

LA FRECUENCIA DE LAS PALABRAS EN LOS PROCESOS DE VARIACIÓN Y CAMBIO*

MANUEL ALMEIDA
Universidad de La Laguna

RESUMEN

La frecuencia de las palabras desempeña un papel importante en la organización y funcionamiento de las lenguas. En esta investigación se analiza la relación entre frecuencia y dos tipos de variación fonética: el mantenimiento y pérdida de /d/ intervocálica y el uso de las variantes laxa y tensa de /tʃ/. Los resultados indican que la frecuencia posee un papel activo en la organización alofónica de ambos fonemas, pero también muestran que este factor puede verse modificado (es decir, reforzado o debilitado) por la intervención de otras variables.

Palabras clave: frecuencia de la palabra, variación alofónica, contextos de variación.

ABSTRACT

Word frequency plays an important role in the organisation and functioning of languages. In this research the relationship between frequency and two kinds of phonetic alternations (the maintenance and loss of intervocalic /d/ and the use of the lax and tense variants of /tʃ/) will be analysed. Results show that frequency plays an active role in the allophonic organisation of both phonemes, but they also show that this role can be modified (that is to say, strengthened or weakened) by the action of other variables.

Key Words: word frequency, allophonic variation, contexts of variation.

RECIBIDO: 18/09/2012

APROBADO: 17/12/2012

1. INTRODUCCIÓN

Cuando en el discurso hablado una forma lingüística (sonido, palabra, categoría, estructura) se usa con más frecuencia que otra, es muy probable que ambas acaben adquiriendo propiedades diferentes tanto en el nivel lingüístico como en el cognitivo. Estudios llevados a cabo

* El autor agradece los comentarios hechos por dos evaluadores de la revista y por el profesor Carmelo P. Vidal a una anterior versión del artículo.

sobre el papel de la frecuencia constatan que los hablantes rentabilizan de muy diversos modos esta variable en áreas como la variación y el cambio lingüísticos, el aprendizaje de una segunda lengua, los juicios de aceptabilidad, la gramaticalización, el acceso al lexicón mental, etc. Por ejemplo, con respecto al modo en que los individuos llevan a cabo la selección léxica durante la planificación lingüística, se ha demostrado que el acceso a palabras de frecuencia alta suele ser más rápido que el acceso a palabras de frecuencia baja. El estudio de Chambers y Forster 1975, p. 551, demostró que el tiempo medio que empleaban los sujetos para acceder a palabras de alta frecuencia era de 649 msecs., mientras que el tiempo medio que empleaban para acceder a palabras de frecuencia baja era de 687 msecs. Estos resultados indican que, a fuerza de repetirse, las formas lingüísticas más frecuentes terminan por estar más reforzadas en la memoria y se vuelven por ello más prominentes, lo que permite que los individuos puedan acceder más fácilmente a ellas. A conclusiones semejantes se ha llegado a través del estudio de las pausas. Una de las funciones de las pausas es permitir a los hablantes seleccionar las palabras que han de ocupar las estructuras sintácticas del mensaje que se va a emitir a continuación. De acuerdo con lo que se acaba de exponer, parece verosímil que la selección de palabras menos frecuentes requiera un mayor número de pausas que la selección de palabras más frecuentes. Esta tesis ha sido corroborada en gran medida por Klatt. En un estudio con sujetos afásicos, este investigador observó que dichos individuos hacían más pausas ante los sustantivos, adjetivos y verbos menos frecuentes que ante los más frecuentes. Por ejemplo, los sustantivos menos frecuentes fueron precedidos por pausas en el 81,5% de los casos, mientras que ante sustantivos más frecuentes el número de pausas descendió al 52% (Klatt 1980, p. 116).

El modo en que los hablantes rentabilizan el poder organizador de la frecuencia se manifiesta en todos los niveles de la lengua. En el plano fónico se ha observado que los sonidos más complejos desde el punto de vista articulatorio tienden a ser menos frecuentes que los sonidos más simples. Por ejemplo, en la correlación aspirada/no aspirada que caracteriza a las oclusivas sordas de determinadas lenguas (p^h/p , t^h/t , k^h/k), las realizaciones aspiradas suponen un gasto de energía mayor que el de las no aspiradas correspondientes, ya que a los movimientos articulatorios característicos de las oclusivas «puras» hay que añadir los gestos correspondientes a la aspiración. Pues bien, en una lengua como el danés, cuyo sistema fonológico contiene este tipo de oposición, la frecuencia de las oclusivas aspiradas fue 1,05%, para $/p^h/$, 3,47% para $/t^h/$ y 2,10% para $/k^h/$, mientras que la frecuencia de las no aspiradas correspondientes fue 1,49%, 5,42% y 3,05% (Zipf 2002, p. 69).

En el plano sintáctico, un efecto de la productividad de la frecuencia se puede observar en los procesos de gramaticalización, que tanto

abundan en la historia de las lenguas. Rojo 2011 menciona al respecto la construcción *ir a* + infinitivo del español, que, en su origen, estuvo constituida por un verbo de movimiento seguido de la preposición *a* y de un infinitivo que expresaba la finalidad del movimiento (*fui a buscar el libro*). Cuando esta construcción se fue volviendo más frecuente el significado de movimiento real se fue perdiendo y la perífrasis fue adquiriendo el valor de «posterioridad», tras pasar por una fase intermedia en que expresaba «intencionalidad». Este cambio de significado de la construcción fue acompañado por un aumento en el número de verbos diferentes que pasaron a formar parte de ella. También se ha observado en el nivel sintáctico que cuando un elemento gramatical se vuelve más frecuente tiende a ir adquiriendo un mayor número de funciones sintácticas. Berkenfield 2001 cita al respecto el elemento gramatical inglés *that*. La información proporcionada por los estudios diacrónicos sugiere que esta partícula comenzó funcionando como pronombre demostrativo (*I want that*), pero actualmente puede desempeñar otras funciones: adjetivo (*that car, that boy*), marcador de cláusula relativa (*the book that I bought*) y elemento introductor de un complemento (*I thought that he had left*). Estas nuevas funciones puede decirse que son más gramaticalizadas que la primera, ya que en estos casos *that* tiene una menor carga semántica y una menor libertad de movimiento en el interior de la estructura sintáctica.

En el nivel semántico parece existir una relación directa entre la frecuencia de uso de una palabra y el carácter polisémico de esta, de modo que las palabras más frecuentes tienden a tener un mayor número de acepciones que las palabras menos frecuentes. A propósito del español, Almela y otros 2005, p. 27, mencionan que *cabeza*, una palabra frecuente en español, llega a tener hasta cien acepciones en el *Gran diccionario de usos del español actual*, mientras que *tacón*, con una frecuencia de uso baja, solo registra dos acepciones. También se observa que cuando una palabra se vuelve más frecuente es normal que vaya adquiriendo matices o sentidos cada vez más abstractos, como se acaba de ver en el caso del inglés *that*. No obstante, el significado básico suele ser el más frecuente. Así parece ocurrir en español, donde *cabeza*, con el significado primario «en las personas y animales, parte superior del cuerpo, donde están el cerebro y los principales órganos de los sentidos», registra el 58% de todas las acepciones de la palabra (Almela y otros 2005, p. 31). Este tipo de información es interesante en campos como la enseñanza de las lenguas, ya que permitiría que el aprendiz de una L2 conozca antes las acepciones más frecuentes (y, por tanto, las que le van a ser más necesarias) que las menos frecuentes.

