

## Disponibilidad léxica de los estudiantes hispanos de Redwood City, California

Sequoia Union High School District Redwood City, California.  
rverdeses@gmail.com

Verdeses-Mirabal, R. T. (2012). Disponibilidad léxica de los estudiantes hispanos de Redwood City, California. *Revista Nebrija de Lingüística Aplicada* (2012) 11.

### RESUMEN

Disponibilidad léxica de los estudiantes hispanos en Redwood City, California presenta el estudio de una comunidad hispanohablante en contacto con el inglés. El objetivo central es realizar el diagnóstico de la disponibilidad léxica de los estudiantes hispanos de duodécimo grado de educación general de la ciudad de Redwood City, California. Se adhiere al Proyecto Panhispánico de Léxico Disponible del español dirigido por Humberto López Morales.

Permite analizar manifestaciones de "variaciones geográficas" y "sociolingüística" del español en esta parte del mundo. Entre las aportaciones de este estudio se encuentran aspectos novedosos en la metodología con el manejo de nuevas variables independientes y el empleo de procedimientos estadísticos. Los resultados que se derivan de la presente investigación ofrecen el léxico disponible de la comunidad hispana de Redwood City. Posibilitan la aplicación práctica a la enseñanza de la lengua materna y de lengua extranjera. Al mismo tiempo permiten hacer comparaciones con otras sintopias ya estudiadas.

Palabras clave: disponibilidad, léxico, palabra, vocablo

### ABSTRACT

*This study of a Spanish-speaking community in contact with English had as its main objective the research and catalog the lexical availability of twelfth grade Hispanic students in general education in the city of Redwood City, California. This research adheres to the Panhispanic Lexical Availability Project led by Humberto López Morales.*

*It allows an examination of aspects of "geographical variations" and "sociolinguistics" of Spanish in this part of the world. Among the contributions of this study are novel aspects in the methodology with the use of new independent variables and the use statistical procedures. The results derived from this research analyze the lexical available to the Hispanic community in Redwood City. They make possible practical applications to the teaching of mother tongue and foreign language. At the same time they allow for comparisons with other communities previously studied.*

*Keywords: availability, lexicon, word*

### INTRODUCCIÓN

El estudio del léxico disponible de estudiantes preuniversitarios en una comunidad estadounidense resulta relevante y novedoso dentro de la lingüística aplicada. Los resultados obtenidos ofrecen posibilidades de aplicación práctica a la enseñanza de la lengua materna y de lengua extranjera.

Este trabajo se adhiere a los estudios que hace más de cincuenta años surgieron en Francia. Se inscribe dentro de la tendencia de estudios léxico-estadísticos.

Los primeros trabajos en el mundo hispano pertenecen a Humberto López Morales. Comenzaron en la década del 70, pero *El léxico disponible de Puerto Rico* y se enriquece enormemente los estudios generales de léxico disponible. En él se ofreció el concepto de léxico disponible, definiéndolo de la siguiente manera:

...existe en el lexicón mental una serie de términos que no se actualizan a menos que sea necesario para comunicar una información muy específica. Se trata de un léxico 'disponible', cuyo estudio no puede emprenderse manejando frecuencias, porque este factor es pertinente solo en el caso de las actualizaciones efectivas, no de las potenciales (López Morales, 1999:11).

El estudio del léxico disponible tiene aplicaciones prácticas que fueron "*entreabiertas*" por los investigadores franceses hace más de cinco décadas. El caso que aquí se presenta permite analizar manifestaciones de "*variaciones geográficas*" y "*sociolingüística*" del español en esta parte del mundo. También posibilita hacer reflexiones sobre el bilingüismo y analizar las manifestaciones del fenómeno de lenguas en contacto vinculadas al creciente flujo migratorio que en las últimas décadas se ha venido produciendo en la frontera mexicana de Estados Unidos. Y en ese marco se presenta: Disponibilidad léxica de los estudiantes hispanos de Redwood City, California.

Las características de la región y la vida sociocultural tienen repercusión en el habla de la comunidad hispana que aquí se asienta.

El trabajo se propuso realizar el diagnóstico de la disponibilidad léxica de los estudiantes hispanos de duodécimo grado de educación general de la ciudad de Redwood City, California a partir de una muestra representativa de tres planteles educativos. Para lograr este propósito general se plantearon los siguientes objetivos específicos:

Conocer la disponibilidad léxica de los hablantes hispanos de la ciudad de Redwood City mediante la aplicación de la metodología léxico estadística.

Determinar la estratificación sociolingüística de la disponibilidad léxica tomando en consideración variables sociales tradicionales (sexo, lengua habitual, nivel sociocultural y generación de inmigrantes) y novedosas como lugar de nacimiento y nivel de español.

- Contribuir al conocimiento de la variedad (diversidad) dentro de la unidad de la lengua española, en un ámbito de lenguas en contacto mediante la obtención de inventarios de los listados de vocabulario disponible de los hablantes hispanos de Redwood City.
- Posibilitar una aplicación pedagógica de los resultados como material útil para la planificación léxica en el proceso educativo de nuestros estudiantes de preuniversitario.

Las hipótesis de esta investigación están apoyadas tanto en el conocimiento sobre la comunidad de Redwood City California como en las conclusiones de investigaciones anteriores.

1. Las mujeres rendirán mejores promedios de respuestas que los hombres. Se intuye que existe una diferencia marcada entre los géneros especialmente en lo relativo a centros de interés específicos.
2. La variable "*lengua habitual*" presentará valor discriminante o diferenciador, dado que se trata de informantes que tienen distinto nivel de uso del español. Se intuye que los estudiantes cuya lengua materna es el español ofrecerán mayor productividad léxica.
3. El nivel de español alcanzado por los estudiantes será determinante en la productividad léxica. Se intuye que a mayor nivel de español habrá mayor productividad léxica.
4. Los estudiantes nacidos en países hispanos ofrecerán mayores promedios de respuestas que los nacidos en Estados Unidos.
5. Existe una correlación entre la producción léxica y las generaciones de inmigrantes. Se parte de la idea que la primera generación de inmigrantes es muy susceptible a la penetración de los vocablos de la lengua de prestigio.
6. La disponibilidad léxica estará correlacionada con el nivel sociocultural de los hablantes. Existen ya investigaciones que señalan que existe una menor disponibilidad en los hablantes de niveles socioculturales bajos. Se debate si se debe a deficiencias cognitivas derivadas del contacto. hablantes de niveles socioculturales bajos. Se debate si se debe a deficiencias cognitivas según la teoría del déficit Bernstein (1973) o intervienen además otros factores (López Morales, 1999).

## ASPECTOS METODOLÓGICOS. DISEÑO DE LA MUESTRA Y OBTENCIÓN DEL MATERIAL

Para la elección de la población se tuvieron en cuenta los principios del Proyecto Panhispánico y los objetivos del estudio. Se trabajó con sujetos que reunían las siguientes características: estudiantes de duodécimo grado que aún no están influidos por los tecnicismos adquiridos en escuelas profesionales y facultades universitarias o en la vida laboral. Lengua materna español o inglés. Nacidos en Estados Unidos o un país hispanohablante y diferentes niveles socioculturales.

Las directrices metodológicas de carácter general que se emplean en esta investigación son las mismas que comparten los trabajos de disponibilidad léxica que se han venido realizando en el mundo hispano en las últimas décadas.

El trabajo investigativo de disponibilidad léxica "consta de 3 fases: I) diseño de la muestra y obtención del material léxico por medio de cuestionarios; II) homogeneización y criterios de edición de las listas de léxico; III análisis de los resultados." (Ávila, 2006:37).

El diseño de la muestra consta de dos partes. Primero se realizó un estudio de la realidad de la ciudad y las escuelas participantes. Este estudio toma en cuenta los datos estadísticos censales (Censo 2000) de la realidad social del área. Se considera el conocimiento de las características sociolingüísticas de la población estudiantil.

Se ha optado por el empleo del sistema norteamericano de las puntuaciones numéricas. Los datos e informaciones que aquí se presentan utilizan el punto como separador decimal.

El método principal es la recogida de datos sociolingüísticos y del léxico de los informantes mediante una encuesta de léxico disponible elaborada para tal efecto. La encuesta se ha venido aplicando con pequeñas variaciones en los diferentes estudios de disponibilidad léxica realizada en España e Hispanoamérica.

La aplicación de la encuesta de disponibilidad contiene información para trabajar seis variables sociolingüísticas independientes: *sexo, lengua habitual, nivel de español, generación de inmigrantes, lugar de nacimiento y nivel sociocultural*. Se busca la descripción del léxico disponible mediante el establecimiento de la relación entre estas variables y los dieciséis centros del estudio del léxico disponible: 16 centros de interés: Centro 01: El cuerpo humano, Centro 02: La ropa, Centro 03: Partes de la casa (sin muebles), Centro 04: Los muebles de la casa, Centro 05: Comidas y bebidas, Centro 06: Objetos colocados en la mesa para la comida, Centro 07: La cocina y sus utensilios, Centro 08: La escuela: muebles y materiales, Centro 09: Calefacción, iluminación y medios de airear un recinto, Centro 10: La ciudad, Centro 11: El campo, Centro 12: Medios de transporte, Centro 13: Trabajos del campo y el jardín, Centro 14: Los animales, 15: Juegos y distracciones, Centro 16: Profesiones y oficios.

## SELECCIÓN DE LA MUESTRA

La muestra de este estudio estuvo constituida por 518 estudiantes hispanos de duodécimo grado (último grado de la enseñanza general) grado precedente a los estudios universitarios del distrito escolar Sequoia Union High School District de la ciudad de Redwood City en California.

La encuesta se llevó a cabo durante los cursos 2008/09 y 2009/10. El corpus de este estudio está formado por una muestra de la población hispana de duodécimo grado del distrito escolar Sequoia. No se ha tenido en cuenta el factor centro escolar o tipo de escuela los cuales se consideran irrelevantes para el área que se ha explorado.

Factor sexo. La muestra de este estudio está representada por 245 hombres y 273 mujeres.

Factor lengua habitual. La muestra fue de 320 informantes cuya lengua habitual es español y 198 lengua habitual inglés.

Según nivel de español la muestra es la siguiente: Nivel bajo 181, nivel medio bajo 173, nivel medio alto 96 y nivel alto 68. Lugar de nacimiento 256 nacidos en Estados Unidos y 262 nacidos en países hispanos. Generación de inmigrantes. 235 Primera generación, 269 segunda generación y 14 tercera generación. Nivel sociocultural. Nivel bajo 271, nivel medio bajo 111, nivel medio alto 111 y nivel alto 25.

La segunda parte consiste en una prueba asociativa que constituye un requerimiento de toda investigación de léxico disponible. El cuadernillo consta de 18 páginas, las dos primeras recogen los datos sociológicos. Las 16 páginas restantes están dedicadas a los 16 centros de interés del Proyecto Panhispánico de Léxico Disponible. Los informantes ofrecen una producción de listas abiertas para cada uno de los centros de interés.

Los centros de interés son: 01. *El cuerpo humano*. 02. *La ropa*. 03. *Partes de la casa (sin muebles)*. 04. *Los muebles de la casa*. 05. *Comidas y bebidas*. 06. *Objetos colocados sobre la mesa para la comida*. 07. *La cocina y sus utensilios*. 08. *La escuela: muebles y materiales*. 09. *Calefacción, iluminación y medios de airear un recinto*. 10. *La ciudad*. 11. *El campo*. 12. *Medios de transporte*. 13. *Trabajos del campo*. 14. *Animales*. 15. *Juegos y diversiones*. 16. *Profesiones y oficios*.

Estos enunciados son breves y se plantean al principio de cada centro de interés:

- Escribe todas las palabras que te vengan a la mente sobre el tema propuesto.
- Si no la recuerdas en español sino en inglés escríbela en inglés.
- Escribe sin tener en cuenta la ortografía, aunque tengas dudas, escribe la palabra.
- No te preocupes si la palabra te resulta malsonante o vulgar, escríbela igual.
- Esta prueba no afectará tus notas.
- Tienes solo dos minutos para cada tema.
- Trabaja con un solo tema. No cambies de página.
- Si necesitas más espacio, escribe en la parte en blanco de la hoja.
- Espera la señal del profesor.

