera muy intuitiva cómo se comportan los datos y, en consecuencia, tomar decisiones metodológicas y teóricas, de manera de ajustar los datos lo más posible a los objetivos buscados (Le Guen, 2003). “Un diagrama de mosaico es básicamente una visualización proporcional al área de frecuencias (típicamente observadas), compuesta de mosaicos (correspondientes a las celdas) creados por divisiones recursivas verticales y horizontales de un rectángulo. Por lo tanto, el área de cada mosaico es proporcional a la entrada de celda correspondiente dadas las dimensiones de las divisiones anteriores” (p. 1).

Estas representaciones permiten adicionar al gráfico el estadístico *chi-cuadrado*, que tiene como objetivo demostrar que existe independencia entre las variables: valores por debajo de 0,05 indican que se rechaza la hipótesis nula de independencia y se asume que las variables están relacionadas significativamente. Sin embargo, este estadístico es muy

sensible al tamaño muestral, es decir, si la muestra es grande, diferencias pequeñas entre las frecuencias observadas y teóricas de las tablas de contingencia pueden dar como resultados chi-cuadrados altos (o estadísticamente significativos).

Es por ello conveniente analizar los residuos corregidos tipificados: estos permiten conocer qué casillas de la tabla de contingencia que está detrás del análisis -o qué cruce de variables- contienen más o menos casos de los que cabría esperar si las variables fueran independientes. Esto se logra mediante la estandarización de las diferencias entre frecuencias observadas y las teóricas. Este estadístico expresa una relación más o menos fuerte entre las variables (mayor o igual que 2 o menor o igual que -2 son los valores de referencia para una asociación fuerte).

Los gráficos de mosaico muestran en colores las casillas donde los cruces de variables tienen residuos que exceden los valores antes mencionados: los mosaicos en rosado representan los cruces donde hay significativamente menos casos que los esperados (residuos negativos), y los mosaicos pintados en azul expresan las relaciones donde hay más casos de los esperados (residuos positivos). Estos colores pueden ser más saturados de acuerdo con el tamaño del residual: los residuos grandes están más saturados y los residuos pequeños, menos (Hornik, Zeileis & Meyer, 2006, p. 27).

El capítulo tiene la siguiente organización interna: primeramente, se expone de manera breve cuál es la tendencia general del orden del sujeto en el corpus con que se trabajó; en segundo lugar, se presentan y discuten los resultados obtenidos para la variable extralingüística del estudio; en tercer lugar se encuentran los hallazgos de las variables relacionadas con el predicado; en cuarto lugar, se comenta la influencia en de las variables referidas al sujeto en el orden de palabras; por último, se recoge el comportamiento de la variable pragmático-discursiva del estudio: el valor pragmático del enunciado interrogativo.

### **3.1.1. El orden predominante en las interrogativas totales del corpus**

El Gráfico 2 muestra el comportamiento general de las interrogativas contenidas en el corpus. Como puede observarse, el 60% (426 casos) presenta un orden VS, mientras que el 40% (274 casos) muestra una secuencia SV. Este resultado es coherente con los postulados de obras, tanto normativas como empíricas, sobre orden de palabras más frecuente en la

interrogación: en efecto, parece haber una preferencia por el orden verbo-sujeto. Parece curioso, además, que las proporciones de orden de nuestros datos se comporten de manera inversa a las encontradas por Delbecque (2011) para la declaración: para 6 casos de posposición, hay 4 de anteposición.

Sin embargo, teniendo en cuenta los referentes teóricos citados en el capítulo 1, que apuntan a la existencia de un orden no marcado VS (Escandell, 1999), sería esperable encontrar una proporción muy alta de posposición del sujeto. Nuestros datos, no obstante, muestran otra tendencia: una alta proporción de sujetos antepuestos (40%) puede estar sugiriendo una alternancia de orden, más que un orden no marcado o neutral. En lo sucesivo, intentaremos explicar qué elementos lingüísticos se pueden asociar con qué secuencia, y si existe una restricción contextual en uno u otro orden. Con ello podremos saber si la secuencia VS es, efectivamente, la no marcada en la interrogación total.

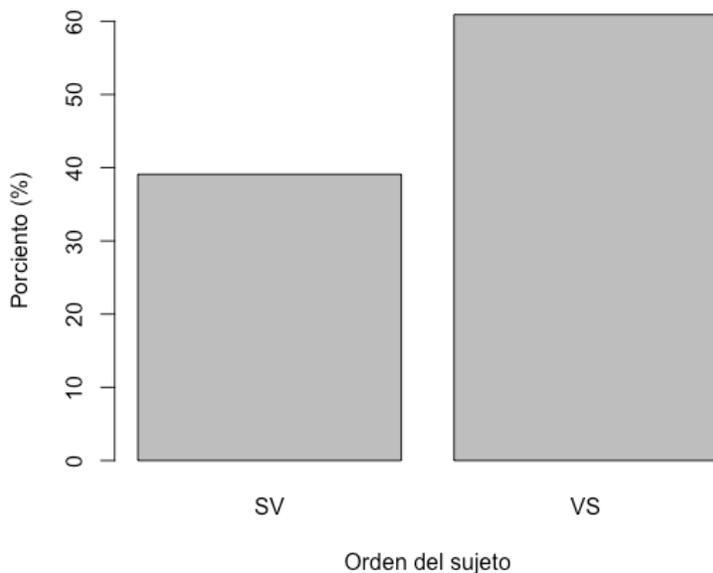


Gráfico 2. Proporciones de interrogativas con anteposición y posposición del sujeto en el corpus

### **3.1.2. Variable extralingüística**

#### **3.1.2.1. El orden del sujeto respecto del verbo según zona geográfica**

El Gráfico 3 muestra el comportamiento de los casos contenidos en el corpus, agrupados por zona geográfica. Recordemos que cada zona cuenta con 100 casos. España es la zona que tiene mayor número de casos de posposición del sujeto, 78% del total. A esta zona le sigue México y Centroamérica con 66% de los casos. Las Antillas es la tercera zona con mayor cantidad de casos en este orden (62%). Este último resultado, si bien no muestra un comportamiento categórico, no estaría en consonancia con la tendencia teórica a la anteposición del sujeto en interrogativas parciales del Caribe (Toribio, 2000; Morales, 1989), aunque sí con la propuesta de Brown y Rivas (2011) sobre el uso prevalente del orden VS en este tipo de interrogativas. Dicho esto, y como se verá en el apartado relativo a la morfología del sujeto, estos resultados deben ser matizados, pues parece haber una diferencia clara según el tipo de sujeto: con los pronombres personales de primera y segunda personas, el orden mayoritario para la zona es el de SV.

Por último, el Caribe Continental no muestra una preferencia por un orden determinado (50% casos responden a uno y otro orden), mientras Río de la Plata apenas sobrepasa la mitad de los casos con orden VS (53% versus 47%). Respecto de esta última zona, una tendencia similar había sido postulada por Lapesa (1992) para las interrogativas parciales.

Podría decirse, frente a este panorama, que existe una tendencia marcada a la posposición del sujeto en España y México y Centroamérica. La tendencia al orden VS en las interrogativas del corpus es menos acusada, aunque prevalente, en Chile y las Antillas (60% y 62% de los casos, respectivamente). Las zonas Andina y Río de la Plata prefieren también este orden (57% y 53%, respectivamente), si bien no se trata de una tendencia categórica, pues la proporción de casos es más balanceada, especialmente en el caso de la segunda zona. Por último, el Caribe Continental no muestra preferencia por ningún orden (50% de los casos para cada variante), por lo que se puede afirmar que es neutral. Este último resultado sí es coherente con las evidencias a favor de una tendencia al orden SV en los países caribeños.

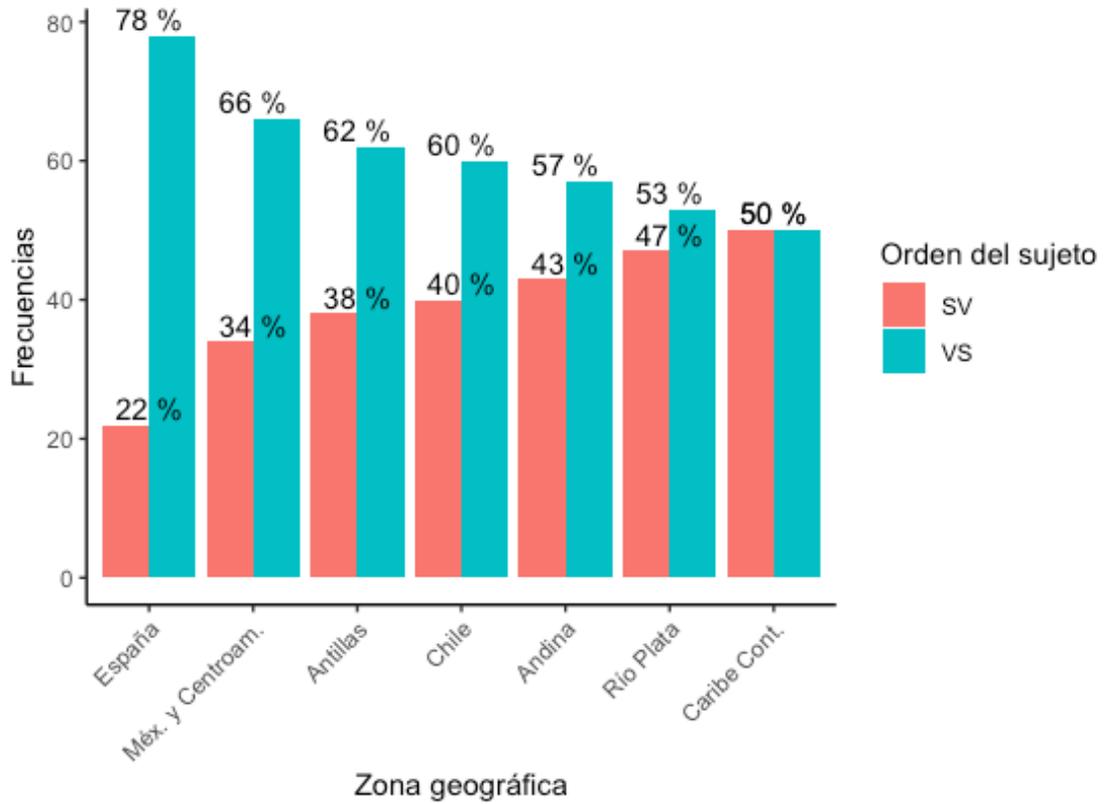


Gráfico 3. Casos con orden SV y VS, agrupados por zona geográfica. Se presentan las zonas en orden decreciente tomando en cuenta la proporción de casos con orden VS

Respecto de la asociación entre las variables zona geográfica y orden de palabras, el Gráfico 4 muestra que existe una sola categoría estadísticamente significativa: España. Tanto para el orden VS como para el SV la zona muestra valores que exceden las frecuencias esperadas, por lo que su relación con el orden es estadísticamente significativa; es decir, que los resultados no se deben al azar. España, por tanto, desfavorece significativamente el orden SV (residuos de -2,7), mientras favorece de manera muy fuerte el orden VS (residuos de 2,2).

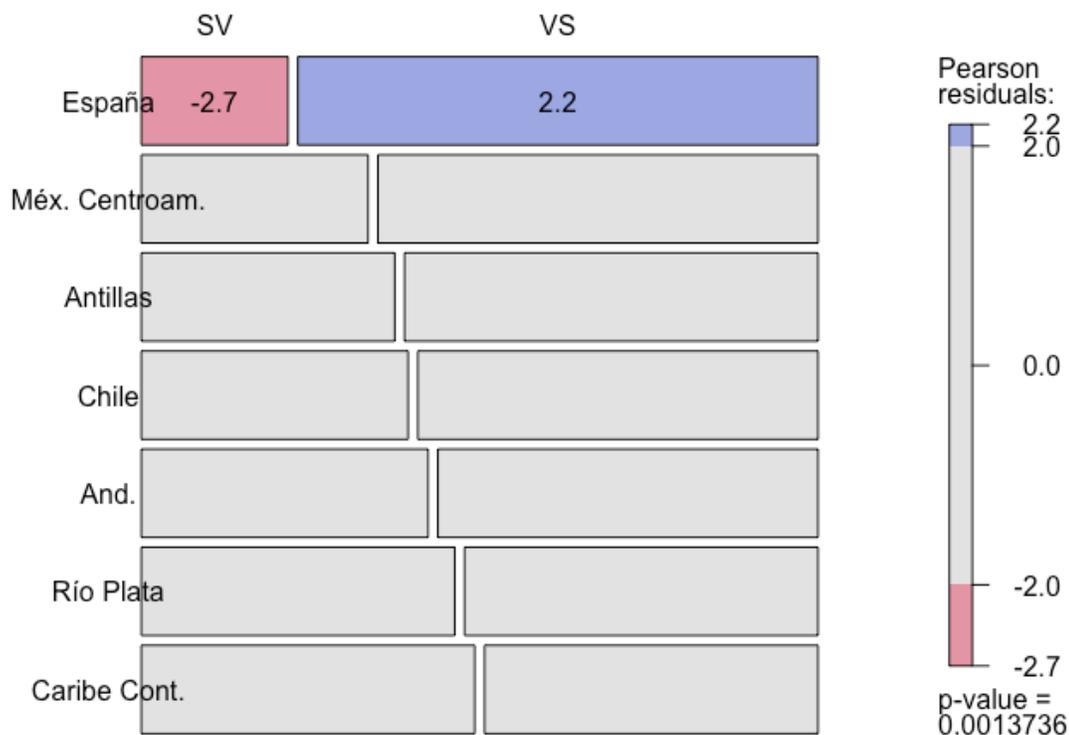


Gráfico 4. Representación de mosaicos de las variables Orden del sujeto y Zona geográfica

Para el resto de las zonas, las diferencias observadas no son estadísticamente significativas. Puede verse que, aún cuando en la mayoría de las zonas predomina el orden VS, como vimos anteriormente, las relaciones entre los cruces muestran solo uno con peso estadístico considerable.

### 3.1.3. Variables referidas al predicado

#### 3.1.3.1. El orden del sujeto respecto del verbo según semántica verbal

Como puede observarse en el Gráfico 5, los verbos de tipo relacional (tales como *ser*, *estar*, *quedar*, *resultar*, *equivaler* y *tener*) son los que prevalecen en el corpus: ocupan el 52,9% de los casos, con 370 recuentos. Les siguen los verbos mentales (*gustar*, *temer*, *desear*, *ver*), que representan un 16,3%, con 114 casos; los materiales (*mover*, *romper*, *comer*), que ocupan el 13,6% del corpus, con 95 casos; los existenciales (*haber*, *ocurrir*, *nacer*, *existir*), que promedian un 12,4%, con 87 casos; los procesos verbales (*decir*, *pedir*, *felicitar*), que constituyen un 3,1%, con 22 casos; y los de modulación (*incitar*, *dejar*, *permitir*), que son el

1,7% del corpus, con 12 casos. Puede apreciarse que existe un desbalance importante entre la categoría de procesos relacionales y el resto de las categorías; además, las categorías verbal y modulación presentan proporciones muy bajas respecto del total del corpus, lo que debe tenerse en cuenta para los análisis inferenciales. La prevalencia de procesos verbales de tipo relacional (cuyos verbos más frecuentes son *ser* o *estar*, los cuales suman 251 dentro del total de la categoría antes mencionada, para un 67,8%) es coherente con lo apuntado por Brown y Rivas (2011), quienes encontraron que estas formas están altamente relacionadas con la expresión de sujeto gramatical. En tanto todas nuestras interrogativas tienen sujeto explícito, es de esperar que sean estos verbos los más abundantes en el corpus.

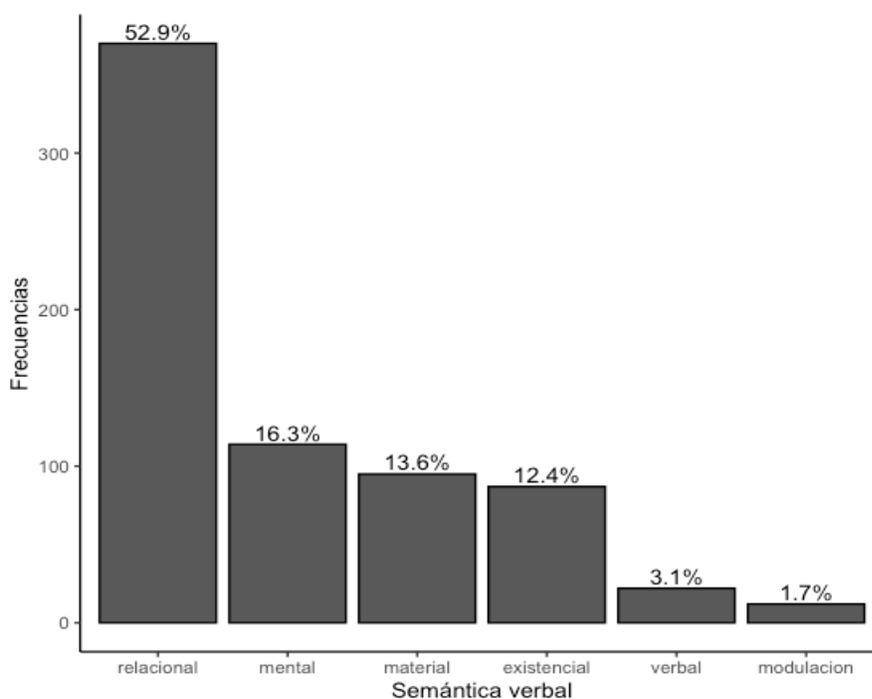


Gráfico 5. Representación de las proporciones de las categorías pertenecientes a la variable Semántica verbal en el corpus

Respecto del orden del sujeto, puede verse en el Gráfico 6 que la secuencia VS es mayoritaria en los verbos existenciales (74,7% del total), los predicados con verbos relacionales (60,5%), los materiales (60%), los mentales (57%) y los verbales (54,5%). Solo los verbos de modulación exhiben una tendencia contraria, con un 75% de orden SV, pero por representar una proporción muy baja del corpus, no se puede interpretar este resultado en términos de tendencia.

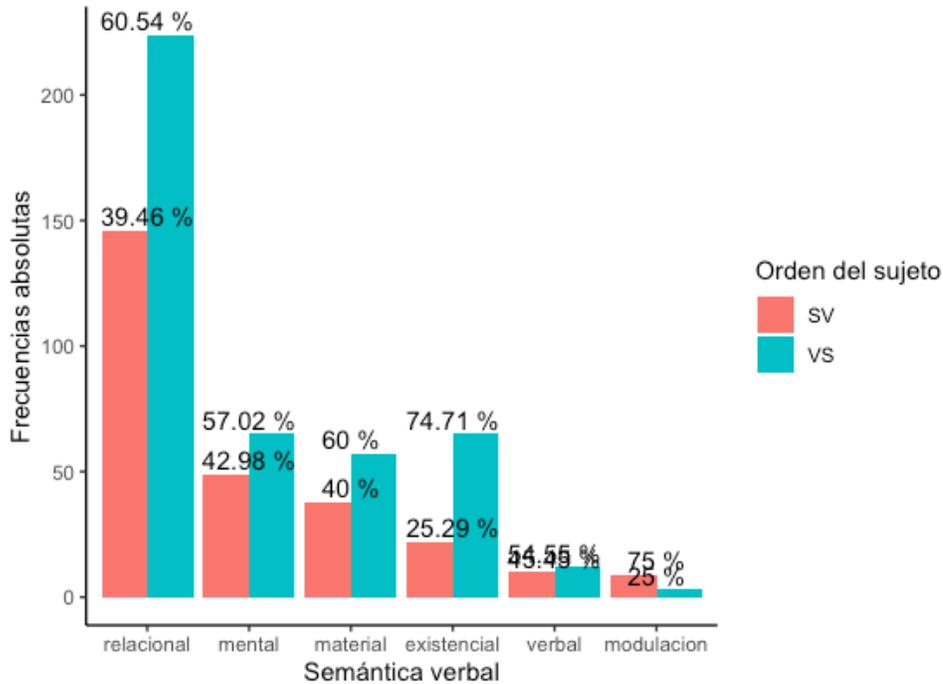


Gráfico 6. Casos con orden SV y VS, agrupados por semántica verbal. Se presentan las categorías en orden decreciente según su número total de casos. Las barras expresan la proporción de casos para cada orden

A tono con investigaciones previas sobre el rol de la semántica verbal en la determinación del orden de constituyentes (Contreras, 1978; Fernández Soriano, 1993; Ocampo, 1991; Sánchez, 2008; Brown y Rivas, 2011 y 2013; Rosemeyer, 2018), los verbos relacionales y existenciales tienen un alto porcentaje de sujetos pospuestos. Tal como señalan Delbecque (1987) y López Meirama (2006), la posposición del sujeto en verbos existenciales puede estar motivada por el bajo valor informativo de estos verbos, lo que a su vez corrobora la propuesta de Bolinger (1954) de que los sujetos pospuestos al verbo son más informativos que este último.

Además, la menor proporción de posposición del sujeto en verbos mentales parece responder, como afirma Morales (1997), a que estos verbos expresan acciones propias de los seres humanos, lo que tiende a favorecer el orden SV, dada su agentividad. También Brown y Rivas (2011 y 2011a) y Rosemeyer (2018) habían encontrado que los verbos mentales favorecen la anteposición.

La Tabla 8 muestra el cruce entre los verbos mentales, agrupados por orden, y la animacidad del sujeto. Como puede observarse (en **negrita**) del total de verbos mentales, el

70% coocurre con un sujeto humano. A su vez, el 55% de estos sujetos aparece antepuesto, versus 45% pospuesto al verbo. Así sucede en los siguientes ejemplos:

*¿Nuestro técnico conoce a los rivales? (And25)*

*¿Usted cree que se puedan cumplir esas aspiraciones ciudadanas? (Car1)*

Puede comprobarse, entonces, que dentro de esta categoría los sujetos con rasgo +humano son los únicos que tienden a anteponerse: en efecto, los sujetos con rasgo +animado y +inanimado aparecen pospuestos con una proporción mucho mayor.

| <i>Animacidad</i>     | <i>Mental “SV”</i> |                                 | <i>Mental “VS”</i> |                                 | <i>Total</i> | <i>Prop. (%)</i> |
|-----------------------|--------------------|---------------------------------|--------------------|---------------------------------|--------------|------------------|
|                       | <i>Frec.</i>       | <i>% Dentro de la categoría</i> | <i>Frec.</i>       | <i>% Dentro de la categoría</i> |              |                  |
| <i>animado</i>        | 2                  | 22%                             | 7                  | 78%                             | 9            | 8%               |
| <i>humano</i>         | 44                 | <b>55%</b>                      | 36                 | <b>45%</b>                      | 80           | <b>70%</b>       |
| <i>inanimado</i>      | 3                  | 12%                             | 22                 | 88%                             | 25           | 22%              |
| <i>Total de casos</i> | 49                 |                                 | 65                 |                                 | 114          | 100%             |

*Tabla 8. Frecuencias totales y absolutas de los cruces entre verbos mentales y animacidad del sujeto. Los datos dentro de la categoría verbos mentales se han agrupado por orden (SV y VS)*

Por último, respecto de los verbos de dicción o procesos verbales, también Brown y Rivas (2011) y Rosemeyer (2018) habían observado una predisposición a la anteposición del sujeto. En nuestro caso, los enunciados con este tipo de verbo son los que menos se posponen de todas las categorías –sin tener en cuenta los verbos de modulación, francamente facilitadores del orden SV–. De cualquier forma, la disparidad de datos totales presentes en el corpus hace que estos deban ser observados con cautela.

De las 12 instancias de verbos de movimiento encontradas en el corpus (*salir, entrar, ir, volver, caminar, etc*), hay 10 (o sea, el 83%) que presentan el sujeto pospuesto al verbo. Esto está en consonancia con los hallazgos de Bentivoglio y Weber (1986) para la declarativa, quienes postularon que los sujetos de estos verbos tienden a adoptar esta posición<sup>54</sup>, producto de su naturaleza presentativa. También coherentemente con lo planteado por estos autores, y según una tendencia del español, los verbos pseudoimpersonales (*gustar, preocupar,*

<sup>54</sup> Teniendo en cuenta este referente investigativo, Rivas (2013) incluyó los verbos de movimiento en su investigación sobre los factores que condicionan el orden en las oraciones principales y subordinadas. El autor no encontró un peso estadístico considerable de este tipo de verbo para el primer tipo de estructura, pero sí para el segundo.

*convenir*, etc), que suman 26 casos dentro del corpus, presentan el sujeto pospuesto en una alta proporción de casos (96%).

Con respecto a la asociación entre las variables, el Gráfico 7 muestra que existe un solo cruce categorial cuya relación es estadísticamente significativa: se trata de los verbos existenciales y el orden SV. Sin embargo, puede verse que, si bien este tipo de verbo aparece como un desfavorecedor de la anteposición del sujeto (el residuo es de -2,1), no hay datos suficientes para afirmar que sea un poderoso favorecedor del orden inverso, pues el residuo no alcanza el nivel crítico. Los verbos relacionales ocupan una parte importante de la muestra (repárese en el tamaño de los mosaicos para esta categoría), mientras los de modulación tienen poca presencia en el corpus<sup>55</sup>.

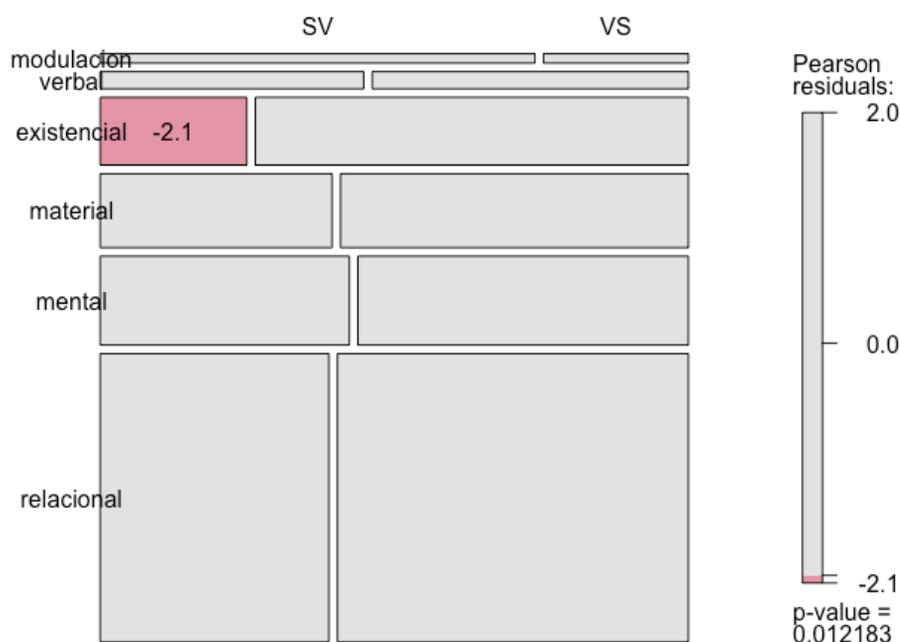


Gráfico 7. Representación de mosaicos de las variables Orden del sujeto y Semántica verbal

En resumen, si bien la posposición del sujeto es generalizada en los procesos verbales estudiados en esta investigación (solo los verbos de modulación tienen el sujeto antepuesto en la mayor parte de los casos), son los verbos existenciales, relacionales y materiales los que

<sup>55</sup> Se realizó una nueva reestructuración de la variable, incluyendo los verbos de modulación dentro de los materiales. Los resultados de la prueba chi-cuadrado, sin embargo, no sufrieron grandes modificaciones, y la categoría existencial continuó siendo la única cuya influencia es significativa en el orden.

tienen una mayor proporción de orden VS, siendo el primer tipo de proceso el único que muestra una relación estadísticamente significativa con una de las dos secuencias de constituyentes estudiadas. Los procesos mentales y verbales muestran una tendencia menos acusada a la posposición.

### 3.1.3.2. El orden del sujeto respecto del verbo según tiempo verbal

La variable en cuestión presenta un gran desbalance, tal como muestra el Gráfico 8. El tiempo presente ocupa el 69% del corpus (con 483 casos), mientras que el resto de las categorías no alcanza más del 10%. El pretérito perfecto compuesto representa un 3,14% del corpus, con 22 casos, el pretérito imperfecto un 4,71%, con 33, el condicional simple un 5,29%, con 37, el futuro simple un 7,43%, con 52 y el pretérito perfecto simple un 9%, con 63. El pluscuamperfecto de subjuntivo, el pretérito imperfecto de subjuntivo, el condicional compuesto de indicativo y el futuro tienen menos del 1% de los casos cada uno.

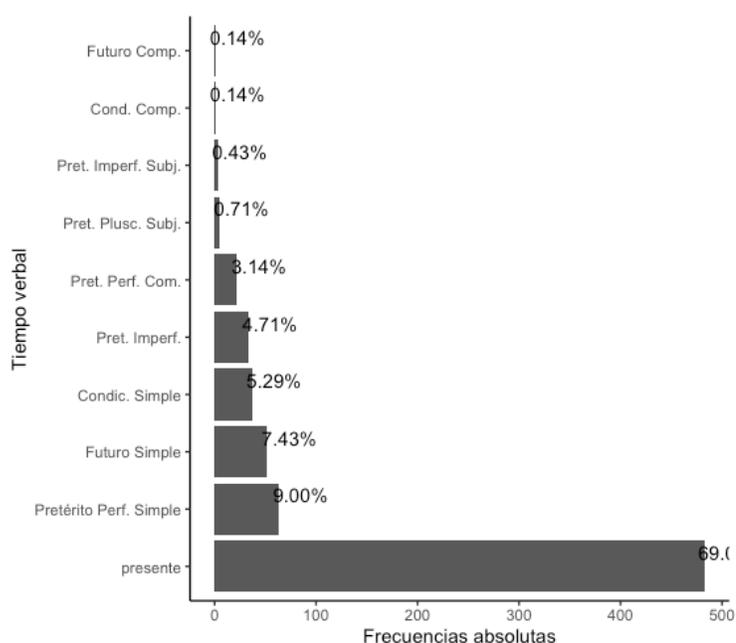


Gráfico 8. Proporción de casos por categoría de la variable tiempo verbal

Respecto del orden del sujeto, exceptuando los casos marginales antes mencionados, el orden VS es mayoritario en todas las categorías. El Gráfico 9 muestra las proporciones según el orden. Como puede verse, el pretérito imperfecto es el que exhibe un mayor índice

(75,8%), seguido del futuro simple (71,1%), el condicional simple (67,6%), el pretérito perfecto simple (62%), el pretérito perfecto compuesto (59%), y el presente (58%).