Los investigadores suelen distinguir entre frecuencia de las formas lingüísticas (*token frequency*) y frecuencia del tipo de estructura en que aparecen dichas formas (*type frequency*) (Bybee 2003, pp. 11-13). La dife-

rencia entre estos dos modos de abordar el estudio de la frecuencia se ve claramente en la investigación de Díaz-Campos y Ruiz-Sánchez 2008 sobre elisión de /r/ final de sílaba en Andalucía y Caracas. Los autores analizaron en primer lugar cómo la elisión de /r/ implosiva estaba relacionada con la frecuencia de la palabra en que aparecía la consonante (*token frequency*) y, en segundo lugar, cómo esta relación entre elisión y frecuencia variaba en función de las diferentes categorías gramaticales: verbos, pronombres, preposiciones, adverbios, etc. (*type frequency*). Según Bybee, la simple frecuencia de las unidades lingüísticas tiene dos efectos en la lengua: el efecto reductor y el efecto conservador (véase sección 2), mientras que la frecuencia del tipo de estructura está más relacionada con la productividad de las construcciones lingüísticas. De este modo, cuantas más unidades estén presentes en un determinado esquema, más reforzado estará dicho esquema y más posibilidades tendrá este de encontrar nuevas aplicaciones.

2. LA FRECUENCIA EN LA PRODUCCIÓN LINGÜÍSTICA

Desde el punto de vista de la producción, dos son las consecuencias más evidentes de la frecuencia de las unidades lingüísticas: el efecto reductor y el efecto conservador (Bybee 2002, 2003, pp. 11-13, 2006).

2.1. *El efecto reductor*

La frecuencia puede ejercer un efecto reductor sobre el significante de las formas lingüísticas, es decir, puede provocar la pérdida de cierta cantidad de material fonético. De hecho, fenómenos como los de debilitamiento y elisión de sonidos tienden a producirse en las palabras y estructuras más frecuentes, como demuestran diversos estudios. Hooper 1976, por ejemplo, llevó a cabo una investigación sobre elisión de *schwa* en el inglés americano en contextos *schwa* átona + resonante (*scenery* > *scen(e)ry*) y pudo comprobar que la elisión era más alta en palabras más frecuentes (como *every* o *memory*) que en palabras menos frecuentes (como *mammory*, *cursorry* o *artillery*). Jurafsky y otros 2001 también observaron que en palabras inglesas terminadas en /t/ o /d/ (*want*, *just*, *good*, *kind*) las formas más frecuentes eran más breves y elidían la consonante final en más ocasiones que las formas menos frecuentes. Tan sistemáticos son estos ajustes que algunos investigadores han llegado a plantearlos en términos universales, elevándolos a la categoría de leyes. A este respecto, Zipf 2002, p. 389, ha propuesto la Ley de Abreviación, que, en síntesis, afirma que «[...] as the relative frequency of a

word increases, it tends to diminish in magnitude». En el mismo sentido, Jurafsky y otros 2001, p. 229, han formulado la Hipótesis de Reducción Probabilística, que los autores resumen en los siguientes términos: «word forms are reduced when they have a higher probability».

En el caso de las unidades o secuencias de unidades lingüísticas complejas, a medida que se van volviendo más frecuentes no solo se reducen, sino que también se hacen cada vez más independientes de las palabras originarias. Esto puede conducir a que dichas formas pierdan también su estructura interna y, con ello, quede abierta la vía libre para los procesos de gramaticalización. Bybee 2006 menciona al respecto el caso del inglés *I'm going to*. En el siglo XVI el significado de «propósito» se expresaba en esta lengua con *be* más el gerundio de verbos como *travel*, *journey* o *go*, entre otros. De este conjunto de verbos, *go* se empezó a usar cada vez más en este tipo de construcciones, llegando a desplazar a las restantes formas. Este proceso terminó por provocar que *I'm going to* se fuera volviendo más independiente de la construcción originaria, perdiera su asociación con el significado de propósito, pasara a ser una simple marca de futuro, y, finalmente, se viera afectada por un proceso de reducción fonética: [aimənə].

La razón de que los sonidos tiendan a elidirse más en las palabras que más se usan tiene que ver, según Bybee 2002, con la relación que se establece entre la actividad neuromotora y el comportamiento lingüístico: las palabras más frecuentes se ven sometidas a un proceso de automatización, lo que implica, entre otras cosas, que los movimientos de los articuladores no sean tan precisos como en las palabras menos predecibles, que se solapen estos movimientos, etcétera.

2.2. *El efecto conservador*

Determinados procesos lingüísticos no se producen en las palabras más frecuentes, sino, precisamente, en las palabras menos frecuentes, mientras que las palabras más frecuentes se resisten al cambio. Bybee 2002, p. 269, ha analizado este efecto en el caso de los verbos regulares e irregulares del inglés. En la historia de esta lengua se ha observado la tendencia de algunos verbos con pasado irregular a tomar la forma de pasado *-ed*, característica de los verbos regulares. En el conjunto de verbos irregulares se encuentran formas como *weep*, *leap*, *creep*, *keep*, *sleep* y *leave*, cuyos pasados tradicionales son *wept*, *lept*, *crept*, *kept*, *slept* y *left*, respectivamente. De estos seis verbos, los tres primeros (*weep*, *leap* y *creep*), que tienen una frecuencia media baja (37), están sufriendo un proceso de nivelación más acelerado hacia las formas en *-ed*: *creeped*, *leaped* y *weeped*. Este pasado regular es más extraño en los tres últimos verbos

(*keep, sleep y leave*), que tienen una frecuencia media más alta (485). La razón del diferente comportamiento de los dos grupos de verbos se basa en el hecho de que las estructuras o secuencias de frecuencia más alta están más reforzadas en la memoria, lo que hace que se vuelvan más resistentes a los cambios analógicos (Bybee 2006).

A juicio de Bybee 2002, 2007, pp. 27-31, los cambios lingüísticos que afectan primero a las palabras de alta frecuencia tienen su origen en la automatización del discurso, mientras que los cambios que afectan antes a las palabras de baja frecuencia son debidos al aprendizaje imperfecto. La tesis del aprendizaje imperfecto la toma la autora de Kiparsky 1970, quien defiende que la analogía puede deberse a una sobregeneralización por parte de los niños que aprenden la lengua materna. Para Bybee, las semejanzas entre el tipo de generalización que hacen los niños y los cambios analógicos son notables: si los niños producen estas generalizaciones a gran escala, es normal que lo hagan en palabras poco frecuentes, donde no se notará tanto. Añade además que esta hipótesis no tiene por qué ser restringida a los niños, sino que puede extenderse a los adultos. Por lo que respecta a los paradigmas menos frecuentes, los adultos pueden no estar seguros de todas sus formas, o puede ocurrir que la simple infrecuencia de un paradigma supletivo haga más aceptable a una formación analógica.