Estas breves consignas se discuten al comienzo de la prueba. Como hicieron Galloso (2003:24-25) y Hernández (2004:28) no se les proporcionó ninguna otra aclaración durante el transcurso de la prueba para evitar un condicionamiento de las respuestas y confusión en los informantes.

## APLICACIÓN DE LA ENCUESTA

La prueba se aplicó durante las clases de español para hispanohablantes, Inglés como segunda lengua (*English Language Development*) e inglés

Se les ofreció una explicación general sobre el propósito del trabajo; como deben ajustarse al tiempo y la necesidad de no preguntar durante el transcurso de la prueba. Se explicó a los estudiantes las consignas que aparecen al principio de cada centro de interés indicándoles que no se van a repetir estas aclaraciones, pero para ayudarles a recordarlas, se encuentran en el inicio de cada tema. Se advirtió sobre la necesidad de no preguntar durante la realización de la prueba para no perturbar la producción de los demás estudiantes. Se comparte la idea de que no se deben ofrecer a los estudiantes aclaraciones ya que eso puede condicionar las respuestas y por tanto alterar los resultados: "A la hora de pasar la encuesta no se le deben hacer aclaraciones a los informantes sobre ningún centro de interés ya que ello condiciona el tipo de respuesta", (Galloso, 1998:27).

Se usó un cronómetro para lograr la precisión del tiempo ajustado a dos minutos. Al igual que González (2002:18), agotados los dos minutos no se les permitió los informantes volver atrás. Se procuraron pasar las encuestas, en la medida de lo posible, durante el tercer período de clases los jueves (de 10:00 a 11:00 de la mañana) por considerar que es el día en el que la curva de rendimiento es óptima.

---

## PREPARACION Y EDICIÓN DEL MATERIAL. CRITERIOS DE UNIFICACIÓN DEL CORPUS

---

Después de recoger todo el material necesario para este estudio se procedió a la edición de la estandarización y lematización de los materiales. Se han aplicado los criterios generales establecidos por los investigadores de disponibilidad léxica.

Con la aplicación de estos criterios el trabajo se adhiere a los acuerdos de las reuniones internacionales de disposición léxica y se continúa el valioso legado de los investigadores precedentes.

En este trabajo se adoptan, fundamentalmente, los criterios de edición de investigadores que sentaron las bases a los estudios del Proyecto Panhispánico: Samper (1998) y López Morales (1999). Se apoya en estudios precedentes que ofrecen en detalles la justificación de la selección de los distintos criterios (Galloso, 2002; Gómez Molina, J.R. & Goómez Devís, M<sup>a</sup> B., 2004; Hernández, 2004). Las decisiones están basadas, también, en los más recientes trabajos publicados (Prado & Galloso, 2005; Ávila, 2006; Ayora, 2006). Se contemplan tanto criterios generales como específicos por centros de interés.

---

## PROCESAMIENTO ESTADÍSTICO DE LOS DATOS

---

La metodología de este estudio se completa con el análisis estadístico que facilita la obtención de conclusiones más fiables. El análisis estadístico nos lleva de la percepción subjetiva e intuitiva a conclusiones de mayor fiabilidad. Al tratarse un trabajo dentro de la léxico estadístico parece conveniente, aprovechar la experiencia precedente en este campo y aplicar al léxico disponible procedimientos existentes alcance de la investigación científica actual. Los recursos tecnológicos cada vez más elaborados ponen hoy ante el investigador herramientas actualizadas y de gran valor. La obtención de los índices de disponibilidad, la frecuencia de aparición, el porcentaje de aparición y la frecuencia acumulada que se ofrecen mediante Displex se ven enriquecidos con la realización de análisis cuantitativos obtenido a través de procesos estadísticos.

Se emplearon dos programas: Displex y *SPSS Statistics 17*. El programa Displex, en línea, permite al investigador obtener en primer lugar los índices de disponibilidad tanto general por cada centro de interés como por cada variable social y los correspondientes estratos de dichas variables.

Por su parte el programa estadístico (Norusis, 2008) facilita el tratamiento de los 518 informantes y las veintidós variables (dieciséis variables dependientes o centros de interés y seis factores sociales o variables independientes).

Con el propósito de procesar los datos estadísticos en el programa SPSS 17 se realizó la codificación que exige este sistema. Básicamente se requiere crear un banco de datos a partir de la información que ya se dispone para Displex. Las variables ya definidas serán las mismas tanto dependientes como independientes, solo que en Displex no se le llama como tal, sino solo centros de interés y factores sociales. Después se procede a la entrada de todos los datos. A partir de entonces predominarán los datos numéricos y los nombres de las variables.

El siguiente paso es procesar los datos para obtener resultados, el investigador está en condiciones para solicitarle al programa los diversos procesos estadísticos, (Norusis, 2008).

En Displex.com se obtuvieron índice de disponibilidad, la frecuencia de aparición de los vocablos, índices cuantitativos total de palabras por cada informante. Se hallaron también los promedios de respuestas por informante, los índices de cohesión el orden de los rasgos de los valores: promedios, total de léxicas y total de vocablos.

En el programa estadístico se han trabajado cálculos que permiten contrastar las hipótesis planteadas. Se muestran los grados de significación estadística de los resultados obtenidos. Así se han podido establecer las relaciones entre las variables tanto dependientes como independientes y el peso que tiene cada una de ellas en la producción léxica de la sintopía. Se hallaron los siguientes valores: Parámetros de tendencia central y medidas de dispersión o centralización, significación de la diferencia entre Medias, análisis factorial simple o análisis de varianza, análisis de regresión lineal múltiple, coeficiente de correlación múltiple y análisis factorial de los componentes principales.

---

## RESULTADOS GENERALES DE LA MUESTRA DE REDWOOD CITY

---

### PALABRAS Y VOCABLOS

En el estudio se siguen los presupuestos de los estudios léxico estadísticos en los cuales se establece la distinción entre palabra y vocablo.

*"Palabra: cada una de las unidades léxicas empleadas, tanto aisladas como unidades complejas o combinación de lexemas para designar un referente, repetidas o no, que integran un texto" (Ayora, 2006:87).*

*"Vocablo: cada una de las unidades léxicas distintas que registra un diccionario de frecuencias. Proporciona información sobre las categorías gramaticales con que se reflejan en el acto discursivo" (Ayora, 2006:87).*

## NÚMERO DE UNIDADES LÉXICAS

Como puede apreciarse en el cuadro 1, el número total de unidades léxicas es de 104,990 (cuadro 1). El promedio por centros de interés es de 6,561.87 unidades léxicas por cada área temática. El promedio del total de palabras recopiladas por cada participante (518) frente a todos los 16 estímulos propuestos es de 202.68.

Centro de interés	Promedio de respuestas			
	Nivel bajo	Medio bajo	Medio alto	Alto
El cuerpo humano	17.21	17.53	17.7	19.2
La ropa	13.21	14.68	14	13.36
Partes de la casa	11.08	12.59	10.97	10.88
Los muebles de la casa	9.16	9.74	9.49	8.88
Comidas y bebidas	18.26	20.51	18.22	20.92
Objetos colocados en ...	9.44	9.34	9.57	9.16
La cocina y sus utensilios	11.18	12.05	11.45	10.76
La escuela...	14.44	16.03	14.68	15.84
Calefacción, iluminación y...	6.27	6.41	6.29	6
La ciudad	13.54	15.34	14.11	14.8
El campo	11.89	12.86	12.31	13.52
Medios de transporte	11.42	12.4	12.04	12.2
Trabajos del campo y ...	7.05	7.32	6.89	7.6
Los animales	19.34	19.7	18.95	20.92
Juegos y distracciones	11.57	13.86	12.29	14.52
Profesiones y oficios	12.16	13.94	13.42	13.16
PROMEDIO GLOBAL	12.33	13.39	12.65	13.23

**CUADRO 1: Resultados generales: total de palabras, rango y distribución porcentual.**

La mayor producción léxica corresponde claramente al centro de interés 14: Los animales. Este centro coincide como el más productivo en la mayoría de los estudios realizados hasta la fecha. El centro 14 Los animales ofrece tres mil cuatrocientas noventa y cuatro palabras más por encima de la producción media del corpus (6,561.87). Siete centros se sitúan por encima de la media. Le sigue el centro 5: Comidas y bebidas con una cantidad también elevada. A continuación de estos dos se ubican en forma descendente: Centro 01: El cuerpo humano, Centro 08: La escuela: muebles y materiales, Centro 10: La ciudad, Centro 02: La ropa, y Centro 16: Profesiones y oficios. Todos ellos por encima de la media.

La distribución porcentual del número total de palabras queda reflejada en el figura 1.

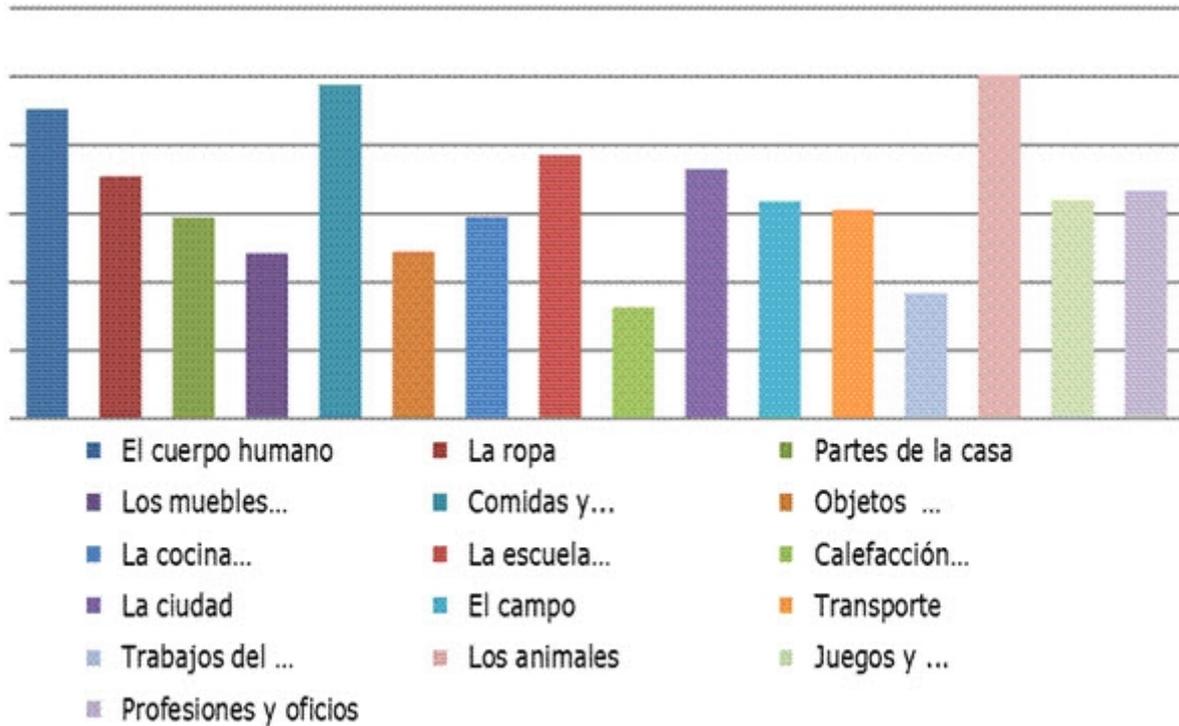


Figura 1: Total de palabras por centros de interés.

De esta primera observación se pueden inferir dos cosas:

- Primero. Existe una relación entre la productividad de los centros de interés y la realidad específicas de la localidad estudiada. Los distintos centros temáticos recogen las particularidades psicosociales y del medio social de la comunidad.
- Segundo. Al comparar los rangos de productividad con otras sintopías se aprecia una marcada similitud en la productividad por centros de interés[1].

Los estímulos con menos distribución porcentual también son los mismos que en la mayoría de las áreas dialectales estudiadas: 'Los muebles de la casa', 'Trabajos del campo y el jardín' y 'Calefacción, iluminación y medios de airear un recinto', aunque no coinciden en la misma posición: cuadro 2, cuadro 3.