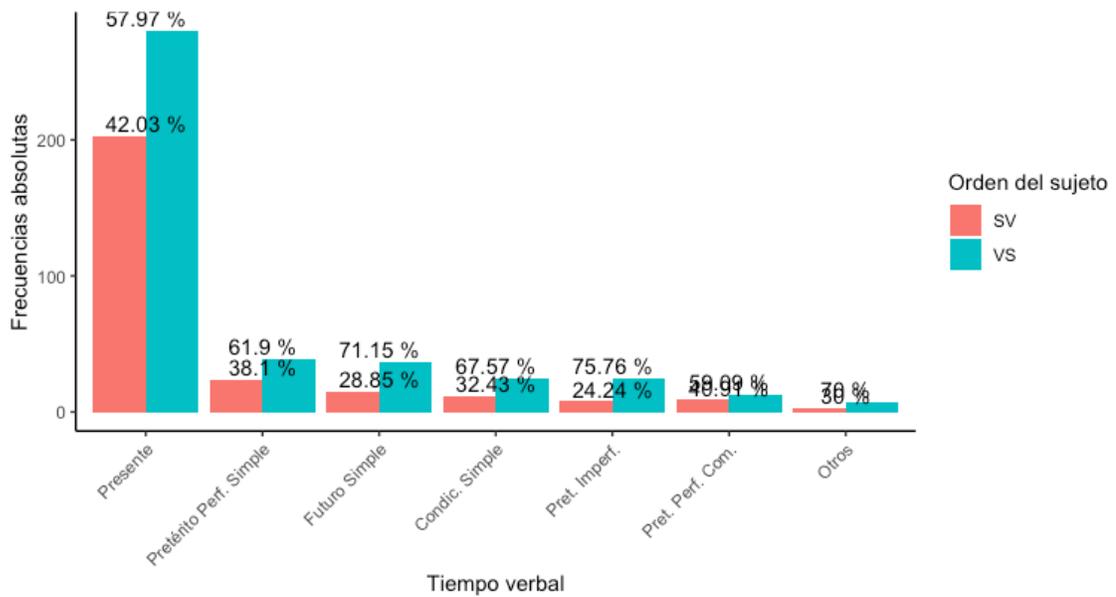


Gráfico 9. Casos con orden SV y VS, agrupados por tiempo verbal. Se presentan las categorías en orden decreciente según su número total de casos. Las barras expresan la proporción de casos para cada orden. Las categorías marginales (menos de 10 casos) se han agrupado en la etiqueta "Otros"

Respecto de la relación entre Tiempo Verbal y Orden del sujeto, el gráfico de mosaico, a continuación, ilustra su asociación:

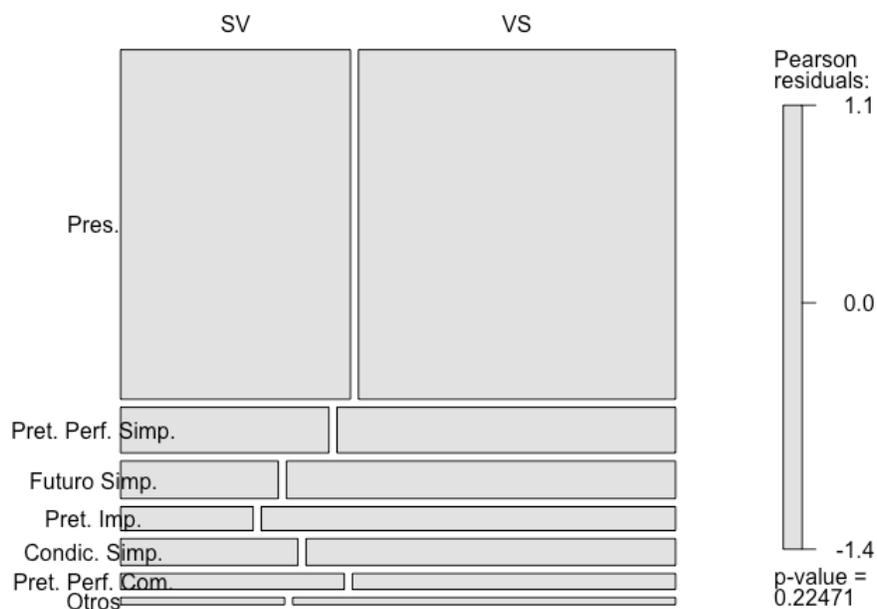


Gráfico 10. Representación de mosaicos de las variables Orden del sujeto y Tiempo verbal

En el eje y se muestran (de arriba hacia abajo) los mosaicos correspondientes a: presente de indicativo (*Pres.*), pretérito perfecto simple (*Pret. Perf. Simp.*), futuro simple (*Futuro Simp.*), pretérito imperfecto (*Pret. Imp.*), condicional simple (*Cond. Simp.*), pretérito perfecto compuesto (*P. Perf. Comp.*), pretérito pluscuamperfecto de subjuntivo (*P. Plusc. Subj.*), pretérito perfecto compuesto (*Pret. Perf. Comp.*), y *Otros*, donde se han agrupado las categorías marginales.

Como puede observarse, no existe una asociación estadísticamente significativa entre estas dos variables (el valor de  $p$  es superior a 0,05), y los mosaicos dan cuenta de ello. Tal como puede constatar, los residuos no superan el valor crítico ( $\pm 2$ ), por lo que ningún mosaico se colorea. Ello indica que ninguno de los cruces categoriales tiene datos suficientes para demostrar una relación fuerte, o bien que estos resultados pueden explicarse por el azar. Nuestros resultados se condicen con la propuesta de Brown y Rivas (2011), quienes no encontraron influencia del tiempo verbal en el orden, y contradicen los hallazgos de Lipski (1977), quien había esbozado la posibilidad de una relación entre el tiempo presente y el orden SV, si bien debe ser dicho que es el tiempo presente el que menos posposición del sujeto presenta en nuestros datos. Si se realiza un reordenamiento de la variable, agrupando los tiempos de manera más general, es decir, según ellos pertenezcan al momento presente, pasado o futuro, la asociación entre las variables es marginalmente significativa ( $p=0.048$ ), por lo que parece deseable un estudio con una mayor cantidad de datos (y más equilibrado) para poder extraer conclusiones atinadas sobre la relación entre el tiempo verbal y el orden del sujeto en la interrogación absoluta.

Respecto de la influencia de los tiempos verbales compuestos en el orden de palabras, Rodríguez (2008) había postulado una tendencia a la anteposición del sujeto. Nuestro análisis, sin embargo, no permite respaldar esta hipótesis, quizás por la escasa disponibilidad de datos. Como puede observarse en la Tabla 9<sup>56</sup>, los tiempos verbales compuestos representan 4% del corpus (solo 29 casos); de ellos, el 62% (18 casos) presentan el sujeto pospuesto. Los tiempos simples, por su parte, constituyen el 96% del corpus y tienen el sujeto

---

<sup>56</sup> El color azul representa las proporciones respecto del corpus, y el color verde los porcentajes al interior de la categoría

pospuesto con una proporción del 61%. En consecuencia, no se puede plantear una diferencia en términos de tendencia a la anteposición del sujeto entre tiempos compuestos y simples.

| <i>Orden del sujeto</i> | <i>Tiempo verbal</i>  |                         | <i>Total</i>           |
|-------------------------|-----------------------|-------------------------|------------------------|
|                         | Compuesto             | Simple                  |                        |
| SV                      | 11<br>4 %<br>37.9 %   | 263<br>96 %<br>39.2 %   | 274<br>100 %<br>39.1 % |
| VS                      | 18<br>4.2 %<br>62.1 % | 408<br>95.8 %<br>60.8 % | 426<br>100 %<br>60.9 % |
| <b>Total</b>            | 29<br>4.1 %<br>100 %  | 671<br>95.9 %<br>100 %  | 700<br>100 %<br>100 %  |

$$\chi^2=0.000 \cdot df=1 \cdot \phi=0.005 \cdot p=1.000$$

Tabla 9. Tabla cruzada tiempo verbal y orden del sujeto. Los niveles de la primera variable han sido reducidos a dos: tiempos simples y tiempos compuestos

### 3.1.3.3. El orden del sujeto respecto del verbo según realización valencial

Como puede observarse en el Gráfico 11, en el corpus predominan los predicados monovalentes, con un total de 402 casos (57,4% del total). Los predicados bivalentes constituyen un 40,3% del total (282 casos) y los trivalentes representan el 2,3% de la muestra (16 casos en total).

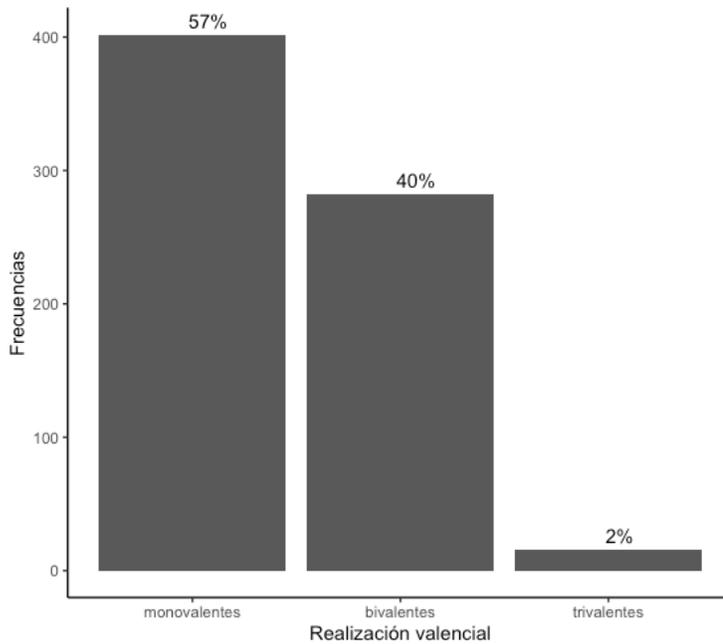


Gráfico 11. Proporción de casos por categoría dentro de la variable Realización valencial

Respecto del orden del sujeto, el Gráfico 12 muestra que la secuencia VS ocurre en una mayor proporción de casos en oraciones con predicados monovalentes (66,9 %), y en un 53,9% de las construcciones con predicados bivalentes. Los predicados trivalentes muestran la tendencia contraria: el orden SV se manifiesta en el 68,8% de los casos.

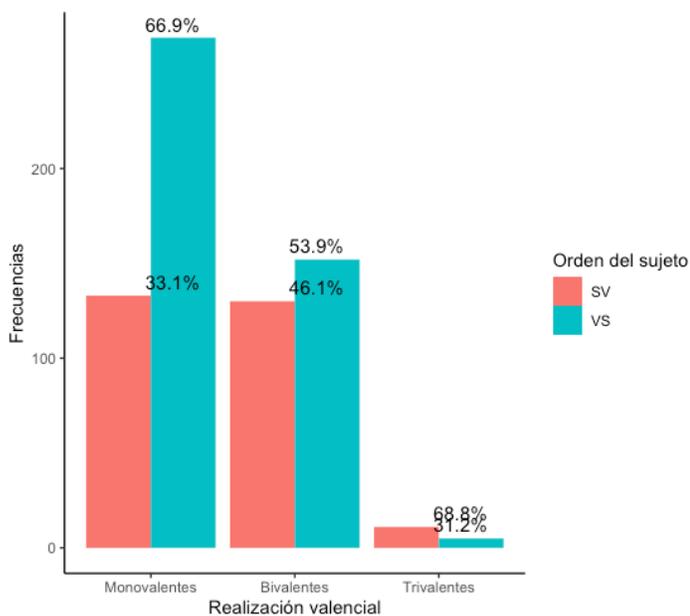


Gráfico 12. Casos con orden SV y VS, agrupados por realización valencial. El eje Y muestra las frecuencias absolutas. Se presentan las categorías en orden decreciente según su número total de casos. Las barras expresan la proporción de casos para cada orden

Como puede apreciarse, parece haber una tendencia entre la cantidad de argumentos de un verbo y el orden adoptado por el sujeto: mientras más argumentos tiene un predicado, más tiende el sujeto a aparecer antepuesto.

El Gráfico 13, sin embargo, muestra que no existe una relación fuerte entre ninguna de las categorías de la variable y el orden adoptado por el sujeto, si bien el valor p es estadísticamente significativo.

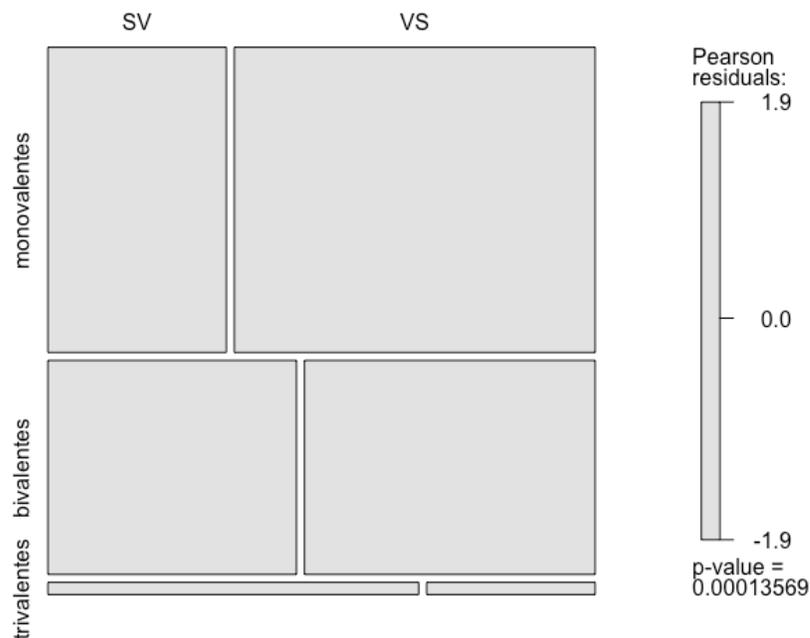


Gráfico 13. Representación de mosaicos de las variables Orden del sujeto y Realización valencial

Puede observarse que, si bien la prueba chi-cuadrado arroja una relación significativa entre las variables –ya habíamos comentado antes que el valor p es sensible al tamaño muestral– no existe ningún cruce de variables cuya asociación sea fuerte. Los predicados monovalentes son los más abundantes en el corpus, y ellos ostentan mayoritariamente un orden VS, pero sin que la relación con ese orden sea estadísticamente significativa, de la misma manera que los verbos con tres argumentos, a pesar de tener un mayor número de sujetos antepuestos que pospuestos, no influyen de manera determinante en ningún orden.

A pesar de los resultados antes obtenidos, se observa en nuestros datos la alta predominancia del orden VS en predicados de un solo argumento, resultado coherente con su comportamiento en la declaración, tal como muestran los estudios de Delbecque (1987) y Rivas (2013)<sup>57</sup>. Para poder clarificar las diferencias entre los verbos de un argumento y los que tienen más de uno, se reagrupó la variable: los verbos bivalentes y trivalentes en la categoría ‘más de un argumento’ y se mantuvo la categoría de ‘monovalentes’. Como puede observarse en la Tabla 10, si bien sigue predominando la posposición del sujeto en predicados

<sup>57</sup> Silva-Corvalán (1982) también encontró una preponderancia del orden VS en predicados con un solo argumento, pero solo en el 54% de los casos, por lo que concluye que no existe una fuerte relación entre los predicados monovalentes y la posposición del sujeto.

cuyos verbos tienen más de un argumento, la proporción de casos es comparativamente menor que en los verbos monovalentes.

| <i>Orden del sujeto</i> | <i>Realización valencial</i> |               | <i>Total</i>  |
|-------------------------|------------------------------|---------------|---------------|
|                         | más de 1 arg.                | monovalentes  |               |
| SV                      | 141<br>47.3 %                | 133<br>33.1 % | 274<br>39.1 % |
| VS                      | 157<br>52.7 %                | 269<br>66.9 % | 426<br>60.9 % |
| <b>Total</b>            | 298<br>100 %                 | 402<br>100 %  | 700<br>100 %  |

$$\chi^2=13.958 \cdot df=1 \cdot \phi=0.144 \cdot p=0.000$$

Tabla 10. Tabla cruzada de las variables realización valencial y orden del sujeto. Los niveles originales –bivalentes y trivalentes– se han fundido en uno solo: más de un argumento.

El análisis de independencia entre las variables permite comprobar que, efectivamente, existe una asociación fuerte entre los predicados de más de un argumento y el orden SV, lo que se condice con los hallazgos de Silva-Corvalán (1982) y Rivas (2013), eso sin perjuicio de que nuestros datos se comporten preponderantemente de forma inversa, lo que puede estar explicado por la influencia de la fuerza ilocutiva. Obsérvese en la Tabla 11 que los residuos tipificados corregidos para el cruce antes mencionado tienen un valor positivo de 2,25, lo que significa que estos predicados suelen favorecer de manera significativa la anteposición del sujeto.

| <i>Valencia</i> | <i>Orden del sujeto</i> |       |
|-----------------|-------------------------|-------|
|                 | SV                      | VS    |
| más de 1 arg.   | 2,25                    | -1,81 |
| monovalentes    | -1,94                   | 1,56  |

Tabla 11. Residuos corregidos tipificados para los cruces de las variables realización valencial y orden del sujeto. Los niveles originales –bivalentes y trivalentes– se han fundido en uno solo: más de un argumento

### 3.1.3.4. El orden del sujeto respecto del verbo según transitividad

El Gráfico 14 muestra la proporción de las categorías de la variable en el corpus.

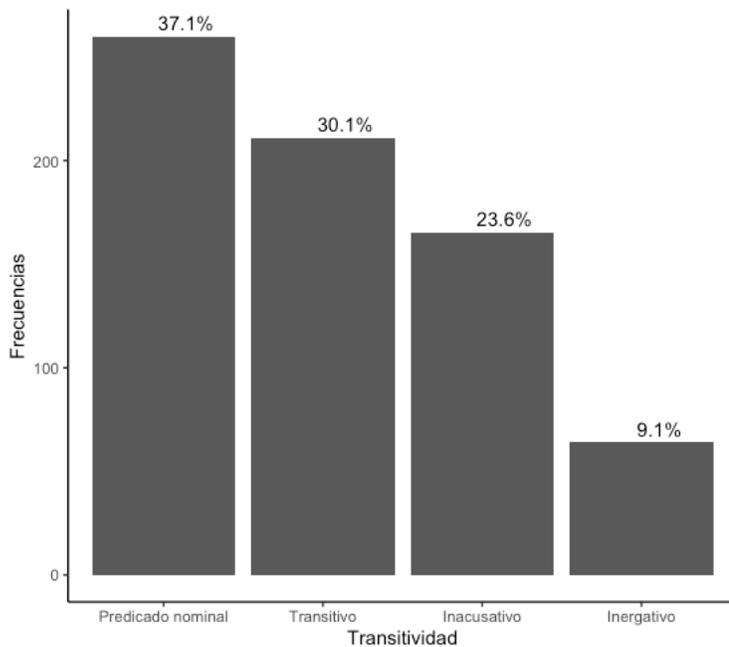


Gráfico 14. Proporción de casos por categoría dentro de la variable Transitividad

Como puede observarse, los verbos transitivos constituyen el 30,1% de la muestra (211 casos), los verbos inacusativos representan el 23,6% (165 casos) del total, mientras que los inergativos alcanzan el 9,1% del corpus, con 64 casos<sup>58</sup>.

---

<sup>58</sup> Los predicados nominales, con un gran porcentaje, no nos interesan por el momento, pues ellos no pueden ser susceptibles de ser analizados respecto de la estructura argumental. Hemos incluido este nivel solo con el fin práctico de poder realizar la descripción de todos los casos contenidos en el corpus, sin dejar ninguna celda vacía.

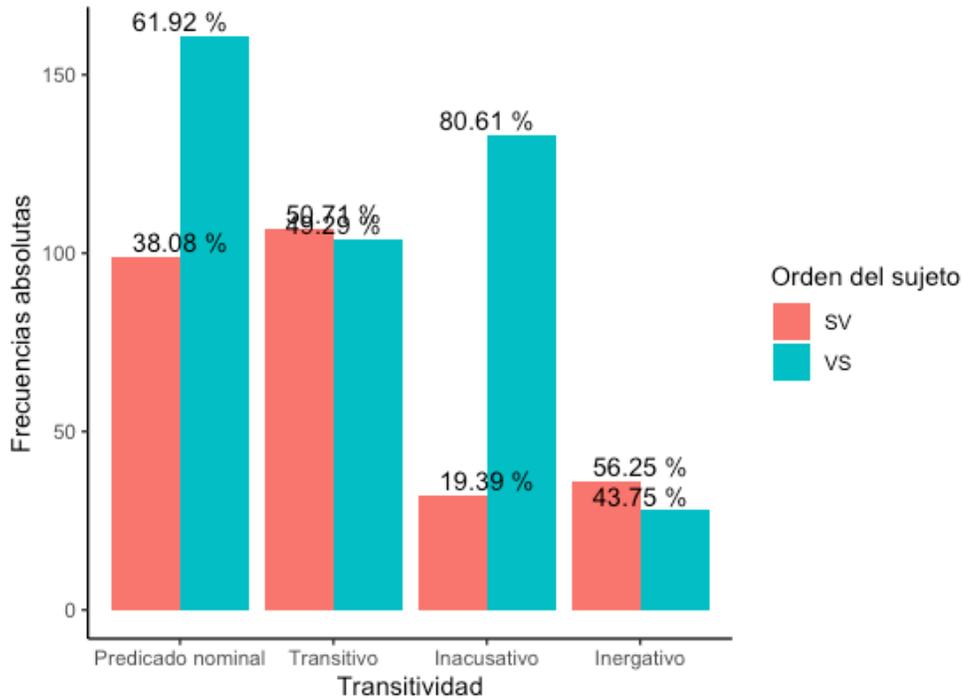


Gráfico 15. Casos con orden SV y VS, agrupados por transitividad verbal. Se presentan las categorías en orden decreciente según su número total de casos. Las barras expresan la proporción de casos para cada orden

Respecto de la posición del sujeto, puede verse en el Gráfico 11 que los predicados inacusativos muestran una clara tendencia a la posposición del sujeto, con un 80,6% de los casos (133), mientras que los inergativos tienen el comportamiento contrario: 56,2% de los casos tiene el sujeto antepuesto, con 36 casos. Los verbos transitivos, por su parte, no parecen preferir uno u otro orden: 49,3% de los casos tiene el sujeto pospuesto (104), frente al 50,7% con el sujeto antepuesto (107).

Frente a estos primeros resultados descriptivos puede señalarse que los sujetos de verbos transitivos, que presentan una ligera tendencia a la anteposición, se comportan de manera semejante a como lo harían en la declaración, donde el orden SV es más probable, tal como lo muestra la literatura (Silva-Corvalán, 1982; Delbecque, 1987; Rivas, 2013; Leonetti, 2017; Bogard, 2019).

| <i>Animacidad</i> | <i>Transitividad</i> |              |                   |               | <i>Total</i>  |
|-------------------|----------------------|--------------|-------------------|---------------|---------------|
|                   | Inacusativo          | Inergativo   | Predicado nominal | Transitivo    |               |
| animado           | 8<br>4.8 %           | 16<br>25 %   | 9<br>3.5 %        | 21<br>10 %    | 54<br>7.7 %   |
| humano            | 14<br>8.5 %          | 43<br>67.2 % | 47<br>18.1 %      | 121<br>57.3 % | 225<br>32.1 % |
| inanimado         | 143<br>86.7 %        | 5<br>7.8 %   | 204<br>78.5 %     | 69<br>32.7 %  | 421<br>60.1 % |
| <b>Total</b>      | 165<br>100 %         | 64<br>100 %  | 260<br>100 %      | 211<br>100 %  | 700<br>100 %  |

$$\chi^2=234.265 \cdot df=6 \cdot \text{Cramer's } V=0.409 \cdot p=0.000$$

Tabla 12. Tabla cruzada de las variables transitividad y animacidad

Tal como se ha planteado en la bibliografía, la razón de este orden radica en que “los sujetos transitivos prototípicos formalizan entidades cuyas propiedades características solo se encuentran, en condiciones normales, en individuos preferentemente humanos” (Bogard, 2019, p. 15) y, como tal, son típicamente agentivos, y por eso se anteponen (De Miguel, 1989; López Meirama, 1997; Bentivoglio, 2003; Alonso Cortés, 2001). Un acercamiento a los datos nos revela que, efectivamente los verbos transitivos coocurren con sujetos humanos en el 57,3% de los casos (121 en total) de la categoría (ver Tabla 12). A su vez, el sujeto aparece antepuesto en 57% de los casos (69 casos), como puede verse en la Tabla 13.

|               | <i>Transitivo “SV”</i> | <i>Transitivo “VS”</i> | Total |
|---------------|------------------------|------------------------|-------|
| Humano (frec) | 69                     | 52                     | 121   |
| Humano (prop) | 57%                    | 43%                    | 100%  |

Tabla 13. Frecuencias absolutas y relativas de sujetos humanos en predicados transitivos, agrupados por orden del sujeto (SV y VS)

Por otra parte, como se constata en la Tabla 14, y según será comentado en el apartado relativo a la variable animacidad, es la categoría humano la única donde prevalece el orden SV: en ella los casos de anteposición constituyen el 52,4%, frente a un 37% en la categoría de sujetos animados y un 32,3% en los sujetos inanimados.

| <i>Orden del sujeto</i> | <i>Animacidad</i> |               |               | <i>Total</i>  |
|-------------------------|-------------------|---------------|---------------|---------------|
|                         | animado           | humano        | inanimado     |               |
| SV                      | 20<br>37 %        | 118<br>52.4 % | 136<br>32.3 % | 274<br>39.1 % |
| VS                      | 34<br>63 %        | 107<br>47.6 % | 285<br>67.7 % | 426<br>60.9 % |
| <b>Total</b>            | 54<br>100 %       | 225<br>100 %  | 421<br>100 %  | 700<br>100 %  |

$$\chi^2=25.078 \cdot df=2 \cdot \text{Cramer's } V=0.189 \cdot p=0.000$$

Tabla 14. Tabla cruzada de las variables animacidad y orden del sujeto en las interrogativas

Por último, un análisis de residuos corregidos tipificados (ver Tabla 15) comprueba que las estructuras transitivas y los sujetos humanos tienen una relación fuerte, dado el valor positivo elevado (6,46) de este cruce, de la misma manera que este tipo de sujeto repele fuertemente las construcciones inacusativas.

| <i>Animacidad del sujeto</i> | <i>Transitividad verbal</i> |            |                   |            |
|------------------------------|-----------------------------|------------|-------------------|------------|
|                              | Inacusativo                 | Inergativo | Predicado nominal | Transitivo |
| Animado                      | -1,33                       | 4,98       | -2,47             | 1,17       |
| Humano                       | -5,36                       | 4,95       | -4,00             | 6,46       |
| Inanimado                    | 4,39                        | -5,40      | 3,81              | -5,14      |

Tabla 15. Residuos corregidos tipificados de la prueba de independencia (chi-cuadrado) entre las variables animacidad del sujeto y transitividad verbal

Respecto de los sujetos de verbos inacusativos e inergativos, estos tienen un comportamiento similar al de la declaración, es decir, el primer tipo de predicado tiende a presentar el sujeto pospuesto, mientras el segundo, antepuesto (Perlmutter, 1978; Beletti, 1987; Fernández Soriano, 1993). Nuestros datos no respaldan la hipótesis de la ausencia de un predominio numérico en términos de orden del sujeto, propuesta por Sánchez (2008): efectivamente, aun cuando no estamos analizando oraciones declarativas, nuestros datos responden a las mismas tendencias –más atenuadas para el orden SV o acentuadas para el VS, por tratarse de enunciados interrogativos– de la declaración.

En resumen, podría afirmarse de manera preliminar que el orden predominante de este tipo de enunciados, al menos en nuestros datos, no se debe tanto a la fuerza ilocutiva como al tipo de verbo. Sin embargo, la fuerza ilocutiva podría estar actuando en dos sentidos:

primero, la proporción de casos de anteposición del sujeto en enunciados con verbos transitivos es inferior a la que se esperaría en la declaración, y lo mismo sucede con los sujetos de verbos inergativos; segundo, la proporción de sujetos pospuestos en construcciones con verbos inacusativos es mucho mayor en nuestro corpus, si se comparan nuestros hallazgos con los de Silva-Corvalán (1982), Alfonso y Melis (2010) y Ortiz-López (2009), autores que niegan una preferencia por el VS en este tipo de verbos.

Respecto de la hipótesis planteada por Silva-Corvalán (1982), Alfonso y Melis (2010) y Melis y Chantal (2013), relativa a la preponderancia del orden VS de manera más acusada cuando existe un circunstante en la primera posición de la oración, nuestros datos no permiten corroborar esta tendencia, probablemente por la escasez de casos que cumplan este requisito. Como puede observarse en la Tabla 16, si bien existe una alta proporción de casos con orden VS en las oraciones cuyos circunstantes están emplazados antes del verbo (80%), también la posposición es mayoritaria en las estructuras sin circunstantes (60,9%) o con circunstantes intercalados (59,4%) o pospuestos (61,8%).

| <i>Orden del sujeto</i> | <i>Presencia de CC</i> |                |               |              | <i>Total</i>  |
|-------------------------|------------------------|----------------|---------------|--------------|---------------|
|                         | CC antep.              | CC intercalado | CC posp.      | sin CC       |               |
| SV                      | 3<br>20 %              | 13<br>40.6 %   | 65<br>38.2 %  | 193<br>40 %  | 274<br>39.1 % |
| VS                      | 12<br>80 %             | 19<br>59.4 %   | 105<br>61.8 % | 290<br>60 %  | 426<br>60.9 % |
| <b>Total</b>            | 15<br>100 %            | 32<br>100 %    | 170<br>100 %  | 483<br>100 % | 700<br>100 %  |

$$\chi^2=2.531 \cdot df=3 \cdot \text{Cramer's } V=0.060 \cdot p=0.470$$

Tabla 16. Tabla cruzada de las variables presencia de complemento circunstancial y orden del sujeto

La prueba de independencia rechazó la hipótesis alternativa, lo cual muestra que no existe una asociación entre la presencia de circunstante y el orden ( $p=0.470$ ). Sin embargo, la diferencia, en promedio, de 20 puntos porcentuales entre la categoría circunstante antepuesto y el resto podría estar afirmando en alguna medida esta tendencia expuesta por las autoras. Debe ser aclarado, además, que de los circunstantes que aparecieron en posición

inicial de frase (15), 6 coocurrieron con verbos existenciales, y en todos los casos el orden fue VS, por lo que parecería que es la semántica verbal la que favorece en primera instancia el orden. Por último, un solo caso de complemento circunstancial fue de tipo locativo. Por tanto, ante la falta de datos para realizar un contraste, se espera poder contar en el futuro con investigaciones más enfocadas en la influencia de esta variable en el orden de enunciados interrogativos totales.

El Gráfico 16 ilustra la relación entre las variables orden del sujeto y la transitividad del predicado.

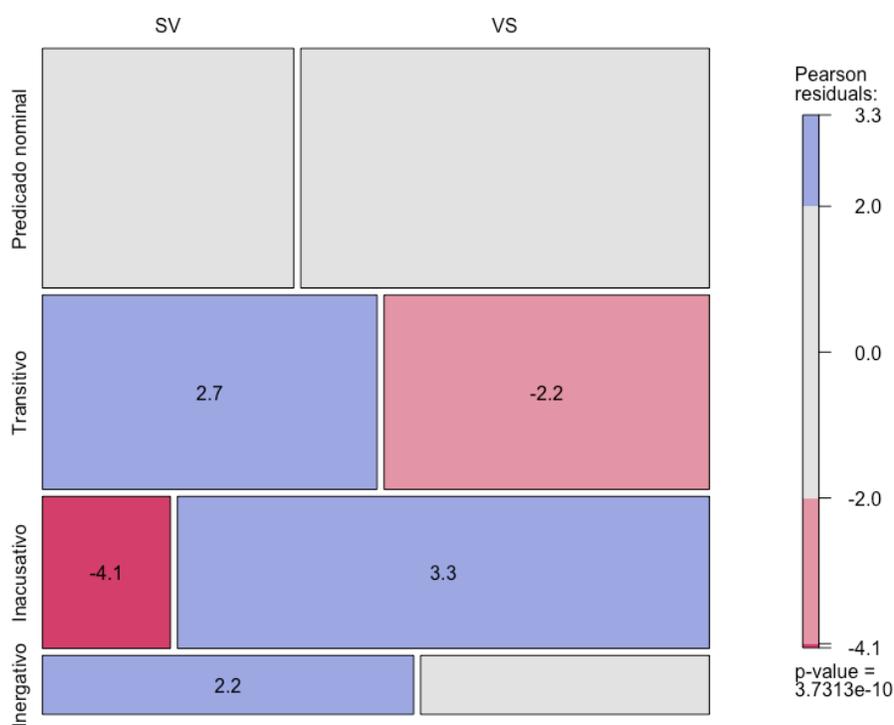


Gráfico 16. Representación de mosaicos de las variables Orden del sujeto y Transitividad verbal

Como puede comprobarse, los predicados inacusativos propician de manera fuerte el orden VS (residuo de 3,3), a la vez que son poderosos inhibidores del orden SV (el residuo es muy alto: -4,1); los verbos inergativos favorecen de manera significativa la anteposición del sujeto (residuos de 2,2). Por último, los verbos transitivos tienen un comportamiento contrario a los verbos inacusativos: como indican los colores y los residuos presentes en los mosaicos, los predicados transitivos determinan fuertemente el orden SV (residuos de 2,7),

mientras inhiben significativamente la posposición del sujeto (-2,2); ello a pesar de que las frecuencias relativas son bastante similares para uno y otro orden.

### 3.1.4. Variables referidas al sujeto

#### 3.1.4.1. El orden del sujeto respecto del verbo según referencialidad del sujeto

Respecto de la referencialidad del sujeto, las dos variables incluidas en el estudio se describen a continuación.

##### 3.1.4.1.1. Animacidad

El Gráfico 17 muestra las frecuencias (recuentos) totales y relativas de la variable animacidad. Como puede observarse, los casos con sujetos inanimados representan un 60,1% de la muestra (con 421 casos). Les siguen los sujetos humanos, con una proporción del 32,1% (con 225 casos) y por último los sujetos animados, que representan el 7,7% (con un total de 54 casos).