Phillips 1981, 1984, 2001 ha analizado también determinados fenómenos del inglés implicados en procesos de cambio que se suelen manifestar más en palabras de menor frecuencia. Uno de estos fenómenos es la elisión de *glide* tras oclusiva alveolar inicial de palabra (*tune, duke, news*) en el inglés americano. Según Phillips, en este caso no puede hablarse de un proceso natural (como en el de la elisión de *schwa*), ya que se produce en una posición donde normalmente los sonidos tienden a ser reforzados. La autora observa que la elisión afecta más a aquellas palabras que tienen una frecuencia más baja. De ese modo, la elisión es más intensa (con porcentajes en torno al 70-75%) en palabras como *nude, neutron* o *tutor*, que tienen una menor frecuencia de uso, y menos intensa (con porcentajes entre el 40-55%) en palabras que se usan más, como *knew, during* o *new* (Phillips 1981, pp. 74-75). La razón que proporciona Phillips para explicar el efecto conservador es que, a diferencia de otros fenómenos (como la elisión de *schwa*), donde la regla se aplica de un modo más o menos automático, en los casos de la elisión de *glide* nos hallamos con ciertas restricciones en el momento de aplicación de la regla. Este tipo de restricción figura como información almacenada en el lexicon para cada unidad léxica afectada, lo que implica que antes de practicar la elisión los hablantes tienen que acceder a la información subyacente con el fin de conocer si deben o no aplicar la regla.

3. OBJETIVOS Y METODOLOGÍA

3.1. *Objetivos*

En esta investigación se tratará de determinar de qué modo la frecuencia de las palabras interfiere en las realizaciones fonéticas de dos fonemas, /d/ intervocálica y /tʃ/, en el español hablado en la ciudad de Santa Cruz de Tenerife (Islas Canarias). Dado el distinto estatus funcional de ambos fonemas y las diferentes características fonéticas de los alófonos correspondientes, los análisis que se llevarán a cabo en cada fonema serán también diferentes.

3.2. *Metodología*

3.2.1. Aspectos sociodemográficos

La muestra que sirve de base a esta investigación fue obtenida en Santa Cruz de Tenerife entre 1994-1998. El número de informantes seleccionados fue de 36, repartidos del siguiente modo según categorías sociales:

- a) según la variable social sexo-género, 18 hombres y 18 mujeres;
- b) según el grupo generacional, 12 individuos de 20-35 años, 12 de 36-55 años y 12 de más de 55 años, y
- c) según la clase social, 12 individuos de clase trabajadora, 12 de clase media y 12 de clase media-alta.

En el actual análisis, la única variable social que se utilizó fue la edad de los individuos. La incorporación de esta variable tuvo como finalidad determinar si las variantes de los dos fonemas que se analizaban se encontraban en una situación de estabilidad o de cambio en marcha¹.

La selección de los informantes se hizo de modo semi-aleatorio, siguiendo la técnica de «bola de nieve» (una persona conocida remite a otra conocida por ella, que a su vez remite a otra, etc.). Las entrevistas

¹ Una de las aportaciones de la Sociolingüística ha sido proponer que, a través de un análisis generacional de una muestra de habla obtenida en un momento histórico concreto, puede determinarse si una variable lingüística se halla sometida a simple variación o se halla bajo los efectos de un cambio en marcha. De ese modo, si una de las manifestaciones de la variable es más frecuente en la generación más joven y menos frecuente en la generación más vieja puede aventurarse que la forma en cuestión se está afianzando en la comunidad y puede progresar en el futuro. Y a la inversa: una forma lingüística que es más frecuente en la generación más vieja y menos frecuente en la generación más joven puede indicar que dicha forma se encuentra en retroceso. En ambos casos las diferencias generacionales deben ser significativas desde el punto de vista estadístico. Los investigadores han diseñado toda una serie de estrategias complementarias que permiten llegar a conclusiones más fiables sobre la existencia o no de un cambio en marcha (Almeida 2003, pp. 259-264).

mantenidas trataron de temas relacionados con la vida cotidiana (trabajo, familia, asuntos de la comunidad, etc.). Estas entrevistas duraban una hora, por lo que el corpus que sirve de base a la investigación consta de 36 horas.

3.2.2. Las variables lingüísticas

Las dos variables fónicas que se analizan en esta investigación son /d/ intervocálica y /tʃ/. /d/ intervocálica registra básicamente tres tipos de alófonos en Santa Cruz de Tenerife: aproximantes con articulación plena, que representan la mayoría de casos (73,2%), aproximantes relajadas (17,3%) y elisiones (9,4%). En este trabajo se analizará la alternancia entre el mantenimiento de la consonante (como forma plena o debilitada) y la elisión. Los resultados del análisis en tiempo aparente revelan que la elisión es un fenómeno que actualmente se encuentra estabilizado en la comunidad, ya que no se han registrado diferencias importantes entre los tres grupos generacionales: 10,2% en la generación joven, 8,8% en la generación intermedia y 9,3% en la generación más vieja. Estas diferencias son no significativas ($p > 0,05$)².

Por lo que respecta al fonema palatal africado sordo /tʃ/, las dos realizaciones más frecuentes en Santa Cruz de Tenerife son las articulaciones laxa y tensa, que representan, respectivamente, el 74,3% y el 25,7% del total de /tʃ/ registradas en la muestra. La primera, que representa a la variante tradicional en todas las comunidades lingüísticas del archipiélago canario, se caracteriza porque manifiesta una tendencia a la sonorización y a la reducción de los momentos oclusivo y fricativo (especialmente de este último) con respecto a la afrificada sorda castellana. La variante tensa, en cambio, que ha sido introducida en la comunidad seguramente como imitación del sonido africado sordo del castellano, experimenta una mayor duración del momento fricativo frente a la forma vernácula, se percibe como más estridente y tiene normalmente carácter sordo (Almeida 1994-1995). Todo ello puede interpretarse como

² En análisis donde se trate de determinar el efecto de unas variables sobre otras (como en este caso, donde se ha estudiado cómo el mantenimiento y elisión de /d/ pueden estar condicionados por la edad de los individuos) debe calcularse, usando el test de Pearson, la probabilidad de que la variación se deba al azar. «p» mide esta probabilidad y, por tanto, cuanto más bajo sea el valor de p menos influencia tendrá el azar sobre la muestra objeto de estudio, es decir, más fuerte será la relación entre las variables estudiadas, y a la inversa. El valor de p que se suele tomar como referencia para medir estos efectos es 0,05 (o 5%), lo que significa (entre otras interpretaciones) que, de cada 100 veces que midamos la influencia de unas variables sobre otras, el azar solo interviene en 5 casos, por lo que la relación entre variables puede considerarse significativa. Cualquier valor de p superior a 0,05 se interpreta como una intervención excesiva del azar (aunque es cierto que en algunas investigaciones el valor de referencia se sitúa en el 0,1, o 10%), como ha sucedido esta vez con el cruce de la variable lingüística y la variable social Edad.