AVI	SAL	ZAM	AST	ARA	VAL	CUE	SOR	MAL	CEU	RWC
14	14	14	14	14	14	14	14	14	14	14
5	5	5	5	5	5	5	5	5	1	5
1	1	1	1	1	1	8	8	1	5	1

?

Cuadro 2: Correspondencia de los 3 primeros lugares, según centros de interés de las distintas investigaciones del mundo hispánico y Redwood City

AVI	SAL	ZAM	AST	ARA	VAL	CUE	SOR	MAL	CEU	RWC
4	4	4	4	4	4	13	4	6	4	4
9	9	9	9	9	9	1	9	9	9	13
13	13	13	13	13	13	9	13	13	13	9

*Cuadro 3: Correspondencia de los 3 últimos lugares, según centros de interés en las distintas investigaciones del mundo hispánico y Redwood City*

## NÚMERO DE VOCABLOS

Comparar los vocablos que generan cada centro de interés es otro criterio fundamental para estudiar las diferencias entre los distintos estímulos (Samper et al 2003). Se puede comprobar en el cuadro 4 que del total de 104,990 unidades léxicas solamente 11,866 constituyen vocablos o palabras diferentes (12.5%). Esta cantidad representa un promedio de 741.62 vocablos por área temática (16). El promedio de palabras diferentes recopiladas por cada participante (518) frente a todos los estímulos propuestos (16) es de 22.90 vocablos por informantes.

Centro de interés	Total de vocablos	Rango	%
El cuerpo humano	373	16	3.96
La ropa	658	8	5.49
Partes de la casa	567	12	4.73
Los muebles de la casa	506	14	4.22
Comidas y bebidas	901	6	7.52
Objetos colocados en la mesa...	380	15	3.17
La cocina y sus utensilios	554	13	4.62
La escuela (muebles y materiales)	662	7	5.53
Calefacción, iluminación y medios...	574	10	4.79
La ciudad	1204	1	10.06
El campo	1088	3	9.09
Medios de transporte	585	9	4.88
Trabajos del campo y del jardín	1087	4	9.08
Los animales	570	11	4.76
Juegos y distracciones	1097	2	9.16
Profesiones y oficios	1060	5	8.85
<b>TOTAL DE VOCABLOS</b>	<b>11,866</b>		<b>100%</b>

*Cuadro 4: Total de vocablos, rango y distribución porcentual*

Una vez más se aprecian marcadas diferencias numéricas. Hay solo tres centros que se encuentran por encima de la media, se destacan como más productivos: Centro 10: La ciudad, Centro 05: Comidas y bebidas y Centro 16: Profesiones y oficios. Ninguno de los centros más productivos logra duplicar la media de vocablos.

El centro 01: El cuerpo humano, uno de los más productivos en la cantidad de palabras registra casi la mitad de la media en cuanto a los vocablos (373). Se sitúa en el lado opuesto en la productividad.

La mayor o menor riqueza de vocablos no se corresponde con la productividad de unidades léxicas en las diferentes áreas temáticas. Se procede a comparar los rangos de productividad del total de palabras y del total de vocablos al igual que han hecho otros investigadores (Gómez Molina, J.R. & Gómez Devís, M<sup>a</sup>.B.(2004).

Existe discrepancia entre los rangos de palabras y vocablos en casi la totalidad de los estímulos. Solo el centro 04 *Los muebles de la casa* se encuentra en la misma posición tanto para los rangos de palabras como para los rangos de vocablos. Se considera que tres áreas temáticas mantienen cierta correspondencia separada solo por una posición entre las palabras y los vocablos: 03 *Las partes de la casa*, 07 *La cocina y sus utensilios* y 12 *Medios de transporte*. Los rangos de productividad entre palabras y vocablos se alejan una posición en los Centros 03 *Las partes de la casa (sin muebles)*, 07 *La cocina y sus utensilios*, 12 *Medios de transportes*. Hay una diferencia hasta de 11 posiciones en el centro 11 *El campo*. El centro 14: *Los animales* que ocupa la posición 1 en la productividad de palabras, en los vocablos ocupa el rango 11.

El rango de palabras y vocablos se encuentra a solo una posición de distancia entre los centros 03: *Partes de la casa*, 04: *Muebles de la casa*, 07: *La cocina y sus utensilios* y 12: *Medios de transporte*.

Los datos obtenidos cobran mayor relevancia cuando se les compara con los de otras sintopías estudiadas en el Proyecto Panhispánico de Léxico Disponible. La jerarquización de los valores muestra una diferencia entre las similitudes entre las diferentes sintopías estudiadas en cuanto a la productividad de palabras y la de vocablos. En ello hay que considerar que tal vez la edición de los materiales es un factor importante que puede incidir en la diferenciación. No obstante, este estudio reafirma la tesis antes expuesta en otras investigaciones (Gómez Molina, J.R. & Gómez Devís, M<sup>a</sup>.B.2004) de que las diferencias léxicas son un reflejo de las realidades psicosociales y los entornos que se estudian.

Algunos centros con más alto número de vocablos en este estudio tienen correspondencia con otras zonas dialectales estudiadas: '*La ciudad*', '*Juegos y distracciones*', '*Comidas y bebidas*', '*La escuela: muebles y materiales*', '*La cocina y sus utensilios*'.

En el extremo opuesto se encuentran, como en otras zonas estudiadas, los estímulos que han generado menos productividad y '*El cuerpo humano*', '*Calefacción, iluminación y medios de airear un recinto*'.

Un análisis más profundo de la jerarquización de las diferentes áreas temáticas, permite establecer lo siguiente:

- Existe un área nocional en la que la amplitud de vocablos se encuentra por encima de la media. Está integrada por los centros: '*La ciudad*', '*Juegos y distracciones*', '*El campo*', '*Trabajos del campo y el jardín*', '*Profesiones y oficios*'.
- No existen áreas temáticas cuyos valores se encuentren próximos de la media tanto por encima como por debajo. Como ocurre en otros estudios, los valores más cercanos a la media están separados por 80 y 84 vocablos: '*La escuela: materiales y muebles*' y '*La ropa*' respectivamente.
- Hay estímulos que se mantienen en los más alejados de la media: '*Los animales*' '*Partes de la casa*', '*La cocina y sus utensilios*', '*El cuerpo humano*', '*Objetos colocados en la mesa para la comida*', '*Los muebles de la casa*'.

La figura 2 permite visualizar el total de vocablos:

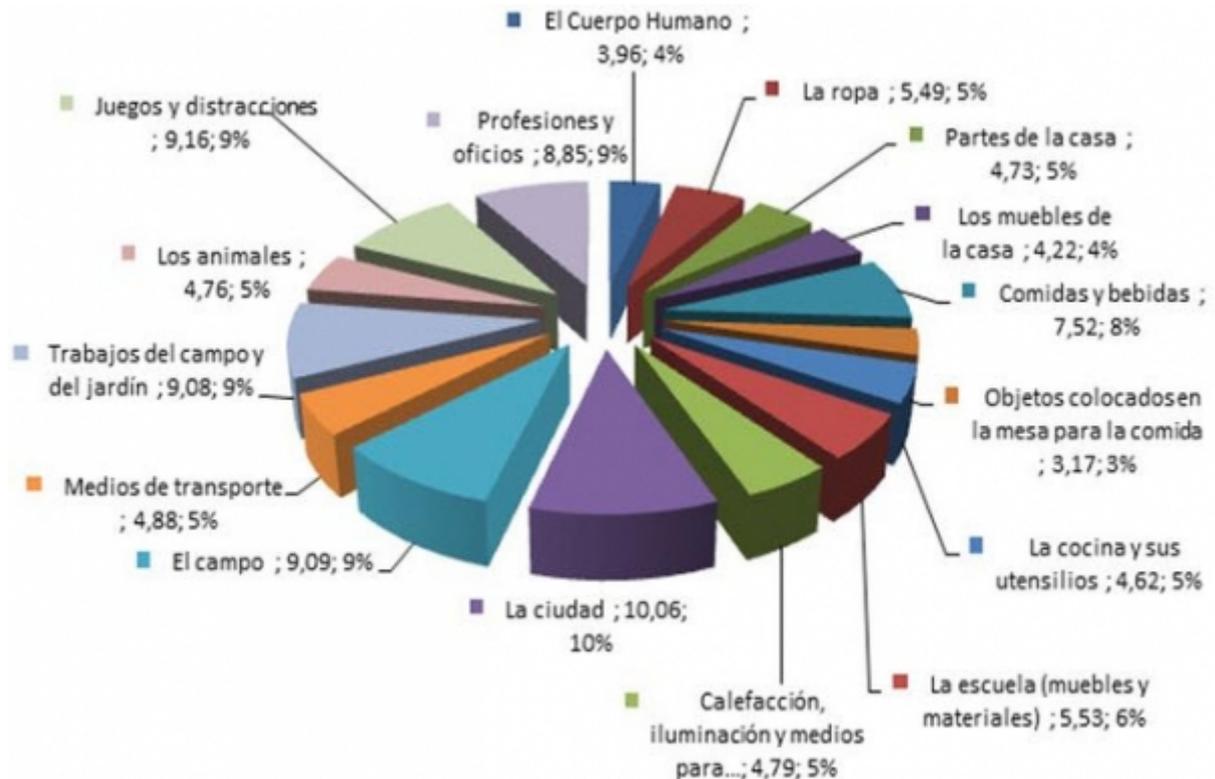


Figura 2: Total de vocablos

## PROMEDIOS DE RESPUESTAS

Para medir la productividad léxica se han venido utilizando con gran preferencia los promedios de respuestas por informantes.

Los resultados obtenidos permiten clasificar los centros de interés en tres grupos bien delimitados. No obstante, "se debe tener en cuenta que el índice de cohesión está condicionado por el número de vocablos y por consiguiente, por las normas de edición" (Bartol, 2004:51)

1) Grupos cuyo campo semántico es más compacto por lo tanto se establece mayor concreción semántica y la asociación con el concepto es más precisa: 'El cuerpo humano' y 'Los animales'.

2) Áreas temáticas con un grado intermedio en la coincidencia de las respuestas. En este grupo se sitúan la mayoría de los centros (9): 'La ropa', 'Partes de la casa', 'Los muebles de la casa', 'Comidas y bebidas', 'Objetos colocados en la mesa para la comida', 'La cocina y sus utensilios', 'La escuela (muebles y materiales)', 'Medios de transporte', 'Trabajos del campo y del jardín', 'Juegos y distracciones'.

3) Centros cuyos campos semánticos son más difusos y el nivel de cohesión es muy bajo: 'Calefacción, iluminación y medios para airear un recinto', 'Medios de transporte', 'Juegos y distracciones', 'Profesiones y oficios'.

Una interpretación sobre los centros más difusos podría atribuirse a las condiciones sociales y el medio en que se desarrollan los informantes. Por una parte el contexto socio histórico del entorno no propicia el reconocimiento inmediato y la amplia asociación de términos relacionados con la calefacción, la iluminación y medios para airear un recinto. Al mismo tiempo, la asociación con juegos y distracciones es abundante. Del mismo modo, son amplios los límites de los medios de transportes y profesiones y oficios. Los informantes hacen una asociación extensa provocando resultados con una menor cohesión.

Es de interés conocer índices de cohesión en estudios anteriores en otras áreas del mundo hispano. Como han hecho otros autores, se recopilaron los datos ofrecidos por los estudios precedentes (2)

Al comparar los datos obtenidos con otras sintopías estudiadas se puede concluir que existen las siguientes relaciones:

- Aquellos estímulos con mayor productividad de respuestas que sobrepasan siempre el nivel promedio: '*Los animales*', '*Comidas y bebidas*' y '*El cuerpo humano*'.
- Un segundo grupo lo forman los campos nocionales con valores uno o dos puntos por encima de la media: '*La ciudad*', '*La ropa*', y '*La escuela: muebles y materiales*'.
- Un tercer grupo lo forman tres campos nocionales con valores ligeramente cercanos (menos de un punto) a la media: '*El campo*', '*Profesiones y oficios*' y '*Juegos y diversiones*'.
- Áreas nocionales que siempre quedan por debajo de la media: '*Medios de transporte*', '*Partes de la casa*', '*La cocina y sus utensilios*', '*Objetos colocados en la mesa para la comida*', '*Los muebles de la casa*'.
- Existen dos áreas nocionales como las menos productivas: '*Calefacción, iluminación y medios de airear un recinto*' y '*Trabajos del campo y el jardín*'.