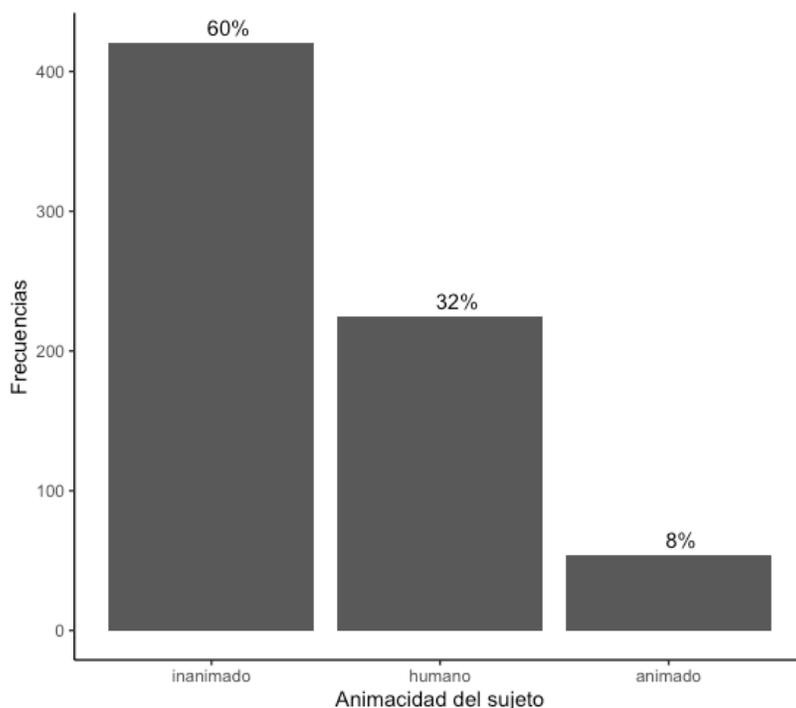


Gráfico 17. Proporción de casos por categoría dentro de la variable Animacidad del sujeto

El Gráfico 18 ilustra la proporción de casos con orden SV y VS por categoría dentro de la variable animacidad. Como puede comprobarse, el sujeto con referentes inanimados se pospone de manera mayoritaria (67,7%), al igual que los sujetos animados (63%). En los sujetos humanos, sin embargo, el orden SV es ligeramente prevaleciente: 52,4% de los casos. Podría decirse, de manera preliminar, que mientras más asciende el sujeto en la escala de animacidad, es decir, mientras más es prototípicamente animado (rasgo humano), más propende a aparecer antepuesto al verbo. Estos hallazgos comprueban una tendencia planteada por algunos autores para la declaración sobre la anteposición de los sujetos con referentes humanos, dada su agentividad (De Miguel, 1989; López Meirama, 1997; Bentivoglio, 2003; Alonso Cortés, 2001) y la posposición de los sujetos inanimados, dada su pasividad (Hatcher, 1956). Asimismo, respaldan los hallazgos previos en la interrogativa parcial, donde se ha demostrado que la animacidad aumenta las probabilidades de que el sujeto aparezca antepuesto al verbo (Rosemeyer, 2016).

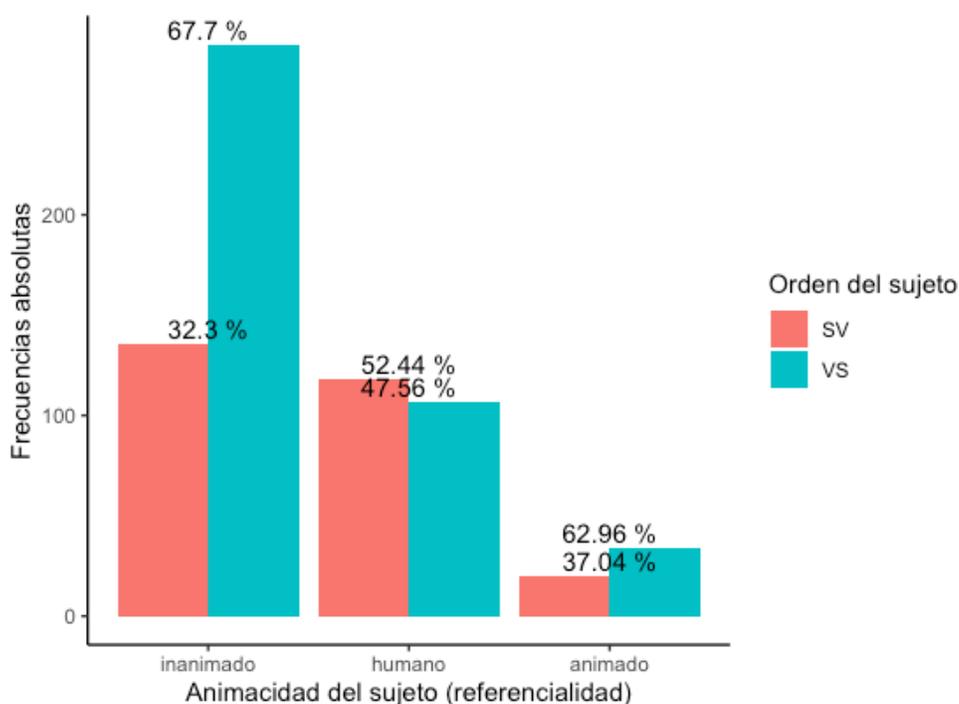


Gráfico 18. Casos con orden SV y VS, agrupados por animacidad. El eje Y muestra las frecuencias absolutas. Se presentan las categorías en orden decreciente según su número total de casos. Las barras expresan la proporción de casos para cada orden

Como ha sido destacado en el acápite relativo a la variable transitividad (ver Tabla 13), los sujetos humanos son prevalecientes en los predicados transitivos (57,3%). Además, salta

a la vista que la diferencia principal entre los verbos inacusativos e inergativos es que los primeros coocurren mayormente con sujetos inanimados (86,7%), mientras que los segundos ostentan principalmente sujetos humanos (67,2%) y animados (25%), y los inanimados con solo el 8%. Los predicados nominales, por su parte, se comportan como los predicados inacusativos: los sujetos en su gran mayoría son inanimados (78,5%).

Estos resultados podrían deberse, como afirma Bogard (2019), a que típicamente el sujeto de los verbos transitivos “formaliza el argumento instigador, causa o agente, el cognoscente, el receptor, el perceptor, el experimentante, el poseedor, etc.”, y se codifica normalmente por sustantivos humanos, mientras que “los sujetos intransitivos que derivan de los verbos inacusativos, así como de los que provienen de predicados verbales pasivos, anticausativos, estativos y resultativos, que formalizan principalmente el tema, el paciente, el estímulo” (p. 33), prototípicamente son codificados con sustantivos inanimados y, por lo tanto, tenderían a aparecer pospuestos al verbo.

Lo anterior puede comprobarse consultando la Tabla 15 del acápite antes mencionado. El análisis de residuos corregidos tipificados (medida que permite ver los cruces que son significativos en la asociación entre dos variables) muestra que los sujetos humanos se correlacionan fuertemente con los predicados transitivos e inergativos (residuos positivos de 6,46 y 4,95, respectivamente) y repelen los predicados nominales e inacusativos (residuos negativos de -4 y -5,36, respectivamente), mientras que los sujetos inanimados tienen un comportamiento contrario: tienden a aparecer en predicados inacusativos (residuo positivo de 4,39) y nominales (residuo positivo de 3,81), y repelen los predicados inergativos (residuo negativo alto: -5,40) y transitivos (residuo negativo alto: -5,14). Los sujetos animados (no humanos) tienen una asociación fuerte con predicados inergativos (4,98) y muy débil con los nominales (-2,47).

Respecto de la asociación entre las variables, obsérvese en el Gráfico 19 que la categoría de humano es un factor influyente en la anteposición del sujeto (los residuos son de 3.2) e inhibe el orden VS (residuo de -2.6).

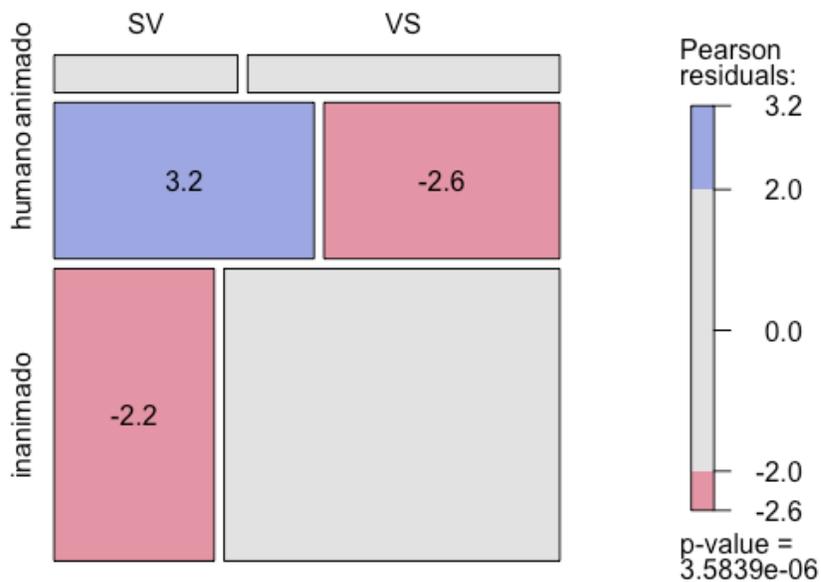


Gráfico 19. Representación de mosaicos de las variables Orden del sujeto y Animación del sujeto

Por el contrario, la categoría sujetos inanimados inhibe de manera significativa el orden SV (residuos de -2.2). Dicho de otra manera, y según los datos con los que se ha trabajado en este corpus, los sujetos con referentes humanos facilitan la aparición del orden SV, mientras los referentes inanimados lo inhiben. No hay datos suficientes, sin embargo, para afirmar que estos últimos sean un factor importante en el orden VS. Para la categoría animado, las diferencias observadas no son estadísticamente significativas.

Con el objetivo de evitar el desbalance de los datos, y partiendo de la suposición de que hay un comportamiento diferenciado entre los sujetos con referentes animados respecto de los que tienen referentes inanimados, se decidió reordenar la variable, dejando solo dos niveles, los antes mencionados. El gráfico de mosaico obtenido para la nueva variable en su relación con el orden del sujeto (ver Gráfico 20) muestra que es razonable semejante organización dicotómica, pues los residuos no sufren grandes variaciones. Se mantiene así nuestro hallazgo inicial: los referentes animados potencian la anteposición del sujeto (residuos positivos), mientras los inanimados la repelen (residuos negativos).

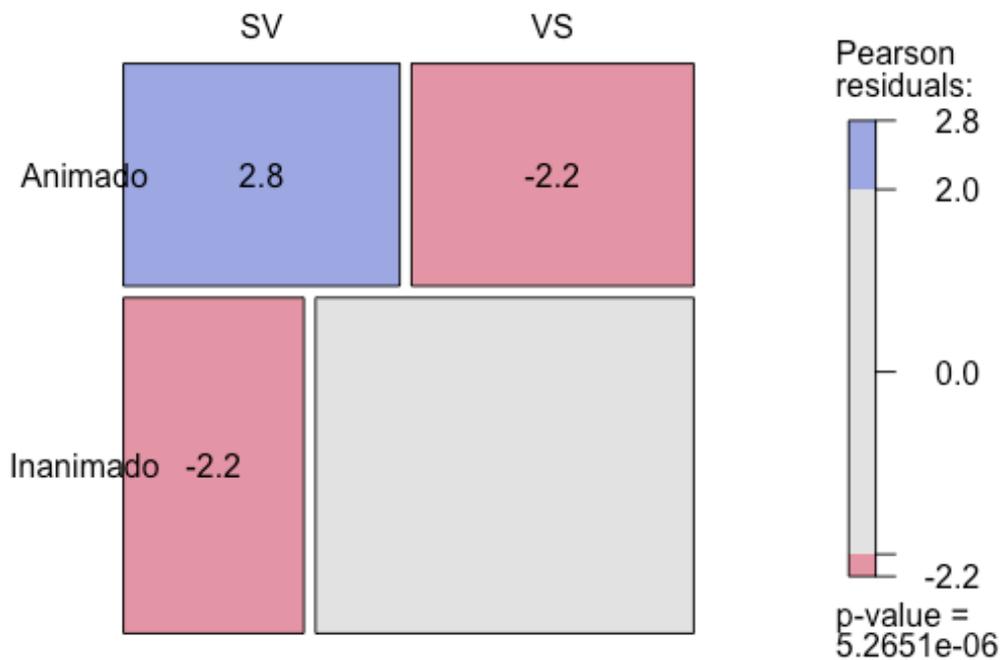


Gráfico 20. Representación de mosaicos de las variables Orden del sujeto y Animación del sujeto (con solo dos niveles: animado e inanimado)

### 3.1.4.1.2. Definitud

Como puede observarse en el Gráfico 21, existe un desbalance en los datos obtenidos para cada una de las categorías de la variable: los sujetos definidos son el 79,1% del total, con 554 casos; los escuetos por su parte, constituyen un 12,6% de la muestra (88 casos), mientras los indefinidos representan solo el 8,3% (58 casos). La distribución de los datos en el corpus responde a una tendencia normal en español: investigaciones previas en oraciones declarativas demuestran que existe una baja prevalencia de sustantivos no definidos o no identificables (Delbecque, 1988; Rivas, 2008 y 2013).

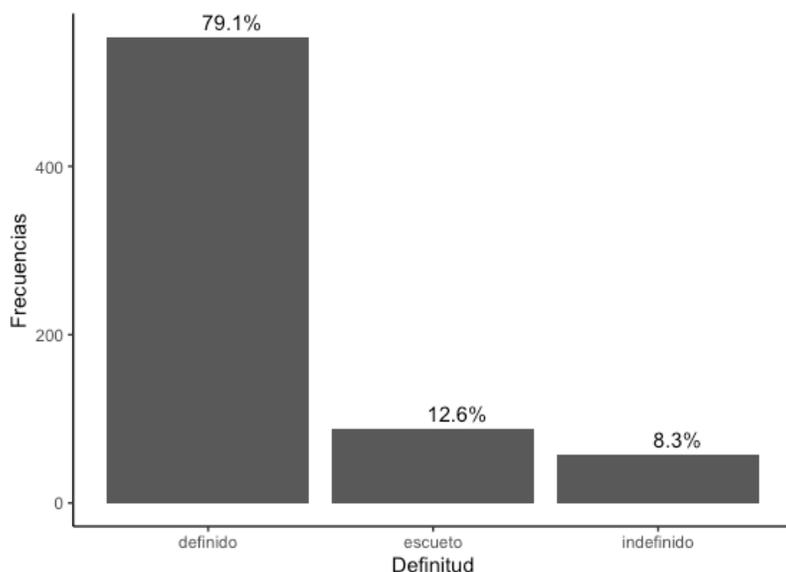


Gráfico 21. Proporción de casos por categoría dentro de la variable Animacidad del sujeto

El Gráfico 22 muestra las proporciones de casos por orden en la variable definitud. Como puede verse, son los sujetos escuetos los que más se posponen (89% de los casos), seguidos de los indefinidos (69%), mientras los definidos tienen una mayor proporción de anteposición al verbo (56% de posposición). Se observa una tendencia clara a la anteposición del sujeto en la medida en que sus referentes pasan de ser escuetos a definidos; dicho de otro modo, mientras más definido es el sujeto, más se antepone. Se comprueba así la existencia de una jerarquía también para la interrogación total (véase Figura 2). En la declaración ya se había demostrado que los sujetos definidos o identificables tienen más probabilidades de aparecer antepuestos que los indefinidos, tal como muestran Rivas (2008 y 2013) y Ocampo (1995). Además, Rivas (2008) tuvo una alta proporción de sujetos escuetos en posición posverbal, a tono con la tendencia general del español discutida en el acápite teórico de esta investigación.

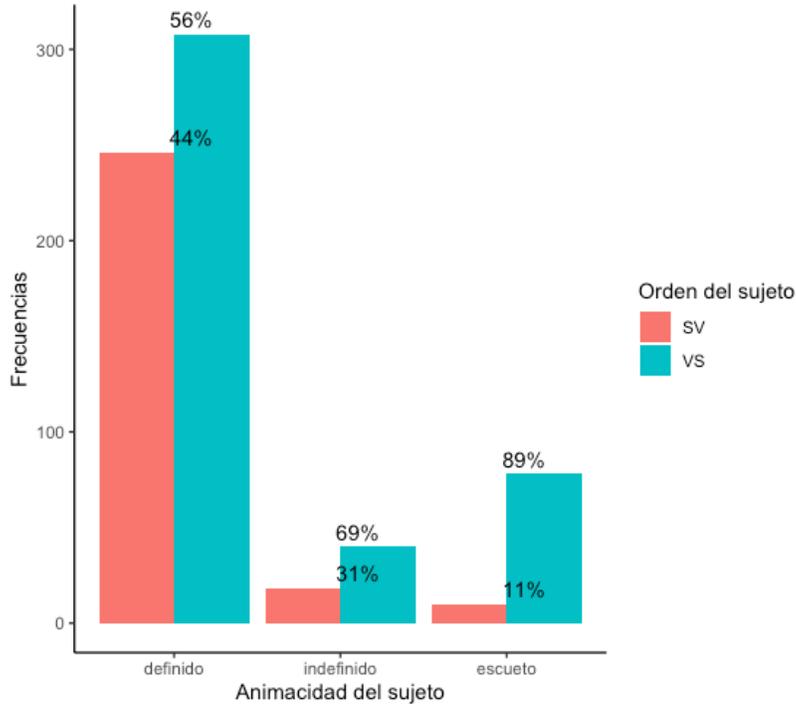


Gráfico 22. Casos con orden SV y VS, agrupados por definitud. El eje Y muestra las frecuencias absolutas. Se presentan las categorías en orden decreciente según su número total de casos. Las barras expresan la proporción de casos para cada orden

Los resultados de la asociación entre orden del sujeto y definitud son mostrados en el Gráfico 23.

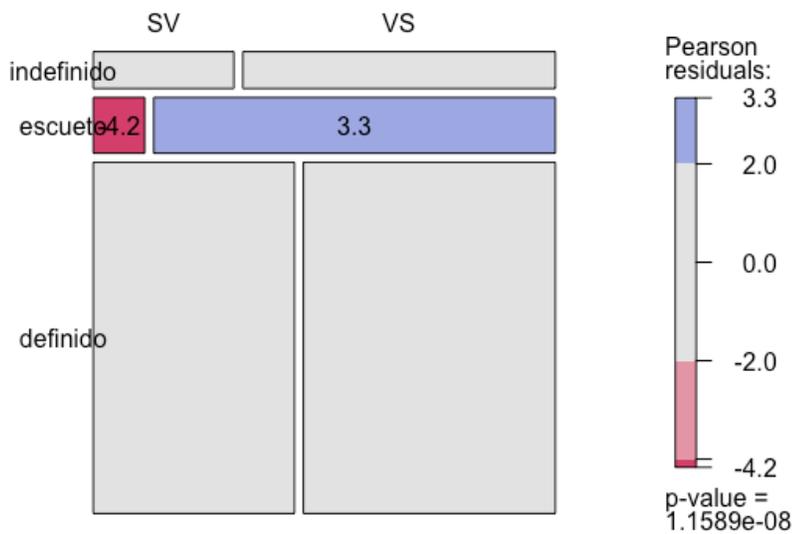


Gráfico 23. Representación de mosaicos de las variables Orden del sujeto y Definitud

Salta a la vista que la mayor cantidad de casos contenidos en el corpus corresponde a sujetos definidos, dado el tamaño de los mosaicos correspondientes a esta categoría. Sin embargo, este nivel de la variable no influye significativamente en el orden, como tampoco lo hace la categoría *indefinido*. La categoría de *escueto*<sup>59</sup>, no obstante, tiene un residuo altamente significativo para el orden sujeto-verbo (-4,2), pero dado que su valor es negativo, podemos interpretarlo como un desfavorecedor poderoso de este orden. Por otra parte, los sujetos escuetos favorecen de manera considerable la posposición del sujeto.

Ante estos resultados, y buscando una interpretación un poco más general de los datos, se reestructuró la variable, eliminando un nivel. Se entiende que podría haber un comportamiento opuesto entre los sujetos definidos y los indefinidos, y que esta distinción podría arrojar luz sobre cómo se comporta el fenómeno del orden adoptado por el sujeto. Así, se fusionaron los niveles indefinido y escueto, y se mantuvo el nivel definido.

En efecto, y como era de esperar, el Gráfico 24 muestra que los sujetos con referente indefinido repelen de manera significativa la anteposición del sujeto. Los sujetos con referente definido siguen sin tener una relación fuerte con ninguno de los dos órdenes.

---

<sup>59</sup> Los sujetos escuetos del corpus constituyen 88 casos en total. Dentro de ellos, hay 15 casos de sustantivos escuetos que debieron haber sido removidos, dada su escasa probabilidad (aunque no nula, según se puede consultar en las gramáticas normativas del español) de aparecer en posición preverbal. Esto los invalida, por definición, para un análisis de orden alternante. Fueron conservados hasta una etapa avanzada de la investigación y se decidió mantenerlos en el análisis estadístico para no perjudicar los resultados ya obtenidos reelaborando un nuevo corpus, lo que habría traído como consecuencia un innecesario retraso del cronograma elaborado. No obstante, debe ser aclarado que estos sustantivos constituyen el 17% de la categoría y el 2,14% del corpus, por lo que consideramos que su presencia no tiene ninguna consecuencia estadística importante. Obviamente, estos casos están también presentes, implícitamente, en variables como Morfología del sujeto y Complejidad del sujeto, pero no revisten mayor importancia desde el punto de vista cuantitativo por las razones antes expuestas.

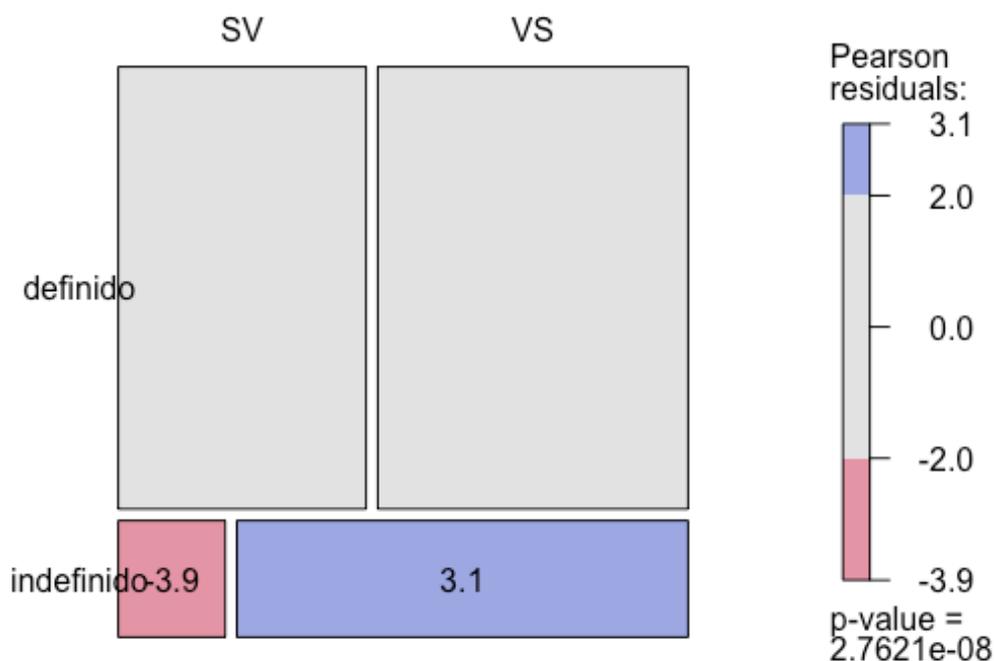


Gráfico 24. Representación de mosaicos de las variables Orden del sujeto y Definitud (con dos niveles: definido e indefinido)

Respecto de la hipótesis planteada por Rivas (2008) sobre la mayor capacidad de la definitud sobre la animacidad para predecir el orden, nuestros datos exhiben una tendencia dispar. La distribución de los casos en nuestro corpus muestra que los parámetros animado y definido no presentan un comportamiento sensiblemente diferente, tanto si se trata de la anteposición como de la posposición del sujeto. Así, el 49% de las interrogativas con sujeto preverbal posee un referente animado y un 44% es definido; por otro lado, las interrogativas con sujeto posverbal tienen un 51% de referentes animados y un 56% de definidos. Para los parámetros de inanimidad e indefinitud, sin embargo, las diferencias son más marcadas: mientras el 19% de los sujetos preverbiales tienen referentes indefinidos, el 32% es inanimado; a la vez que el 81% de los sujetos pospuestos es indefinido y el 68% es inanimado.

Los residuos corregidos tipificados del test chi-cuadrado de la asociación entre las dos variables<sup>60</sup> (los Gráficos 20 y 24 ilustran la relación) y el orden del sujeto aparecen plasmados en la Tabla 17. Respecto del rol que desempeña la referencialidad del sujeto en la

<sup>60</sup> La distribución original tripartita de los casos se ha sustituido por otra dicotómica. Nos interesa saber si existe una variación importante en los resultados ordenando los sujetos según sus referentes sean definidos/indefinidos o animados/inanimados.

determinación de la secuencia de constituyentes, la variable animacidad mantiene una relación fuerte, tanto en la categoría de animado como en la de inanimado, con el orden: los sujetos animados favorecen la anteposición y los inanimados la inhiben; mientras que la variable definitud tiene una asociación importante con el orden solo cuando los sujetos son indefinidos, potenciando la posposición de manera fuerte.

| <i>Referencialidad del sujeto</i> | <i>Orden del sujeto</i> |           |
|-----------------------------------|-------------------------|-----------|
|                                   | <b>SV</b>               | <b>VS</b> |
| <i>definido</i>                   | 1,98                    | -1,59     |
| <i>indefinido</i>                 | -3,86                   | 3,09      |
|                                   | <b>SV</b>               | <b>VS</b> |
| <i>animado</i>                    | 2,76                    | -2,21     |
| <i>inanimado</i>                  | -2,24                   | 1,80      |

*Tabla 17. Residuos corregidos tipificados para la asociación de las variables referidas a la referencialidad del sujeto (animacidad y definitud) y orden de palabras. Los residuos por encima de +2 o por debajo de -2 muestran cruces estadísticamente significativos*

El Gráfico 25, a continuación, refleja estas tendencias. El eje horizontal muestra las agrupaciones posibles entre las categorías de las variables (recuérdese que se han reagrupado, conformándose dos grupos dentro de cada una de ellas) y el eje vertical representa las proporciones de esas agrupaciones en el corpus. Como puede observarse, el orden SV sigue una tendencia clara y esperada para el cruce de estos parámetros, es decir, los sujetos definidos, tanto animados como inanimados, se anteponen más, mientras los indefinidos lo hacen en una menor proporción; el rasgo de definitud, en consecuencia, prevalecería en la determinación de este orden. Para la secuencia VS es la animacidad la que parece tener un rol más central: los sujetos definidos e inanimados son más frecuentes que para el orden SV, mientras que los sujetos indefinidos e inanimados se posponen en un gran número de casos.

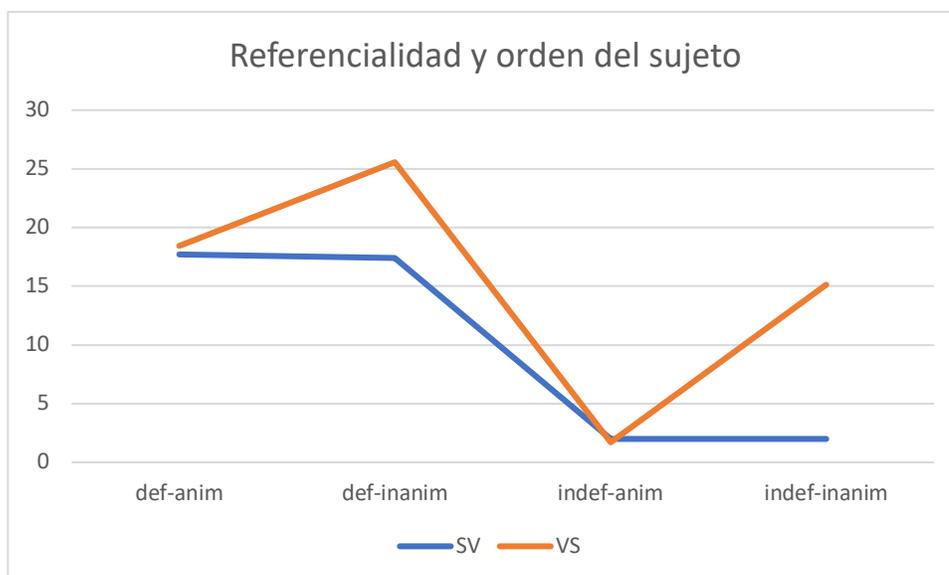


Gráfico 25. Tendencias de las combinaciones entre los niveles de las variables definitud, animacidad y el orden del sujeto. Los números del eje vertical expresan las proporciones de casos en cada una de las agrupaciones representadas en el eje horizontal

### 3.1.4.2. El orden del sujeto respecto del verbo según morfología del sujeto

El Gráfico 26 recoge las frecuencias relativas de la variable morfología del sujeto. Primeramente, debe ser mencionado que, de las 15 categorías incluidas en el estudio, hubo una que no contó con ningún caso: el pronombre personal de la tercera persona del plural (*ellos*). Además, existen otras categorías con frecuencias relativas menores al 2% (menos de 10 casos). Ello ocurre con los pronombres personales de la primera y segunda personas del plural (*nosotros* y *ustedes*), con 2 y 8 casos, respectivamente; los pronombres personales de la primera y tercera persona del singular (*yo* y *él/ella*), con 2 y 1 caso, respectivamente; los *numerales*, con 8 casos; y los *sujetos clausulares*, con 7.

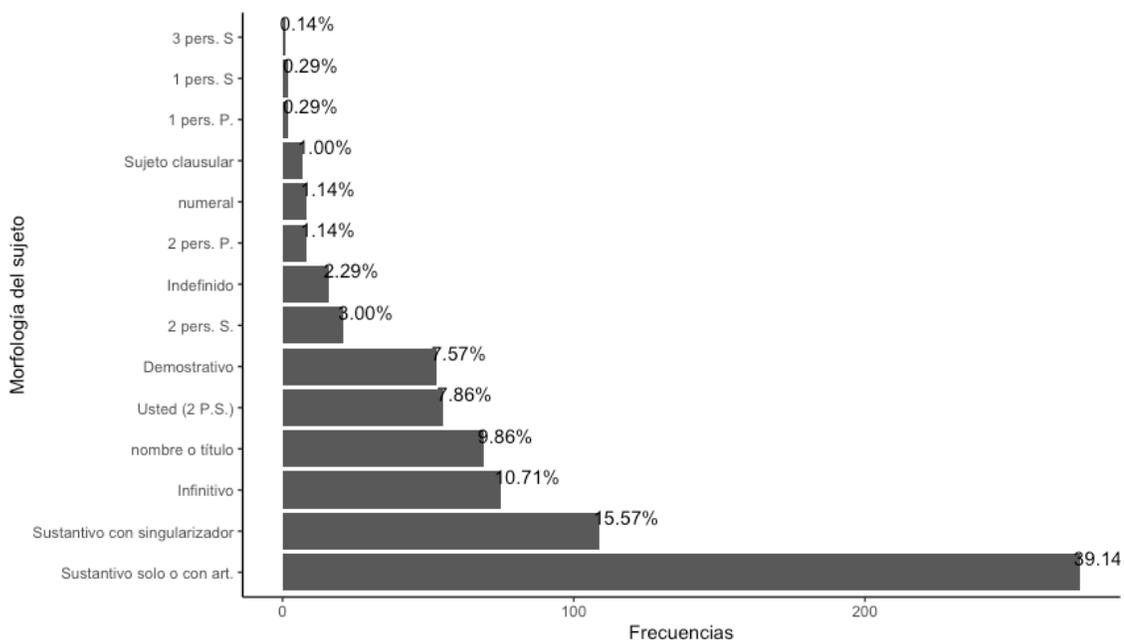


Gráfico 26. Proporción de casos por categoría dentro de la variable Morfología del sujeto

Respecto de la distribución de las proporciones, los sujetos compuestos por sustantivos solos o con artículo constituyen el 39,1% del corpus, con 274 casos. Los sustantivos con singularizador (modificador indefinido, numeral, etc.)<sup>61</sup> componen el 15,6% de la muestra, con 109 casos; los sujetos compuestos por un infinitivo representan el 10,7%, con 75 casos; los sujetos compuestos por nombre o título constituyen el 9,9%, con 69 casos; los sujetos con *usted* representan el 7,9%, con 55 casos; los integrados por demostrativos son el 7,6% del corpus (53 casos); los sujetos con *tú* constituyen el 3% de los datos (21 casos) y los indefinidos tienen una proporción del 2,3% (16 casos). La diferencia encontrada entre las frecuencias de los sujetos constituidos por frases nominales cuyo núcleo es un sustantivo común y el resto de las construcciones morfológicas es reflejo de una tendencia existente en la lengua. Según habíamos comentado en el apartado teórico, hay evidencia de que la segunda persona del singular puede aparecer de manera explícita con una frecuencia más alta que el resto de los pronombres personales (Morales, 1997). Asimismo, como comenta Bogard

<sup>61</sup> Como fue declarado en el acápite teórico, se estableció una separación entre sustantivos escuetos o con artículos y sustantivos con otro tipo de singularizador, en el supuesto de que podría haber una diferencia en el comportamiento de ambos tipos de construcciones. Nuestros datos, en efecto, muestran que el segundo tipo de sujeto es mucho menos frecuente en el habla, aunque desde el punto de vista del orden adoptado por el sujeto las diferencias no parecen respaldar la hipótesis de un comportamiento diferente de ambas estructuras.