signo de una mayor tensión articulatoria. No obstante, el sonido resultante no alcanza el grado de tensión de la consonante castellana, debido a que cada uno de los dos momentos de su articulación, el oclusivo y el fricativo, suelen ser aún más breves en la variante innovadora canaria que en la variante estándar. En este sentido la forma tensa puede ser catalogada como una realización «híbrida» o «interdialectal» (Berruto 2005, Britain 2009, Trudgill 1986, pp. 57-66), es decir, como una variante que reúne características fonéticas de dos formas dialectales. Por un lado, se ha restaurado la fricción y la sordez del sonido castellano, pero, por otro, la duración de los momentos oclusivo y fricativo es menor que en la variante estándar castellana.

Los datos sociolingüísticos revelan que esta variable se halla inmersa en un cambio en marcha, marcado por una clara estratificación generacional: los jóvenes usan la nueva variante el 38,1% de los casos, los miembros de la generación intermedia el 25,9%, mientras que los individuos de más edad la usan solo el 14%. Estas diferencias son estadísticamente significativas ($p < 0,000$).

En primer lugar, se analizará la relación entre los alófonos de /d/ y /tʃ/ y la frecuencia de las palabras en que aparecen dichos fonemas (*token frequency*). Este análisis sigue el mismo tipo de procedimiento que el llevado a cabo en investigaciones semejantes, por lo que no requiere comentarios adicionales. En segundo lugar, se tratará de determinar también si la relación entre mantenimiento y elisión de /d/ y la frecuencia de la palabra se mantiene o no constante en diferentes contextos morfológicos (*type frequency*). Con este propósito se analizará la alternancia [ð]/[ø] en los morfemas terminados en *-ado(s)*, *-ada(s)*, *-ido(s)* e *-ida(s)* –por ejemplo, en palabras como *salado(s)*, *callada(s)*, *afligido(s)* o *salida(s)*–³. No obstante, con el fin de determinar si el comportamiento observado es exclusivo o no de las terminaciones morfológicas, se llevarán a cabo estos mismos análisis en el grupo de los lexemas que tienen estas mismas terminaciones –*mercado(s)*, *abogada(s)*, *marido(s)* o *vida(s)*–. Este análisis está inspirado fundamentalmente en el que llevó a cabo Bybee 2002, 2003, aunque se han introducido algunas diferencias en la metodología empleada. Las razones de estas diferencias entre las dos investigaciones tienen que ver, sobre todo, con el hecho de que Bybee ha pasado por alto aspectos de tipo coarticulatorio que quizás hubiera sido conveniente haber controlado con el fin de alejar sospechas sobre posibles sesgos en los resultados. A este respecto, Bybee contó

³ Bybee 2002 y 2003, pp. 148-153, se refiere a los dos primeros como morfemas de la primera conjugación y a los dos segundos como morfemas de la segunda/tercera conjugación. Aunque en el análisis llevado a cabo en Santa Cruz de Tenerife se mantiene a veces esta misma distinción, conviene recordar que dicho análisis no queda limitado a los participios y adjetivos, sino que también incluye los sustantivos –*parada(s)*, *salida(s)*, etcétera–.

como un solo grupo de morfemas las terminaciones *-ado(s)* y *-ada(s)* por un lado y las terminaciones *-ido(s)* e *-ida(s)* por otro, mientras que en el análisis realizado en la capital tinerfeña se analizará por separado cada uno de los cuatro tipos de terminación. Esta decisión está justificada por el hecho de que la vocal posconsonántica no es la misma en los morfemas de cada uno de los dos grupos, *o* en un caso y *a* en el otro, una circunstancia que podría provocar efectos diferentes sobre la articulación de /d/. Lo mismo cabe decir de los análisis donde se contrastan los morfemas de la primera y de la segunda/tercera conjugación con las palabras que contienen /d/ intervocálica no morfemática. En este caso Bybee confronta los resultados de elisión de /d/ en los morfemas *-ado(s)* y *-ada(s)* vs. palabras lexemáticas terminadas en *-ad* + vocal, y morfemas en *-ido(s)* e *-ida(s)* vs. palabras lexemáticas terminadas en *-id* + vocal. El problema aquí es que tampoco se controla el tipo de vocal que sigue a /d/. La decisión que se ha tomado en el análisis de la muestra de Santa Cruz de Tenerife viene avalada por algunos investigadores que han apuntado la posibilidad de que el contexto vocálico pueda influir en la retención y elisión de /d/. A este respecto, Menéndez Pidal 1980, pp. 99-101, al explicar que la elisión de /d/ intervocálica en español es más frecuente en las terminaciones en *-ado* que en otros contextos, apunta varias razones, una de las cuales se refiere a la mayor dificultad de articular /ado/ frente a otras terminaciones, como /ada/. A su juicio, durante la articulación de /ad/ la parte anterior de la lengua adopta una posición cóncava y adelantada, lo que hace que sea más difícil pasar a la postura opuesta (convexa y retraída) que requiere la articulación de /o/. Por esta razón, en la pronunciación rápida se evita la articulación adelantada de /d/. Es cierto que la interpretación de Menéndez Pidal requiere una confirmación por procedimientos más experimentales y menos impresionistas, pero tampoco se pueden descartar del todo los posibles efectos coarticulatorios.

3.2.3. El tratamiento de la frecuencia: ¿palabras o lemas?

Cuando se trata de analizar la frecuencia de las unidades lingüísticas puede utilizarse como referencia la palabra o el lema (Almela y otros 2005). En el primer caso se consideran entradas diferentes alternancias como masculino/femenino (*todo* vs. *toda*), singular/plural (*todo* vs. *todos*), cada una de las formas verbales (*quedo*, *quedó*, *quedaría*), etc. En el segundo caso se considera que forman parte de una misma unidad todas las variantes en masculino y femenino o singular y plural de las formas lingüísticas susceptibles de recibir estas marcas morfológicas, así como los diferentes tiempos, modos y personas del mismo verbo. De ese modo, en los ejemplos anteriores tendríamos un lema *todo* y un lema *quedar*. En esta investigación se ha optado por considerar como unidad

de frecuencia la palabra y no el lema, en la línea de Bybee 2002. Esta decisión está sustentada en el hecho de que, de haber tomado la opción del lema, se hubiera tenido que haber agrupado bajo una misma unidad léxica manifestaciones fonéticas muy diferentes. Por ejemplo, habría que haber considerado bajo el lema *quedar* las formas verbales *quedo* y *quedó* (entre otras muchas), es decir, una forma de presente, donde /d/ aparece en una sílaba átona, y una forma de pretérito, donde /d/ aparece en una sílaba tónica. En algunas investigaciones sobre mantenimiento y elisión de /d/ intervocálica llevadas a cabo en el mundo hispánico se ha observado que /d/ se suele perder más en sílaba átona que en sílaba tónica (Almeida 2011). Por tanto, si se hubiera optado por tomar como referencia el lema en vez de la palabra, posiblemente se habría creado un grupo de unidades léxicas muy heterogéneo con respecto al estatus acentual de /d/, una situación que, tal vez, habría provocado una distorsión en los resultados, ya que afecta a un gran número de casos (por ejemplo, a verbos frecuentes como *poder*, *quedar*, etc., que tienen formas donde /d/ puede pertenecer tanto a una sílaba átona como a una sílaba tónica).