Se observan dos centros nocionales que ocupan la misma posición en todos los estudios que se han comparado: '*Los animales*' y '*El campo*' (3). El centro nocional '*Comidas y bebidas*' (4) difiere en una sola sintopía (Ceuta). Por otra parte, los centros nocionales '*Trabajos del campo y el jardín*' y '*Calefacción, iluminación y medios de airear un recinto*' ocupan diferentes posiciones sólo en dos sintopías (Cuenca y Redwood City) las cuales se encuentran en la posición 16. Así mismo el área temática '*Los muebles de la casa*' también ocupa un rango distinto sólo en dos estudios (Cuenca y Málaga).

Estas observaciones reafirman el criterio de que el léxico disponible es un reflejo de la realidad socio psicológica del entorno ambiental de los informantes. Las posibilidades asociativas de un centro particular se ven afectadas por el conocimiento y el uso cotidiano de los estímulos propuestos. Por ejemplo los informantes en este estudio no están familiarizados con el término *calefacción* tal vez porque no es un elemento importante en la vida de una comunidad que goza de un clima moderado.

Como ha quedado demostrado en los estudios precedentes, se presentan también en Redwood City los mismos ámbitos de asociación expuestos por Samper et al (2003) Se pueden distinguir tres grupos: el primero constituido por centros muy productivos en los cuales se ubican los centros 14:*Los animales*, 05:*Comidas y bebidas* y 01:*El cuerpo humano*. Un segundo grupo formado por "campos léxicos" que se encuentran por encima de la media. En este se ubican: 08*La escuela: muebles y materiales*, 10*La ciudad* y 02*La ropa*. Sin embargo, los resultados de Redwood City marcan una diferencia entre los estudios de léxico disponible. Los centros 11*El campo* y 16 *Profesiones y oficios* alcanzan promedios por debajo de la media.

Al igual que en los estudios precedentes del *Proyecto Panhispánico de Léxico Disponible*, en Redwood City, tampoco alcanzan a superar la media de respuesta los centros: 03 *Las partes de la casa*, 04 *Los muebles de la casa*, 06 *Objetos colocados en la mesa para la comida*, 12 *Medios de transporte*, 15 *Juegos y distracciones* y 16 *Profesiones y oficios*.

También en Redwood City los campos asociativos menos productivos son 09*Calefacción, iluminación y medios de airear un recinto* y 13*Trabajos del campo y el jardín*. Se toman prestadas las palabras de Samper et al (2003) para concluir que en este caso estas son "áreas léxicas más cerradas y poco familiares para los informantes de este tipo de estudios.

## COHESIÓN DE LOS DATOS

El índice de cohesión es considerado por algunos investigadores como un valor con mayor elaboración (Bartol, 2001); (Carcedo, 2001). Una vez obtenidos ambos valores se comparan para determinar el grado de coincidencia de las respuestas dadas por los informantes "*El índice de cohesión es un indicador del grado de coincidencia en las respuestas*" (Urzúa C, P; Saez, K.; Echeverría, M. S. 2006).

Algunos centros de interés logran altos grados de cohesión semántica pues se trata de centros con asociaciones lógicas muy precisas mientras otros centros tienen una asociación lógica muy amplia. Esto hace que el grado de cohesión sea muy "débil", (Gómez Molina, J.R. & Gómez Devís, M<sup>a</sup>.B.2004:82).

Existen discrepancias en cuanto al uso del criterio densidad léxica. Algunos estiman que este criterio es inadecuado porque está condicionado por el número de participantes (Bartol, 2001). Se advierte, por tanto, una relación directa entre la densidad léxica y el número de participantes: a mayor cantidad de informantes mayor cantidad de palabras y por lo tanto "*mayor índice de densidad*". Sin embargo, el mismo Bartol (2004:50) reconoce el valor del índice de cohesión en la comparación entre diferentes sintopías. La utilización del cálculo del índice de cohesión "...se hace indispensable para lograr una caracterización verdaderamente válida y adecuada de la realidad de los dialectos estudiados" (Alba, 1998:304).

A continuación se presentan los índices de cohesión reflejados por los informantes de Redwood City.

Como puede observarse en el cuadro 5 existe plena correspondencia entre los rangos alcanzados por la densidad y el índice de cohesión.

CENTROS DE INTERÉS	DENSIDAD	RNG	IC	RNG
El cuerpo humano	24.27	1	0.046	1
La ropa	10.78	6	0.02	6
Partes de la casa	10.39	9	0.02	9
Los muebles de la casa	9.56	10	0.018	10
Comidas y bebidas	10.84	5	0.02	5
Objetos colocados en la ...	12.85	3	0.024	3
La cocina y sus utensilios	10.66	7	0.02	7
La escuela ...	11.65	4	0.022	4
Calefacción, iluminación ...	5.67	15	0.016	15
La ciudad	6.06	12	0.011	12
El campo	5.83	14	0.011	14
Medios de transporte	10.44	8	0.02	8
Trabajos del campo y del jardín	3.38	16	0.006	16
Los animales	17.64	2	0.034	2
Juegos y distracciones	5.83	13	0.011	13
Profesiones y oficios	6.28	11	0.012	11

**Cuadro 5: Densidad, Índice de cohesión y sus rangos**

Se han clasificado los centros de interés de forma decreciente en tres grupos, (Gómez Molina, J.R. & Gómez Devís, M<sup>a</sup>.B.2004:83):

1. Los campos semánticos más compactos con la mayor asociación conceptual son: *'El cuerpo humano', 'Los animales', 'Objetos colocados en la mesa para la comida'*.
2. Los grupos semánticos que alcanzan un grado medio de correspondencia en las respuestas son: *'La escuela: muebles y materiales'* y con igual índice *'La ropa', 'Partes de la casa', 'Comidas y bebidas', 'La cocina y sus utensilios', 'Medios de transporte'*.
3. Los campos semánticos más dispersos lo integran aquellos centros que alcanzaron los valores más bajos en los índices de cohesión y densidad: *'Los muebles de la casa', 'La ciudad', 'El campo', 'Trabajos del campo y del jardín' y 'Juegos y distracciones' y 'Profesiones y oficios'*.

Los campos semánticos con mayor grado de cohesión (son las áreas más cerradas, más compactas) a través de los diez estudios que se están comparando son: *'El cuerpo humano', 'Los animales' y "Objetos colocados en la mesa para la comida'*

- A. Se registran un grupo de centros de interés con índice de cohesión medio: *'La cocina y sus utensilios', 'Medios de transporte' y 'Partes de la casa (sin muebles)'*.
- B. Los campos de más bajos grados de cohesión en Redwood City son: *'El campo', 'Calefacción, iluminación y medios de airear un recinto', Trabajos del campo y el jardín' ([5])*.

## VALORACIÓN DE LOS RESULTADOS GENERALES

Los distintos índices a disposición del investigador son las herramientas necesarias para poder interpretar y explicar los resultados contenidos en los inventarios del léxico disponible.

La nómina léxica de una comunidad se puede conocer mediante índices parciales como el número total de unidades léxicas y la densidad léxica.

Para analizar la correspondencia que se presentan entre los resultados se utilizan: el índice de cohesión, el promedio de respuestas y el número de vocablos. Es necesario puntualizar que los distintos valores de un centro de interés están relacionados con múltiples aspectos.

El cuadro 6 contiene, jerarquizados en forma decreciente, los valores: índice de cohesión, promedio de respuestas y número de vocablos de los centros de interés lo cual permite distinguir la similitud de las respuestas.

Centro de interés	IC	Promedio de respuestas	Vocablos Total
El cuerpohumano	0.046	17.85	373
Los animales	0.034	19.28	570
Objetos colocados ...	0.024	9.3	380
La escuela...	0.022	14.73	662
Comidas y bebidas	0.02	18.95	901
La ropa	0.02	13.55	658
Partes de la casa	0.02	11.3	567
Medios de transporte	0.02	11.69	585
La cocina y ...	0.02	10.71	554
Los muebles ...	0.018	9.19	506
Calefacción, ...	0.016	6.23	574
Profesiones y oficios	0.012	12	1060
La ciudad	0.011	13.98	1204
El campo	0.011	12.11	1088
Juegos y distracciones	0.011	12.2	1097
Trabajos del campo...	0.006	7.02	1087

**Cuadro 6: Indicadores generales: Índice de cohesión, promedios de respuestas y número total de vocablos.**

Observando atentamente el cuadro 6 se puede subrayar lo siguiente:

- Como en el Léxico valenciano (Gómez Molina, J.R.& Gómez Devís, M<sup>a</sup>.B.2004) existe en Redwood City una divergencia entre los centros de interés cuando se ofrece una "gradación jerarquizada".
- Las áreas temáticas que se presentan como las más compactas presentan un número de vocablos inferior a la media del estudio. Solo el área nocional '*Comidas y bebidas*' supera la media del conjunto.
- De los cinco campos semánticos más dispersos, es decir de menor concreción conceptual, cuatro registran mayor cantidad de vocablos que la media total. El área nocional '*Calefacción, iluminación y medios de airear un recinto*' alcanza un total de vocablos inferior a la muestra.

Ha quedado establecido en el léxico disponible de Redwood City la probable relación inversa "...y en cierta medida lógica, entre estos valores de los centros de interés: a mayor índice de cohesión, menor número de vocablos y a menor índice de cohesión, mayor número de palabras diferentes (Gómez Molina, J.R.& Gómez Devís, M<sup>a</sup>.B. (2004:87).

Este trabajo toma en consideración las propuestas anteriores y considera el índice en  $\geq 0.1$ . Se incluye además entre paréntesis, como han hecho otros investigadores, (Gómez Molina, J.R.& Gómez Devís, M<sup>a</sup>.B.2004:89) el I.D propuesto por Bartol (2002).

Los resultados obtenidos mediante el programa Dispollex permiten distinguir cuatro valores generales predominantes: Índice de disponibilidad, frecuencia relativa, por ciento de aparición y frecuencia acumulada. A partir de ahí se ha seguido la idea de investigaciones precedentes (Gómez Molina, J.R & Gómez Devís, M<sup>a</sup>.B.2004) según la cual se modificó el criterio de Bartol (2002) situando el I.D  $> 0.1$  y se mantiene en paréntesis el criterio utilizado por el de Bartol (de I.D  $> 0.2$  Se añade también los vocablos mencionados por más del 25% de los encuestados (Gómez Molina, J.R.& Gómez Devís, M<sup>a</sup>.B. (2004:88), cuadro 7.

Centro de interés	Nº de vocablos con ID .> 0.1	Nº de vocablos mencionados por más del 25% de los informantes	Nº de vocablos con frecuencia acumulada del 75%
El cuerpo humano	25 (15)	24	29
La ropa	18 (13)	13	41
Partes de la casa	16 (9)	12	35
Los muebles de la casa	15 (5)	8	35
Comidas y bebidas	25(10)	12	92
Objetos colocados...	9(7)	7	21
La cocina y ...	17(9)	9	46
La escuela...	18(13)	13	49
Calefacción, ...	9(7)	7	55
La ciudad	20(12)	11	133
El campo	14(6)	6	156
Medios de transporte	19(10)	13	28
Trabajos ...	7(2)	2	272
Los animales	29(11)	19	55
Juegos y distracciones	16(4)	5	154
Profesiones y oficios	17(5)	7	111

*Cuadro 7: Vocablos con mayor ID en cada centro*

A continuación se presentan las unidades léxicas que alcanzaron los más altos índices de disponibilidad, de modo individual, como se ha hecho en estudios anteriores (Gómez Molina, J.R. & Gómez Devís, Ma.B.2004); (Ávila 2006).

**LISTADO DE LOS VOCABLOS MÁS DISPONIBLES DEL CORPUS**

--  
--  
-----  
--  
--  
--  
--  
--  
--  
--  
--  
--

**INFLUENCIA CUANTITATIVA DE LOS FACTORES SOCIALES**

El presente trabajo se ha propuesto entre sus objetivos conocer la influencia de factores sociales ya establecidos en el Proyecto Panhispánico de Léxico Disponible: sexo, lengua habitual y nivel sociocultural. Se han agregado: nivel de español, lugar de nacimiento y generación de inmigrantes.