(2019), los sustantivos comunes ostentan una gran diversidad léxica, lo cual “expande las posibilidades temáticas a las que puede recurrir el usuario de la lengua en el flujo comunicativo en el que interviene” (p.10). No es extraño, entonces, que sean muy abundantes en el uso.

En lo referente al orden de palabras, el Gráfico 27 ilustra la proporción de casos con orden SV y VS por categoría. Se han omitido en la representación todos los niveles con menos de 10 casos. El sujeto antepuesto es mayoritario en los sujetos con *tú* (76,2%) y los sujetos constituidos por indefinidos (68,8%); los sujetos con *usted* no muestran una tendencia definida (50,9% de casos con orden VS frente a 49,1% de orden SV). Los sujetos constituidos por infinitivos se posponen en el 84,9% de los casos; a esta frecuencia relativa le sigue la de los sustantivos con singularizador (64,2%), sustantivos solos o con artículos (60,9%), demostrativos (60,4%), y por último los nombres o títulos (56,5%).

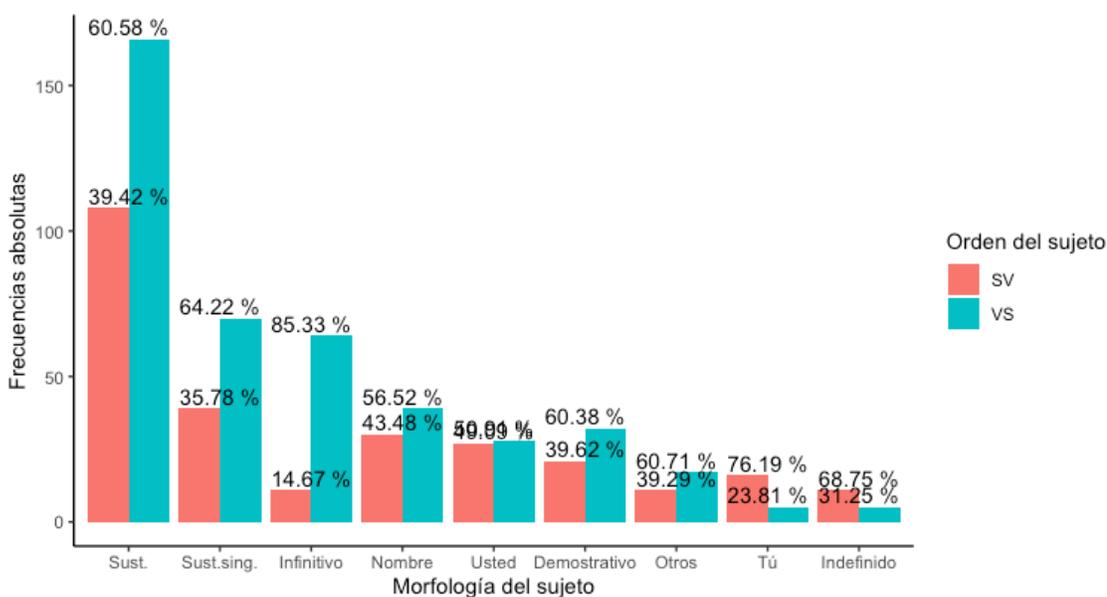


Gráfico 27. Casos con orden SV y VS, agrupados por morfología del sujeto. Se presentan las categorías en orden decreciente según su número total de casos. Las barras expresan la proporción de casos para cada orden. Las categorías con casos marginales (menos de 10) se han agrupado en la etiqueta “Otros”

A continuación, comentaremos el gráfico obtenido para la asociación de las variables:

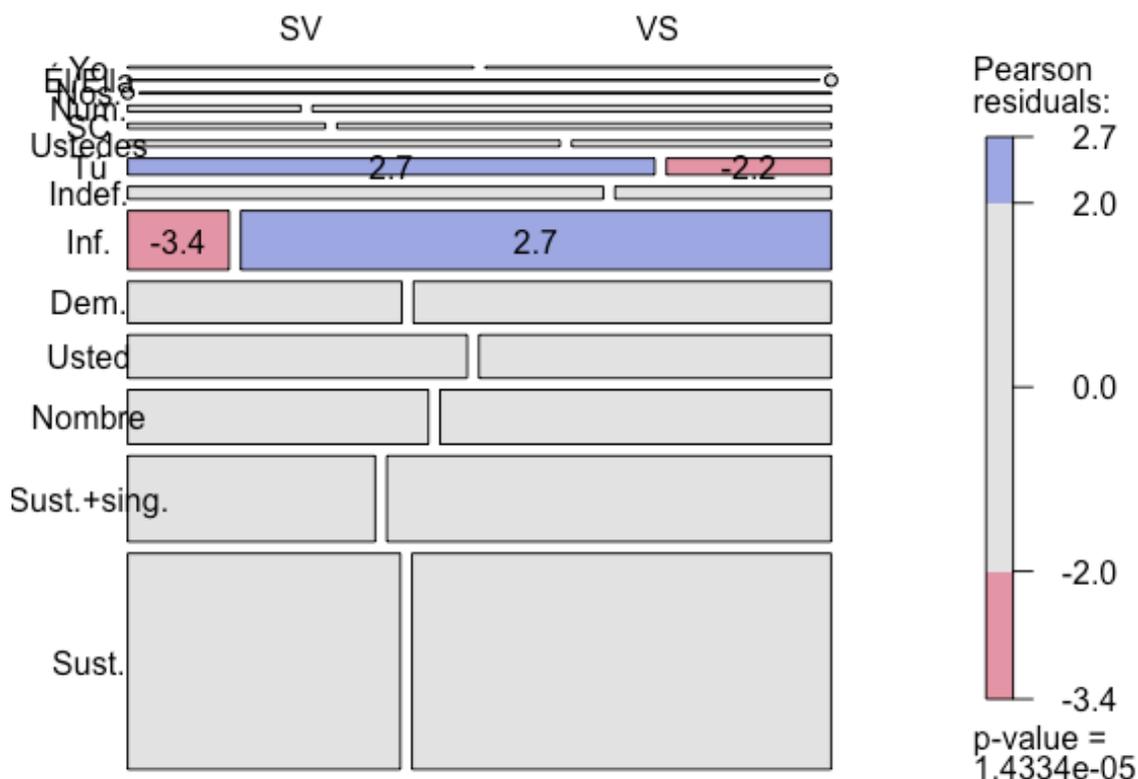


Gráfico 28. Representación de mosaicos de la relación entre las variables Orden del sujeto y Morfología del sujeto. En el eje vertical se muestran (de abajo hacia arriba) los mosaicos correspondientes a: sustantivo, solo o con artículo (Sust.), sustantivo con singularizador (Sust.+Sing.), nombre propio o título (Nombre), Usted, pronombre posesivo o demostrativo (Dem.), infinitivo (Inf.), indefinido (Indef.), Tú, Ustedes, sujeto clausular (SC), pronombre numeral (Num.), nosotros (Nos.), Él o Ella, y Yo.

A tono con lo comentado anteriormente respecto de la escasa presencia en el corpus de algunas categorías, el gráfico muestra valores marginales (marcados con un punto en el extremo izquierdo o derecho del mosaico que les corresponde) para la primera persona tanto del singular como del plural, así como para la tercera persona del singular. En principio, ambas variables están relacionadas ( $p=1.43 \text{ e-}05$ ), pero de la extensa lista de categorías incluidas en el estudio, solo dos mostraron una relación fuerte respecto del orden del sujeto: ellas son la categoría de *infinitivo* –que tiene un residuo positivo alto (2,7) para la secuencia VS, lo cual indica que es un factor favorecedor de este orden, mientras que exhibe residuo negativo alto (-3,4) para SV, lo que significa que inhibe este orden–; y el pronombre de segunda persona del singular, *tú* –potenciador del orden SV (residuo de 2,7), e inhibidor de manera significativa del orden VS (residuo de -2,2).

Brown y Rivas (2011) habían apuntado en su estudio cuantitativo sobre interrogativas parciales portorriqueñas que los sujetos pospuestos estaban constituidos sobre todo por frases nominales, es decir, que los sujetos pronominales y lexicales tenían comportamientos contrarios en cuanto al orden que adoptaban respecto del verbo. Nuestros resultados apoyan esta hipótesis: como puede verse, los sujetos compuestos por los pronombres *tú, usted* e *indefinido* tienden más a la anteposición que los sujetos lexicales (sustantivos con artículo u otro tipo de determinante, infinitivo y nombre o título).

Teniendo en cuenta este referente investigativo, y para poder observar de manera más precisa cómo se comportan nuestros datos, se realizó una redistribución de la variable y se mantuvieron solo 3 niveles: sujeto léxico, infinitivo y sujeto clausular y sujeto pronominal. La Tabla 18 recoge las frecuencias totales y relativas de las tres categorías en las que se reagruparon los datos. Como puede comprobarse, los sujetos léxicos tienden a la posposición de manera mayoritaria (61% de los casos); los sujetos compuestos por infinitivo o por oraciones (SC), lo hacen de manera acusada (84% de los casos), mientras que los sujetos pronominales muestran una ligera preferencia por la anteposición (50,6%).

| <i>Orden del sujeto</i> | <i>Morfología del sujeto</i> |                       |                        | <i>Total</i>           |
|-------------------------|------------------------------|-----------------------|------------------------|------------------------|
|                         | Suj. lex.                    | Inf. y SC             | Suj. Pron.             |                        |
| SV                      | 177<br>39.2 %                | 13<br>15.9 %          | 84<br>50.6 %           | 274<br>39.1 %          |
| VS                      | 275<br>60.8 %                | 69<br>84.1 %          | 82<br>49.4 %           | 426<br>100 %<br>60.9 % |
| <b>Total</b>            | 452<br>64.6 %<br>100 %       | 82<br>11.7 %<br>100 % | 166<br>23.7 %<br>100 % | 700<br>100 %<br>100 %  |

$$\chi^2=27.822 \cdot df=2 \cdot \text{Cramer's } V=0.199 \cdot p=0.000$$

Tabla 18. Tabla cruzada de las variables *Morfología del sujeto* y *Orden del sujeto*. La primera de ellas se ha simplificado, dejando solo tres niveles: sujetos léxicos, infinitivos y sujetos clausulares, y sujetos pronominales

El análisis de residuos (Gráfico 29) arroja que, si bien la asociación entre las variables es significativa ( $p=0.000$ ), el cruce de la categoría sujeto lexical con el orden no es estadísticamente significativo. Los sujetos pronominales, por su parte, favorecen de manera

fuerte el orden SV (residuo de 2,4), mientras que los infinitivos y sujetos clausulares exhiben un alto residuo negativo para el orden SV (-3,4) y un residuo positivo por encima del valor crítico para el orden VS (2,7), por lo que se puede afirmar que son potenciadores de la posposición del sujeto. Estos resultados se condicen con las tendencias mostradas en literatura científica para la declaración (Delbecque, 1988; Bentivoglio, 2003).

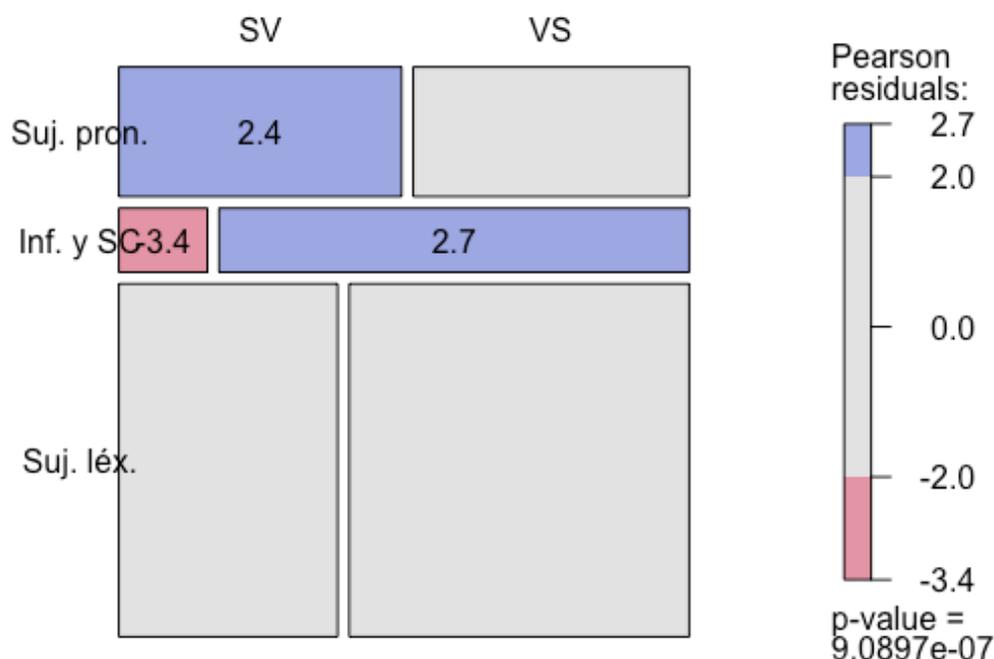


Gráfico 29. Representación de mosaicos de las variables Orden del sujeto y Morfología del sujeto (se han reducido los niveles a tres: sujeto pronominal, infinitivo y sujeto clausular y sujeto léxico)

Respecto de la categoría pronombre personal *tú*, su comportamiento es coherente con una tendencia normal en la declarativa española: al ser una lengua de sujeto nulo, la expresión del pronombre suele ser focal y, en consecuencia, ocupar la primera posición de la oración<sup>62</sup>.

Una mirada un poco más específica a los sujetos pronominales en las zonas Antillas y Caribe Continental puede encontrarse en la Tabla 19. Las observaciones aquí recogidas son coherentes con algunos de los resultados de investigaciones previas (Lipski, 1977; Lantolf, 1980; Morales, 1997). Básicamente, se ha señalado que en el Caribe las dos primeras

<sup>62</sup> Algo similar ocurre con el pronombre de segunda persona del plural (ustedes). También en nuestro corpus los sujetos constituidos por este pronombre están en su mayoría antepuestos al verbo (62,5%). Sin embargo, la baja cantidad de casos presentes en el corpus impide que se pueda hacer una generalización.

personas del discurso son las que aparecen de manera más frecuente cuando el sujeto tiene expresión fónica y es un pronombre, y que tienden a aparecer antepuestas al verbo en oraciones declarativas.

Nuestros datos también corroboran este hallazgo en la interrogación, aunque solo para los pronombres *tú/usted*, y no para las primeras personas del singular y el plural (*yo* y *nosotros*) y para la segunda del plural (*ustedes*), dada la escasez de datos pertenecientes a esta categoría en el corpus. En efecto, las Antillas es la zona que más casos presenta con estos sujetos *tú* y *usted*: la zona cuenta con 8 casos de *tú* (38% de los casos totales de este pronombre en el corpus), y de ellos 7 (el 87,5%) presentan el orden SV; respecto del pronombre *usted*, hay 12 en esta zona (22% del total para este pronombre), aunque no existe una preferencia por un orden, pues los casos están distribuidos en igual número para la anteposición y la posposición del sujeto. La zona Caribe Continental tiene 9 casos de sujeto con pronombre *usted* (16% del total), y si bien no es la única con este número, sí se destaca por tener una alta proporción de anteposición (89%, es decir, solo un caso tiene el orden VS).

| <b>Zona geográfica</b> |      |          |               |       |        |                   |           |
|------------------------|------|----------|---------------|-------|--------|-------------------|-----------|
| <b>Orden SV</b>        |      |          |               |       |        |                   |           |
| <i>Sujeto pron.</i>    | And. | Antillas | Car.<br>Cont. | Chile | España | Méx.<br>Centroam. | Río Plata |
| <i>Yo</i>              | 0    | 0        | 0             | 0     | 0      | 0                 | 1         |
| <i>Tú</i>              | 0    | 7        | 2             | 2     | 1      | 1                 | 3         |
| <i>Él/Ella</i>         | 0    | 0        | 0             | 0     | 1      | 0                 | 0         |
| <i>Nosotros</i>        | 0    | 0        | 0             | 0     | 0      | 0                 | 0         |
| <i>Ustedes</i>         | 0    | 0        | 1             | 3     | 0      | 0                 | 1         |
| <i>Usted</i>           | 3    | 6        | 8             | 3     | 1      | 2                 | 4         |

| <b>Zona geográfica</b> |      |          |               |       |        |                   |           |
|------------------------|------|----------|---------------|-------|--------|-------------------|-----------|
| <b>Orden VS</b>        |      |          |               |       |        |                   |           |
| <i>Sujeto pron.</i>    | And. | Antillas | Car.<br>Cont. | Chile | España | Mex.<br>Centroam. | Río Plata |
| <i>Yo</i>              | 0    | 0        | 1             | 0     | 0      | 0                 | 0         |
| <i>Tú</i>              | 1    | 1        | 0             | 1     | 1      | 1                 | 0         |
| <i>Él/Ella</i>         | 0    | 0        | 0             | 0     | 0      | 0                 | 0         |
| <i>Nosotros</i>        | 0    | 0        | 0             | 0     | 0      | 1                 | 1         |
| <i>Ustedes</i>         | 0    | 0        | 0             | 0     | 1      | 1                 | 1         |
| <i>Usted</i>           | 6    | 6        | 1             | 3     | 5      | 3                 | 4         |

Tabla 19. Frecuencias absolutas de pronombres personales en las interrogativas del corpus, según zona geográfica

El Gráfico 30 muestra de manera más clara la tendencia antes declarada. Si bien el número total de casos no es alto, se puede ver que hay un comportamiento diferente, tanto respecto del tipo de pronombre como de su posición en la cláusula, en el Caribe Continental y las Antillas.

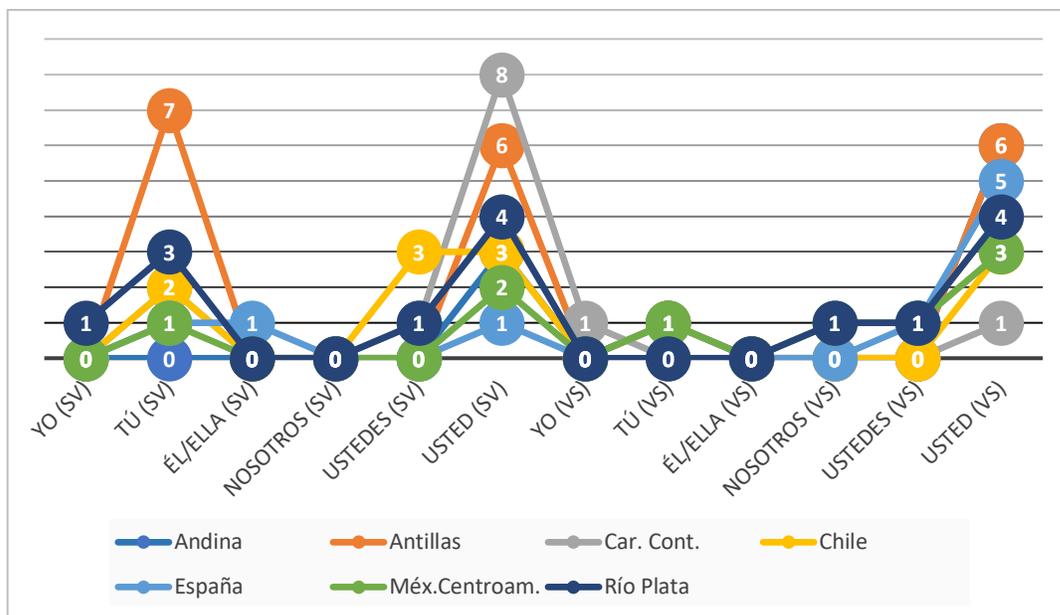


Gráfico 30. Representación de la distribución de frecuencias absolutas de sujetos con pronombre personal, según zona geográfica

A modo de resumen, la categoría más abundante en el corpus es la de sustantivo con o sin artículo. La posposición del sujeto es mayoritaria en todas las categorías, salvo para los sujetos constituidos por *tú* e *indefinido*. Es en la categoría de infinitivo donde se manifiesta más acusadamente el orden VS. No obstante, la única categoría que cubre más del 20% del corpus es la de sustantivo solo o con artículo, por lo que los resultados expuestos no pueden ser expresados en términos de tendencia, salvo para la categoría antes mencionada.

### 3.1.4.3. El orden del sujeto respecto del verbo según complejidad del sujeto

El Gráfico 31 resume las proporciones de las categorías de complejidad del sujeto en el corpus. La unidad mínima (UM), es decir, aquellos sujetos constituidos por un sustantivo sin determinante, un pronombre, un nombre o título ocupan el 31,4% de la muestra, con 220 casos. Los sujetos con estructura de sustantivo con determinante (det+SN) tienen una frecuencia relativa de 27,1% (190 casos), mientras que las construcciones de sustantivo con determinante y frase preposicional (det+SN+fprep) representan el 14,3% del corpus (100 casos cada una). Las cláusulas sustantivas (CS) constituyen el 11,7% del corpus (82 casos), y les siguen en orden descendente los sujetos constituidos por sintagmas nominales con determinante nominal, modificados por un adjetivo o un adverbio, o dos modificadores nominales coordinados (det+SN+mod), que representan el 9,1% de la muestra, con 64 casos.

Los elementos nominales sin determinante, acompañados de uno o varios complementos de nombre (SN+compl) son el 2,6% de la muestra (18 casos), mientras que los sustantivos con determinante nominal más una cláusula de relativo (det+SN+CR) tienen constituyen el 1,6% del corpus (con apenas 11 ocurrencias). Las categorías de elementos nominales sin determinante, yuxtapuestos o coordinados (Sncoord), los sintagmas nominales con estructuras complejas (SNCompl.) y los sintagmas nominales con determinante nominal, coordinados, que contengan además complementos preposicionales o subordinadas adjetivas (det+SNcoord) tienen ocurrencias marginales, y constituyen menos del 2 %, con 9, 3 y 3 casos, respectivamente<sup>63</sup>.

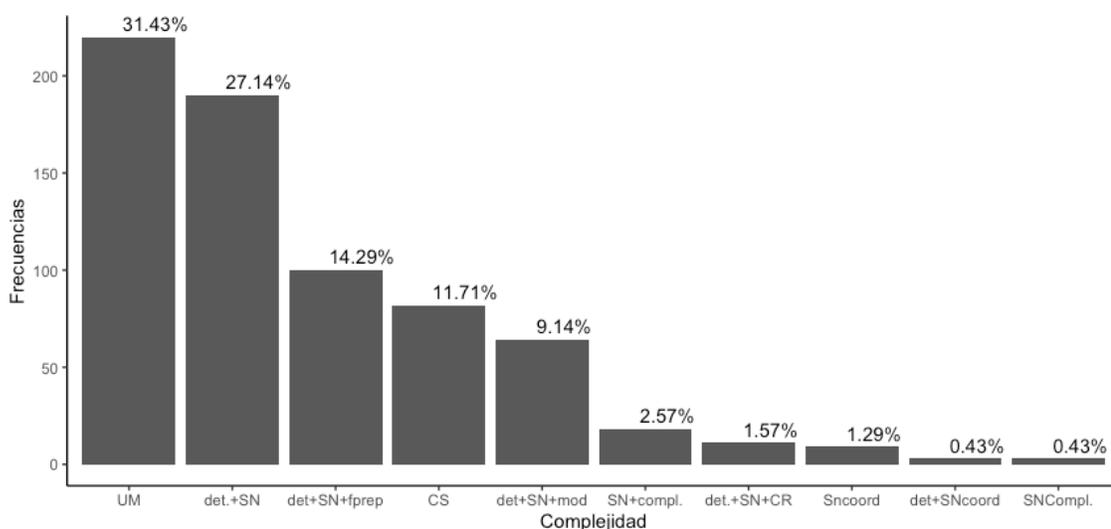


Gráfico 31. Proporción de casos por categoría dentro de la variable Complejidad del sujeto

Respecto del orden del sujeto (ver Gráfico 32), si bien en todos los casos la posposición de este respecto del verbo es mayoritaria, son las categorías CS y SN+compl. las que mayor proporción exhiben para este orden, con 84,1% y 77,8%, respectivamente. Las estructuras det+SN se posponen en un 62,1% de los casos y las det+SN+fprep tienen un 59% de los casos con este orden. Las categorías det+SN+mod, det+SN+CR y UM muestran una ligera preferencia por la posposición del sujeto, con 54,7%, 54,5% y 52,3%, respectivamente.

<sup>63</sup> Omitiremos la descripción relativa al orden del sujeto, correspondiente a estas variables, por constituir casos marginales dentro del corpus.

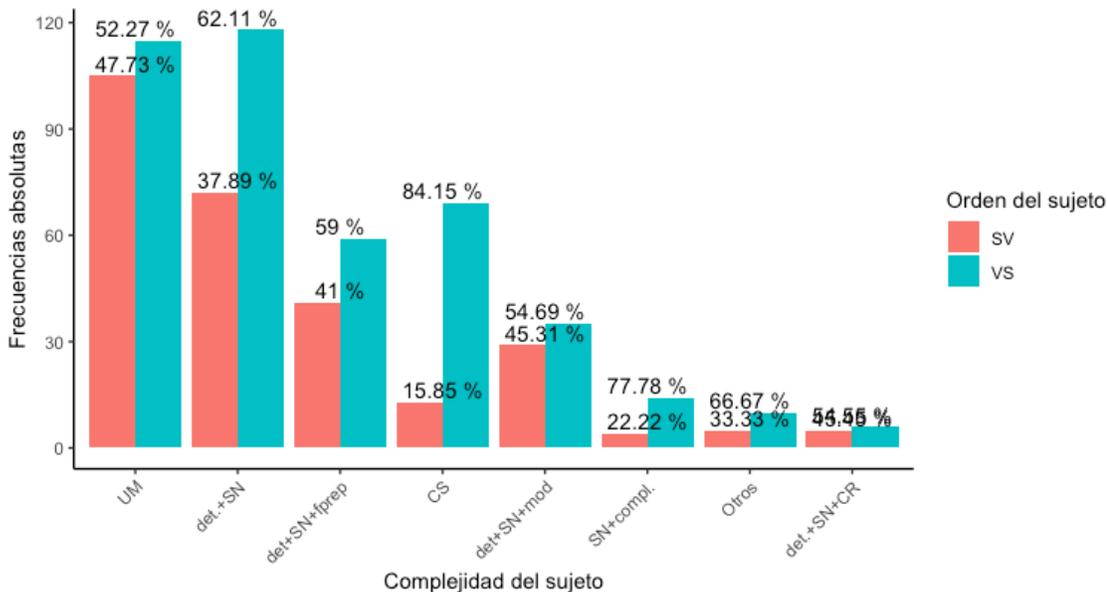


Gráfico 32. Casos con orden SV y VS, agrupados por complejidad del sujeto. El eje Y muestra las frecuencias absolutas. Se presentan las categorías en orden decreciente según su número total de casos. Las barras expresan la proporción de casos para cada orden. Las categorías con casos marginales (menos de 10) se han agrupado en la etiqueta "Otros"

En consonancia con estudios previos (Bentivoglio, 2003), los sujetos agrupados en la categoría UM –mayormente compuestos por sustantivos comunes o propios y por pronombres– tienden a la anteposición (obsérvese que la posposición apenas sobrepasa la mitad de los casos). Ordóñez y Olarrea (2006) habían encontrado que para las interrogativas parciales caribeñas el sujeto constituido por una frase nominal suele estar antepuesto; los resultados de esta investigación ofrecen evidencia sobre una tendencia semejante en la interrogación absoluta. Más concretamente, en la zona Caribe Continental, el orden de sujetos pertenecientes a la categoría UM está antepuesto con una proporción del 84% (21 casos de 25) y en las Antillas alcanza el 48% (17 casos de 35).

También respaldando los resultados encontrados en la literatura (Delbecque, 1988; Bentivoglio, 2003; Bogard, 2017), los sujetos clausulares o los sintagmas con núcleo nominal y varios modificadores tienen una marcada tendencia a la posposición.

Los sujetos constituidos por sustantivos con determinantes, los que son modificados por una frase preposicional, un adjetivo o un adverbio, tienen proporciones más balanceadas, lo que es concordante con la tendencia encontrada por Delbecque (1988), quien postuló que este tipo de estructura es menos proclive a la posposición.

A continuación, se muestra el gráfico de asociación entre complejidad y orden del sujeto.

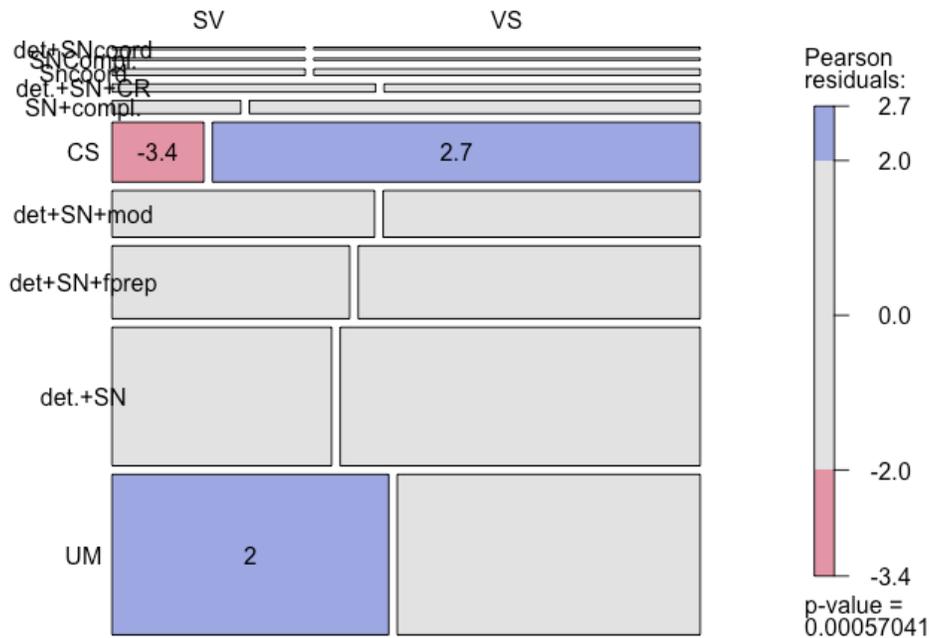


Gráfico 33. Representación de mosaicos de las variables Orden del sujeto y Complejidad del sujeto

Puede observarse que las categorías que exhiben una relación significativa respecto del orden del sujeto son unidad mínima (*UM*) y cláusula sustantiva (*CS*). La primera de ellas solo favorece de manera fuerte el orden SV (tiene un residuo de positivo 2), mientras que la segunda tiene un comportamiento contrario: es un factor inhibitor el orden SV (residuo de -3,4) y favorecedor del orden VS (residuo de 2,7).