Para determinar el grado de frecuencia de las palabras de la muestra se tomó como referencia el Corpus CUMBRE, que recoge la lista de las 10.000 palabras más frecuentes del español (Almela y otros 2005, pp. 117-298) y de algunas combinaciones de palabras⁴. Este corpus presenta algunos puntos discutibles, como su carácter tan heterogéneo. Esta heterogeneidad se manifiesta, fundamentalmente, en dos aspectos. Por un lado, los textos que se estudian proceden tanto de España como de Hispanoamérica; de ellos, los textos españoles tienen casi el doble de peso que los hispanoamericanos (65% vs. 35%; cf. Almela y otros 2005, p. 8). Por otro lado, para elaborar el recuento de las unidades lingüísticas (palabras, lemas o combinaciones de palabras) se utilizaron tanto textos escritos (donde se incluían libros, revistas, folletos de información, tebeos) como orales (incluyendo programas de radio y televisión, conversaciones en situaciones reales)⁵. A pesar de estas limitaciones (y otras que podrían ser mencionadas) se consideró que, dado que se trata

⁴ Este listado puede consultarse también en Internet: <<http://www.um.es/lacell/proyectos>> (consulta 20/05/2010).

⁵ Problemas semejantes se pueden encontrar en cualquier corpus de frecuencias del español. Tal vez por eso algunos investigadores han optado por crear sus propios listados de frecuencias. Por ejemplo, Díaz Campos 2004 tomó como referencia el corpus de frecuencias de Julliard y Chang-Rodríguez 1964, basado en textos literarios, pero también las formas léxicas del corpus obtenido por el autor en Caracas. Por su parte, González 2006 ideó un procedimiento más complicado: elaboró una lista de 39 palabras que un grupo de jueces debía evaluar de acuerdo a los siguientes aspectos: muy rara, rara, frecuente, bastante frecuente y muy frecuente. Este procedimiento resultó ser bastante realista, ya que la evaluación de los diferentes jueces registró una gran coincidencia entre ellos.

de un corpus reciente y de fácil manejo, el listado de frecuencias de Almela y otros era adecuado para la finalidad de esta investigación.

En el Corpus CUMBRE las palabras fueron clasificadas en cinco grupos con respecto a su frecuencia:

- a) baja: hasta tres ocurrencias por millón de palabras en el corpus,
- b) moderada: entre 4-10 ocurrencias,
- c) notable: entre 11-25 ocurrencias,
- d) alta: entre 26-75 ocurrencias, y
- e) muy alta: más de 75 ocurrencias (Almela y otros 2005, p. 16).

En la investigación sobre Santa Cruz de Tenerife se agruparon los tres primeros niveles por un lado y los dos últimos por otro, de modo que se trabajará con dos grupos: palabras de frecuencia baja vs. palabras de frecuencia alta.

3.2.4. Hipótesis

En el caso de las dos variantes de /d/ que se estudian se espera que la elisión de la consonante se produzca más en las palabras más frecuentes. Investigaciones llevadas a cabo en diferentes dialectos hispánicos, tanto americanos como europeos, se han referido al efecto reductor de la frecuencia en el caso del mantenimiento y elisión de /d/ intervocálica. Menéndez Pidal 1980, pp. 99-101, de una forma un tanto intuitiva, señalaba la frecuencia como el factor responsable de la elisión de /d/ intervocálica en español. Según este autor, el hecho de que la consonante se elida normalmente en *lado* pero no en *vado* se debe a que la primera es una palabra de uso frecuente, mientras que la segunda es una palabra que raramente se usa. Análisis posteriores han acabado por confirmar estas intuiciones. Así, en el español hablado en Nuevo México se observó que la elisión de /d/ intervocálica alcanzaba el 21% en palabras de frecuencia alta y el 8,6% en palabras de frecuencia baja (Bybee 2002, p. 66, 2003, pp. 148-153). La misma tendencia ha sido descrita en niños de entre 42-71 meses de Caracas (Díaz Campos 2004, p. 229). En este caso se manejaron dos tablas de frecuencias: el diccionario de frecuencias del español de Jullian y Chang-Rodríguez 1964 y las palabras del propio corpus. En el primer caso, la elisión representó el 46% de las palabras de frecuencia alta, el 20% de las palabras de frecuencia media y el 12% de las palabras de frecuencia baja. En el segundo caso se contemplaron dos niveles: palabras de frecuencia alta y baja, que arrojaron valores de elisión de 55% y 7%, respectivamente. Este mismo patrón de variación ha sido observado en el español hablado en la provincia de Castellón, tanto en el grupo de hablantes que

tenía como lengua dominante el castellano como en el grupo que tenía como lengua dominante el valenciano. En el primer grupo el porcentaje de elisiones fue del 65% en las palabras más frecuentes y del 48% en las palabras menos frecuentes, mientras que en el segundo grupo los porcentajes respectivos fueron 29% y 11% (Blas Arroyo 2006, p. 17).

Son muy pocas las investigaciones que no respaldan la tendencia reductora de los sonidos en las palabras más frecuentes. Una de ellas es la investigación de González 2006 sobre elisión de /d/ final de palabra en español (*edad*, *virtud*, etc.). La autora estableció dos grupos de palabras, frecuentes y poco frecuentes, y las introdujo en frases marco que eran leídas por los informantes. A continuación analizó el papel de la frecuencia en aspectos como la sonoridad, duración y modo de articulación (oclusiva, fricativa o aproximante) de la consonante. Los resultados mostraron que el efecto fue nulo. Según la autora, la no funcionalidad de la frecuencia en este caso puede estar relacionado con el tipo de tarea diseñada (la lectura de textos).