Los tres factores sociales incorporados responden a la realidad de Redwood City. Como ha quedado explícito en páginas anteriores se estudia una comunidad hispanohablante en contacto con la comunidad angloparlante del norte de California. La incidencia de la lengua habitual se ha venido investigando en estudios precedentes. En el caso particular que aquí se presenta, es necesario agregar el factor nivel de español pues se busca establecer la incidencia del dominio de la lengua española frente a la lengua inglesa porque se trata de una relación de lenguas en contacto. Resulta relevante también establecer el papel que desempeña en la producción léxica el lugar de nacimiento y la generación de inmigrantes. Se considera que el análisis de las correlaciones de los distintos factores sociales permitirá acercarse a una adecuada descripción de la comunidad hispanohablante de Redwood City.

## FACTOR SEXO

Algunas investigaciones precedentes han establecido la incidencia del factor social sexo. No existe unanimidad en la influencia del factor sexo en la disponibilidad léxica ([6]).

La muestra está formada por 245 hombres y 273 mujeres. Se procede a analizar el promedio de respuestas. En la sintopía de Redwood City las mujeres ejercen una influencia distintiva en la producción léxica. Las mujeres superan a los hombres en el promedio de respuestas en todos los centros temáticos, cuadro 8.

Centro de interés	Promedios de respuestas		Índice de cohesión	
	Hombres	Mujeres	Hombres	Mujeres
El cuerpo humano	16.82	18.07	0.045	0.048
La ropa	12.6	14.68	0.019	0.022
Partes de la casa	10.57	12.1	0.018	0.021
Los muebles...	8.25	10.32	0.016	0.02
Comidas y bebidas	17.33	20.23	0.019	0.022
Objetos colocados...	8.89	9.92	0.023	0.026
La cocina y sus...	10.04	12.63	0.018	0.022
La escuela...	13.27	16.36	0.02	0.024
Calefacción, ...	5.53	6.97	0.009	0.012
La ciudad	12.6	15.46	0.01	0.012
El campo	10.8	13.58	0.009	0.012
Medios de transporte	11.06	12.47	0.018	0.021
Trabajos...	6.44	7.7	0.005	0.007
Los animales	18.7	20.05	0.032	0.035
Juegos y distracciones	11.54	13.08	0.01	0.012
Profesiones y oficios	11.51	14.08	0.01	0.013
PROMEDIO GLOBAL	11.62	13.61		

**Cuadro 8: Promedios de respuestas e índices de cohesión según el factor sexo**

## FACTOR LENGUA HABITUAL

La inclusión del factor lengua habitual en la descripción de la disponibilidad léxica de Redwood City cobra especial importancia por tratarse de una comunidad inmersa en ámbito de contacto con el inglés. Este estudio facilita el análisis de la incidencia de la lengua que más usan los estudiantes por eso se decide usar el término de lengua habitual. En la comunidad estudiada se da un bilingüismo activo. El contacto lingüístico del español y el inglés en esta sintopía afecta todas las esferas de la vida de los inmigrantes hispanos.

Los promedios de respuestas de los encuestados ofrecen ligera diferencia entre los que poseen el español como lengua habitual y aquellos que usan el inglés. En todos los centros de interés la diferencia favorece al subgrupo español. No obstante, solo tres áreas temáticas, 'La ropa', 'La escuela: muebles y materiales', y 'Los animales' tienen más de dos puntos mientras que el resto no alcanza dos puntos. Según el promedio general ambos subgrupos se encuentran separados por un punto, cuadro 9.

Centro de interés	Promedios de respuestas		Índice de cohesión	
	Español	Inglés	Español	Inglés
El cuerpo humano	17.9	16.81	0.047	0.045
La ropa	14.14	12.99	0.021	0.019
Partes de la casa	11.77	10.74	0.020	0.018
Los muebles de la casa	9.73	8.72	0.019	0.017
Comidas y bebidas	19.4	18	0.021	0.019
Objetos colocados...	9.68	9.04	0.025	0.023
La cocina y sus utensilios	11.85	10.68	0.021	0.019
La escuela...	15.53	13.88	0.023	0.020
Calefacción, s...	6.51	5.93	0.011	0.010
La ciudad	14.3	13.8	0.011	0.011
El campo	12.59	11.73	0.011	0.010
Medios de transporte	12.21	11.14	0.020	0.019
Trabajos del campo...	7.33	6.73	0.006	0.006
Los animales	20.17	18.19	0.035	0.031
Juegos y distracciones	12.61	11.94	0.011	0.010
Profesiones y oficios	13.19	12.33	0.012	0.011
PROMEDIO GLOBAL	13.06	12.04		

**Cuadro 9: Promedio de respuestas factor: Lengua habitual**

Los resultados indican superioridad del subgrupo lengua habitual español sobre el subgrupo lengua habitual inglés en cuanto a los índices de cohesión sugiere que el subgrupo español alcanza mayor grado de concreción semántica en todas las áreas temáticas.

## FACTOR NIVEL DE ESPAÑOL

En este caso se trata de ver como la instrucción recibida en español y el estudio formal de la lengua incide en la producción del léxico disponible. Se piensa que este factor sociológico tiene una incidencia de peso en la producción del léxico disponible "...es normal que las personas más instruidas hagan mayor uso de las variantes que son consideradas como más prestigiosas o que más se ajustan a la norma"(Moreno, 2004:61). Se realizó una escala que permite constatar los niveles de español divididos en cuatro subgrupos: nivel bajo, nivel medio bajo, nivel medio alto y nivel alto.

En este estudio se ha podido constatar que, en general, hay una progresión en los valores de los promedios a través de los distintos niveles, es decir una relación directa entre nivel de español y producción léxica. Sin embargo, en el nivel bajo, los centros: 03: Partes de la casa (sin muebles) y centro 05: Comidas y bebidas superan al nivel medio bajo. También el nivel medio alto supera al nivel alto en los centros 5: Comidas y bebidas y Centro 16: Profesiones y oficios. Además, el nivel medio bajo supera al nivel medio alto en los centros 6: Objetos colocados sobre la mesa para la comida y centro 11: El campo cuadro 10.

Centro de interés	Promedio de respuestas			
	Nivel bajo	Medio bajo	Medio alto	Alto
El cuerpo humano	16.41	17.16	18.74	19.38
La ropa	12.67	13.86	14.63	14.71
Partes de la casa	11.41	11.07	11.81	11.44
Los muebles de la casa	8.39	9.46	10.41	10.07
Comidas y bebidas	19.03	18.5	19.4	18.59
Objetos colocados ...	9.05	9.6	9.47	9.97
La cocina y sus utensilios	11.24	11.46	11.69	11.28
La escuela...	13.98	14.94	16.11	15.49
Calefacción, iluminación y ...	5.72	6.36	6.89	6.76
La ciudad	13.02	14.36	14.82	15.35
El campo	11.47	12.47	12.26	13.84
Medios de transporte	10.8	12.23	12.66	12.18
Trabajos del campo y...	6.87	7.01	7.17	7.88
Los animales	18.56	18.92	20.77	21.01
Juegos y distracciones	11.53	12.58	12.96	13.12
Profesiones y oficios	11.45	13.4	13.93	13.74
PROMEDIO GLOBAL	11.98	12.71	13.36	13.43

*Cuadro 10: Promedios de respuestas factor nivel de español.*

Al igual que se ha visto en los promedios por niveles, también los índices de cohesión muestran, en general, valores que van progresando a medida que aumentan los niveles de español. Sin embargo, el nivel medio bajo supera al nivel medio alto en 'Objetos colocados en la mesa para la comida'. Los índices se corresponden en trece áreas. En todos los centros de interés el nivel alto posee los mayores grados de concreción semántica, es decir que la mayor afinidad en las respuestas léxicas corresponde al nivel alto, cuadro 11.

Centro de interés	Índices de cohesión			
	Nivel bajo	Medio bajo	Medio alto	Alto
El cuerpo humano	0.043	0.046	0.050	0.051
La ropa	0.019	0.021	0.022	0.022
Partes de la casa	0.020	0.019	0.020	0.020
Los muebles de la casa	0.016	0.018	0.020	0.019
Comidas y bebidas	0.021	0.020	0.021	0.020
Objetos colocados en la mesa...	0.023	0.025	0.024	0.026
La cocina y sus utensilios	0.020	0.020	0.021	0.020
La escuela (muebles y materiales)	0.021	0.022	0.024	0.023
Calefacción, iluminación y ...	0.009	0.011	0.012	0.011
La ciudad	0.010	0.011	0.012	0.012
El campo	0.010	0.011	0.011	0.012
Medios de transporte	0.018	0.020	0.021	0.020
Trabajos del campo y del jardín	0.006	0.006	0.006	0.007
Los animales	0.032	0.033	0.036	0.036
Juegos y distracciones	0.010	0.011	0.011	0.011
Profesiones y oficios	0.010	0.012	0.013	0.012

*Cuadro 11: Índice de cohesión nivel de español*

## GENERACIÓN DE INMIGRANTES

Los estudios realizados en California sobre el mantenimiento del español a través de las distintas generaciones (Silva-Corvalán, 2001) sugieren que si bien por una parte la continuidad del español es un factor innegable, el cambio hacia el inglés es una tendencia persistente. Este estudio recoge tres generaciones de inmigrantes. En once centros de interés los promedios de la primera generación superan a la segunda generación, excepto en las áreas nocionales: '*Calefacción, iluminación y medios de airear un recinto*', '*La ciudad*', '*El campo*', '*Medios de transporte*' y '*Trabajos del campo y el jardín*'. Entre la segunda y tercera generación, todos los promedios son superiores en la tercera generación. Sin embargo, diez áreas temáticas recogen promedios superiores en la primera generación de inmigrantes respecto a la tercera generación: '*El cuerpo humano*', '*La ropa*', '*Partes de la casa (sin muebles)*', '*Los muebles de la casa*', '*Comidas y bebidas*', '*La cocina y sus utensilios*', '*La escuela: muebles y materiales*', '*Los animales*' y '*Profesiones y oficios*'. Los promedios globales, en todos los centros de interés, indican que los valores son superiores en la primera y tercera generación e inferiores en la segunda generación, cuadro 12.

Centro de interés	Promedio de respuestas		
	Primera	Segunda	Tercera
El cuerpo humano	18.49	16.64	16.86
La ropa	14.40	13.06	14.14
Partes de la casa	12.00	10.80	11.93
Los muebles ...	10.00	8.75	9.57
Comidas y bebidas	19.38	18.43	18.50
Objetos...	9.59	9.28	9.86
La cocina y ...	11.85	11.00	11.71
La escuela ...	15.82	14.06	15.43
Calefacción, ...	6.91	5.71	7.00
La ciudad	14.59	13.60	15.79
El campo	13.06	11.51	13.29
Medios de transporte	12.23	11.32	13.93
Trabajos y...	7.57	6.64	8.21
Los animales	20.73	18.24	19.79
Juegos y ...	12.82	11.90	13.36
Profesiones y oficios	13.89	11.91	13.86
PROMEDIO GLOBAL	13.33	12.05	13.33

**Cuadro 12: Promedios de respuesta según el factor: generación de inmigrantes**

Al igual que se ha visto en los promedios por las distintas generaciones, también los índices de cohesión muestran lo siguiente: entre la primera y segunda generación los valores son más altos en la primera generación excepto en el centro 13:Trabajos del campo y el jardín donde aparecen iguales (0,006) Entre la segunda y la tercera generación, todos los índices de cohesión son superiores en la tercera generación, excepto el centro 05:Comidas y bebidas donde se presentan los mismos índices. La segunda generación es la que tiene los índices más bajos. Los centros con índices de cohesión más altos en las tres generaciones son: '*El cuerpo humano*', '*Los animales*', '*Objetos colocados en la mesa para la comida*' y '*La escuela: muebles y materiales*' Los centros con índices más bajos en las tres generaciones son: '*Trabajos del campo y el jardín*', '*Juegos y distracciones*' y '*Profesiones y oficios*', cuadro 13.