Con el objetivo de evitar el desbalance en los datos (téngase en cuenta que hay cinco variables muy poco productivas, que no superan el 3% del corpus, en tanto la variable más representada ocupa la tercera parte de él), hemos redistribuido los datos al interior de la variable, a partir del comportamiento general descrito anteriormente. Así, se decidió reagrupar los datos en tres categorías<sup>64</sup>, a saber:

<sup>64</sup> Se han usado letras para diferenciar los niveles en lugar de etiquetas más explícitas porque ellas serían demasiado largas (especialmente en los casos de las dos últimas categorías, debido a la diversa naturaleza de las estructuras nominales agrupadas) y entorpecerían la visualización de los datos en los gráficos de mosaico.

A. Estructuras de complejidad baja: corresponde a las categorías unidad mínima (pronombre, nombre propio, sustantivo o infinitivo escuetos y títulos) y elemento nominal con determinante (det+SN).

B. Estructuras de complejidad media: agrupa a los elementos nominales sin determinante, yuxtapuestos o coordinados (SNcoord), los que son modificados por una frase preposicional (det+SN+prep), un adjetivo o un adverbio (det+SN+mod) y los elementos nominales sin determinante, acompañados de uno o varios complementos de nombre (SN+compl).

C. Estructuras de complejidad alta: concentra a sujetos nominales con estructuras complejas (SNCompl.), cláusulas sustantivas (CS), y sintagmas nominales con determinante nominal, coordinados, que contengan además complementos preposicionales o subordinadas adjetivas (det+SNcoord) y los elementos nominales determinados con una cláusula de relativo (det+SN+CR)

El análisis de residuos, recogido en el Gráfico 34, muestra que, una vez redistribuidos los datos una sola relación se muestra como fuerte, la que se establece entre los sujetos con estructuras más complejas y ambos órdenes, es decir, los sujetos extensos (o más complejos cognitivamente) estimulan la aparición del orden VS e inhiben el orden SV. Ello ofrece evidencia acerca del principio de emplazamiento de los sujetos más extensos (*heavy*) al final de la oración, reconocido para el español (López Meirama, 2006). El análisis de regresión posterior podrá aportar otros elementos al análisis de las probabilidades con que aparecen cada una de estas categorías en ambos órdenes.

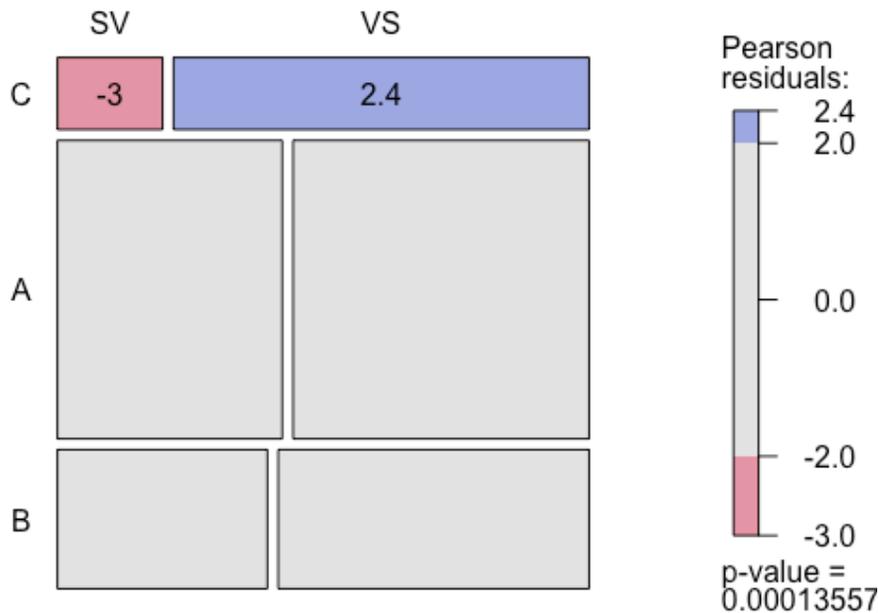


Gráfico 34. Representación de mosaicos de las variables Orden del sujeto y Complejidad del sujeto (se han reducido los niveles a 3, según su grado de complejidad)

### 3.1.5. Variable semántico-pragmática

#### 3.1.5.1. El orden del sujeto respecto del verbo según la función pragmática del enunciado

El Gráfico 35 muestra la distribución de las proporciones de las categorías de función pragmática en el corpus. Las preguntas reales son las más abundantes: ocupan un 34% del corpus, con 238 casos. Les siguen, en orden descendente, las interrogativas *subjectio* (17,1% del corpus con 120 casos), las problemáticas (12,4% con 87 casos) y las retóricas (11,4% con 80 casos). Las interrogativas referidas y anticipativas ocupan el 5% del corpus, con 36 casos cada una, mientras que el resto de las categorías (interrogativas deliberativas, confirmativas, de foco conversacional, de información compartida, repetitivas, clarificadoras, fáticas y peticiones u ofrecimientos) tienen porcentajes que no alcanzan el 2%. Esta distribución de frecuencias es coherente con la naturaleza del corpus, el cual está compuesto por textos escritos, en su mayoría provenientes de libros y publicaciones periódicas. Las preguntas

*subjectio*, problemáticas y retóricas son recurrentes en el discurso de la argumentación, muy común en textos periodísticos, de ahí su alta proporción en el corpus: sumadas, constituyen el 41% del total. En consecuencia, el resto de las categorías, típicas de la conversación, está presente en mucha menor proporción: sumadas representan el 14,3% del corpus (obviando la categoría N/A). Es lógico, por demás, que sea la pregunta real la más frecuente, pues es la manifestación prototípica de la interrogación.

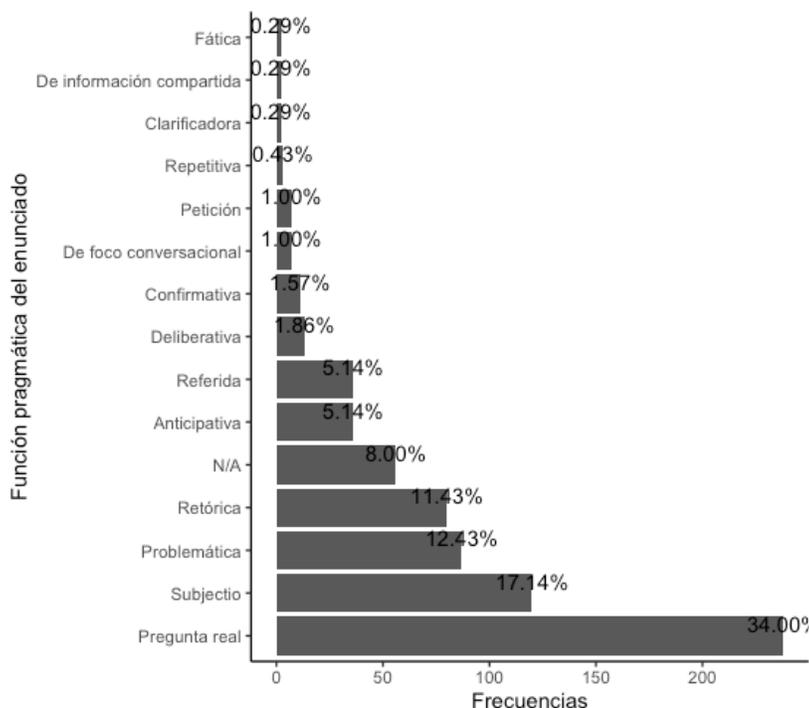


Gráfico 35. Frecuencias relativas de las 14 categorías de función pragmática incluidas en el análisis

El Gráfico 36 recoge las frecuencias relativas de los dos órdenes estudiados en las categorías más abundantes del corpus. Las interrogativas que ocupan menos del 2% del corpus, coincidentes, en su mayoría, con los tipos funcionales de interrogativas habla y relacionales de Freed (1994) y las atribuidas de Escandell (1999), fueron agrupadas en la categoría “Otras”. Como puede observarse, la posposición del sujeto es mayoritaria (salvo en “Otras”). Llama la atención, sin embargo, que las preguntas reales son las que tienen porcentajes más equilibrados, es decir, el orden VS sobrepasa al orden SV solo en 11 puntos porcentuales (55% y 44%, respectivamente). Esto contradice la hipótesis de la existencia de un orden no marcado (VS) para enunciados de este tipo. En nuestros datos, las preguntas parecen tender más al orden alternante.

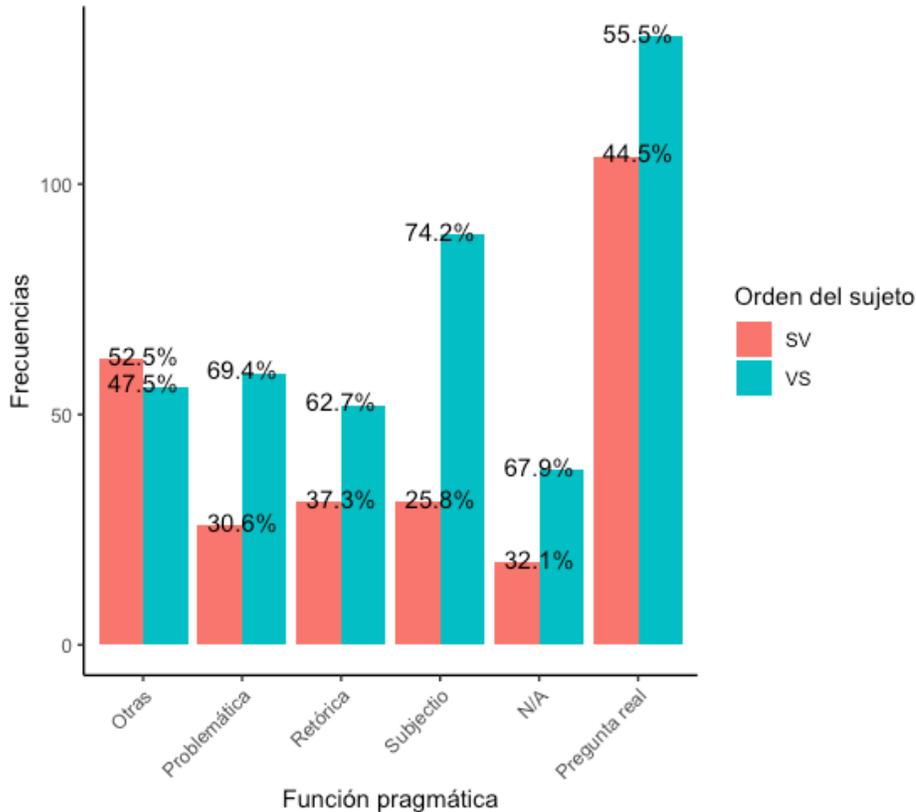


Gráfico 36. Frecuencias relativas para anteposición y posposición del sujeto, según tipos funcionales de Freed (1994)

El porcentaje de anteposición del sujeto presenta una tendencia ligeramente diferente, sin embargo, si se tiene en cuenta la morfología del sujeto, los sujetos constituidos por pronombres personales están antepuestos con una proporción del 63% (26 casos de 41), en tanto los sujetos con elementos nominales no pronominales aparecen en esta posición en el 40,6% de los casos (80 de 197).

Aijón y Serrano (2012) proponen que la anteposición del sujeto, cuando este es pronominal, es más probable cuando el emisor espera cierta respuesta, o cuando la da por consabida, es decir, la pregunta constituiría una toma de postura. A pesar de que el foco de esta investigación no es el análisis cognitivo de los enunciados, una observación superficial de los datos revela que, si bien existen casos con sujetos antepuestos que hablan a favor de una modalización, como en el ejemplo (1), existen también enunciados con sujetos antepuestos cuyo contexto de emisión no contradice un uso neutral de la pregunta (2).

(1) Podemos inferir que el presidente Allende respiró más aliviado aquel atardecer del domingo 9. Las «acciones en curso» debieron de parecerle contundentes. En lo político, creía que realmente avanzaba el principio de acuerdo con el presidente del Partido Demócrata Cristiano, el senador Patricio Aylwin, y su propuesta de plebiscito iba a ser un paso clave. En lo militar, el general Pinochet le daba amplia confianza. Pero a esa misma hora, Pinochet —en su casa— estaba tomando la lapicera para firmar el pacto golpista.

— *¿Tú has pensado en que esto nos puede costar la vida a nosotros y a muchos más?* —le dijo al general Leigh, manteniendo la punta del lápiz aún en el aire. (Chi55)

(2) —¿Cuánto se invertirá en eso?

—El presupuesto es de cerca de 300 millones de dólares.

—*¿Usted quisiera quedarse en el Ministerio hasta que vea resultados en la exploración?*

—No. Voy a completar 5 años, he tenido el enorme orgullo de pertenecer a este gobierno pero obligaciones personales, que son de mi vida y que debo cumplir, me impiden continuar. Además, esto es un trabajo agotador. (Car97)

Sucede algo parecido con los sujetos no pronominales. Obsérvese en el ejemplo (3) que el contexto previo a la pregunta y la respuesta indican que el objetivo del emisor es obtener una respuesta sobre algo que desconoce.

(3) — ¿Cómo han andado los nuevos hoteles?

— Tierra Atacama muy bien, excelente. Tierra Patagonia partió en diciembre y perdimos casi inmediatamente porque tuvimos que evacuar a los pasajeros por el incendio.

Habíamos inaugurado apenas ocho días antes. Pero, bueno, la idea era diversificar nuestras operaciones.

— *¿Katari, la sociedad que crearon, piensa construir más hoteles?*

— Es posible, pero primero queremos estar seguros de que los dos nuevos anden bien. Suponiendo que no haya más terremotos, incendios o cosas así, tras un par de años debíamos hacer algo más. (Chi12)

Se impone, entonces, un análisis pormenorizado de otros elementos semióticos no contemplados en esta investigación para poder llegar a conclusiones más precisas sobre la alternancia de orden en las preguntas reales.

Las interrogativas *subjectio* y problemáticas, por su parte, tienen altos porcentajes de posposición del sujeto (74% y 70%, respectivamente), tendencia que sigue, de manera menos acusada, la interrogación retórica (62,7%).

Por otra parte, se realizó una reagrupación de los datos con el objetivo de comprobar si podía establecerse una correlación entre forma y función, tomando como eje las categorías funcionales de Freed (1994), agrupadas por la autora en un *continuum*, que va desde la interrogación externa (estimada como prototípica y por tanto asociada a un orden específico, a saber, VS para la interrogación parcial) hasta la interrogación expresiva (cuyo propósito se presenta como marcado o menos prototípico, lo que motivaría la prominencia de un orden no canónico). La hipótesis de partida, para las interrogativas totales, es que el polo izquierdo del *continuum* está principalmente asociado al orden VS, y que el polo derecho se corresponde mayoritariamente con el orden SV.

La Tabla 20 muestra que las interrogativas externas (pregunta real, interrogativa problemática y peticiones u ofrecimientos) son las más frecuentes (64,6%). Aunque este tipo funcional es el que tiene un mayor promedio de posposición (62,8%), también las interrogativas expresivas (interrogativas *subjectio*, deliberativa, retórica y referida) ostentan el sujeto pospuesto en el 60% de los casos. Las interrogativas relacionales (fáticas, de foco conversacional, de información compartida y de elaboración) muestran una ligera tendencia a la posposición del sujeto (54,5%), mientras que las de habla (interrogativas clarificadora, repetitiva, confirmativa y anticipativa) presentan el sujeto antepuesto en el 60% de los casos. Sin embargo, estas dos últimas categorías están desbalanceadas respecto del resto (cuentan con 52 y 11 casos, respectivamente), por lo que los resultados para ellas deben ser mirados con reserva.

| <i>Orden del sujeto</i> | <i>Categorías funcionales</i> |                     |                   |                 |            | <i>Total</i> |
|-------------------------|-------------------------------|---------------------|-------------------|-----------------|------------|--------------|
|                         | <i>De habla</i>               | <i>Relacionales</i> | <i>Expresivas</i> | <i>Externas</i> | <i>N/A</i> |              |
| SV                      | 31                            | 5                   | 52                | 168             | 18         | 274          |
|                         | 11.3 %                        | 1.8 %               | 19 %              | 61.3 %          | 6.6 %      | 100 %        |
|                         | 59.6 %                        | 45.5 %              | 40.3 %            | 37.2 %          | 32.1 %     | 39.1 %       |
| VS                      | 21                            | 6                   | 77                | 284             | 38         | 426          |
|                         | 4.9 %                         | 1.4 %               | 18.1 %            | 66.7 %          | 8.9 %      | 100 %        |
|                         | 40.4 %                        | 54.5 %              | 59.7 %            | 62.8 %          | 67.9 %     | 60.9 %       |
| <b>Total</b>            | 52                            | 11                  | 129               | 452             | 56         | 700          |
|                         | 7.4 %                         | 1.6 %               | 18.4 %            | 64.6 %          | 8 %        | 100 %        |
|                         | 100 %                         | 100 %               | 100 %             | 100 %           | 100 %      | 100 %        |

$$\chi^2=11.299 \cdot df=4 \cdot \text{Cramer's } V=0.127 \cdot \text{Fisher's } p=0.023$$

Tabla 20. Tabla cruzada de categorías funcionales de Freed (1994) y Orden del sujeto

Los resultados obtenidos se condicen solo parcialmente con los hallazgos de Brown y Rivas (2011a) para la interrogación parcial. Las autoras encontraron que las interrogativas externas, destinadas a obtener información desconocida por el emisor, tienen de manera mayoritaria un orden VS. Sin embargo, la categoría de interrogativas expresivas en nuestro corpus tiene un comportamiento contrario a los resultados obtenidos por estas autoras, es decir, la secuencia que prima es, aunque en proporción un poco menor, la de VS. Ello implica, desde luego, que los datos de nuestro corpus se comportan de una manera no esperada: la suposición inicial, de que los polos opuestos del *continuum* traían aparejados un orden mayoritario –vale decir, VS para las interrogativas externas y de habla; y SV para las relacionales y expresivas– no se ha podido comprobar.

Una segunda reagrupación y análisis de datos tuvo como objetivo conocer cómo se comporta el orden del sujeto según la tipología establecida por Escandell (1999). Se partió de la hipótesis de que el orden no marcado para la pregunta epistémica es el de VS. Los resultados, recogidos en el Gráfico 37, muestran que este tipo de interrogativa (que agrupa las preguntas reales, las interrogativas problemáticas, las *subjectio* y las deliberativas) tiene un orden predominante verbo-sujeto (62,45%), si bien la anteposición es más alta de lo que cabría para un orden que la autora plantea como marcado (37,5%). También las interrogativas

atribuidas (interrogativas repetitivas, clarificadoras, confirmativas, referidas y anticipativas) mantienen el comportamiento esperado según la propuesta teórica de Escandell, pues el sujeto aparece antepuesto en el 54% de los casos. Sin embargo, tanto la anteposición relativamente alta del sujeto (37%) en las preguntas epistémicas como su posposición considerable en las interrogativas atribuidas (46,5%), podría estar hablando a favor de la existencia de tendencias poco categóricas, por lo que la hipótesis del orden no marcado no parece comprobarse. Por su parte, las interrogativas directivas presentan el sujeto antepuesto en 4 de los 7 casos (57%), pero la baja frecuencia de casos no permite llegar a conclusiones definitivas. En cualquier caso, parecen comprobar la hipótesis de que la anteposición puede responder a contextos interactivos relacionados con la cortesía verbal Aijón y Serrano (2012). La interrogación retórica, por su parte, tiene un comportamiento diferente al planteado por Escandell, pues la secuencia mayoritaria es VS (62,6% de los casos).

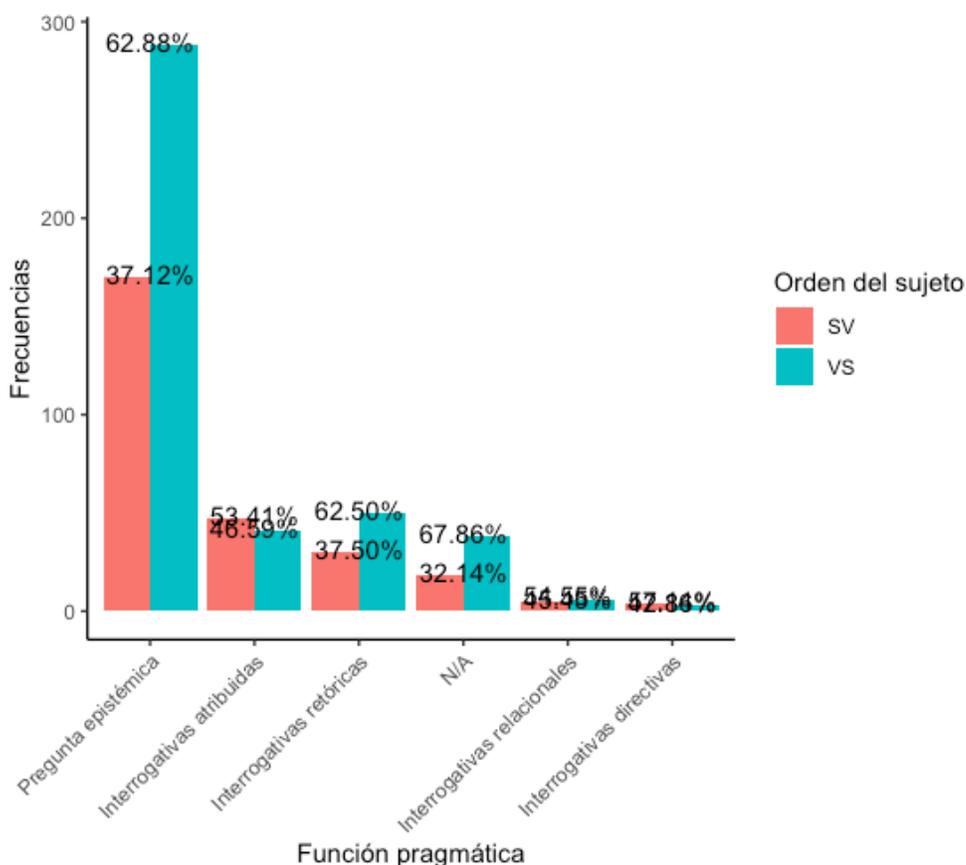


Gráfico 37. Frecuencias relativas para anteposición y posposición del sujeto, según tipos generales propuestos por Escandell (1999). Se ha mantenido en el gráfico el tipo Interrogativa relacional de Freed (1994) por no existir en la propuesta de la autora

Respecto de la aseveración anteriormente hecha sobre la inexistencia de un orden no marcado en las interrogativas que buscan información (preguntas epistémicas), debe ser señalado que la proporción más baja de posposición se debe de manera importante al comportamiento de las preguntas reales y las deliberativas. En el primer tipo, tal como se había mencionado en el comentario del Gráfico 36, la anteposición alcanza una proporción más alta de lo esperable (44,5%); en el segundo tipo, el mismo orden se manifiesta en el 46,2% de los casos. Las otras dos categorías que pertenecen a este macrotipo (interrogativas problemáticas y *subjectio*) tienen una tendencia marcada a la posposición del sujeto.

Un análisis del orden VS según la zona geográfica arroja elementos adicionales interesantes para poder complementar y entender estas tendencias. Como se puede observar en la Tabla 21, la posposición del sujeto es muy acusada en España en todos los tipos de interrogativa (excepto en las retóricas). Ello podría ayudar a entender por qué en todas las gramáticas normativas acuñan la afirmación de que el orden no marcado en la interrogación total o absoluta del español es el de VS: dado que la mayoría de estas obras ha sido concebida y editada en la península, se entiende que la variedad lingüística haya influido de manera decisiva en estas aseveraciones. Del mismo modo, el sujeto de las interrogativas retóricas de España tiende comparativamente más a la anteposición que el del resto de las zonas geográficas, y esto puede estar en la base de la presunción de que este tipo de enunciado favorece la anteposición del sujeto.

Por otra parte, el Caribe Continental y las Antillas (esta última zona en menor proporción) exhiben el comportamiento esperado: el sujeto tiende a la anteposición tanto en las preguntas reales como en las retóricas. Los resultados obtenidos para Chile corroboran los hallazgos previos de la autora de esta investigación respecto del orden predominante en la interrogación absoluta con valor de pregunta de ese país. Asimismo, el comportamiento de las preguntas reales en las zonas Andina y Río de la Plata pueden estar indicando que en el orden no marcado VS para la interrogación absoluta no es un fenómeno generalizado en todo el entorno hispanohablante, sino que se concentra en las variantes de la península ibérica y México y Centroamérica.

| <i>Función pragmática</i>         | <i>Zona geográfica</i> |          |            |       |        |      |           | <b>Total</b> |
|-----------------------------------|------------------------|----------|------------|-------|--------|------|-----------|--------------|
|                                   | And.                   | Antillas | Car. Cont. | Chile | España | Méx  | Río Plata |              |
| <b>Otras</b>                      | 0.25                   | 0.45     | 0.37       | 0.64  | 0.73   | 0.64 | 0.35      | 0.48         |
| <b>Interrogativa problemática</b> | 0.85                   | 0.61     | 0.85       | 0.70  | 0.82   | 0.54 | 0.64      | 0.69         |
| <b>Interrogativa retórica</b>     | 0.53                   | 0.60     | 0.40       | 0.67  | 0.64   | 0.80 | 0.75      | 0.62         |
| <b>Interrogativa subjectio</b>    | 0.71                   | 0.79     | 0.65       | 0.85  | 0.86   | 0.72 | 0.59      | 0.74         |
| <b>N/A</b>                        | 0.20                   | 0.70     | 0.55       | 0.67  | 0.67   | 0.87 | 0.89      | 0.68         |
| <b>Pregunta real</b>              | 0.58                   | 0.61     | 0.43       | 0.47  | 0.81   | 0.60 | 0.44      | 0.55         |
| <b>Total</b>                      | 0.57                   | 0.62     | 0.50       | 0.60  | 0.78   | 0.66 | 0.53      | 0.60         |

Tabla 21. Proporción de casos con orden VS según zona geográfica

El análisis de la correlación entre la variable función pragmática del enunciado y orden del sujeto se realizó usando como criterio de reagrupación el de categorías más frecuentes en el corpus. El gráfico a continuación presenta el resultado de la prueba chi-cuadrado. Como puede observarse, las interrogativas más comunes en la conversación espontánea, agrupadas en la categoría “Otras” –vale decir, las interrogativas deliberativas, confirmativas, de foco conversacional, de información compartida, repetitivas, clarificadoras, fáticas y peticiones u ofrecimientos, que llamaremos de manera general interrogación no epistémica, siguiendo los criterios de Ecanell (1999)–, favorecen fuertemente la anteposición del sujeto (su residuo es positivo y mayor que 2), mientras que las interrogativas *subjectio* inhiben fuertemente este orden (su residuo es negativo y menor que -2). El resto de las categorías no tiene una relación estadísticamente significativa con el orden del sujeto, si bien la secuencia predominante es VS en todas.

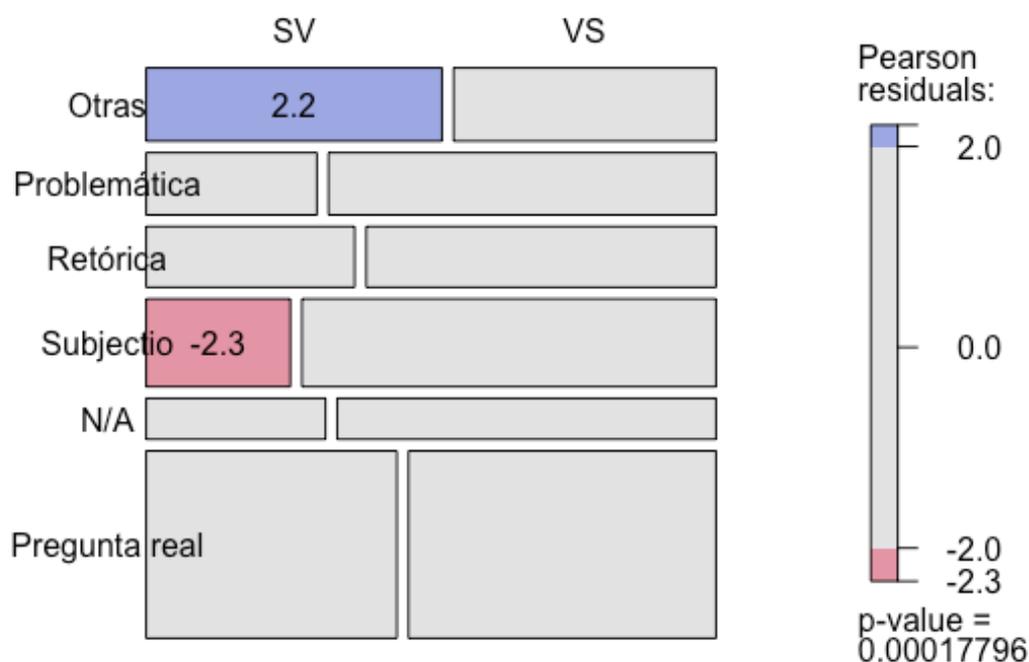


Gráfico 38. Gráfico de mosaico para función pragmática del enunciado interrogativo y Orden del sujeto. Se han agrupado las 14 categorías originales según los tipos más frecuentes en el corpus

A modo de resumen, el análisis de la función pragmática del enunciado interrogativo permitió conocer que las preguntas reales, las interrogativas problemáticas, las interrogativa *subjectio* y las retóricas, los tipos más abundantes en el corpus, favorecen la posposición del sujeto, si bien la tendencia para el primer tipo es menos acusada de lo que cabría esperar, dada la hipótesis teórica de un orden no marcado VS. Las interrogativas que Escandell (1999) clasifica como atribuidas y las de habla y relacionales de Freed (1994), que no buscan una información sobre el mundo, sino que están íntimamente relacionadas con el intercambio comunicativo mismo, tienden a la anteposición, aunque tampoco se trata de una tendencia categórica.

Si bien estos resultados no divergen de las tendencias encontradas para el resto de las variables lingüísticas –es decir, se ha podido observar que el orden VS es mayoritario, pero en muchos casos, dista de ser una tendencia categórica–, debe ser aclarado que los resultados estadísticos obtenidos para la función pragmática del enunciado parten del establecimiento de categorías funcionales en las que se obvió uno de los componentes que interviene en el análisis pragmático (el fonológico) por ser una información no existente para los casos

pertenecientes a textos escritos en el Corpes XXI. Las tendencias encontradas en esta investigación deberán ser confirmadas con un estudio comparativo de la lengua hablada.

### **3.1.6. Recapitulación de la estadística descriptiva y del análisis de asociación entre variables**

Los resultados aquí relatados, pertenecientes al análisis descriptivo y a la primera parte del análisis inferencial, han permitido comenzar a esbozar las respuestas a las tres preguntas de investigación que se plantearon en el acápite teórico. En primer lugar, se ha determinado que el orden VS es, efectivamente, el predominante en la interrogación total española, si bien se ha comprobado la existencia de la variación, pues el orden SV está presente, y en algunos casos, según las categorías recogidas en cada variable, puede ser prevaeciente.

En segundo lugar, de las variables lingüísticas recogidas en el estudio, solo *tiempo verbal* careció de una relación estadísticamente significativa con el orden del sujeto. El resto de los factores semánticos y pragmáticos tienen una influencia estadísticamente demostrada en la posposición del sujeto. Las categorías que se perfilan como más influyentes son: los verbos existenciales, los predicados monovalentes, especialmente si tienen un verbo inacusativo, los sujetos con referentes no animados e indefinidos, los sujetos de infinitivo o constituidos por cláusulas sustantivas; los sujetos extensos, y la función pragmática de interrogación *subjectio*. Respecto de la zona geográfica, es España la región dialectal que muestra una tendencia fuerte a posponer el sujeto.

## **3.2. Análisis de regresión logística**

### **3.2.1. Objetivos**

La regresión logística tiene como fin predecir la probabilidad de que ocurra un evento, representado por la variable respuesta (output) o dependiente (que asume valores entre 0 y 1), dada una configuración de valores de la o las variables predictoras (predictor) (Levshina, 2015). Mediante esta técnica se puede determinar el peso de las categorías o niveles de las variables en el aumento o disminución de la probabilidad de que ocurra tal evento.

En el caso de la presente investigación, el análisis estadístico de regresión tiene los siguientes objetivos:

- 1) Determinar cuál es la influencia de las categorías (o variantes) de las variables independientes (o predictoras) en la determinación del orden del sujeto (variable predecida). Dicho de otra manera, queremos saber cuáles de los niveles de las variables son más probables dado cierta variante de orden (VS o SV).
- 2) Determinar qué variables (o qué combinación de ellas) predice mejor la posición del sujeto.

Dado que nuestra variable dependiente es dicotómica, el análisis más pertinente es el de la regresión logística binomial.