En lo que respecta a la variable /tʃ/, se espera que la realización tensa esté más presente en las palabras menos frecuentes que en las más frecuentes. Ya se ha considerado la realización tensa como una variante «interdialectal» o «híbrida». Esta realización seguramente se ha constituido como consecuencia de un fallo durante el proceso de adquisición de un segundo dialecto (el estándar) en una situación de contacto interdialectal (Trudgill 1986, p. 60, se refiere a estos casos como ejemplos de acomodación incompleta). De hecho, es muy difícil para un hablante canario, donde lo normal es que las consonantes se vean afectadas por procesos de debilitamiento (y, en los casos extremos, también de elisión), lograr una articulación como la de [tʃ] africana del castellano estándar. En este sentido, aunque no se trate de un cambio analógico, sí comparte con este el hecho del «aprendizaje imperfecto» al que se ha referido Bybee 2002. Por tanto, lo mismo que ocurre con los cambios analógicos, es posible que los porcentajes más altos de variantes tensas se registren en las palabras menos frecuentes. Esto no quiere decir que se estén poniendo al mismo nivel los cambios analógicos (que se producen en el nivel gramatical y que conducen a la aparición de estructuras morfológicas regulares) y la tensión de /tʃ/ (un fenómeno exclusivamente fonético).

Por último, del análisis del tipo de frecuencia se espera que la elisión de /d/ intervocálica sea más alta en aquellos morfemas que pertenecen a palabras más frecuentes. Bybee 2002, 2003, pp. 148-153, analizó estas cuestiones a partir del análisis de una muestra de hablantes de Nuevo México y de materiales extraídos del *Corpus Oral de Referencia del Español Contemporáneo* (COREC). Los resultados de Nuevo México demostraron que los porcentajes de elisión eran más frecuentes en los

morfemas de la primera conjugación *-ado(s)* y *-ada(s)* que en los morfemas de la segunda/tercera conjugación *-ido(s)* e *-ida(s)*: 58% vs. 28,6% (Bybee 2003, p. 150). Se observó también que la elisión alcanzaba porcentajes más altos en los morfemas de la primera conjugación que en las demás palabras del corpus que contenían una /d/ intervocálica después de [a] (58% vs. 41%), y en los morfemas de la segunda/tercera conjugación frente a las restantes palabras que llevaban una /d/ intervocálica después de [i] (29% vs. 9%) (Bybee 2003, p. 151). El hecho de que los porcentajes más altos de elisión se hayan registrado en los morfemas de la primera conjugación puede tener que ver con el hecho de que estos son más frecuentes que los morfemas de la segunda/tercera conjugación: 91 casos de *-ado* (y variantes) frente a 52 casos de *-ido* (y variantes). La tesis de la frecuencia se ve reforzada por el hecho de que en las palabras terminadas en *-ado* (y variantes) la elisión alcanza el 64% en palabras de frecuencia alta y el 50% en palabras de frecuencia baja (Bybee 2003, p. 152). En el análisis del Corpus del COREC se comprobó que los morfemas de la primera conjugación registraban un 21,7% de elisión, mientras que los demás morfemas registraban un 2,4% en las palabras de baja frecuencia y un 8,7% en las palabras de alta frecuencia. En las restantes palabras la elisión se situó en el 5,6% (Bybee 2002, p. 285). Estos resultados sugieren que los morfemas de la primera conjugación comenzaron a ser tratados como palabras de las que se ha borrado cualquier tipo de estructura interna, y, por extensión, dadas las conexiones gramaticales entre los participios de la primera conjugación con los de la segunda/tercera conjugación, es probable que este comportamiento se haya extendido al resto de los morfemas (Bybee 2002, 2003, pp. 148-153).

4. RESULTADOS

En esta sección se tratarán en primer lugar los aspectos relacionados con la frecuencia de las palabras y posteriormente se expondrán los resultados relacionados con el tipo de frecuencia.

4.1. *Variación fonética y frecuencia de las palabras*

4.1.1. Mantenimiento y elisión de /d/

De acuerdo con la hipótesis inicial, la elisión de /d/ habría de afectar en mayor medida a las palabras más frecuentes que a las menos frecuentes, una tesis que aparece confirmada por los resultados. El análisis de los datos mediante tablas de contingencia (tabla 1) revela que de un total de 1118 palabras de baja frecuencia que contienen /d/ intervocá-

lica la elisión ha afectado a 82 (el 7,3%), mientras que entre las 2674 palabras de frecuencia alta el número de elisiones ha sido de 276 (el 10,3%). Estas diferencias son estadísticamente significativas: $\chi^2 = 8,228$, $p = 0,004^6$. Este resultado no hace sino confirmar las tesis que relacionan la frecuencia con los procesos de debilitamiento y elisión tanto en español como en otras lenguas.

	Frecuencia baja		Frecuencia alta	
	N	%	N	%
Conservación	1036	92,7	2398	89,7
Elisión	82	7,3	276	10,3
TOTAL	1118		2674	

TABLA 1: Frecuencias absolutas y relativas de mantenimiento y elisión de /d/ intervocálica según la frecuencia de la palabra

Ahora bien, los resultados indican que, a pesar de que los porcentajes registrados en los dos grupos de palabras son estadísticamente significativos, la diferencia de elisión entre palabras de frecuencia alta y palabras de frecuencia baja es de tres puntos, lo que lleva a pensar que el alcance explicativo de la frecuencia es más limitado de lo que en principio se esperaba. Por tanto, a la hora de analizar el modo en que está organizada la variación de los dos alófonos de /d/ habrá que tener en cuenta el efecto de otras variables (lingüísticas, sociales, actitudinales) no contempladas aquí. Por esta razón, parece útil aportar los porcentajes de elisión de algunas palabras concretas pertenecientes a cada una de las dos categorías (baja y alta frecuencia) y determinar sus características particulares. En la tabla 2 se recogen algunas unidades léxicas del corpus con indicación de su frecuencia (T), número de elisiones (N) y porcentaje de elisiones (%). Este último índice (el porcentaje) solo se ha incorporado en el caso de las palabras de alta frecuencia, ya que su incorporación en el grupo de palabras de frecuencia baja hubiera producido una gran distorsión de los resultados. Por ejemplo, hubiera llevado a afirmar que en *profesorado* las elisiones representan el 25% de las manifestaciones de /d/, un valor que no puede ser tomado en serio, ya que en el corpus solo se obtuvieron 8 casos de esta palabra y, de ellos, la elisión solo afectó a 2.

⁶ El valor χ^2 es el resultado de diversas operaciones donde se contrastan los valores reales de cada una de las casillas con los valores ideales. Por su parte, el valor $p = 0,004$ obtenido en el cruce de las dos variables indica que el efecto del azar en este caso ha sido del 4 por mil, muy por debajo del 0,05 de referencia.