Centro de interés	Índice de cohesión Generación de inmigrantes		
	Primera	Segunda	Tercera
El cuerpo humano	0.049	0.044	0.045
La ropa	0.021	0.019	0.021
Partes de la casa	0.021	0.019	0.021
Los muebles de la casa	0.019	0.017	0.018
Comidas y bebidas	0.021	0.020	0.020
Objetos colocados en la mesa...	0.025	0.024	0.025
La cocina y sus utensilios	0.021	0.019	0.021
La escuela (muebles y materiales)	0.023	0.021	0.023
Calefacción, iluminación y...	0.012	0.009	0.012
La ciudad	0.012	0.011	0.013
El campo	0.022	0.019	0.022
Medios de transporte	0.020	0.019	0.023
Trabajos del campo y del jardín	0.006	0.006	0.007
Los animales	0.036	0.016	0.034
Juegos y distracciones	0.011	0.010	0.012
Profesiones y oficios	0.013	0.011	0.013

**Cuadro 13: Índice de cohesión según factor generación de inmigrantes**

En el contraste de hipótesis, los índices de probabilidad hallados revelan la existencia de cinco valores no significativos, cuadro 17. Estas cinco diferencias podrían estar ocasionadas por diversas causas. Aparecen once diferencias significativas y esto permite refutar la hipótesis nula y aceptar la hipótesis alternativa. Los datos sugieren que existe una correlación entre la producción léxica y las generaciones de inmigrantes. Esta correlación presenta particularidades relacionadas con las lenguas en contacto. En la comunidad hispana de Redwood City como se ha señalado la mayor producción se encuentra en el subgrupo de la primera generación. La segunda es la que posee los valores más bajos tanto en los promedios de respuestas como en los índices de cohesión. Asimismo llama la atención el hecho que la tercera generación posee valores más altos que la segunda generación.

## FACTOR LUGAR DE NACIMIENTO

El factor lugar de nacimiento resulta relevante en este estudio por tratarse de una población de inmigrantes. El lugar de origen de los informantes revela el aporte que cada dialecto ofrece al léxico que se recupera en este estudio.

Según los resultados alcanzados en este estudio, los promedios de los informantes nacidos en Estados Unidos son inferiores a los encuestados que nacieron en países hispanos en todas las áreas temáticas, cuadro 14.

Separado por dos puntos a favor de los informantes nacidos en países hispanos se reflejan los centros temáticos', 'El cuerpo humano', 'Los muebles de la casa', 'La escuela: muebles y materiales' y 'Los animales'.

Centro de interés	Promedio de respuestas		Indices de cohesión	
El cuerpo humano	18.49	16.64	0.045	0.048
La ropa	14.40	13.06	0.019	0.021
Partes de la casa	12.00	10.80	0.019	0.020
Los muebles ...	10.00	8.75	0.017	0.019
Comidas y bebidas	19.38	18.43	0.020	0.021
Objetos...	9.59	9.28	0.024	0.025
La cocina y ...	11.85	11.00	0.019	0.021
La escuela ...	15.82	14.06	0.021	0.024
Calefacción, ...	6.91	5.71	0.009	0.012
La ciudad	14.59	13.60	0.011	0.011
El campo	13.06	11.51	0.010	0.011
Medios de transporte	12.23	11.32	0.019	0.020
Trabajos y...	7.57	6.64	0.006	0.006
Los animales	20.73	18.24	0.032	0.035
Juegos y ...	12.82	11.90	0.010	0.011
Profesiones y oficios	13.89	11.91	0.011	0.013
PROMEDIO GLOBAL	13.33	12.05		

**Cuadro 14: Promedio de respuestas, factor Lugar de nacimiento**

Los índices de cohesión muestran superioridad en la concreción semántica en casi todas las áreas temáticas excepto en las áreas '*La ciudad*' y '*Trabajos del campo y el jardín*' donde los índices son iguales.

## FACTOR NIVEL SOCIOCULTURAL

Es considerado uno de los factores más importantes en los estudios sociolingüísticos. La medición de este parámetro a través de la encuesta de disponibilidad léxica contempla los estudios de los padres y las labores que desempeñan. Para este estudio, se obtuvieron 271 (52.30%) para el nivel bajo, 111 (21.40%) para el nivel medio bajo, 111 (21.40%) para el nivel medio alto y 25 (4.80%) para el nivel alto.

Los promedios de respuestas no son proporcionales a los niveles en todas las áreas. Los promedios entre los niveles bajos favorecen al nivel medio alto. Sin embargo, entre los niveles medio bajo y medio alto solo las áreas: '*El cuerpo humano*' y '*Objetos colocados en la mesa para la comida*' presentan valores superiores en el nivel medio alto. Al comparar los valores entre los niveles medio alto y alto se puede concluir que no todos los promedios son superiores en el nivel alto. Las áreas '*La ropa*', '*Partes de la casa*', '*Los muebles de la casa*', '*Objetos colocados en la mesa para la comida*', '*La cocina y sus utensilios*', '*La ciudad*' y '*Profesiones y oficios*' '*Calefacción, iluminación y medios de airear un recinto*' presentan promedios superiores en el nivel medio alto, cuadro15.

Centro de interés	Promedio de respuestas			
	Nivel bajo	Medio bajo	Medio alto	Alto
El cuerpo humano	17.21	17.53	17.7	19.2
La ropa	13.21	14.68	14	13.36
Partes de la casa	11.08	12.59	10.97	10.88
Los muebles de la casa	9.16	9.74	9.49	8.88
Comidas y bebidas	18.26	20.51	18.22	20.92
Objetos colocados en ...	9.44	9.34	9.57	9.16
La cocina y sus utensilios	11.18	12.05	11.45	10.76
La escuela...	14.44	16.03	14.68	15.84
Calefacción, iluminación y...	6.27	6.41	6.29	6
La ciudad	13.54	15.34	14.11	14.8
El campo	11.89	12.86	12.31	13.52
Medios de transporte	11.42	12.4	12.04	12.2
Trabajos del campo y ...	7.05	7.32	6.89	7.6
Los animales	19.34	19.7	18.95	20.92
Juegos y distracciones	11.57	13.86	12.29	14.52
Profesiones y oficios	12.16	13.94	13.42	13.16
PROMEDIO GLOBAL	12.33	13.39	12.65	13.23

**Cuadro 15: Promedio de respuestas según el factor Nivel sociocultural**

Hay cuatro áreas en las cuales los promedios del nivel alto son más bajos que el nivel bajo: 'La ropa', 'Objetos colocados sobre la mesa para la comida', 'La cocina y sus utensilios', 'Calefacción, iluminación y medios de airear un recinto', cuadro 16.

Estos resultados sugieren que el nivel sociocultural no ejerce influencia en la producción léxica de la población de Redwood City. Sin embargo, se continúan explorando los resultados de los índices de cohesión y las significaciones estadísticas con el análisis de otros procedimientos estadísticos.

Al analizar los índices de cohesión se aprecia una supremacía en el nivel alto. El nivel bajo muestra asociación semántica más divergente. Existe igualdad en los índice de cohesión semántica para los niveles medio excepto para el área Comidas y bebidas en la cual el nivel medio bajo supera al nivel medio alto. Esto sugiere que el sociolecto medio manifiesta similar convergencia asociativa. Se destacan como los menos compactos, en general, los centros de interés 'Calefacción, iluminación y medios de airear un recinto', 'La ciudad', 'El campo', 'Trabajos del campo y el jardín', 'Juegos y distracciones' y 'Profesiones y oficios', cuadro 16.

Centro de Interés	Bajo	Medio bajo	Medio alto	Alto
El cuerpo humano	0.05	0.09	0.09	0.19
La ropa	0.03	0.05	0.05	0.16
Partes de la casa	0.06	0.10	0.10	0.22
Los muebles de la casa	0.06	0.09	0.09	0.22
Comidas y bebidas	0.06	0.11	0.09	0.24
Objetos colocados...mesa...	0.08	0.11	0.11	0.26
La cocina y sus utensilios	0.06	0.09	0.09	0.20
La escuela ...	0.06	0.10	0.10	0.27
Calefacción, iluminación y ...	0.02	0.03	0.03	0.09
La ciudad	0.02	0.03	0.03	0.09
El campo	0.02	0.03	0.03	0.07
Medios de transporte	0.03	0.05	0.05	0.14
Trabajos del campo y ...	0.01	0.02	0.02	0.06
Los animales	0.05	0.07	0.07	0.15
Juegos y distracciones	0.02	0.03	0.03	0.08
Profesiones y oficios	0.02	0.03	0.03	0.08

**Cuadro 16: Índices de cohesión según el factor nivel sociocultural**

Al interpretar los grados de significación entre los promedios de respuestas se encuentran 10 valores estadísticos no significativos, es decir diez áreas temáticas cuyos valores del estadístico de contraste se hallan próximos a la afirmación establecida en  $H_0$ . Compatible con los datos estadísticos recogidos se sugiere que las causas de la producción del léxico disponible probablemente han estado condicionadas por el azar o por causas accidentales en la mayoría de los centros de interés por lo tanto se plantea como regla de decisión la aceptación de la hipótesis nula: los valores de frecuencia obtenidos no se deben a la influencia específica del nivel sociocultural.

Solo las áreas temáticas '*La ropa*', '*Partes de la casa (sin muebles)*', '*Comidas y bebidas*', '*La ciudad*', '*Juegos y distracciones*' y '*Profesiones y oficios*' presentan valores significativos ( $p$  igual o menor de 0.05). De esta manera se rechaza la hipótesis alternativa que indica que la disponibilidad léxica estará correlacionada con el nivel sociocultural de los hablantes.

Los hallazgos favorecen el presupuesto de que en el léxico disponible intervienen una serie de factores no solamente el nivel sociocultural (López Morales, 1999).

## VALORACIÓN ESTADÍSTICA DE LOS RESULTADOS

Se estima oportuno en este punto del análisis realizar una breve consideración general de los resultados, antes de entrar al análisis de las correlaciones que se exploraran en el siguiente apartado agrupando los centros de interés de acuerdo a su significación estadística y se han establecidos tres grupos, (Gómez Molina, J.R. & Gómez Devís, M<sup>a</sup>. B. 2004).

- Primero los centros con mayor estratificación social significativa lo cual facilita extraer conclusiones estadísticas para describir el léxico disponible la población estudiantil de Redwood City: '*El cuerpo humano*', '*La ropa*', '*Partes de la casa (sin muebles)*', '*Los muebles de la casa*', '*La escuela: muebles y materiales*', '*Medios de transporte*', '*Los animales*' y '*Profesiones y oficios*'.
- Segundo centros de interés con mayor inestabilidad que no facilitan hacer inferencias estadísticas sobre la estratificación sociodialectal de la comunidad de Redwood City, ellos son: '*Comidas y bebidas*', '*La cocina y sus utensilios*', '*Calefacción, iluminación y medios de airear un recinto*', '*La ciudad*' y '*El campo*'.
- Tercero. Este bloque lo forman las tres áreas temáticas con menor significación estadística: '*Objetos colocados en la mesa para la comida*', '*Trabajos del campo y el jardín*' y '*Juegos y distracciones*'. En estas áreas hay mayor incidencia de causas accidentales por lo que no permiten hacer inferencias estadísticas para el conjunto de los estudiantes.

preuniversitarios de Redwood City.

Se generalizan además la incidencia de cada uno de los factores para determinar el peso que tienen en los inventarios léxicos recogidos. Los factores nivel de español y generación de inmigrantes, con el mismo peso, son los que presentan mayor incidencia. Le siguen los factores sexo y lengua habitual también con el mismo peso. Queda con menor proporción de inferencia estadística el factor nivel sociocultural.

Los valores de 'p' obtenidos mediante el procesamiento estadístico ANOVA sugieren que todos los centros de interés poseen valores significativos con lo cual se refuta la hipótesis en cinco de las variables estudiadas. Solo se acepta la hipótesis nula en la variable nivel sociocultural cuadros 17 y 18. Estos valores asumen que se desestima la hipótesis nula admitiendo la contraria o alternativa. Se establece entonces que la asociación es cierta en un 95% en las variables: sexo, lengua habitual, nivel de español, lugar de nacimiento y generación de inmigrantes.