Nuestro reporte de resultados contará con dos partes: en primer lugar, se describirán los hallazgos de la regresión con un solo predictor, y en segundo lugar se detallará los resultados obtenidos en la regresión múltiple.

### **3.2.2. Herramienta de análisis**

La sociolingüística ha usado y desarrollado durante décadas una herramienta de análisis estadístico, denominada la regla variable, destinada a medir cuantitativamente las dependencias o correlaciones entre variables de naturaleza lingüística y extralingüística. Si bien existen programas informáticos de larga data para este análisis, como Goldbarv o

Varbrul, en la actualidad está disponible una versión en software libre, denominada Rbrul<sup>65</sup>, cuyo análisis es más robusto (Johnson, 2009), a la vez que cuenta con una interfaz basada en texto, lo cual lo convierte en una herramienta muy fácil de usar.

### 3.2.3. Variables y categorías de análisis

Las variables que fueron incluidas tanto en el análisis de regresión simple como múltiple<sup>66</sup> fueron las que mostraron una relación fuerte con la variable dependiente, ello mediante la realización de la prueba de independencia chi-cuadrado. Con el objetivo de alcanzar un mayor balance en los datos, algunas variables fueron transformadas, mediante la fusión de niveles.

A continuación, se presentan las nuevas categorías de que constan las variables que han sido modificadas:

Semántica verbal: los verbos de modulación se fundieron con los verbos materiales, dado su número marginal y la cercanía semántica del verbo principal con los procesos materiales. Aún así, como se apuntará oportunamente, la categoría de procesos verbales, que se mantuvo tal cual, tiene solo 22 casos. Sin embargo, una redistribución adicional de niveles (además de la fusión de verbos materiales y de modulación) para evitar un desbalance en los datos habría implicado una arbitrariedad teórica, por lo que se conservaron las siguientes categorías: existencial, material, mental, relacional y verbal.

Morfología del sujeto: las 14 categorías originales fueron reagrupadas en **tres niveles**: niveles sujeto léxico, sujeto pronominal e infinitivos y sujetos clausulares.

Extensión del sujeto: las 10 categorías originales de la variable se reconfiguraron en función de la extensión relativa de las construcciones, quedando solo **tres niveles**: A (complejidad mínima), donde se agrupan los pronombres, nombres propios, sustantivos o infinitivo escuetos, y títulos; B (complejidad media), que reúne a los elementos nominales

---

<sup>65</sup> <http://www.danielezrajohnson.com/rbrul.html>

<sup>66</sup> La cantidad de predictores que intervienen en la regresión logística es importante para medir la calidad del análisis. Se conoce que el número de parámetros debe ser igual o menor que la frecuencia del resultado esperado menos frecuente dividida por 10 (Levshina, 2015). En nuestro caso, la variante SV es la de menor frecuencia, con 274 casos. Eso significa que nuestro modelo podría tener aproximadamente 27 parámetros. Como se verá, la cantidad de predictores es bastante menor.

sin determinante, yuxtapuestos o coordinados, sintagmas nominales con determinante modificados por una frase preposicional, un adjetivo o un adverbio; y C (complejidad alta), que recoge las cláusulas sustantivas, los sintagmas nominales con determinante, coordinados, que contengan complementos preposicionales o subordinadas adjetivas; los elementos nominales con una cláusula de relativo y las estructuras complejas no incluidas en ninguna de las descripciones anteriores.

#### **3.2.4. Sobre la interpretación de los datos**

El valor  $p$  arrojado por el análisis de regresión da cuenta de las categorías (si el modelo tiene solo una variable predictora) o de las variables (si la regresión es múltiple) que tienen un efecto significativamente estadístico en la ocurrencia de un sujeto posverbal (orden VS). Como es convención en este tipo de análisis, interpretaremos como significativos los valores de  $p$  menores que 0,05.

A su vez, los *logodds* se interpretan como las probabilidades logarítmicas de que aparezca el orden que se intenta predecir dado determinado factor. De manera más simple, este estadístico muestra la fuerza de la relación entre el factor y la respuesta que se está prediciendo (Daleszynska, 2016). Cuando el valor de este estadístico es positivo, se interpreta que el nivel especificado aumenta las posibilidades de aparición del orden esperado (orden VS); de manera contraria, el valor negativo indica que dicho nivel disminuye la probabilidad de aparición de la variante que se observa. Una manera más intuitiva de saber cuán probables son determinados factores en la manifestación de cierto orden es traducir los *logodds* a probabilidades simples<sup>67</sup> mediante la función *exp()* de R.

El valor de los *logodds* debe ser reportado junto a los pesos de las categorías (*centered factor weight*). Este estadístico, que toma valores entre 0 y 1, es una medida de favorecimiento del orden que se esté observando dado la categoría de que se trate. Valores por encima de 0,5 indican categorías favorecedoras y valores bajo este umbral, categorías

---

<sup>67</sup> En nuestro caso, es sería razón entre el número total de casos con orden VS y el número total de observaciones (VS+SV).

desfavorecedoras. Si el valor de los *logodds* es cercano a 0 y el peso categorial ronda el 0,5, se puede decir que este factor es neutro.

Para la regresión múltiple, Rbrul ofrece un análisis estadístico denominado algoritmo de subida y bajada (*setp-up, step-down*). El análisis de subida comienza sin predictores y agrega el grupo de factores más significativo, si lo hay, antes de repetir el procedimiento. El análisis de bajada comienza con todos los predictores posibles y elimina el que menos contribuye al modelo, y luego lo repite hasta que todos los predictores restantes sean significativos (Johnson, 2009, p. 380).

La exclusión de algunas variables durante este proceso deriva de su poca representatividad en la determinación (o favorecimiento) del orden en que se centra el análisis, a partir del valor p obtenido. Si este es mayor que 0,05 se interpreta que la variable en cuestión (es decir, el conjunto de sus categorías) es una débil predictora del orden cuando todas las variables interactúan.

### **3.3. Resultados regresión logística simple**

El análisis de regresión simple está encaminado a responder el primer objetivo estadístico, a saber, determinar la probabilidad de aparición de una determinada categoría de las variables lingüísticas, dada cierta variante de orden (VS o SV). A continuación, se presentan los resultados de análisis de regresión de la variable dependiente Orden del sujeto y su relación con las variables que intervienen en el estudio. Partiendo de la necesidad de conocer cuál es la influencia de las categorías en la disposición del sujeto en la interrogación, el análisis se ha hecho sobre la aplicación del orden VS en cada uno de los casos del corpus. Así, el estadístico central del análisis de regresión logística, el *logit* (o *logodds*), refleja las posibilidades de que aparezca el sujeto pospuesto dada la configuración de los niveles de los predictores. Si el valor del *logodds* es positivo, el nivel especificado aumenta las posibilidades de aparición de la variante en la que se centra la regresión. Si, por el contrario, el coeficiente es negativo, el nivel especificado disminuye las probabilidades de que aparezca dicha variante u orden.

Debe ser aclarado que, aunque la categoría o nivel usado como referencia no tiene implicaciones estadísticas (Levshina, 2015), se ha decidido fijar como valores

predeterminados los que se estima son los más neutros en la declaración. En el caso de las variables referidas al sujeto, los niveles desde los cuales el programa realiza la comparación son: sustantivo con determinante, animado, definido, humano; para las variables relativas al predicado, los niveles referenciales son: verbos transitivos, que designan procesos materiales y que tienen más de un argumento. En el caso de zona geográfica la comparación se realiza con España como nivel base.

### **3.3.1. Variable zona geográfica**

El análisis de regresión para la variable Zona geográfica (Tabla 22) arrojó que es un buen predictor del orden ( $p= 0.000954$ ). Acorde con los resultados obtenidos en el análisis bivariado, es España la zona que mejor predice el orden VS, con probabilidad logarítmica de 0.80 y un peso categorial de 0.70; en términos de simples probabilidades, esto significa que la posposición del sujeto tiene dos veces más probabilidades de manifestarse en España que en resto de las zonas. También México y Centroamérica (probabilidades de 0.20) y las Antillas (0.02) son zonas que manifiestan un peso categorial por encima de 0.50, lo que indica que son regiones en las que el orden favorecido es VS. A pesar de lo anterior, se debe aclarar que, si bien el resultado para las Antillas respalda el análisis descriptivo efectuado en el capítulo anterior –es decir, que el orden VS es mayoritario en la zona, lo que contradice los postulados encontrados en múltiples estudios sobre interrogación parcial y declaración–, tanto el peso de la variable como las probabilidades logarítmicas podrían estar indicando la neutralidad de esta zona respecto de la posposición del sujeto; o dicho, de otro modo, no se puede afirmar categóricamente que la zona privilegie el uso del orden VS.

En el Caribe continental la anteposición del sujeto parece mucho más probable, dado el índice negativo (-0.46) y el bajo peso del factor (0.38) para el orden VS. También las zonas de Chile, Andina y Río de la Plata desfavorecen la posposición del sujeto, siendo Chile la que muestra una tendencia más neutra (probabilidades logarítmicas cercanas a 0 y peso categorial cercano a 0.50). En resumen, a partir de los datos con los que contamos, pudiera afirmarse que el orden del sujeto observa, efectivamente, una variación dialectal. Pudiera decirse, en consecuencia, que existe un *continuum* en esta variación que va del sujeto mayoritariamente pospuesto (con España y México y Centroamérica como las zonas donde

la posposición del sujeto es más probable) al sujeto mayoritariamente antepuesto (con las zonas Andina, Río de la Plata y Caribe Continental como las que más favorecen la anteposición del sujeto). En el centro de este *continuum* las zonas Antillas y Chile aparecen como las más neutrales (sus probabilidades están en torno al valor central, por lo que no tienen una tendencia marcada a expresar uno u otro orden).

---

ONE-LEVEL ANALYSIS OF RESPONSE Orden del sujeto WITH PREDICTOR(S): Zona geográfica  
(0.000954)

**Zona geográfica**

| factor             | logodds | tokens | VS/VS+SV | centered factor weight |
|--------------------|---------|--------|----------|------------------------|
| España             | 0.805   | 100    | 0.78     | 0.691                  |
| México y Centroam. | 0.202   | 100    | 0.66     | 0.55                   |
| Antillas           | 0.029   | 100    | 0.62     | 0.507                  |
| Chile              | -0.055  | 100    | 0.60     | 0.486                  |
| Andina             | -0.179  | 100    | 0.57     | 0.455                  |
| Río de la Plata    | -0.341  | 100    | 0.53     | 0.416                  |
| Caribe Continental | -0.461  | 100    | 0.50     | 0.387                  |

---

\$misc.1

| n   | df | intercept | overall proportion | centered input | prob |
|-----|----|-----------|--------------------|----------------|------|
| 700 | 7  | 0.461     | 0.609              | 0.613          |      |

\$misc.2

| log.likelihood | AIC     | AICc    | Dxy   | R2    |
|----------------|---------|---------|-------|-------|
| -457.283       | 928.565 | 928.727 | 0.197 | 0.044 |

---

Tabla 22. Análisis de regresión logística entre orden del sujeto y zona geográfica

### 3.3.2. Variable semántica verbal

El análisis de regresión para la variable Semántica verbal (Tabla 23) arrojó que el orden VS es más probable en estructuras con verbos existenciales (probabilidades de 0.63). Nótese, además, que esta categoría es la única que tiene un peso categorial considerable (0.65). Los procesos relacionales no muestran una tendencia determinada (su valor de probabilidades es muy cercano a 0 y su peso ronda el valor medio 0.50, por lo que no favorece ninguna de las dos opciones de la variable respuesta). Los procesos mentales, materiales y verbales no

potencian las probabilidades de un orden VS, si bien estos los resultados para estos últimos verbos deben ser mirados con cautela, dada la baja cantidad de datos en el corpus: como es sabido, las categorías con pocos tokens pueden dar lugar a coeficientes poco confiables. Debe ser aclarado que, aun cuando el modelo es estadísticamente significativo ( $p=0.0491$ ), esta significatividad está muy cercana al valor umbral. A su vez, el valor de  $R^2$  indica que este modelo explica solo el 2% de los datos. Podríamos concluir, entonces, que si bien para nuestros datos el modelo parece predecir de manera fehaciente el orden, estas tendencias deben ser comprobadas en el futuro con otros *corpora*, idealmente más equilibrados.

---

ONE-LEVEL ANALYSIS OF RESPONSE Orden del sujeto WITH PREDICTOR(S): Semántica verbal  
(0.0491)

| <b>Semántica</b> |                |               |                 |                               |
|------------------|----------------|---------------|-----------------|-------------------------------|
| <b>factor</b>    | <b>logodds</b> | <b>tokens</b> | <b>VS/VS+SV</b> | <b>centered factor weight</b> |
| Existencial      | 0.639          | 87            | 0.747           | 0.655                         |
| Relacional       | -0.016         | 370           | 0.605           | 0.496                         |
| Mental           | -0.162         | 114           | 0.570           | 0.46                          |
| Material         | -0.200         | 107           | 0.561           | 0.45                          |
| Verbal           | -0.262         | 22            | 0.545           | 0.435                         |

---

\$misc.1

| n   | df | intercept | overall | proportion | centered | input | prob |
|-----|----|-----------|---------|------------|----------|-------|------|
| 700 | 5  | 0.444     | 0.609   | 0.609      |          |       |      |

\$misc.2

| log.likelihood | AIC     | AICc    | Dxy   | R2    |
|----------------|---------|---------|-------|-------|
| -463.802       | 937.605 | 937.691 | 0.106 | 0.019 |

---

Tabla 23. Análisis de regresión logística entre orden del sujeto y semántica verbal

### 3.2.3. Variable realización valencial

Los resultados arrojaron que la variable Realización valencial es un buen predictor del orden ( $p=0.000145$ ). Como muestra la Tabla 24, los verbos monovalentes muestran la mayor probabilidad logarítmica (0.68) y el mayor peso categorial (0.66). Quiere ello decir que este tipo de verbos condiciona de manera fuerte el orden VS. Los verbos bivalentes también

favorecen la posposición del sujeto, pero los valores de las probabilidades y el peso del factor indican que su influencia es bastante menor, es decir, no puede afirmarse categóricamente, al menos con los datos de que disponemos, que sea un factor condicionante del orden VS. Por último, los trivalentes muestran una probabilidad negativa, lo que los coloca como verbos que desfavorecen la posposición del sujeto. Sin embargo, tanto el valor de las probabilidades como del peso categorial deben ser vistos con reserva por la baja cantidad de casos existentes en el corpus (solo 16).

---

ONE-LEVEL ANALYSIS OF RESPONSE Orden del sujeto WITH PREDICTOR(S): Realización valencial (0.000145)

---

| <b>Realización valencial</b> |                |               |                 |                               |
|------------------------------|----------------|---------------|-----------------|-------------------------------|
| <b>factor</b>                | <b>logodds</b> | <b>tokens</b> | <b>VS/VS+SV</b> | <b>centered factor weight</b> |
| Monovalentes                 | 0.680          | 402           | 0.669           | 0.664                         |
| Bivalentes                   | 0.132          | 282           | 0.539           | 0.533                         |
| Trivalentes                  | -0.813         | 16            | 0.312           | 0.307                         |

---

\$misc.1

| n   | df | intercept | overall | proportion | centered | input | prob |
|-----|----|-----------|---------|------------|----------|-------|------|
| 700 | 3  | 0.024     | 0.609   | 0.506      |          |       |      |

\$misc.2

| log.likelihood | AIC     | AICc    | Dxy   | R2    |
|----------------|---------|---------|-------|-------|
| -459.726       | 925.452 | 925.486 | 0.155 | 0.032 |

---

Tabla 24. Análisis de regresión logística entre orden del sujeto y estructura argumental

### 3.2.4. Variable transitividad verbal

La regresión de la Transitividad verbal y el Orden arrojó que el modelo utilizado es estadísticamente significativo ( $p=1.23 \cdot 10^{-10}$ ). Según se observa en la Tabla 25, los verbos inacusativos favorecen de manera fuerte la aparición del sujeto pospuesto, dada la probabilidad logarítmica de 1.01 y el peso de 0.73. En términos de probabilidades simples, estos verbos tienen casi tres veces más chances de coocurrir con sujetos pospuestos que el resto de las categorías. Los predicados nominales también favorecen el orden VS, mientras que los verbos transitivos e inergativos tienen una tendencia contraria: son factores

desfavorecedores de este orden, dadas las probabilidades de signo negativo y el peso factorial por debajo de 0.50.

---

ONE-LEVEL ANALYSIS OF RESPONSE Orden del sujeto WITH PREDICTOR(S): Transitividad  
(1.23e-10)

**Transitividad**

| factor            | logodds | tokens | VS/Vs+SV | centered factor weight |
|-------------------|---------|--------|----------|------------------------|
| Inacusativo       | 1.017   | 165    | 0.806    | 0.734                  |
| Predicado nominal | 0.078   | 260    | 0.619    | 0.52                   |
| Transitivo        | -0.436  | 211    | 0.493    | 0.393                  |
| Inergativo        | -0.659  | 64     | 0.438    | 0.341                  |

---

\$misc.1

| n   | df | intercept | overall | proportion | centered | input | prob |
|-----|----|-----------|---------|------------|----------|-------|------|
| 700 | 4  | 0.408     | 0.609   | 0.601      |          |       |      |

\$misc.2

| log.likelihood | AIC     | AICc    | Dxy   | R2    |
|----------------|---------|---------|-------|-------|
| -444.008       | 896.016 | 896.074 | 0.287 | 0.093 |

---

Tabla 25. Análisis de regresión logística entre orden del sujeto y transitividad verbal

### 3.3.5. Variable referencialidad del sujeto

#### 3.3.5.1. Animacidad

Respecto de la variable Animacidad, los resultados de la regresión muestran que la variable es una buena predictora del orden ( $p=4.05 \text{ e-}06$ ); asimismo, puede comprobarse acá la tendencia comentada en el acápite anterior de que mientras más asciende el sujeto en la escala de animacidad –es decir, mientras más el sujeto es prototípicamente animado (es el caso de los sujetos humanos)–, más tiende a aparecer antepuesto. Obsérvese en la Tabla 26 que los sujetos con referentes inanimados tienen la más alta probabilidad logarítmica de aparecer pospuestos al verbo (0.35). En el extremo opuesto están los sujetos con referentes

humanos, con una probabilidad negativa (-0.49). Los sujetos con referentes animados, por su parte, si bien tienen una probabilidad positiva, esta no dista mucho del 0, del mismo modo que su peso no está muy alejado de 0.50, por lo que son favorecedores del orden VS en menor medida que los sujetos inanimados. Estos resultados se condicen con los hallazgos de Rosemeyer (2018), quien encontró, para la interrogación parcial, que los sujetos animados aumentan fuertemente la probabilidad de anteposición del sujeto.

---

ONE-LEVEL ANALYSIS OF RESPONSE Orden del sujeto WITH PREDICTOR(S): Animacidad (4.05e-06)

| <b>Animacidad</b> |                |               |                 |                               |
|-------------------|----------------|---------------|-----------------|-------------------------------|
| <b>factor</b>     | <b>logodds</b> | <b>tokens</b> | <b>VS/VS+SV</b> | <b>centered factor weight</b> |
| Inanimado         | 0.349          | 421           | 0.677           | 0.586                         |
| Animado           | 0.140          | 54            | 0.630           | 0.535                         |
| Humano            | -0.489         | 225           | 0.476           | 0.38                          |

---

\$misc.1

| n   | df | intercept | overall proportion | centered input | prob |
|-----|----|-----------|--------------------|----------------|------|
| 700 | 3  | 0.391     | 0.609              | 0.596          |      |

\$misc.2

| log.likelihood | AIC     | AICc    | Dxy   | R2    |
|----------------|---------|---------|-------|-------|
| -456.151       | 918.303 | 918.337 | 0.189 | 0.043 |

---

Tabla 26. Análisis de regresión logística entre orden del sujeto y su animacidad

### 3.3.5.2. Definitud

Los resultados de la regresión para las variables Definitud y Orden del sujeto son significativos ( $p=7.89 \text{ e-}10$ ). Como se muestra en la Tabla 27, las probabilidades logarítmicas muestran que los sujetos escuetos son los únicos que permiten predecir la posposición del sujeto (probabilidades de 1.02 y peso de 0.73), es decir, tienen casi tres veces más chances que los demás tipos de sujeto de aparecer pospuestos. Los sujetos con referentes indefinidos parecen favorecer el orden SV (probabilidades -0.22 y peso categorial de 0.44 para el orden VS), si bien de manera menos fuerte que los sujetos definidos: estos tienen una probabilidad logarítmica bastante menor que los indefinidos (-0.8) y lo mismo sucede con su peso (0.31).

Pudiera afirmarse, en consecuencia, que la posposición del sujeto está fuertemente condicionada por la inexistencia de un determinante. Estos resultados son coherentes con el comportamiento normal de los sintagmas escuetos también en la declaración. La interrogación, por tanto, no parece jugar un rol importante en la determinación del orden en estos casos. Además, los sintagmas nominales con un elemento determinante no se correlacionan de manera fuerte con el orden VS en las interrogativas del corpus, por lo que cabría concluir que favorecen la anteposición del sujeto, si bien en el caso de los elementos nominales definidos esta tendencia es mucho más clara.

---

ONE-LEVEL ANALYSIS OF RESPONSE Orden del sujeto WITH PREDICTOR(S): Definitud (7.89e-10)

**Definitud**

| <b>factor</b> | <b>logodds</b> | <b>tokens</b> | <b>VS/VS+SV</b> | <b>centered factor weight</b> |
|---------------|----------------|---------------|-----------------|-------------------------------|
| Escueto       | 1.028          | 88            | 0.886           | 0.737                         |
| Indefinido    | -0.227         | 58            | 0.690           | 0.443                         |
| Definido      | -0.801         | 554           | 0.556           | 0.31                          |

---

\$misc.1

n df intercept overall proportion centered input prob  
700 3 1.026 0.609 0.736

\$misc.2

log.likelihood AIC AICc Dxy R2  
-447.607 901.215 901.249 0.183 0.101

---

Tabla 27. Análisis de regresión logística entre orden del sujeto y definitud del sujeto

### 3.3.6. Variable morfología del sujeto

La regresión de las variables Morfología del sujeto y Orden arrojó que la relación es significativa ( $p=2.85 \text{ e-}07$ ). Según se observa en la Tabla 28, la categoría de infinitivo y sujetos clausulares condiciona de manera importante el orden (probabilidades logarítmicas de 0.97 y peso de 0.72). Los sujetos léxicos y pronominales reducen las probabilidades la

aparición del orden VS, si bien son los segundos lo hacen de manera considerable: (*logodds* de -0.79 y peso de 0.32).

---

ONE-LEVEL ANALYSIS OF RESPONSE Orden del sujeto WITH PREDICTOR(S): Morfología del sujeto (2.85 e-07)

---

| <b>Morfología</b> |                |               |                 |                               |
|-------------------|----------------|---------------|-----------------|-------------------------------|
| <b>factor</b>     | <b>logodds</b> | <b>tokens</b> | <b>VS/V5+SV</b> | <b>centered factor weight</b> |
| Inf_y_SC          | 0.974          | 82            | 0.841           | 0.726                         |
| Suj_lex           | -0.255         | 452           | 0.608           | 0.437                         |
| Suj_pron          | -0.719         | 166           | 0.494           | 0.328                         |

---

| \$misc.1 |    |           |                    |                     |
|----------|----|-----------|--------------------|---------------------|
| n        | df | intercept | overall proportion | centered input prob |
| 700      | 3  | 0.695     | 0.609              | 0.667               |

| \$misc.2       |         |         |       |       |
|----------------|---------|---------|-------|-------|
| log.likelihood | AIC     | AICc    | Dxy   | R2    |
| -453.497       | 912.995 | 913.029 | 0.188 | 0.065 |

---

Tabla 28. Análisis de regresión logística entre orden del sujeto y morfología del sujeto

### 3.3.7. Variable complejidad del sujeto

La regresión de las variables Complejidad del sujeto y Orden del sujeto (Tabla 29) muestra que las estructuras más complejas (categoría C), entre las que se encuentran i) las cláusulas sustantivas; ii) los sintagmas nominales con determinante, coordinados, que contengan además complementos preposicionales o subordinadas adjetivas; y iii) los elementos nominales con una cláusula de relativo y estructuras complejas en general en función de sujeto, exhiben una probabilidad logarítmica de 0.86 y un peso categorial de 0.70. Los sintagmas con complejidad, que podemos caracterizar como ‘media’ (categoría B), es decir i) elementos nominales sin determinante, yuxtapuestos o coordinados; ii) sintagmas nominales con determinante modificados por una frase preposicional, un adjetivo o un adverbio; y iii) elementos nominales sin determinante, acompañados de uno o varios complementos, tienen probabilidades con signo negativo, lo que indica que no favorecen el

orden observado (VS) y que disminuyen la probabilidad de posposición del sujeto (*logodds* de -0,28 y peso de 0.42). Por último, los sintagmas nominales conformados por sustantivos propios, títulos o infinitivos (sin modificadores) y los sustantivos con determinante, es decir, las estructuras más simples (categoría A), tienen la probabilidad más baja (-0.40 y peso de 0.4) de aparición posverbal, lo que significa que son los que más favorecen la anteposición del sujeto. En resumen, a mayor complejidad o extensión del sujeto, mayores son las probabilidades de que aparezca pospuesto.

---

ONE-LEVEL ANALYSIS OF RESPONSE Orden del sujeto WITH PREDICTOR(S): Complejidad del sujeto (6.68e-05)

| <b>Complejidad</b> |  |                |               |                 |                               |
|--------------------|--|----------------|---------------|-----------------|-------------------------------|
| <b>factor</b>      |  | <b>logodds</b> | <b>tokens</b> | <b>VS/VS+SV</b> | <b>centered factor weight</b> |
| C                  |  | 0.693          | 99            | 0.798           | 0.667                         |
| B                  |  | -0.288         | 191           | 0.597           | 0.429                         |
| A                  |  | -0.405         | 410           | 0.568           | 0.4                           |

---

\$misc.1

| n   | df | intercept | overall | proportion | centered | input | prob |
|-----|----|-----------|---------|------------|----------|-------|------|
| 700 | 3  | 0.68      | 0.609   | 0.664      |          |       |      |

\$misc.2

| log.likelihood | AIC     | AICc    | Dxy   | R2    |
|----------------|---------|---------|-------|-------|
| -458.954       | 923.909 | 923.943 | 0.132 | 0.041 |

---

Tabla 29. Análisis de regresión logística entre orden del sujeto y complejidad del sujeto

### 3.3.8. Variable función pragmática del enunciado interrogativo

El análisis estadístico de regresión mostró que la variable función pragmática predice orden del sujeto ( $p= 0.000197$ ). Las interrogativas *subjectio* son las que favorecen de manera más fuerte la posposición del sujeto, dada su probabilidad logarítmica de 0.51 y su peso factorial de 0.62. Las interrogativas problemáticas también son potenciadoras del orden VS, aunque en con una fuerza menor (*logodds* de 0.25 y peso de 0.56). La interrogativa retórica presenta una tendencia neutra, pues su probabilidad de coocurrencia con el orden VS (0.03) está cercana a 0 y su peso factorial está próximo al valor medio (0.49). El valor de

probabilidades para la categoría *pregunta real* indica que esta favorece la anteposición del sujeto, lo cual es un resultado sorprendente, dada la hipótesis teórica de un orden VS como no marcado para este tipo de enunciado. Su peso categorial, sin embargo, no dista mucho de 0.5, por lo que podríamos decir que es un factor que favorece el orden contrario al observado de manera moderada. Las interrogativas agrupadas en la categoría Otras, es decir, las interrogativas no epistémicas, son las que más inhiben el orden VS (tienen una probabilidad de -0.62 y un peso categorial de 0.34), por lo que se interpreta que son francas favorecedoras del orden SV. De manera más sucinta, las interrogativas *subjectio* favorecen la posposición del sujeto, mientras las interrogativas no relacionadas directamente con una duda auténtica la inhiben. Las interrogativas retóricas muestran un comportamiento neutro. Mirados los datos en su conjunto, la pregunta real favorece el orden SV, aunque de manera moderada.

---

ONE-LEVEL ANALYSIS OF RESPONSE Orden.del.sujeto WITH PREDICTOR(S): Función pragmática del enunciado (0.000197)

**Función pragmática**

| factor                         | logodds | tokens | VS/VS+SV | centered factor weight |
|--------------------------------|---------|--------|----------|------------------------|
| Interrogativa <i>subjectio</i> | 0.514   | 120    | 0.742    | 0.626                  |
| Interrogativa problemática     | 0.257   | 87     | 0.690    | 0.564                  |
| N/A                            | 0.206   | 56     | 0.679    | 0.551                  |
| Interrogativa retórica         | -0.030  | 80     | 0.625    | 0.492                  |
| Pregunta real                  | -0.322  | 238    | 0.555    | 0.42                   |
| Otras                          | -0.625  | 119    | 0.479    | 0.349                  |

---

\$misc.1

| n   | df | intercept | overall | proportion | centered | input | prob |
|-----|----|-----------|---------|------------|----------|-------|------|
| 700 | 6  | 0.541     | 0.609   | 0.632      |          |       |      |

\$misc.2

| log.likelihood | AIC     | AICc    | Dxy   | R2    |
|----------------|---------|---------|-------|-------|
| -456.458       | 924.916 | 925.037 | 0.212 | 0.044 |

---

Tabla 30. Análisis de regresión logística entre orden del sujeto y complejidad del sujeto

### 3.3.9. Recapitulación análisis de regresión

En esta segunda fase de análisis estadístico se ha podido dar respuesta con mayor precisión a la pregunta de cuáles son los factores lingüísticos o extralingüísticos que influyen en el orden que adopta el sujeto en la interrogación total.

El análisis de regresión logística permitió conocer las probabilidades de aparición de la variante observada (orden VS) dada cierta configuración de las categorías que se insertan en el modelo de predicción.

La Tabla 31, a continuación, resume las categorías que mejor predicen el orden VS. Según nuestros datos, pudiera afirmarse que los sujetos sin determinante, extensos, constituidos por infinitivos o cláusulas, y con referentes inanimados, que aparezcan en enunciados con valor pragmático de interrogativa *subjectio* cuyos predicados sean existenciales, inacusativos y monovalentes aumentarán considerablemente las posibilidades de que el orden del enunciado interrogativo total sea VS. A estos factores se añade la influencia de la zona geográfica, siendo España la que favorece fuertemente la anteposición (probabilidades de 0.80 y peso categorial de 0.70). Estos resultados confirman los hallazgos comentados en 3.1.6.

| <i>Categoría</i>                                 | <i>Logodds</i> | <i>Peso</i> |
|--|----------------|-------------|
| Sujetos escuetos                                 | 1.02           | 0.73        |
| Verbos inacusativos                              | 1.01           | 0.73        |
| Estructuras de infinitivos y sujetos clausulares | 0.97           | 0.72        |
| Sujetos complejos (extensos)                     | 0.69           | 0.66        |
| Predicados monovalentes                          | 0.68           | 0.66        |
| Procesos existenciales                           | 0.63           | 0.65        |
| Función pragmática int. <i>subjectio</i>         | 0.51           | 0.62        |
| Sujetos inanimados                               | 0.34           | 0.58        |

Tabla 31. Análisis de regresión logística entre orden del sujeto y morfología del sujeto

### 3.4. Resultados regresión logística múltiple

El análisis de regresión simple está encaminado a responder el segundo objetivo estadístico, a saber, determinar cuáles variables (o qué combinación de ellas) predicen mejor la posición del sujeto (con especial atención al orden VS). El análisis de subida y bajada arrojó un resultado coherente, es decir, en ambos recorridos las variables que mejor predicen el orden VS coinciden. Estas son: Transitividad ( $p=4.57 \text{ e-}05$ ), Valor pragmático ( $p=0.000154$ ), Definitud ( $p=0.0165$ ) y Realización valencial ( $p=0.0494$ ).