Frecuencia baja		Frecuencia alta		
	N/T		N/T	%
profesorado	2/8	cambiado	6/15	40,0
delegada	1/4	todas	23/75	30,7
parado	1/5	todavía	9/39	23,1
pescado	1/6	todos	33/166	19,9
quedo	1/7	nada	27/242	11,2
quedé	1/11	cada	5/75	6,7
vídeo	0/14	queda	1/26	3,8
virtudes	0/6	comunidad	0/10	0,0
viuda	0/6	edad	0/37	0,0

TABLA 2: Frecuencias absolutas y relativas de elisión de /d/ intervocálica en algunas unidades léxicas según su frecuencia

En la tabla 2 se observa que entre las seis unidades léxicas de baja frecuencia que registran algún tipo de elisión figuran tres que terminan en *-ado* (un contexto que, según Moreno Fernández 2008, p. 1000, favorece las elisiones de /d/ en español): *profesorado*, *parado* y *pescado*. En las demás palabras de este grupo es probable que los bajos o nulos valores de elisión tengan que ver simplemente con su baja frecuencia de uso. Entre las palabras de frecuencia alta existen fuertes discrepancias en lo que respecta al porcentaje de elisión de /d/. Un pequeño grupo registra porcentajes que pueden considerarse importantes, ya que se sitúan por encima del 25% (sobre todo si se tiene en cuenta que el porcentaje medio de elisión de este fonema en la comunidad ronda el 9%): *cambiado* y *todas*. Otras palabras, como *todavía* y *todos*, registran porcentajes destacados de elisión (en torno al 20%). Palabras como *nada* y *cada* se sitúan en torno a la frecuencia media de elisión (9,4%), mientras que en otras como *comunidad* o *edad* las elisiones o son muy bajas o no se producen. Resulta muy difícil explicar el porcentaje de elisión en cada uno de estos casos, pero puede intentarse alguna explicación para determinadas palabras. Por ejemplo, hay dos ítems, *comunidad* y *edad*, donde no se registraron elisiones. Es probable que en este caso el efecto del acento haya sido más poderoso que el de la frecuencia, ya que los sonidos tienden a conservarse más en posición tónica (Laver 1995, p. 511). En la misma comunidad, Almeida 2011, p. 42, observó que en posición tónica /d/ solo se elidía un 1,6%, frente al 11,2% de elisiones en posición átona. Más difícil resulta determinar las razones por las que otras unidades léxicas registran porcentajes aparentemente discrepantes con la tesis de la frecuencia. A este respecto

conviene recordar dos cosas. En primer lugar, aunque la frecuencia de las palabras influye significativamente en los usos de determinados alófonos, dicha variable solo explica una parte de la variación alofónica que se ha analizado. En segundo lugar, no conviene olvidar que la clasificación de frecuencias de las palabras de Almela y otros 2005 contemplaba cinco niveles (desde el bajo al muy alto), mientras que en esta investigación solo se han manejado dos niveles (bajo y alto). Esto quiere decir que en cada uno de los dos grupos de frecuencia contemplados en este estudio hay una cierta heterogeneidad, lo que tal vez puede influir en las diferencias porcentuales registradas en cada grupo.

4.1.2. Realizaciones laxas y tensas de /tʃ/

Por lo que respecta al fonema africado sordo /tʃ/, los resultados indican que las realizaciones tensas se producen más en el grupo de palabras de alta frecuencia que en el de palabras de baja frecuencia. De 975 palabras de frecuencia baja registradas en la muestra, hubo 218 casos de variantes tensas, lo que representa el 22,4%, mientras que en las 2290 palabras de frecuencia alta los alófonos tensos fueron 622, es decir, el 27,2% (tabla 3). Igual que en el caso de /d/, también esta vez las diferencias registradas en la tabla son estadísticamente significativas: $\chi^2 = 8,254$, $p = 0,004$. Este resultado obliga a rechazar la hipótesis inicial, que predecía un porcentaje más alto de variantes tensas en las palabras menos frecuentes, y a elaborar una hipótesis alternativa. A la vista de los resultados cabe pensar que, dado que la variante tensa de /tʃ/ es una forma interdialectal, cuya articulación es probable que produzca una cierta inseguridad lingüística en los individuos, estos traten de usarla más en aquellos contextos en que se sienten más seguros, es decir, en el de las palabras más frecuentes.

A pesar del carácter significativo de las diferencias, no puede pasarse por alto el hecho de que solo hay cinco puntos de diferencia en los usos de los alófonos tensos, lo que sugiere que puede existir un

	Frecuencia baja		Frecuencia alta	
	N	%	N	%
/tʃ/ laxa	757	77,6	1668	72,8
/tʃ/ tensa	218	22,4	622	27,2
TOTAL	975		2290	

TABLA 3: Frecuencias absolutas y relativas de variantes laxas y tensas de /tʃ/ según la frecuencia de la palabra

importante grado de variación alofónica dentro de cada grupo de palabras. Por eso, igual que se procedió en el caso de /d/, con el fin de entender mejor estas excepciones se consideró oportuno representar la relación entre frecuencia y tensión en algunas de las palabras del corpus (tabla 4).

Igual que ocurre con el fonema /d/, no resulta fácil explicar los porcentajes de variantes tensas de /tʃ/ en cada una de las palabras, por lo que me ocuparé explícitamente de algunos, comenzando por los más claros. Las cuatro primeras palabras del grupo de baja frecuencia donde se producen algunos casos de variantes tensas (entre 2-3) son *choca*, *chino*, *cochito* y *cochino*. En las dos primeras unidades la consonante aparece en posición inicial de palabra y es, además, tónica, por lo que podría pensarse que ambos factores pueden estar condicionando las realizaciones tensas. Sin embargo, el análisis estadístico mostró que ninguna de las dos variables (la posición inicial o interior de palabra y el carácter tónico o átono de la sílaba) consideradas aisladamente resultó determinante en esta muestra, por lo que cabría pensar que la distribución de la tensión podría ser debida al azar. Ahora bien, cuando se toman en cuenta las dos variables juntas sí se observan diferencias relevantes. El análisis estadístico mostró que en posición tónica inicial de palabra el porcentaje de variantes tensas de la africada se sitúa en el 30,6%, mientras que en posición átona inicial la frecuencia de esta variante desciende al 18,2%. Estas diferencias son altamente significativas ($\chi^2 = 12,181$, $p < 0,000$). Por tanto, puede concluirse que el efecto del acento asociado al de la posición inicial puede explicar los casos de tensión en estas dos palabras. En el caso de *cochito* (otra palabra de frecuencia baja), las realizaciones tensas que registra pueden estar relacionadas con el hecho de que los hablantes la asocian con *coche*, una palabra de frecuencia alta (que registra un 35% de alófonos innovadores). Algo parecido puede haber ocurrido con *chiquita*: es probable que los hablantes la relacionen con *chica* o *chicas*, que son palabras de frecuencia alta y que tienen porcentajes de variantes tensas del 24% y 42,5%, respectivamente. Bybee 2006 demostró que una forma lingüística de baja frecuencia puede tener un comportamiento que es más bien propio de una forma de alta frecuencia solo por el hecho de que ambas se asemejen. En un estudio sobre el grado de aceptabilidad de las expresiones españolas *quedarse inmóvil* (con una frecuencia alta), *quedarse parado* (una construcción de baja frecuencia, pero relacionada con la anterior) y *quedarse orgullósísimo* (una construcción de baja frecuencia y no relacionada con la primera) Bybee observó que la construcción más frecuente fue valorada como más aceptable y las dos menos frecuentes fueron valoradas como menos aceptables. Sin embargo, observó también que estas dos últimas construcciones recibieron valoraciones de aceptabilidad diferentes, de modo que *quedarse parado* fue considerada

más aceptable que *quedarse orgullosísimo*, posiblemente por la relación que la primera tiene con *quedarse inmóvil*.