Centro de interés	Significación de las diferencias	
	No significativas	Significativas
El cuerpo humano	1	5
La ropa	0	6
Partes de la casa	1	5
Los muebles ...	1	5
Comidas y bebidas	3	3
Objetos...	4	2
La cocina y ...	3	3
La escuela ...	1	5
Calefacción, ...	2	4
La ciudad	3	3
El campo	2	4
Medios de transporte	1	5
Trabajos y...	4	2
Los animales	1	5
Juegos y ...	4	2
Profesiones y oficios	1	5
	32	64

**Cuadro 17: Significación de las diferencias según los centros de interés.**

Variables sociológicas	No significativas	Significativas
Sexo	0	16
Lengua habitual	6	10
Nivel de español	7	9
Generación de inmigrantes	5	11
Lugar de nacimiento	7	9
Nivel sociocultural	10	6
Total	35	61

**Cuadro 18: Significación de las diferencias según los factores sociológicos.**

Se puede inferir que en el comportamiento del léxico disponible de los estudiantes hispanos de Redwood City inciden múltiples factores sociológicos.

## CORRELACIÓN ENTRE LÉXICO DISPONIBLE Y FACTORES SOCIALES

Este trabajo ha querido aportar un análisis valorativo de la incidencia de las variables sociológicas en la estratificación sociodialectal de una comunidad hispana en contacto con otra lengua. De esta manera, se sigue la idea ya explorada en la caracterización de la comunidad valenciana, donde el español también está en contacto con otra lengua (catalán). Este tipo de análisis complementario será un aporte para describir el léxico disponible de los hispanos en Estados Unidos. Desde el punto de vista sociolingüístico es fundamental establece una relación entre las lexías que se han ofrecido en cada área temática y los factores sociológicos establecidos.

### ANÁLISIS FACTORIAL SIMPLE

En los cálculos se han analizado las correlaciones específicas entre cada variable social y todas las áreas temáticas. Se comprobó que:

- La variable *sexo* ofrece valores de F y su significación correspondiente (valores de '*p*') por debajo de 0.05 en los dieciséis centros temáticos con lo cual se advierte una diferencia estadísticamente significativa en todas las áreas temáticas quedando rechazada la hipótesis nula.
- La variable *lengua habitual* ofrece valores de F y su significación correspondiente (valores de '*p*') por debajo de 0.05 en diez centros temáticos lo cual indica diferencia estadísticamente significativa en esas diez áreas temáticas. Por lo tanto se rechaza la hipótesis nula.
- La variable *lugar de nacimiento* indica valores de F y su significación correspondiente (valores de '*p*') por debajo de 0.05 en diez áreas temáticas queda entonces rechazada la hipótesis nula.
- La variable *nivel de español* indica valores de F y su significación correspondiente (valores de '*p*') por debajo de 0.05 también en nueve áreas temáticas rechazada así la hipótesis nula.
- La variable *generación de inmigrantes* ofrece de F y su significación correspondiente (valores de '*p*') por debajo de 0.05 en once áreas temáticas con esto se está rechazando también la hipótesis nula.
- La variable *nivel sociocultural* ofrece valores de F y su significación correspondiente (valores de '*p*') por debajo de 0.05 solo en seis áreas temáticas con lo cual acepta la hipótesis nula.
- Tomando en cuenta la jerarquización de la razón F y de su correspondiente significación en cada área temática según las variables sociales se concluye que:
- Existe correlación entre cada una de las variables sociales y los centros temáticos en la producción léxica de los encuestados.
- La variable social más significativa es *sexo* que influye en los dieciséis centros temáticos, le siguen en orden *generación de inmigrantes*, *lengua habitual* y con la misma incidencia *lugar de nacimiento* y *nivel de español*.
- La variable *nivel sociocultural* tiene menor influencia en esta sintopía incidiendo solo en seis de las áreas temáticas.

### ANÁLISIS DE REGRESIÓN LINEAL MÚLTIPLE

Al partir de una relación causal entre las variables explicativas y las explicadas, la aplicación de este instrumento estadístico, ha permitido explicar el comportamiento de cada uno de los centros léxicos de este estudio.

A continuación se relacionan las variables independientes, factores sociológicos que mayor incidencia tienen en los inventarios léxicos recopilados según los quince centros de interés restantes.

- En el área nocional '*La ropa*' son: *sexo*, *lengua habitual* y *lugar de nacimiento*.
- En el área nocional '*Partes de la casa (sin muebles)*' son: *sexo*, *lengua habitual* y *generación de inmigrantes*.
- En el área nocional '*Los muebles de la casa*' son: *sexo*, *lengua habitual* y *lugar de nacimiento*.
- En el área nocional '*Comidas y bebidas*' son: *sexo*, *lengua habitual* y *generación de inmigrantes*.

- En el área nocional '*Objetos colocados en la mesa para la comida*' son: *sexo, lengua habitual y nivel de español*.
- En el área nocional '*La cocina y sus utensilios*' son: *sexo, lengua habitual y lugar de nacimiento*.
- En el área nocional '*La escuela: muebles y materiales*' son: *sexo, lengua habitual y lugar de nacimiento*.
- En el área nocional '*Calefacción, iluminación y medios de airear un recinto*' son: *sexo, lugar de nacimiento y lengua habitual*.
- En el área nocional '*La ciudad*' son: *sexo, nivel de español y lengua habitual*.
- En el área nocional '*El campo*' son: *sexo, lengua habitual y nivel de español*.
- En el área nocional '*Medios de transporte*' son: *sexo, lengua habitual y nivel de español*.
- En el área nocional '*Trabajos del campo y el jardín*' son: *sexo, lengua habitual y nivel de español*.
- En el área nocional '*Los animales*' son: *sexo, generación de inmigrantes y lengua habitual*.
- En el área nocional '*Juegos y distracciones*' son: *sexo, lengua habitual y nivel sociocultural*.
- En el área nocional '*Profesiones y oficios*' son: *sexo, lugar de nacimiento y lengua habitual*.

Finalizada la interpretación total de los resultados para cada una de los centros de interés y los factores sociolingüísticos de este estudio se concluye que:

- Todas las variables han demostrado incidencia en la producción léxica de la sintopía.
- Las correlaciones entre las variables independientes o factores sociolingüísticos y las variables dependientes o centros de interés ofrecen tres tipos de relaciones que se han denominado alta, media y baja. Las relaciones altas marcan una incidencia relevante de dos factores sociolingüísticos en la producción del léxico disponible de la comunidad estudiada. La relación media baja indica una vinculación menos fuerte, pero considerable de cómo dos factores influyen en la producción léxica. Por otro lado, la relación baja sugiere que dos de los factores considerados ejercen una influencia baja e insignificante para la producción del léxico en la sintopía estudiada.
- Las variables sociolingüísticas con mayor valor de predicción en el léxico disponible obtenido son las siguientes: las variables *sexo* y *lengua habitual* que inciden en los dieciséis centros de interés.
- Según su peso específico la variable *lugar de nacimiento* incide en seis de los centros de interés; le sigue la variable *nivel de español* que contribuye en cinco centros de interés.
- Un peso más bajo tiene la variable *generación de inmigrantes* que solo incide en la producción léxica de cuatro centros de interés.
- La incidencia más baja la aporta la variable *nivel sociocultural* que incide solo en un centro temático (*Juegos y distracciones*).

## COEFICIENTE DE CORRELACION MULTIPLE: SPEARMAN

El cálculo del coeficiente Spearman se ha podido comprobar que, en general, las relaciones entre las variables son despreciables. La relación entre las variables independientes muestra una sola correlación: entre *lugar de nacimiento* y *lengua habitual*. La correlación entre las variables dependientes indica que según los coeficientes R obtenidos no se presenta una relación marcada entre ninguno de las áreas temáticas.

- Hay relación entre diez centros y las variables sociales. La relación se mantiene predominantemente entre dos factores sociales. Solamente dos centros se relacionan con más de un factor sociológico. El área nocional '*El cuerpo humano*' está correlacionada con *lengua habitual* y *lugar de nacimiento*.
- El área nocional '*Partes de la casa*' y *lugar de nacimiento*.
- El área nocional '*Comidas y bebidas*' está correlacionada con *lengua habitual*.
- El área nocional '*Objetos colocados en la mesa para la comida*' está correlacionada con *lengua habitual* y *nivel de español*.
- El área nocional '*La cocina y sus utensilios*' está correlacionada con *generación de inmigrantes, lengua habitual* y *lugar de nacimiento*.
- El área nocional '*Calefacción, iluminación y medios de airear un recinto*' está correlacionada con *nivel de español*.

- El área nocional 'La ciudad' está correlacionada con *nivel sociocultural*.
- El área nocional 'EL campo' está correlacionada con *lugar de nacimiento*.
- El área nocional 'Medios de transporte' está correlacionada con *lengua habitual* lugar de nacimiento.
- El área nocional 'Trabajos del campo y del jardín' está correlacionada con *generación de inmigrantes*.
- El área nocional 'Juegos y distracciones' está correlacionada con *generación de inmigrantes*.

Los valores de '*p*' calculados para determinar la significación estadística de las correlaciones múltiples entre ambas magnitudes indican que se rechaza la hipótesis nula con un elevado número de valores significativos: 192 correlaciones significativas frente a 38 correlaciones no significativas.

Hay cinco centros de interés que no ofrecen correlación alguna con ninguna de las variables: '*La ropa*', '*Los muebles de la casa*', '*La escuela: mueble y materiales*', '*Los animales*' y '*Profesiones y oficios*'.

En cuanto a la relación entre las variables independientes solo se da relación entre *lengua habitual* y *lugar de nacimiento*.

La correlación entre las variables independientes es la siguiente: la variable *nivel de español* está correlacionada con cinco centros de interés, la variable social *generación de inmigrantes* establece correlación con tres centros de interés y la variable social *nivel sociocultural* solo se correlaciona con un centro de interés: '*Juegos y distracciones*'. La variable independiente *sexo* ofrece una correlación nula, no está correlacionada con ninguna otra variable dependiente o independiente.

Los grados de asociación que reflejan los índices de correlación de Spearman no son muy sustanciales o marcado, en ningún caso los valores superan 0.50.

Aparecen dos variables como las más homogeneizadoras: *lengua* y *lugar de nacimiento* que establecen correlación con cinco centros de interés y entre ellas mismas. Al tratarse de dos factores relacionados entre sí es sugerente la correlación entre ellos. Se trata de dos variables que tienen algo común entre ellas: rasgos comunes compartidos, productividad léxica. Es indicador sugerente del peso que tiene en la estratificación sociodialectal la *lengua habitual* y el lugar donde nacieron los informantes pues ello determina el predominio de una de las dos lenguas en el uso cotidiano y el dominio del léxico.

## ANÁLISIS FACTORIAL DE COMPONENTES PRINCIPALES

Se parte de la idea de que las variables que integran este estudio están correlacionadas por lo tanto el propósito fundamental del análisis de componentes principales es encontrar componentes o factores que sucesivamente expliquen la mayor parte de la varianza total. Se busca hallar combinaciones lineales de variables originales que sirvan para comprender la mayoría de la variación total.

- El primer factor explica el 37.42% de la varianza total. Resulta interesante que todos los centros de interés (variables dependientes) tienen valores por encima de (.588) La contribución (comunalidad) más alta corresponde al área temática '*El campo*' (.790) seguido de, '*La ciudad*' (.769) y '*Medios de transporte*' (.759), Los valores más bajos recaen en las áreas temáticas: '*Profesiones y oficios*', '*El cuerpo humano*' y '*Trabajos del campo y el jardín*'.
- El segundo factor explica el 45.04% de la varianza total. Las contribuciones más altas corresponden a las variables sociales: *lugar de nacimiento* (.796), *generación de inmigrantes* (-.790) y *nivel de español* (.509).
- El tercer factor explica el 50.62% del total de la varianza y contribuyen las variables independientes: *nivel de español*, *nivel sociocultural* y *lengua habitual* que al mismo tiempo ejercen la mayor contribución.

---

## CONCLUSIONES

Se ha obtenido un corpus del léxico disponible de los estudiantes hispanos de Redwood City en California en el cual se ha comprobado que no existe riqueza léxica como en otras sintopías estudiadas. Sin embargo, los vocablos recogidos reflejan un español patrimonial con mínima incidencia de préstamos lingüísticos.