A continuación, la Tabla 32 presenta los resultados de los valores de probabilidades para las categorías, una vez ejecutada la regresión múltiple con un modelo máximo, es decir, donde interactúan todas las variables lingüísticas (se han omitido las variables no significativas).

| <b>Variables</b>             | <b>Factores</b>            | <b>logodds</b> | <b>tokens</b> | <b>VS/Vs+SV</b> | <b>centered factor weight</b> | <b>vif</b> |
|------------------------------|----------------------------|----------------|---------------|-----------------|-------------------------------|------------|
| <b>Transitividad</b>         | Inacusativo                | 0.953          | 165           | 0.806           | 0.722                         | >10        |
|                              | Transitivo                 | -0.051         | 211           | 0.493           | 0.487                         | >10        |
|                              | Pred. nominal              | -0.334         | 260           | 0.619           | 0.417                         | >10        |
|                              | Inergativo                 | -0.568         | 64            | 0.438           | 0.362                         | >10        |
| <b>Valor pragmático</b>      | <i>Subjectio</i>           | 0.512          | 120           | 0.742           | 0.625                         |            |
|                              | N/A                        | 0.252          | 56            | 0.679           | 0.563                         |            |
|                              | Problemática               | 0.245          | 87            | 0.690           | 0.561                         |            |
|                              | Retórica                   | 0.086          | 80            | 0.625           | 0.521                         |            |
|                              | Pregunta real              | -0.528         | 238           | 0.555           | 0.371                         |            |
|                              | Otras (int. no epistémica) | -0.567         | 119           | 0.479           | 0.362                         |            |
| <b>Definitud</b>             | Escueto                    | 0.879          | 88            | 0.886           | 0.707                         | >2.5       |
|                              | Indefinido                 | -0.186         | 58            | 0.690           | 0.454                         | >2.5       |
|                              | Definido                   | -0.694         | 554           | 0.556           | 0.333                         | >2.5       |
| <b>Realización valencial</b> | Monovalente                | 0.699          | 402           | 0.669           | 0.668                         | >2.5       |
|                              | Bivalente                  | 0.063          | 282           | 0.539           | 0.516                         | >2.5       |
|                              | Trivalente                 | -0.762         | 16            | 0.312           | 0.318                         | >2.5       |

Tabla 32. Resultados de la regresión logística múltiple. Variables estadísticamente significativas

A partir de estos resultados, puede concluirse que el fenómeno de la posposición del sujeto respecto del verbo es determinado, en gran medida, por la interacción de más de una

variable (más concretamente, por el conjunto de categorías que la integran); del mismo modo, se corroboran las probabilidades obtenidas en el análisis de regresión simple para las categorías que integran las variables. Se puede plantear, en base al análisis estadístico, que existe una escala de influencia (de mayor a menor) en el orden antes mencionado, la cual se resume en la figura 3:

Transitividad > Función pragmática > Definitud > Realización valencial

Figura 3. Variables más significativas en la determinación del orden sujeto-verbo. Escala de influencia

Teniendo en cuenta tanto la anteposición como la posposición del sujeto, tal escala de influencia de las categorías podría representarse de la manera siguiente (Figura 4):

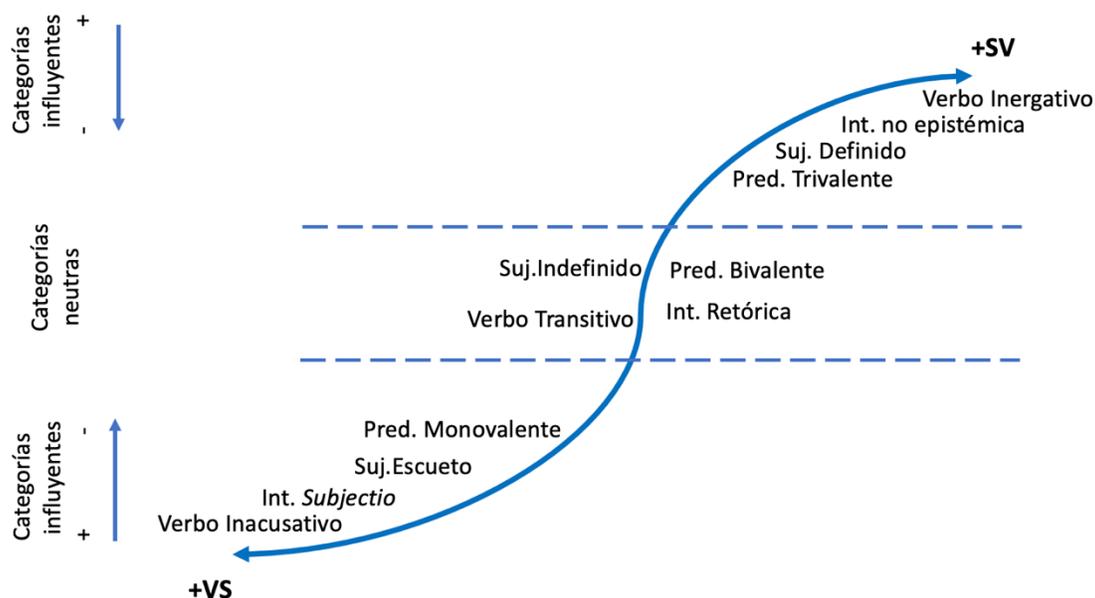


Figura 4. Representación de la influencia de factores lingüísticos en el orden sujeto-verbo

A modo de recapitulación, podrían realizarse los siguientes planteamientos, referentes del orden sujeto-verbo en la interrogación total del español:

(1) en presencia de oraciones con verbos inacusativos y predicados de un solo argumento; con sujetos compuestos por un elemento nominal escueto; y cuya función pragmática sea de interrogativa *subjectio*, el orden de palabras tenderá, típicamente, a ser VS.

(2) en presencia de oraciones con verbos inergativos y predicados trivalentes; con sujetos compuestos por un elemento nominal definido; y cuya función pragmática sea

diferente de la obtención de información (interrogación no epistémica), el orden de palabras tenderá, típicamente, a ser SV.

#### **Capítulo 4. Conclusiones**

La presente investigación tuvo como objetivo comprender el comportamiento del orden adoptado por el sujeto respecto del verbo en enunciados interrogativos totales o absolutos del entorno hispanico. Para ello, se procedió, mediante métodos cuantitativos, a la descripción de un corpus de interrogativas totales de lengua escrita, obtenido del Corpes XXI. Posteriormente, se analizó la relación entre las variables que intervinieron en el estudio y se determinó la influencia de estas, de manera individual y conjunta, en el fenómeno estudiado.

El primer resultado que debe destacarse es que, en términos de frecuencia, el orden predominante en la interrogación absoluta del español es el de VS. Sin embargo, la tendencia a la posposición del sujeto no es categórica, dada la importante proporción de casos con sujetos en posición antepuesta. Una primera conclusión de nuestro estudio es, entonces, que el orden no marcado propuesto por las gramáticas normativas no existe de manera general en el español (aunque, como se verá a continuación, sí puede plantearse para algunas zonas). A este respecto, la interrogación absoluta parece diferenciarse de la parcial, donde el sujeto se presume en una posición fija, a saber, después del verbo; ello, independientemente de cualquier circunstancia discursiva.

También pudo comprobarse que la posición del sujeto muestra, como habíamos hipotetizado, una variación dialectal. El orden VS es mayoritario en España, México y Centroamérica. El área caribeña –considerando las Antillas y el Caribe Continental–, que ha concentrado abundante investigación sobre el orden de palabras, especialmente en el ámbito de la interrogación parcial, no muestra una tendencia uniforme, pues el orden VS es mayoritario en la primera zona, mientras la segunda no tiene preferencia por ninguna de las dos secuencias. No obstante, este resultado varía si se considera el tipo de sujeto como covariable: la tendencia a colocar el sujeto pronominal en la primera posición de la oración es muy clara para las dos zonas dialectales, en comparación con el resto, por lo que el orden prevaleciente en estos casos es el SV. Por consiguiente, y en base a los datos de la presente investigación, se comprueba la hipótesis de la anteposición del sujeto de manera mayoritaria en el área caribeña, siempre que los sujetos sean pronominales.

Las zonas Chile, Andina y Río de la Plata tienen también una proporción mayoritaria de sujetos pospuestos, siendo la última de ellas la que más se acerca a la neutralidad. Los resultados de la regresión completan el panorama dialectal: puede postularse, de manera general, la existencia de un comportamiento más o menos uniforme según la siguiente distribución: España y México y Centroamérica son las zonas donde el sujeto tiende de manera considerable a la posposición, mientras que el Caribe Continental, Río de la Plata y Andina propenden a exhibir sujetos antepuestos. Chile y las Antillas tienen un comportamiento más moderado, es decir, no favorecen de manera fuerte ninguno de los dos órdenes. Estos resultados hablan a favor de una diferencia dialectal marcada entre el español de la península y la región de México y Centroamérica y el sur del continente americano.

El análisis bivariado de las variables lingüísticas de la presente investigación arrojó que existe una relación estadísticamente significativa entre estas y el orden, salvo para la variable tiempo verbal. En efecto, si bien el *presente* es el tiempo que más sujetos antepuestos presenta y el *pretérito imperfecto* el que más favorece el orden VS, no se halló una asociación fuerte (por encima del azar) de ninguna de las dos categorías con el orden del sujeto. Sin embargo, la reagrupación de los datos en tres categorías generales (pasado, presente y futuro), nos permitió intuir que, de contarse con un corpus más equilibrado, podría emerger un resultado diferente, pues el análisis de independencia entre las variables mostró un valor marginalmente significativo.

Respecto de la semántica verbal, los resultados indican que, si bien el orden VS es el más frecuente en la mayoría de los procesos verbales recogidos en el estudio (los verbos de modulación tienen la tendencia contraria), son solo los verbos existenciales los que favorecen fuertemente este orden, según quedó demostrado en el análisis de regresión. Los verbos relacionales y materiales tienen también una alta proporción de sujetos pospuestos, mientras los procesos mentales y verbales muestran una tendencia más atemperada al orden VS, probablemente motivada por la prominencia de sujetos con referente humano. Los resultados obtenidos no difieren de las tendencias encontradas en la literatura para la declaración, por lo que podría afirmarse que, si bien el orden mayoritario de la secuencia VS se debe a la acción de la fuerza ilocutiva, esta no es lo suficientemente influyente, dados los valores importantes de anteposición.

El análisis de la realización valencial del verbo mostró que, en términos de frecuencia, el orden adoptado por el sujeto depende de la cantidad de argumentos con que cuenta el verbo. Así, los sujetos de verbos monovalentes tienden a la posposición más que los de verbos bivalentes; y los verbos trivalentes tienen el sujeto antepuesto de manera mayoritaria. Si bien los casos de este último tipo no son suficientes para afirmar que esta es una tendencia categórica, sí se pudo comprobar mediante el test de independencia que existe una relación estadísticamente significativa entre los predicados monovalentes y los de más de un argumento, teniendo estos últimos una asociación fuerte con el orden SV. El análisis de regresión mostró que, efectivamente, en presencia de un predicado monovalente aumentan las probabilidades de aparición del orden VS.

En lo referente a la transitividad verbal, se pudo comprobar la tesis de un comportamiento diferenciado de los verbos inacusativos e inergativos, y las tendencias encontradas en nuestros datos son similares a las que existen también en las oraciones declarativas: los primeros favorecen fuertemente el orden VS, mientras los segundos lo inhiben. Los sujetos de predicados transitivos tienen, en términos de frecuencia, una ligera tendencia a la anteposición, probablemente debido a la predominancia de referentes humanos. El test de independencia y la regresión mostraron que este tipo de predicado aumenta las probabilidades de aparición del orden SV en la interrogación total.

Respecto de la referencialidad del sujeto, el análisis reveló que el orden adoptado por este constituyente oracional responde a un comportamiento escalar en el caso de la variable animacidad. Así, los sujetos con referente humano (los prototípicamente animados), en lo más alto de la escala, tienden fuertemente a aparecer antepuestos al verbo; y los sujetos inanimados, en el extremo opuesto, propenden a la posposición; en el centro de la escala, los sujetos que contienen referentes animados (no humanos) tienen un comportamiento más atemperado respecto del orden. En lo relativo a la definitud del sujeto, también se comprobó un comportamiento jerárquico: los sujetos tienden a la anteposición en la misma medida en que ellos son más definidos (es decir, los indefinidos se posicionan después del verbo en menor cuantía que los escuetos, y los definidos aumentan considerablemente las probabilidades de aparición del orden SV).

La variable morfología del sujeto presentó un desbalance importante al interior de las categorías. No obstante, pudieron esbozarse algunas tendencias respecto del orden. Los sujetos con pronombre personal *tú* y pronombres indefinidos se anteponen de manera mayoritaria. Por otra parte, la anteposición es más acusada, sin ser mayoritaria, también en los sujetos compuestos por sustantivos comunes o propios. Los infinitivos, por su parte, son los que tienen una mayor proporción de posposición en todo el corpus. El análisis de regresión dio cuenta de una tendencia clara si los datos son distribuidos de una manera más general: los sujetos con estructuras de infinitivo y cláusula aumentan las probabilidades de aparición del orden VS, mientras que los sujetos léxicos y pronominales favorecen el orden contrario, siendo el segundo tipo un franco inhibidor de la posposición del sujeto.

El análisis descriptivo e inferencial arrojó, además, que la complejidad del sujeto puede influir de manera significativa en el orden de constituyentes. En efecto, los sujetos más extensos o pesados, tales como las cláusulas sustantivas y los elementos nominales con varios modificadores son fuertes favorecedores de la posposición del sujeto, mientras que las estructuras más simples, como los sustantivos sin modificadores, y las unidades nominales con un solo modificador aumentan las probabilidades de aparición del sujeto antepuesto. Todo ello aun cuando, en términos de frecuencia, el orden VS es el predominante.

Por último, respecto de la función pragmática del enunciado se observó que son las preguntas reales, las interrogativas *subjectio*, las problemáticas y las retóricas los tipos más abundantes en el corpus. Contrario a los postulados por las gramáticas normativas, las preguntas reales, manifestación prototípica de la interrogación, son las que menos proporción de sujetos pospuestos tienen, por lo que no se comprobó la hipótesis de un orden no marcado para este tipo de enunciado. De hecho, el análisis de regresión puso en evidencia que la pregunta real es la que menos favorece el orden VS, teniendo la probabilidad logarítmica más baja. En el extremo contrario se encuentran las interrogativas *subjectio*, que potencian la posposición del sujeto de manera importante. Estos resultados apuntan, al menos, en dos sentidos: primero, el orden de palabras es un fenómeno complejo y multifactorial, y no se puede afirmar que sea la función pragmática la que determine, en primera instancia, la posición del sujeto; segundo, la coexistencia de los dos patrones de orden en las interrogativas que buscan información requiere de un análisis más específico de la concomitancia de otros factores semióticos no contemplados en esta investigación, lo cual permitirá alcanzar

respuestas más generales. Por último, las interrogativas no epistémicas favorecen considerablemente a la anteposición.

No obstante, como fue indicado en la discusión de resultados, las tendencias encontradas en los datos deben ser comprobadas mediante la comparación con un corpus de lengua hablada, pues el presente estudio, en tanto se concentró en casos de lengua escrita, no pudo contemplar la prosodia como elemento complementario en la determinación de la función pragmática del enunciado.

Los principales resultados antes comentados nos llevan, entonces, a concluir que:

1- En la interrogación total del español, según los datos con que se trabajó en esta investigación, no existe un orden no marcado, sino que se observa un continuo respecto del orden. Los factores que explican la posposición del sujeto son los mismos que influyen, al menos de manera teórica, en la secuenciación VS de la declaración (Delbecque, 2011). Sin embargo, puede plantearse que la fuerza ilocutiva, efectivamente, actúa en favor de este orden, pues es el más frecuente.

2- Si bien existe una correlación estadísticamente significativa entre la mayor parte de las variables lingüísticas que conformaron el estudio, el análisis de regresión múltiple develó que, cuando actúan todas juntas, son la transitividad verbal, la función pragmática del enunciado, la definitud del sujeto y la cantidad de argumentos del verbo (en ese orden) las que influyen de manera significativa en la secuenciación de constituyentes. Dicho de otra manera, la posposición del sujeto se puede predecir, en gran medida, a partir de la coaparición de estas variables.

3- El orden del sujeto respecto del verbo en la interrogación total exhibe una variación dialectal, siendo las zonas de España y México y Centroamérica las que tienden de manera importante a presentar un orden VS. El área caribeña y las zonas del sur del continente americano muestran un comportamiento más atemperado, por lo que puede postularse, para ellas, una transición hacia un orden menos rígido en la interrogación. A partir de los datos con que se contó en la presente investigación, se distinguen, por su comportamiento respecto del orden, tres macrozonas: 1) la Península Ibérica, México y Centroamérica, 2) el Caribe (antillano y continental); y 3) el sur de América Latina.

La presente investigación ha constituido un intento de ofrecer una evidencia empírica para un ámbito de la investigación lingüística virtualmente inexplorado. En efecto, los estudios cuantitativos de corpus de habla real para el tema que nos ocupa son escasos en la lingüística hispánica. La mirada variacionista del fenómeno ha permitido trazar, de manera preliminar, tendencias que deberán ser comprobadas o refutadas con estudios más amplios en el futuro.

Como toda exploración doctoral, la presente tiene limitaciones o debilidades que podrían motivar nuevas investigaciones. Entre ellas podemos destacar las siguientes:

- a) Carencia de comparación empírica entre oraciones declarativas e interrogativas absolutas.

Si bien en nuestro marco teórico se hizo referencia en muchas ocasiones al comportamiento general en la declaración de las variables lingüísticas incluidas en el estudio, la imposibilidad de partir de textos completos para la extracción de casos de análisis, es decir, de obras de las que pudieran obtenerse tanto enunciados declarativos como interrogativos, con fines comparativos, ha hecho que nuestro estudio carezca de un punto de referencia empírico para contrastar el comportamiento de los dos tipos de actos comunicativos. En consecuencia, no se ha podido saber en qué medida actúa la fuerza ilocutiva en los enunciados objeto de análisis de nuestro corpus. No obstante, creemos que se han puesto en claro ciertas tendencias en enunciados interrogativos absolutos que podrán ser comprobadas (o eventualmente refutadas) en un estudio que integre los dos tipos de enunciados.

- b) Desbalance de determinadas categorías en el corpus.

Razones tecnológicas –naturaleza del corpus y herramientas de extracción de casos– y de tiempo –importancia de contar con un corpus acotado que permitiera una descripción lo más detallada posible en los tiempos inicialmente pautados por el proyecto– nos impidieron implementar estrategias de “poblamiento” de algunas categorías con presencia minoritaria en el corpus, vale decir, algunos tiempos verbales, algunas estructuras sintácticas (para la variable extensión del sujeto), algunos tipos de sujeto o tipos funcionales de interrogación. No obstante, debe apuntarse que nuestros datos son un reflejo, en última instancia, de las tendencias generales de uso (el tiempo presente es mucho más frecuente que el pretérito imperfecto de subjuntivo, por poner un ejemplo); a pesar de ello, hubiera sido interesante

poder refutar o comprobar algunas hipótesis planteadas en la literatura científica y que, por falta de datos, quedaron para estudios futuros.

c) Falta de análisis integrado de la sintaxis y entonación del español coloquial.

La entonación es un recurso fundamental en la comunicación. Ella forma parte importante del análisis de los propósitos comunicativos de los hablantes (Escandell, 1999; Pinuer, 2007). Un análisis de la interfaz sintaxis-entonación permite llegar a resultados mucho más ricos, pues en la comunicación ambos niveles funcionan al unísono. Por razones de tiempo y de corpus, fundamentalmente, la presente tesis careció del necesario análisis fonológico de una muestra de lengua oral. Nos vimos impedidos de evaluar, por ejemplo, fenómenos tales como el desgajamiento fonético del Gili Gaya (1958 [1943]) o tematización de la interrogación (Escandell, 1999), o la influencia de la entonación en la discriminación de ciertos tipos de valores pragmáticos en las interrogativas eco (Escandell, 1999). Infelizmente, el tiempo acotado de la investigación impidió extender el foco de análisis a la oralidad. De cualquier modo, se cuenta ya con un corpus conformado (que incluye fragmentos de audio), que será objeto de análisis y futuras publicaciones.

Asimismo, ha sido imposible probar la hipótesis propuesta por Escandell (1999), respecto de los patrones entonativos típicos de oraciones interrogativas con y sin inversión. Al analizar un corpus escrito, la sola existencia de los signos de puntuación ha sido asumida como entonación interrogativa, sin hacer hincapié en las variaciones de ese patrón, por lo que no se pudo medir la correlación entre orden del sujeto y patrón melódico.

Además del desafío que imponen estos vacíos a las investigaciones que la autora de esta tesis pretende emprender en el futuro, se han identificado algunos elementos interesantes para el planteamiento de próximos proyectos.

El primero de ello es la indagación en el fenómeno de la correferencialidad, con el fin de conocer qué efectos tiene la estructura informativa en el orden del sujeto en las interrogativas absolutas. Se sabe que existe una correlación entre información nueva y sujeto pospuesto, por un lado, e información dada y sujeto antepuesto, por el otro. Hay algunos referentes interesantes, tanto para la declaración (Delbecque, 1987) como para la interrogación parcial (Rosemeyer, 2018) que pueden constituir un punto de partida.

El segundo elemento, asociado con el tema anterior, corresponde a la “perseverancia” (*priming*). Se postula que puede existir una influencia entre las formas en una determinada porción de texto. Es decir, podría haber una influencia de las estructuras sintácticas presentes en el contexto inmediato (anterior o posterior) al enunciado y la estructura de la interrogativa absoluta. La interrogativa en negrita del fragmento a continuación es un buen ejemplo de esta conjetura:

- ¿Es alto el costo que implica poner en marcha una desaladora?
- El agua desalada necesita una planta desaladora y después un sistema de bombeo por cañerías. En cambio, una extracción de agua subterránea requiere solamente las cañerías, el bombeo, que es normalmente donde no hay tanta diferencia de altitud. Cuando se comienza a operar una planta desalada, el costo de operación para traer agua de una fuente subterránea es multiplicado por cuatro o cinco para una de salada.
- *¿El agua y la energía podrían pesar en contra de desarrollar algunos proyectos?* (Chi50)

Algunos hallazgos para la interrogación parcial indican que el efecto de perseverancia puede tener una influencia en el orden de palabras (Rosemeyer, 2018).

Por último, y esperando disponer de un corpus con más datos, sería conveniente poder conocer cuáles de los factores estudiados en el presente proyecto son preponderantes en las principales zonas geográficas del español para poder profundizar en las causas de la variación sintáctica experimentada en el orden sujeto-verbo de la interrogación absoluta en el entorno hispánico.

## Bibliografía

- Aijón, M. Á. y Serrano, M. J. (2012). La posición del sujeto pronominal en las cláusulas no declarativas. *Onomázein*, 26, 131-164.
- Aissen, J. (2003). Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory*, 21(3), 435-483.
- Alarcos Llorach, E. (1994). *Gramática de la lengua española*. Espasa Calpe.
- Alarcos Llorach, E. (1970). *Estudios de gramática funcional del español*. Gredos.
- Albertuz, F. J. (2007). Sintaxis, semántica y clases de verbos: clasificación verbal en el proyecto ADESSE en *Actas del VI Congreso de Lingüística General, Santiago de Compostela, 3-7 de mayo de 2004 (2015-2030)*. Arco Libros.
- Alcina, J., & Blecua, J. M. (Eds.). (1980). *Gramática española* (Vol. 10). Ariel.
- Alfonso, M., & Melis, Ch. (2011). La complementación de los verbos intransitivos. *Lingüística*, 25, 9-29.
- Alfonso, M., & Melis, Ch. (2010). La posición del sujeto en la oración intransitiva del español. *Semántica, pragmática y prosodia. Reflejos en el orden de palabras en español. México: El Colegio de México*, 39-68.
- Alonso Cortés, A. (2001). Rasgos activos y ergativos del español. *Verba: Anuario galego de filoloxia*, 28, 187-206.
- Barbosa, P. (2001). On inversion in wh-questions in Romance. En A. Hulk & J.-Y. Pollock (orgs.), *Subject inversion in Romance and the theory of Universal Grammar* (pp.20-59). Oxford Press.
- Batsiukova, O. (2004). Sobre la intransitividad y la estructura subeventiva de los verbos de movimiento. *Interlingüística*, 15, 177-188.
- Belletti, A. (1987). Los inacusativos como asignadores de caso. En Demonte, V., & Lagunilla, M. F. (Eds.). *Sintaxis de las lenguas románicas* (pp.167-230). Ediciones El Arquero.
- Bello, A. (1984). *Gramática de la lengua castellana*. Editorial EDAF, S.A.

- Bentivoglio, P., & Weber, E. G. (1986). A Functional Approach to Subject Word Order in Spoken Spanish. En O. Jaeggli & C. Silva-Corvalán (eds.): *Studies in Romance Linguistics* (pp. 23-40). De Gruyter Mouton.
- Bentivoglio, P. (2003). Orden de palabras en español: un análisis sintáctico-semántico-pragmático del sujeto. *Lexis*, 27(1-2), 235-261.
- Bogard, S. (2019). Los sujetos transitivos e intransitivos con núcleo nominal común en español. Estructura y evolución (siglos XIII, XVI y XIX). *Anuario de Letras: Lingüística y filología*, 7(1), 5-59.
- Bogard, S. (2017). Oraciones de complemento en español. Tipos, estructura y función. *Boletín de filología*, 52(1), 11-45.
- Bolinger, D. L. (1954). Meaningful word order in Spanish. *Boletín de Filología*, 8, 45-56.
- Bosque, I. & Gutiérrez-Rexach, J. (2009). *Fundamentos de sintaxis formal*. Ediciones Akal.
- Breñes, E. (2011). La personalización del enunciado. Análisis pragmlingüístico del pronombre personal sujeto en la oralidad. En J.J. Bustos Tovar, R. Cano Aguilar, E. Méndez García de Paredes y A. López Serena (Coords.), *Sintaxis del discurso hablado en español. Homenaje a Antonio Narbona* (Vol.1, pp. 121-136). Sevilla: Grupo de Investigación El español hablado en Andalucía, Secretariado de Publicaciones, Universidad de Sevilla.
- Brown, E. L. y Rivas, J. (2011). Subject-verb word order in Spanish interrogatives: A quantitative analysis of Puerto Rican Spanish. *Spanish in context*, 8 (1), 23-49.
- Brown, E. L. y Rivas, J. (2011a). Correlaciones entre forma y función en las construcciones interrogativas parciales del español de Puerto Rico. *ELUA*, 25, 289-315.
- Bürki, Y. (2010). Mecanismos argumentativos en textos de opinión en la prensa judeoespañola de entre siglos (ss. XIX-XX). En Iliescu, M., Siller-Runggaldier, H. & Danler, P. (Eds.), *Actes du XXVe Congrès International de Linguistique et de Philologie Romanes* (Vol. 5, pp. 357-365). De Gruyter.
- Cameron, R., & Flores-Ferrán, N. (2004). Perseveration of subject expression across regional dialects of Spanish. *Spanish in Context*, 1(1), 41-65.
- Cano-Aguilar, R. (1981). *Estructuras sintácticas transitivas en el español actual*. Gredos.

- Campos, H. (1999). Transitividad e intransitividad. En Bosque, I. y Demonte, V. (Eds.). *Gramática descriptiva de la Lengua Española* (Vol. 2, pp. 1519-1574). Espasa Calpe.
- Contreras, H. (1978). *El orden de palabras en español*. Cátedra.
- Cortés, F. J., González, C., & Jiménez, R. (2012). Las clases léxicas. Revisión de la tipología de predicados verbales. En *El funcionalismo en la teoría lingüística: la Gramática del Papel y la Referencia: Introducción, avances y aplicaciones*, 59-84.
- Casado Velarde, M. (1993). *Introducción a la gramática del texto del español* (Vol. 13). Arco libros.
- Crous, B., & Gràcia, L. (2015). Verbos de movimiento, cambios de estado y categorización de afecciones. *RLA. Revista de lingüística teórica y aplicada*, 53(1), 149-174.
- Cuervo, M. C., & Roberge, Y. (Eds.). (2012). *The End of Argument Structure*. Emerald Group Publishing Limited.
- Daleszyska, A. (2016). Analysing linguistic variation with Rbrul—a step-by-step guide. <https://danielezrajohnson.com>.
- De Miguel, E. (2004). Qué significan aspectualmente algunos verbos y qué pueden llegar a significar. *ELUA. Estudios de Lingüística*. Anexo 2 (2004), 167-206.
- De Miguel, E. (1989). Sujetos invertidos en las construcciones ergativas del castellano: Los conceptos de tema y rema. En *Lenguajes naturales y lenguajes formales: actas del IV congreso de lenguajes naturales y lenguajes formales: Lleida, 19-23 de septiembre de 1988* (pp. 753-766). Promociones y Publicaciones Universitarias, PPU.
- Delbecque, N. (2011). El análisis de corpus al servicio de la gramática cognoscitiva: Hacia una interpretación de la alternancia lineal SV/VS. En Knauer, G., & von Colbe, V. B. (Eds.). *Variación sintáctica en español* (Vol. 494, pp. 51-74). De Gruyter.
- Delbecque, N. (1988). On subject position in Spanish: A variable rule analysis of constraints at the level of the subject NP and of the VP. *Literary and Linguistic Computing*, 3(3), 185-201.
- Delbecque, N. (1987). *Problèmes et méthodes de l'étude de la variation syntaxique: le cas de la position du sujet en espagnol* (Vol. 4). Leuven University Press.

- Delbecque, N., & Lamiroy, B. (1999). La subordinación sustantiva: las subordinadas enunciativas en los complementos verbales. En Bosque, I. y Demonte, V. (Eds.). *Gramática descriptiva de la Lengua Española* (Vol2, pp. 1965-2083). Espasa-Calpe, S.A.
- Demonte, V. (2002). Preliminares de una clasificación léxico-sintáctica de los predicados verbales del español. En Sybille Grobe et al., *Ex oriente lux. Festschrift für Eberhard Gärtner zu seinem 60 Geburtstag* (pp. 121-146). Valentia.
- Dickey, M. W., Wainelboim, A. J., & París, L. (2015). The thematic hierarchy in sentence comprehension: A study on the interaction between verb class and word order in Spanish. *The Quarterly Journal of Experimental Psychology*, 68(10), 1981-2007.
- Dixon, R. M. (1979). Ergativity. *Language*, 59-138.
- Du Bois, J. W. (1980). Beyond definiteness: The trace of identity in discourse. *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*, 3, 203-274.
- Escandell, M. V. (2000). Sintaxis y uso interpretativo. En K.Korta y F. García Murga (comps.): *Palabras. Víctor Sánchez de Zavala in memoriam* (pp. 219-259). Servicio de Publicaciones de la Universidad del País Vasco.
- Escandell, M.V. (1999). Los enunciados interrogativos. Aspectos semánticos y pragmáticos. En Bosque, I. y Demonte, V. (eds.). *Gramática descriptiva de la Lengua Española* (Vol3). Espasa-Calpe, S.A.
- Escandell, M. V. (1988). *La interrogación en español: semántica y pragmática*. Memoria para optar por el grado de Doctor. ISBN:978-84-692-0112-1.
- Fernández de Castro, F. (2003). El lugar de las perífrasis verbales en la descripción de las lenguas: los verbos auxiliares y la determinación del verbo. En Pusch, C. D. (2003). *Verbalperiphrasen in den (ibero-) romanischen Sprachen / Perífrasis verbales en las lenguas (ibero-)románicas* (pp.11-22). Helmut Buske.
- Fernández de Castro, F. (1999). *Las perífrasis verbales en el español actual*. Gredos.
- Fernández Ramírez, S. (1986). *Gramática española. El verbo y la oración*, 4. Arco Libros, S.A.

- Fernández Soriano, O. (1993). Sobre el orden de palabras en español. *Dicenda. Cuadernos de filología hispánica*, 113-152.
- Flores-Ferrán, N. (2007). A bend in the road: Subject personal pronoun expression in Spanish after 30 years of sociolinguistic research. *Language and Linguistics Compass*, 1(6), 624-652.
- Flores-Ferrán, N. (2002). *Subject personal pronouns in Spanish narratives of Puerto Ricans in New York City: A sociolinguistic perspective*. Lincom Europa.
- Fox, B. A. (1995). The category 'S' in English conversation. *Discourse grammar and typology*, 153-178.
- Freed, A. F. (1994). The form and function of questions in informal dyadic conversation. *Journal of Pragmatics*, 21(6), pp. 621-644.
- García Riverón, R. (1987). *La interrogación*. Editorial Científico-Técnica.
- García-Miguel, J. M. (2005). Verbos aspectuales en español. La interacción de significado verbal y significado construccional. *Estudos en homenagem ao Professor Doutor Mário Vilela* (pp. 405-418). Faculdade de Letras da Universidade do Porto.
- García-Miguel, J. M. (1995). *Transitividad y complementación preposicional en español* (Vol. 40). Universidade de Santiago de Compostela.
- García, J. F. (2008). Estudio diacrónico de la oración interrogativa en varios lingüistas. *Revista Copérnico Arbitrada interdisciplinaria*, 4(8), 11-26.
- Gattei, C. A., Dickey, M. W., Wainelboim, A. J., & París, L. (2015). The thematic hierarchy in sentence comprehension: A study on the interaction between verb class and word order in Spanish. *The Quarterly Journal of Experimental Psychology*, 68(10), 1981-2007.
- Genta, F. (2008). *Perífrasis verbales en español: focalización aspectual, restricción temporal y rendimiento discursivo*. Tesis doctoral. Universidad de Granada. <http://hdl.handle.net/10481/2032>
- Gili Gaya, S. (1958 [1943]). *Curso superior de sintaxis española* (6ta ed.). Publicaciones y ediciones SPES, S.A.
- Givón, T. (1995). *Functionalism and grammar*. John Benjamins Publishing.

- Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. *Universals of language*, 2, 73-113.
- Gutiérrez, R. (2008). La identificación de los tópicos y los focos. *Nueva revista de filología hispánica*, 363-401.
- Halliday, M. A. K., & Matthiessen, C. M. (2013). *Halliday's introduction to functional grammar*. Routledge.
- Hatcher, Ann G. (1956). *Theme and Underlying Question. Two Studies of Spanish Word Order*. The Linguistic Circle of New York.
- Hernando, L. A. (2005). El orden de palabras en español. *Revista de filología de la Universidad de La Laguna*, 23, 161-178.
- Hernanz, M. (1999). El infinitivo. En Bosque, I. y Demonte, V. (Eds.). *Gramática descriptiva de la Lengua Española*. Vol 2. Espasa-Calpe, S.A.
- Hornik, K., Zeileis, A., & Meyer, D. (2006). The strucplot framework: visualizing multi-way contingency tables with vcd. *Journal of Statistical Software*, 17(3), 1-48.
- Ibáñez, S., & Melis, C. (2016). Ambivalencia transitiva y estructura argumental: resultados de un estudio de uso. *Anuario de Letras. Lingüística y Filología*, 3(2), 153-198.
- Ji, J., & Liang, M. (2018). An animacy hierarchy within inanimate nouns: English corpus evidence from a prototypical perspective. *Lingua*, 205, 71-89.
- Johnson, D. E. (2009). Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and linguistics compass*, 3(1), 359-383.
- Kahane, H., & Kahane, R. (1950). The position of the actor expression in colloquial Mexican Spanish. *Language*, 26(2), 236-263.
- Laca, B. (1999). Presencia y ausencia de determinante. En Bosque, I. y Demonte, V. (Eds.). *Gramática descriptiva de la Lengua Española* (Vol. 1, pp. 891-928). Espasa-Calpe, S.A.
- Lantolf, J. P. (1980). Constraints on interrogative word order in Puerto Rican Spanish. *Bilingual Review/La Revista Bilingüe*, 7(2), 113-122.

- Lapesa, R. (1992). La interpolación del sujeto en las oraciones interrogativas. En *Actas del II Congreso Internacional de Historia de la Lengua española* (pp. 545-554). Pabellón de España.
- Lavandera, B. (1975). *Linguistic Structure and Sociolinguistic Conditioning in the Use of Verbal endings in si-clauses:(Buenos Aires Spanish)*. University Microfilms.
- Leal, E. (2005). *El orden de las palabras*. Liceus, Servicios de Gestión y Comunicación, SL.
- Le Guen, M. (2003). Tableaux croisés et Diagrammes en Mosaïque, pour Visualiser les probabilités marginales et conditionnelles. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, 77(1), 62-79.
- Lenz, R. (1944 [1920]). *La oración y sus partes* (4ta ed.). Editorial Nascimento.
- Leonetti, M. (2018). Two types of postverbal subject. *Italian Journal of Linguistics*, 30(2), 11-36.
- Leonetti, M. (2017). 24. Basic constituent orders. In *Manual of Romance morphosyntax and syntax* (pp. 887-932). De Gruyter.
- Leonetti, M. (2014). Gramática y pragmática en el orden de palabras. *Linred: lingüística en la Red*, 12, 1-25.
- Leonetti, M. (1999). El artículo. En Bosque, I. y Demonte, V. (Eds.). *Gramática descriptiva de la lengua española* (Vol. 1, pp. 787-890) Espasa Calpe, S.A.
- Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. University of Chicago press.
- Levshina, N. (2015). *How to do linguistics with R: Data exploration and statistical analysis*. John Benjamins Publishing Company
- Liceras, J. M. (1994). La teoría sintáctica y los juicios de gramaticalidad: la posposición del sujeto en español. *Revista canadiense de estudios hispánicos*, 219-255.
- Lipski, J. M. (1977). Preposed subjects in questions: Some considerations. *Hispania*, 60(1), 61-67.
- Lizárraga, G. Z., y Mora-Bustos, A. (2010). Variación en la marcación diferenciada de objeto en español. *Forma y Función*, 23(1), 9-38.

- López Meirama, B. (2006). Semantic and discourse-pragmatic factors in Spanish word order. In *Functional Approaches to Spanish Syntax* (pp. 7-52). Palgrave Macmillan.
- López Meirama, B. (1997). *La posición del sujeto en la cláusula monoactancial en español*. Tesis Doctoral, Universidad de Santiago de Compostela.
- López Meirama, B. (1997a). Aportaciones de la tipología lingüística a una gramática particular: el concepto orden básico y su aplicación al castellano. *Verba*, 24, 45-82.
- Martín Butragueño, P. (1994). Hacia una tipología de la variación gramatical en sociolingüística del español. *Nueva Revista de Filología Hispánica*, 42(1), 29-75.
- Melis, Ch. & Alfonso, M. (2013). Circunstancias, sujetos y orden de palabras en la oración intransitiva del español. *Lingüística*, 29(2), 127-151.
- Melis, C., & Flores, M. (2007). Los verbos pseudo-impersonales del español: una caracterización semántico-sintáctica. *Verba*, 34, 7-57.
- Menéndez Pidal, R. (9na. Ed.). (1952). *Manual de gramática histórica española*. Espasa-Calpe, S.A.
- Mendikoetxea, A. (1999). Construcciones inacusativas y pasivas. En Bosque, I. y Demonte, V. (Eds.). *Gramática descriptiva de la Lengua Española* (Vol. 2, pp. 1575-1630). Espasa-Calpe, S.A.
- Mendikoetxea, A. (1999a). Construcciones con "se": medias, pasivas e impersonales. En Bosque, I. y Demonte, V. (eds.). *Gramática descriptiva de la Lengua Española* (Vol. 2, pp. 1635-1722). Espasa-Calpe, S.A.
- Morales, A. (1997). La hipótesis funcional y la aparición de sujeto no nominal: el español de Puerto Rico. *Hispania*, 153-165.
- Moreno-Fernández, F. (2018). Dialectología:(Dialectology). En *The Routledge Handbook of Spanish Language Teaching* (pp. 377-390). Routledge.
- Moreno-Fernández, F. (1994). Sociolingüística, estadística e informática. *Lingüística*, (6), 95-154.
- Navarro Tomás, T. (1966). *Manual de entonación española*. (3ra.ed.) Colección Málaga, SA.

- Nebrija, E. A. D. (1944). Gramática castellana.
- Núñez, R. A. (1983). Pérdida de trasposición de sujeto en interrogativas pronominales del español del Caribe. *Thesaurus. Boletín del Instituto Caro y Cuervo*, 38(1), 35-58.
- Ocampo, F. (1995). The word order of two-constituent constructions in spoken Spanish. *Word order in discourse*, 30, 425-447.
- Ocampo, F. (1991). Word order in constructions with a one-valency verb, a subject NP and a PP in spoken Rioplatense Spanish. *Hispania*, 74(2), 409-416.
- Ordóñez, F. y Olarrea, A (2006). Microvariation in Caribbean/non Caribbean Spanish Interrogatives. *Probus*, 18, 59-97.
- Ortiz-López, L. (2009). El español del Caribe: orden de palabras a la luz de la interfaz léxico-sintáctica y sintáctico-pragmática. *Revista internacional de lingüística iberoamericana*, 75-93.
- Padilla, X. A. (2005). *Pragmática del orden de palabras*. Publicaciones de la Universidad de Alicante.
- Padilla, X. A. (2001). *El orden de palabras en el español coloquial*. Unpublished doctoral dissertation, Universitat de Valencia. <http://www.cervantesvirtual.com/obra/el-orden-de-palabras-en-el-espanol-coloquial-0/>
- Pavey, E. L. (2010). *The structure of language: An introduction to grammatical analysis*. Cambridge University Press.
- Pericchi (2011). *Posición del sujeto en cláusulas con verbos pseudo-impersonales de reacción psicológica*. Tesis de Maestría. Universidad Central de Venezuela.
- Perlmutter, D. M. (1978). Impersonal passives and the unaccusative hypothesis. En *Annual meeting of the Berkeley Linguistics Society* (Vol. 4, pp. 157-190).
- Pinuer, C. (2007). Análisis sintáctico funcional: principios, perspectivas y casos. *Estudios filológicos*, 42, 175-190.
- Rabanales, A. (1996). La estructura gramatical del enunciado interrogativo español, *Boletín de la Academia Chilena de la Lengua*, 71, 237-67.

Ramos, D. (2013). Estructura sintáctica y esquemas entonativos en enunciados interrogativos absolutos del español hablado en Cuba. *Rastros Rostros*, 15 (29), 59-72.

Real Academia Española (2013). Corpus del español del siglo XXI (CORPES). Descripción del sistema de codificación. Libros y prensa. [Recurso de Internet]. Madrid: Real Academia Española, en [https://www.rae.es/sites/default/files/CORPES\\_Sistema\\_de\\_codificacion\\_12\\_2015.pdf](https://www.rae.es/sites/default/files/CORPES_Sistema_de_codificacion_12_2015.pdf) [consultado el 6 de diciembre de 2019].

Real Academia Española y Asociación de Academias de la Lengua Española (ASALE) (2009). *Nueva gramática de la lengua española*, I. Espasa Libros.

Real Academia Española (1977). *Esbozo de una nueva gramática de la lengua española*. Espasa-Calpe, S.A.

Real Academia Española (1959). *Gramática de la lengua española. Nueva edición reformada, de 1931*. Espasa-Calpe, S.A.

Real Academia Española (1931). *Gramática de la lengua española, nueva edición, reformada, de 1931*. Espasa-Calpe.

Real Academia Española (s/f). Banco de datos (CORPES XXI) [en línea]. *Corpus del español del siglo XXI*. Recuperado el 13 de octubre de 2016 en <<http://www.rae.es>>.

Rivas, J. (2013). Variable subject position in main and subordinate clauses in Spanish: a usage-based approach. *Moenia*, 19, 97-113.

Rivas, J. (2008). La posición del sujeto en las construcciones monoactanciales del español: Una aproximación funcional. *Hispania*, 91(4), 897-912.

Roca Pons, J. (1960). *Introducción a la gramática*. Vergara editorial.

Rodríguez, R. (2008). Del sujeto pospuesto al cambio de orden de palabras en español (Estudio diacrónico basado en textos de los S. XIII y XX). *Avances Cuadernos de Trabajo*, 183, 1-19.

Rosemeyer, M. (2018). El orden de palabras en las interrogativas-Q. Un análisis contrastivo del español caribeño y portugués brasileño. *Revista Internacional de Lingüística Iberoamericana*, 16(2), 135-148.

Sánchez, M. E. (2008). Tipos de cláusula, clases verbales y posición del sujeto en español. *Lexis*, 32(1), 83-105.

Seco, R. (1978). *Manual de gramática española*. Aguilar.

Serra, S. (2010). Estructura argumental y definición lexicográfica. De cómo definir los predicados verbales en un diccionario semasiológico de español. *Lenguas Modernas*, 35, 71-87.

Serra, S. (s/f). Estructura gramatical del enunciado interrogativo indagativo en el español culto de Santiago de Chile. Tesis para optar al grado de Magíster. Universidad de Chile. En [https://www.academia.edu/30057191/Estructura\\_gramatical\\_del\\_enunciado\\_interrogativo\\_en\\_lengua\\_espa%C3%B1ola](https://www.academia.edu/30057191/Estructura_gramatical_del_enunciado_interrogativo_en_lengua_espa%C3%B1ola)

Serrano, M. J. (2009). Cronología del estudio de la variación sintáctica en español: pasado y presente. *Revista de Filología*, 27, 155-170.

Serrano, M. J. (2007). Historia que ya es historia: evolución y actualidad del concepto y la metodología de la variación sintáctica. *Boletín de Lingüística*, 19(28), 102-127.

Siewierska, A. (1988). *Word order rules*. Croom Helm.

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

Silva-Corvalán, C. (1982). Subject expression and placement in Mexican-American Spanish (pp. 93-120). En J. Amaste, & E. Olivares (eds.), *Spanish in the United States: Sociolinguistic aspects*. Cambridge University Press.

Sinclair, A., & Van Gessel, R. (1990). The form and function of questions in children's conversations. *Journal of Pragmatics*, 14(6), 923-944.

Suárez, M. (2016). Las propiedades del sujeto y la inacusatividad de los verbos de movimiento en castellano medieval. *Itinerarios: revista de estudios lingüísticos, literarios, históricos y antropológicos*, 23, 155-174.

Suñer, M. (1982). *Syntax and semantics of syntaxis presentational sentences-types*. Georgetown University Press.

Suñer, M. (1994). V-movement and the licensing of argumental wh-phrases in Spanish. *Natural Language & Linguistic Theory*, 12(2), 335-372.

Tesnière, L. (1959). *Éléments de syntaxe structurale*. Klincksieck.

Toribio, A. J. (2000). Setting parametric limits on dialectal variation in Spanish. *Lingua*, 110 (5), 315-341.

Urra, P. (2010). *El orden de constituyentes en construcciones con verbos de alternancia causativa en una muestra de la lengua española*. Tesis de Magíster en Lingüística mención Lengua Española. Universidad de Chile.

Van Valin Jr, R. D. (1990). Semantic parameters of split intransitivity. *Language*, 221-260.

Villalba, X. (2019). *El orden de palabras en contraste*. Arcos Libros.

Von Stechow, K., & Kaiser, G. A. (2003). The interaction of animacy, definiteness and specificity in Spanish. En *Proceedings of the Workshop: Semantic and Syntactic Aspects of Specificity, Romance Languages* (pp. 41-65).

## Anexos

### Tablas de contingencia

| <i>Orden del sujeto</i> | <i>Zona geográfica</i> |                        |                        |                        |                        |                        |                        | <i>Total</i>           |
|-------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
|                         | Andina                 | Antillas               | Car. Cont.             | Chile                  | España                 | Méx. Y Centroam.       | Río de la Plata        |                        |
| SV                      | 43<br>15.7 %<br>43 %   | 38<br>13.9 %<br>38 %   | 50<br>18.2 %<br>50 %   | 40<br>14.6 %<br>40 %   | 22<br>8 %<br>22 %      | 34<br>12.4 %<br>34 %   | 47<br>17.2 %<br>47 %   | 274<br>100 %<br>39.1 % |
| VS                      | 57<br>13.4 %<br>57 %   | 62<br>14.6 %<br>62 %   | 50<br>11.7 %<br>50 %   | 60<br>14.1 %<br>60 %   | 78<br>18.3 %<br>78 %   | 66<br>15.5 %<br>66 %   | 53<br>12.4 %<br>53 %   | 426<br>100 %<br>60.9 % |
| <b>Total</b>            | 100<br>14.3 %<br>100 % | 700<br>100 %<br>100 %  |

$$\chi^2=21.697 \cdot df=6 \cdot \text{Cramer's } V=0.176 \cdot p=0.001$$

| <i>Orden del sujeto</i> | <i>Semántica verbal</i> |                       |                        |                      |                         |                       | <i>Total</i>           |
|-------------------------|-------------------------|-----------------------|------------------------|----------------------|-------------------------|-----------------------|------------------------|
|                         | existencial             | material              | mental                 | modulacion           | relacional              | verbal                |                        |
| SV                      | 22<br>8 %<br>25.3 %     | 38<br>13.9 %<br>40 %  | 49<br>17.9 %<br>43 %   | 9<br>3.3 %<br>75 %   | 146<br>53.3 %<br>39.5 % | 10<br>3.6 %<br>45.5 % | 274<br>100 %<br>39.1 % |
| VS                      | 65<br>15.3 %<br>74.7 %  | 57<br>13.4 %<br>60 %  | 65<br>15.3 %<br>57 %   | 3<br>0.7 %<br>25 %   | 224<br>52.6 %<br>60.5 % | 12<br>2.8 %<br>54.5 % | 426<br>100 %<br>60.9 % |
| <b>Total</b>            | 87<br>12.4 %<br>100 %   | 95<br>13.6 %<br>100 % | 114<br>16.3 %<br>100 % | 12<br>1.7 %<br>100 % | 370<br>52.9 %<br>100 %  | 22<br>3.1 %<br>100 %  | 700<br>100 %<br>100 %  |

$$\chi^2=14.607 \cdot df=5 \cdot \text{Cramer's } V=0.144 \cdot p=0.012$$



| <i>Orden del sujeto</i> | <i>Realización valencial</i> |                         |                      | <i>Total</i>           |
|-------------------------|------------------------------|-------------------------|----------------------|------------------------|
|                         | bivalentes                   | monovalentes            | trivalentes          |                        |
| SV                      | 130<br>47.4 %<br>46.1 %      | 133<br>48.5 %<br>33.1 % | 11<br>4 %<br>68.8 %  | 274<br>100 %<br>39.1 % |
| VS                      | 152<br>35.7 %<br>53.9 %      | 269<br>63.1 %<br>66.9 % | 5<br>1.2 %<br>31.2 % | 426<br>100 %<br>60.9 % |
| <b>Total</b>            | 282<br>40.3 %<br>100 %       | 402<br>57.4 %<br>100 %  | 16<br>2.3 %<br>100 % | 700<br>100 %<br>100 %  |

$\chi^2=17.810 \cdot df=2 \cdot \text{Cramer's } V=0.160 \cdot p=0.000$

| <i>Orden del sujeto</i> | <i>Transitividad</i>    |                        |                         |                         | <i>Total</i>           |
|-------------------------|-------------------------|------------------------|-------------------------|-------------------------|------------------------|
|                         | Inacusativo             | Inergativo             | Predicado nominal       | Transitivo              |                        |
| SV                      | 32<br>11.7 %<br>19.4 %  | 36<br>13.1 %<br>56.2 % | 99<br>36.1 %<br>38.1 %  | 107<br>39.1 %<br>50.7 % | 274<br>100 %<br>39.1 % |
| VS                      | 133<br>31.2 %<br>80.6 % | 28<br>6.6 %<br>43.8 %  | 161<br>37.8 %<br>61.9 % | 104<br>24.4 %<br>49.3 % | 426<br>100 %<br>60.9 % |
| <b>Total</b>            | 165<br>23.6 %<br>100 %  | 64<br>9.1 %<br>100 %   | 260<br>37.1 %<br>100 %  | 211<br>30.1 %<br>100 %  | 700<br>100 %<br>100 %  |

$\chi^2=46.855 \cdot df=3 \cdot \text{Cramer's } V=0.259 \cdot p=0.000$

| <i>Tiempo verbal</i>    |                     |                       |                     |                       |                       |                       |
|-------------------------|---------------------|-----------------------|---------------------|-----------------------|-----------------------|-----------------------|
| <i>Orden del sujeto</i> | Cond. Comp.         | Condic. Simple        | Futuro Comp.        | Futuro Simple         | presente              | Pret. Imperf.         |
| SV                      | 0<br>0 %<br>0 %     | 12<br>4.4 %<br>32.4 % | 0<br>0 %<br>0 %     | 15<br>5.5 %<br>28.8 % | 203<br>74.1 %<br>42 % | 8<br>2.9 %<br>24.2 %  |
| VS                      | 1<br>0.2 %<br>100 % | 25<br>5.9 %<br>67.6 % | 1<br>0.2 %<br>100 % | 37<br>8.7 %<br>71.2 % | 280<br>65.7 %<br>58 % | 25<br>5.9 %<br>75.8 % |
| <b>Total</b>            | 1<br>0.1 %<br>100 % | 37<br>5.3 %<br>100 %  | 1<br>0.1 %<br>100 % | 52<br>7.4 %<br>100 %  | 483<br>69 %<br>100 %  | 33<br>4.7 %<br>100 %  |

$\chi^2=9.167 \cdot df=9 \cdot \text{Cramer's } V=0.114 \cdot \text{Fisher's } p=0.399$

| <i>Tiempo verbal</i>    |                      |                       |                     |                        |                        |
|-------------------------|----------------------|-----------------------|---------------------|------------------------|------------------------|
| <i>Orden del sujeto</i> | Pret. Imperf. Subj.  | Pret. Perf. Com.      | Pret. Plusc. Subj.  | Pretérito Perf. Simple | <b>Total</b>           |
| SV                      | 1<br>0.4 %<br>33.3 % | 9<br>3.3 %<br>40.9 %  | 2<br>0.7 %<br>40 %  | 24<br>8.8 %<br>38.1 %  | 274<br>100 %<br>39.1 % |
| VS                      | 2<br>0.5 %<br>66.7 % | 13<br>3.1 %<br>59.1 % | 3<br>0.7 %<br>60 %  | 39<br>9.2 %<br>61.9 %  | 426<br>100 %<br>60.9 % |
| <b>Total</b>            | 3<br>0.4 %<br>100 %  | 22<br>3.1 %<br>100 %  | 5<br>0.7 %<br>100 % | 63<br>9 %<br>100 %     | 700<br>100 %<br>100 %  |

$\chi^2=9.167 \cdot df=9 \cdot \text{Cramer's } V=0.114 \cdot \text{Fisher's } p=0.399$

| <i>Orden del sujeto</i> | <i>Animacidad</i>    |                         |                         | <i>Total</i>           |
|-------------------------|----------------------|-------------------------|-------------------------|------------------------|
|                         | animado              | humano                  | inanimado               |                        |
| SV                      | 20<br>7.3 %<br>37 %  | 118<br>43.1 %<br>52.4 % | 136<br>49.6 %<br>32.3 % | 274<br>100 %<br>39.1 % |
| VS                      | 34<br>8 %<br>63 %    | 107<br>25.1 %<br>47.6 % | 285<br>66.9 %<br>67.7 % | 426<br>100 %<br>60.9 % |
| <b>Total</b>            | 54<br>7.7 %<br>100 % | 225<br>32.1 %<br>100 %  | 421<br>60.1 %<br>100 %  | 700<br>100 %<br>100 %  |

$\chi^2=25.078 \cdot df=2 \cdot \text{Cramer's } V=0.189 \cdot p=0.000$

| <i>Orden del sujeto</i> | <i>Definitud</i>        |                        |                      | <i>Total</i>           |
|-------------------------|-------------------------|------------------------|----------------------|------------------------|
|                         | definido                | escueto                | indefinido           |                        |
| SV                      | 246<br>89.8 %<br>44.4 % | 10<br>3.6 %<br>11.4 %  | 18<br>6.6 %<br>31 %  | 274<br>100 %<br>39.1 % |
| VS                      | 308<br>72.3 %<br>55.6 % | 78<br>18.3 %<br>88.6 % | 40<br>9.4 %<br>69 %  | 426<br>100 %<br>60.9 % |
| <b>Total</b>            | 554<br>79.1 %<br>100 %  | 88<br>12.6 %<br>100 %  | 58<br>8.3 %<br>100 % | 700<br>100 %<br>100 %  |

$\chi^2=36.546 \cdot df=2 \cdot \text{Cramer's } V=0.228 \cdot p=0.000$

| <i>Orden del sujeto</i> | <i>Morfología del sujeto</i> |                     |                      |                       |                     |                       |                      |                       | Nombre o título        |
|-------------------------|------------------------------|---------------------|----------------------|-----------------------|---------------------|-----------------------|----------------------|-----------------------|------------------------|
|                         | 1 PP                         | 1 PS                | 2 PP                 | 2 PS                  | 3 PS                | Dem.                  | Indef.               | Inf.                  |                        |
| SV                      | 0<br>0 %<br>0 %              | 1<br>0.4 %<br>50 %  | 5<br>1.8 %<br>62.5 % | 16<br>5.8 %<br>76.2 % | 1<br>0.4 %<br>100 % | 21<br>7.7 %<br>39.6 % | 11<br>4 %<br>68.8 %  | 11<br>4 %<br>14.7 %   | 30<br>10.9 %<br>43.5 % |
| VS                      | 2<br>0.5 %<br>100 %          | 1<br>0.2 %<br>50 %  | 3<br>0.7 %<br>37.5 % | 5<br>1.2 %<br>23.8 %  | 0<br>0 %<br>0 %     | 32<br>7.5 %<br>60.4 % | 5<br>1.2 %<br>31.2 % | 64<br>15 %<br>85.3 %  | 39<br>9.2 %<br>56.5 %  |
| <b>Total</b>            | 2<br>0.3 %<br>100 %          | 2<br>0.3 %<br>100 % | 8<br>1.1 %<br>100 %  | 21<br>3 %<br>100 %    | 1<br>0.1 %<br>100 % | 53<br>7.6 %<br>100 %  | 16<br>2.3 %<br>100 % | 75<br>10.7 %<br>100 % | 69<br>9.9 %<br>100 %   |

| <i>Orden del sujeto</i> | <i>Morfología del sujeto</i> |                      |                               |                            |                       | <b>Total</b>           |
|-------------------------|------------------------------|----------------------|-------------------------------|----------------------------|-----------------------|------------------------|
|                         | Numeral                      | Sujeto clausular     | Sustantivo con singularizador | Sustantivo solo o con art. | Usted (2 P.S.)        |                        |
| SV                      | 2<br>0.7 %<br>25 %           | 2<br>0.7 %<br>28.6 % | 39<br>14.2 %<br>35.8 %        | 108<br>39.4 %<br>39.4 %    | 27<br>9.9 %<br>49.1 % | 274<br>100 %<br>39.1 % |
| VS                      | 6<br>1.4 %<br>75 %           | 5<br>1.2 %<br>71.4 % | 70<br>16.4 %<br>64.2 %        | 166<br>39 %<br>60.6 %      | 28<br>6.6 %<br>50.9 % | 426<br>100 %<br>60.9 % |
| <b>Total</b>            | 8<br>1.1 %<br>100 %          | 7<br>1 %<br>100 %    | 109<br>15.6 %<br>100 %        | 274<br>39.1 %<br>100 %     | 55<br>7.9 %<br>100 %  | 700<br>100 %<br>100 %  |

$\chi^2=45.982 \cdot df=13 \cdot \text{Cramer's } V=0.256 \cdot \text{Fisher's } p=0.000$

| <i>Orden del sujeto</i> | <i>Complejidad del sujeto</i> |                        |                      |                        |                      |
|-------------------------|-------------------------------|------------------------|----------------------|------------------------|----------------------|
|                         | CS                            | det.+SN                | det.+SN+CR           | det+SN+fprep           | det+SN+mod           |
| SV                      | 13                            | 72                     | 5                    | 41                     | 29                   |
|                         | 4.7 %<br>15.9 %               | 26.3 %<br>37.9 %       | 1.8 %<br>45.5 %      | 15 %<br>41 %           | 10.6 %<br>45.3 %     |
| VS                      | 69                            | 118                    | 6                    | 59                     | 35                   |
|                         | 16.2 %<br>84.1 %              | 27.7 %<br>62.1 %       | 1.4 %<br>54.5 %      | 13.8 %<br>59 %         | 8.2 %<br>54.7 %      |
| <b>Total</b>            | 82<br>11.7 %<br>100 %         | 190<br>27.1 %<br>100 % | 11<br>1.6 %<br>100 % | 100<br>14.3 %<br>100 % | 64<br>9.1 %<br>100 % |

| <i>Orden del sujeto</i> | <i>Complejidad del sujeto</i> |                      |                     |                     |                        | <b>Total</b>          |
|-------------------------|-------------------------------|----------------------|---------------------|---------------------|------------------------|-----------------------|
|                         | det+SNcoord                   | SN+compl.            | SNCompl.            | Sncoord             | UM                     |                       |
| SV                      | 1                             | 4                    | 1                   | 3                   | 105                    | 274                   |
|                         | 0.4 %<br>33.3 %               | 1.5 %<br>22.2 %      | 0.4 %<br>33.3 %     | 1.1 %<br>33.3 %     | 38.3 %<br>47.7 %       | 100 %<br>39.1 %       |
| VS                      | 2                             | 14                   | 2                   | 6                   | 115                    | 426                   |
|                         | 0.5 %<br>66.7 %               | 3.3 %<br>77.8 %      | 0.5 %<br>66.7 %     | 1.4 %<br>66.7 %     | 27 %<br>52.3 %         | 100 %<br>60.9 %       |
| <b>Total</b>            | 3<br>0.4 %<br>100 %           | 18<br>2.6 %<br>100 % | 3<br>0.4 %<br>100 % | 9<br>1.3 %<br>100 % | 220<br>31.4 %<br>100 % | 700<br>100 %<br>100 % |

$\chi^2=29.328 \cdot df=9 \cdot \text{Cramer's } V=0.205 \cdot \text{Fisher's } p=0.000$

| <i>Orden del sujeto</i> | <i>Función pragmática del enunciado</i> |                     |                      |                      |                     |                      |                     |                       |
|-------------------------|---|---------------------|----------------------|----------------------|---------------------|----------------------|---------------------|-----------------------|
|                         | Anticipativa                            | Clarif.             | Confirm.             | Foco Conv.           | Inf. Comp.          | Deliberat.           | Fática              | Problem.              |
| SV                      | 19<br>6.9 %<br>52.8 %                   | 2<br>0.7 %<br>100 % | 8<br>2.9 %<br>72.7 % | 4<br>1.5 %<br>57.1 % | 0<br>0 %<br>0 %     | 6<br>2.2 %<br>46.2 % | 1<br>0.4 %<br>50 %  | 27<br>9.9 %<br>31 %   |
| VS                      | 17<br>4 %<br>47.2 %                     | 0<br>0 %<br>0 %     | 3<br>0.7 %<br>27.3 % | 3<br>0.7 %<br>42.9 % | 2<br>0.5 %<br>100 % | 7<br>1.6 %<br>53.8 % | 1<br>0.2 %<br>50 %  | 60<br>14.1 %<br>69 %  |
| <b>Total</b>            | 36<br>5.1 %<br>100 %                    | 2<br>0.3 %<br>100 % | 11<br>1.6 %<br>100 % | 7<br>1 %<br>100 %    | 2<br>0.3 %<br>100 % | 13<br>1.9 %<br>100 % | 2<br>0.3 %<br>100 % | 87<br>12.4 %<br>100 % |

| <i>Orden del sujeto</i> | <i>Función pragmática del enunciado</i> |                      |                        |                        |                       |                      |                         | <b>Total</b>           |
|-------------------------|---|----------------------|------------------------|------------------------|-----------------------|----------------------|-------------------------|------------------------|
|                         | Ref.                                    | Repet.               | Retórica               | Subjectio              | N/A                   | Petición             | Pregunta real           |                        |
| SV                      | 16<br>5.8 %<br>44.4 %                   | 2<br>0.7 %<br>66.7 % | 30<br>10.9 %<br>37.5 % | 31<br>11.3 %<br>25.8 % | 18<br>6.6 %<br>32.1 % | 4<br>1.5 %<br>57.1 % | 106<br>38.7 %<br>44.5 % | 274<br>100 %<br>39.1 % |
| VS                      | 20<br>4.7 %<br>55.6 %                   | 1<br>0.2 %<br>33.3 % | 50<br>11.7 %<br>62.5 % | 89<br>20.9 %<br>74.2 % | 38<br>8.9 %<br>67.9 % | 3<br>0.7 %<br>42.9 % | 132<br>31 %<br>55.5 %   | 426<br>100 %<br>60.9 % |
| <b>Total</b>            | 36<br>5.1 %<br>100 %                    | 3<br>0.4 %<br>100 %  | 80<br>11.4 %<br>100 %  | 120<br>17.1 %<br>100 % | 56<br>8 %<br>100 %    | 7<br>1 %<br>100 %    | 238<br>34 %<br>100 %    | 700<br>100 %<br>100 %  |

$\chi^2=31.539 \cdot df=14 \cdot \text{Cramer's } V=0.212 \cdot \text{Fisher's } p=0.002$