El estudio permitió conocer la disponibilidad léxica de los hablantes hispanos de la ciudad de Redwood City mediante la aplicación de la metodología léxico estadística. Se contribuyó al conocimiento de la variedad (diversidad) dentro de la unidad de la lengua española, en un ámbito de lenguas en contacto mediante la obtención de inventarios de los listados de vocabulario disponible de los hablantes hispanos de Redwood City.

Se ha podido establecer que los centros de interés en su conjunto y de manera particular contribuyen a la mejor descripción del léxico disponible en la comunidad estudiada. La estratificación sociolingüística de la disponibilidad léxica tomando en consideración variables sociales tradicionales (sexo, lengua habitual, nivel sociocultural y generación de inmigrantes) y novedosas como lugar de nacimiento y nivel de español. Existe además una incidencia de cada una de las seis variables sociales. Los valores más altos corresponden a las variables sociales *sexo, lengua habitual, lugar de nacimiento, nivel de español y generación de inmigrantes*. La variable social *sexo* es la que más valor tiene y se ubica en el cuarto factor. La variable nivel sociocultural no es discriminadora en esta sintopía.

Este estudio posibilita la aplicación pedagógica de los resultados como material útil para la planificación léxica en el proceso educativo de nuestros estudiantes de preuniversitario.

Se aspira a que este estudio sea solo el comienzo y una porción del estudio del español de Estados Unidos. Al incorporar una comunidad hispana en contacto con inglés en esta región del mundo, hoy se está en mejores condiciones de continuar describiendo el léxico disponible de los hispanos en Estados Unidos. Se aspira a que este trabajo permita la continuidad de otros estudios de disponibilidad léxica tanto en el ámbito de estudiantes nativos como no nativos. Además, se espera que pueda servir de referencia en los trabajos de estudio de disponibilidad léxica en lenguas en contacto. Asimismo, se quiere que este estudio permita la comparación con otras isoglosas en Estados Unidos. Queda abierto un trecho en el empeño sociolingüístico por describir la variación léxica del español del otro lado del Atlántico.

## REFERENCIAS BIBLIOGRÁFICAS

- AHUMADA, I. (2006). El léxico disponible de los estudiantes preuniversitarios de la provincia de Jaén. Jaén, Publicaciones de la Universidad de Jaén.
- ALBA, O. (1995). El léxico disponible de la República Dominicana. Santiago de los Caballeros, República Dominicana, PUCMM.
- ARNAL, Ma. L. et al (2004). Léxico disponible de Aragón. Zaragoza. España, Libros Pórtico.
- ASOCIACIÓN DE ACADEMIAS DE LA LENGUA ESPAÑOLA. (2010). Diccionario de americanismos. [Lima], Santillana.
- ÁVILA, A. M. (2006). Léxico disponible de los estudiantes preuniversitarios de Málaga. Málaga, Universidad de Málaga, Servicio de Publicaciones e Intercambio Científico.
- AYORA, Ma. C. (2006). Disponibilidad léxica en Ceuta: aspectos sociolingüísticos, Cádiz, Servicio de Publicaciones de la Universidad de Cádiz.
- BARTOL, J. A. (2004). Léxico disponible de Soria: estudio y diccionarios. Colección Beltenebros, 7. Burgos, Fundación Instituto Castellano y Leonés de la Lengua.
- BARTOL, J. A. (2003) "Léxico disponible y norma lingüística", Moreno, F. et al, Lengua, variación y contexto. Homenaje a H. López Morales, Madrid, Arco Libros, 127-144, vol. I.
- BARTOL, J. A. (2001) "Reflexiones sobre la disponibilidad léxica". Bartol, J. A. et al (eds) Nuevas aportaciones al estudio de la lengua española. [Salamanca], Lusu-Española de Ediciones.
- BERNSTEIN, B. (1973). Class, codes and control Vol. 1, Theoretical studies towards sociology of language. St. Albans, Paladin.
- CALIFORNIA DEPARTMENT OF EDUCATION [en línea]. <[www.cde.ca.gov](http://www.cde.ca.gov)> [Consulta: 17 de agosto 2008].
- CARCEDO, A. (2000). Disponibilidad léxica en español como lengua extranjera: el caso finlandés: (estudio del nivel

preuniversitario y cotejo con tres fases de adquisición). Turku, Turun Yliopisto.

- CARCEDO, A. (2000) "La lengua como manifestación de otredad cultural (o convergencia intercultural)" [en línea]. <<http://www.ucm.es/info/especulo/ele/carcedo.html>> [Consulta: 22 de marzo de 2008]
- CARCEDO, A. (2001). Léxico disponible de Asturias. Universidad de Turku, Publicaciones del Dpto. de Lengua Española, Monografías, vol. 1.
- ECHEVERRÍA, M. (1991) "Crecimiento de la disponibilidad léxica en estudiantes chilenos de nivel básico y medio", López Morales, H. (ed): La enseñanza del español como lengua materna. Actas del II Seminario Internacional sobre "Aportes de la Lingüística a la Enseñanza del Español como Lengua Materna". Río Piedras, P.R., Editorial de la Universidad de Puerto Rico.
- ECHEVERRÍA, M. et al. (1987) "Disponibilidad léxica en Educación Media", Revista de Lingüística Teórica y Aplicada, 25, 55-116.
- ECHEVERRÍA, M. et al. (2008) "DispoGrafo: una nueva herramienta computacional para el análisis de relaciones semánticas en el léxico disponible". RLA, Revista de Lingüística Teórica y Aplicada, 46 (1), I Sem. 2008, pp. 81-91. [en línea]. <<http://www.articlearchives.com/>> [Consulta: 23 de septiembre de 2008]
- EDUCATIONAL DATA ABOUT CALIFORNIA K-12. [en línea]. <<http://www.ed-data.k12.ca.us>> [Consulta: 18 de agosto de 2009].
- GALLOSO, M<sup>a</sup>. V. (2004). Léxico disponible de Ávila, Salamanca y Zamora. Fundación Instituto Castellano Leonés de la Lengua.
- GALLOSO, M. V. (2002). El léxico disponible en el nivel preuniversitario: provincia de Zamora: trabajo de grado. Huelva, Universidad de Huelva.
- GALLOSO, M. V. (1998) "Hacia una definición de léxico disponible: posibles aplicaciones", Actas del III Congreso de Lingüística General. Lingüística para el siglo XXI. Salamanca: Universidad de Salamanca, 111-117.
- GÓMEZ MOLINA, J. R., & GÓMEZ DEVÍS, M. B. (2004). La disponibilidad léxica de los estudiantes preuniversitarios valencianos: estudio de estratificación sociolingüística. Quaderns de filología / Universitat de València, 56. Valencia, Universitat de València.
- GOUGENHEIM, G., MICHEA, R., RIVENC, P. y SAUVAGEOT, A. (1956). L'élaboration du français élémentaire. Étude sur l'établissement d'un vocabulaire et d'une grammaire de base, Paris: Didier.
- LÓPEZ, J. (1992): "Alcances panhispánicos del léxico disponible", Lingüística: 4, 26-124.
- LÓPEZ, H. (2009). Enciclopedia del español en los Estados Unidos. Madrid, Instituto Cervantes.
- LÓPEZ, H. Discurso pronunciado con motivo de su investidura como Doctor Honoris Causa en la Universidad de Alicante el 23 de marzo de 2000 [en línea]. <<http://www.ua.es>> [Consulta: 23 de abril de 2009].
- LÓPEZ, H. (2004). Sociolingüística. Biblioteca románica hispánica, 70. Madrid, Editorial Gredos.
- LÓPEZ, H. (1999). Léxico disponible de Puerto Rico, Madrid, Arco Libros.
- LÓPEZ, H. (1995) "Los estudios de disponibilidad léxica: pasado y presente", BFUCH (Homenaje a Rodolfo Oroz), 35,245-259
- LÓPEZ, H. (1983). Estratificación social del español de San Juan De Puerto Rico.
- MORENO DE ALBA, J. G. Y PERISINOTO, G. (1988) "Observaciones sobre el español de Santa Bárbara, California (Nueva Rev. de Filol. Hispánica, XXXVI [1988], pp.171-201)
- MORENO, F. (2008) "Dialectología hispánica de los Estados Unidos" en LÓPEZ MORALES, H. (eds). Enciclopedia del

español en los Estados Unidos. Madrid, Instituto Cervantes.

- MORENO, F. (2006) "Lexidisp. Cálculo de disponibilidad léxica". [en línea]. <<http://linguas.net/Proyectos/LexiDisp/tabid/73/Default.aspx>> [Consulta: 30 de octubre 2007]
- MORENO, F. (2005). Principios de sociolingüística y sociología del lenguaje. Ariel lingüística. Barcelona, Ariel.
- MORENO, F. (2003) "Anglicismos en el léxico disponible de los adolescentes de Chicago" en POTOWSKI, K. & CAMERON, R. (2007) Spanish in contact: policy, social and linguistic inquiries, Amsterdam; Philadelphia, John Benjamins Pub. Co. 41-58.
- MORENO, F. (1986) "Intercorrelaciones lingüísticas en una comunidad rural". Revista Española de Lingüística Aplicada, (2) 87-107
- NORUSÍIS, M. J. (2008). SPSS statistics 17.0 guide to data analysis. Upper Saddle River, N.J. Prentice Hall.
- SAMPER, J. A. (1990). Estudio sociolingüístico del español de Las Palmas de Gran Canaria. Santa Cruz de Tenerife, La Caja de Canarias.
- SAMPER, J. A. (1998) "Criterios de edición del léxico disponible", Lingüística, 10, 311-333.
- SAMPER, J. A. Et al. (2003) "El proyecto de estudio de la disponibilidad léxica en español", Raúl Ávila, José A. Samper y Hiroto Ueda, Pautas y pistas en el análisis del léxico hispano (americano), Frankfurt-Madrid, Vervuert Iberoamericana, 27-140.
- SAMPER, J. A. Y CLARA HERNANDEZ (1997) "El estudio de la disponibilidad léxica en Gran Canaria: datos iniciales y variación sociolingüística". Contribuciones al estudio de la lingüística hispánica, Homenaje al profesor Ramón Trujillo, Santa Cruz de Tenerife: Montesinos, vol: 229-239
- URZUA C, P.; SAEZ, K.; ECHEVERRIA, M. S. "Disponibilidad léxica matemática: análisis cuantitativo y cualitativo". RLA, Concepción, v. 44, n.22, 2006. [en línea]. <<http://www.scielo.org>>

## ANEXO

- [1] Fuentes: Aragón(2004), Valencia (2004), Cuenca(2004), Soria (2004), Málaga(2006) Ceuta(2006).
- [2] Bartol (2004:52) ofrece un cuadro comparativo de los estudios de las sintopías: Cuenca, Asturias, Cádiz, Ávila, Salamanca y Zamora; Hernández (2004:59) compara Cuenca, Ávila, Salamanca y Zamora. Gómez Molina, J.R. & Gómez Devís, M<sup>a</sup>.B. (2004:84) Valencia, Asturias, Madrid, Salamanca y Cádiz; Galloso (2003:77) Compara solo 3 sintopías: Ávila Zamora y Salamanca. Este estudio recoge las comparaciones de estudios anteriores e incorpora los resultados de trabajos posteriores a Asturias.
- [3] Fuente: Samper et al (2003) En los estudios del léxico disponible el área nocional '*Los animales*' difiere en rango solo en las sintopías de Almería y Bilbao donde se sitúa en segundo puesto en ambos estudios. El centro '*El campo*' tiene un rasgo diferente en Puerto Rico(13), República Dominicana(10), Madrid(14) y Almería(5)
- [4] Fuente: Samper et al (2003) El centro nocional '*Comidas y bebidas*' no ocupa el mismo rango en Puerto Rico (3) Almería (1) y Bilbao(1)
- [5] Los resultados de Redwood City y los estudios precedentes del Proyecto Panhispánico de Léxico Disponible difieren en el caso de los centros de interés con índices de cohesión más bajos.
- [6] Fuente: Gómez Molina & Gómez Devís, M<sup>a</sup>.B. (2004:91) No encuentran influencia significativa Gougenheim et al (1956); Alba (1995); Mateo 1998; Carcedo (2001) Reconocen influencia del factor sexo (Román, 1985; Samper. & Hernández (1997); Galloso (2002).