Más difícil resulta explicar los porcentajes de /tʃ/ tensas de otras palabras que se registran dentro de cada uno de los dos grupos de frecuencia (baja vs. alta), pero podría tener que ver con los argumentos expuestos a propósito de /d/ sobre la configuración de los grupos de frecuencias de esta investigación con respecto al corpus de frecuencias que se tomó como referencia.

Frecuencia baja		Frecuencia alta		
	N/T		N/T	%
choca	3/6	derecho	9/22	40,9
chino	2/5	coche	57/161	35,4
cochito	2/6	noche	51/150	34,4
cochino	2/7	mucho	129/431	29,9
chiquita	3/13	leche	4/18	22,2
echan	2/11	ochenta	3/29	10,3
anoche	1/11	muchísimo	3/35	8,6
bachiller	0/20			
capricho	0/5			
chaqueta	0/6			

TABLA 4: Frecuencias absolutas y relativas de /tʃ/ tensa en algunas unidades léxicas según su frecuencia

Una diferencia que se observa entre las dos variables fonéticas que están siendo analizadas es que en el caso de /d/ se podían encontrar palabras de frecuencia alta donde no se registrara ningún caso de elisión, mientras que en el caso de /tʃ/ resulta difícil encontrar palabras de alta frecuencia en las que no se documente algún caso de variante innovadora. Una razón que puede estar detrás de estos comportamientos diferentes es la incidencia de ambos fenómenos en la comunidad: las variantes tensas de /tʃ/ casi triplican a las elisiones de /d/ (25,7% vs. 9,4%). Al haber avanzado más el proceso de tensión que el de elisión es probable que también haya ido ganando en regularidad en el grupo de palabras frecuentes. Otra razón puede ser que la elisión de /d/ está estigmatizada en Santa Cruz (Almeida 2011), mientras que no existen actitudes sociolingüísticas manifiestas hacia la tensión de la consonante africada. Es posible, pues, que la actitud negativa hacia la elisión pueda ser más fuerte que la inercia articulatoria que conduce a la elisión de sonidos frecuentes.

4.2. Elisión y frecuencia de /d/ en los morfemas y lexemas

Un aspecto importante del tipo de variación lingüística que se está tratando es si existe algún tipo de relación entre la elisión de /d/ y la frecuencia de la palabra en función del tipo de terminación y del carácter morfológico o lexemático de esta terminación. Con esta finalidad se analizará en primer lugar el resultado de la elisión de /d/ en las terminaciones *-ado(s)*, *-ada(s)*, *-ido(s)* e *-ida(s)* de los morfemas (tabla 5). Los resultados muestran que el grado de elisión más alto se registra en las palabras terminadas en *-ado(s)*, con el 18,5%. Los porcentajes de elisión en los demás contextos son poco relevantes: 3,5% en *-ada(s)*, 2,7% en *-ido(s)* y 6,5% en *-ida(s)*. Estas diferencias son significativas ($\chi^2 = 51,949$, $p = 0,000$). A partir de estos valores porcentuales pueden crearse dos categorías de morfemas: una, representada por *-ado(s)*, que promueve la elisión, y otra representada por el resto de las terminaciones, que la restringen. Por otro lado, el hecho de que haya una diferencia tan importante de elisiones entre las terminaciones *-ado(s)* y *-ada(s)* justifica que en esta investigación ambas hayan sido tratadas como dos grupos, y no como un solo grupo, tal como había hecho Bybee 2002, 2003.

	-ado(s)		-ada(s)		-ido(s)		-ida(s)	
	N	%	N	%	N	%	N	%
[ð]	432	81,5	137	96,5	219	97,3	58	93,5
[ø]	98	18,5	5	3,5	6	2,7	4	6,5
TOTAL	530		142		225		62	

TABLA 5: Frecuencias absolutas y relativas de mantenimiento y elisión de /d/ intervocálica [+ gram] según el tipo de morfema

El paso siguiente consistió en determinar si las elisiones en estos morfemas están relacionadas con la mayor o menor frecuencia de las palabras en que aparecen. Con esta finalidad, se procedió a determinar la frecuencia de estos cuatro tipos de morfemas. Los resultados (tabla 6) indican que en el grupo de palabras de baja frecuencia el porcentaje más alto de variantes elididas se localiza en los morfemas terminados en *-ado(s)* (59,5%), seguidos, en orden de frecuencias, de los morfemas terminados en *-ada(s)* (22,4%), *-ido(s)* (12,2%) e *-ida(s)* (5,7%). En el grupo de las palabras de frecuencia alta los morfemas más representados son, de nuevo, los terminados en *-ado(s)* (49,2%), seguidos por los morfemas en *-ido(s)* (39,1%). Mucho menos represen-

tados en este grupo están los morfemas en *-ada(s)* (4%) e *-ida(s)* (7,6%). Estas diferencias son significativas: $\chi^2 = 129,205$, $p < 0,000$.

Los resultados parecen confirmar en cierta medida la hipótesis que defiende la existencia de una relación directa entre la frecuencia de la palabra que contiene un morfema y la reducción fónica de dicho morfema (Bybee 2002, 2003). En efecto, *-ado(s)*, el morfema que más aparece entre las palabras de frecuencia alta (49,2%), es también el que registra el valor más alto de elisión (18,5%). Sin embargo, *-ado(s)* es también el morfema más frecuente entre las palabras de frecuencia baja (59,5%), lo que parece no apoyar del todo la tesis de partida. De lo que no cabe ninguna duda es de que *-ado(s)* es el morfema más frecuente en el corpus utilizado en esta investigación, y esto, unido al hecho de que es el morfema más frecuente entre las palabras de frecuencia alta, ha podido influir en el mayor número de elisiones que registra. Al contrario, *-ada(s)* e *-ida(s)*, los morfemas que menos aparecen en el grupo de palabras más frecuentes, figuran también entre los morfemas que menos eliden (3,5% y 6,5%, respectivamente). Además, se trata de los dos morfemas menos frecuentes en el corpus. El único caso que no parece adecuarse a la hipótesis que relaciona la elisión y la frecuencia es el del morfema *-ido(s)*, que aparece situado en segundo lugar en la escala de palabras más frecuentes (39,1%) y que, por tanto, debería haber registrado un porcentaje de elisiones más alto del que figura en la tabla 5 (2,7%). Así y todo, la tesis inicial aparece corroborada por la mayoría de los resultados.

	Frecuencia baja		Frecuencia alta	
	N	%	N	%
-ado(s)	335	59,5	195	49,2
-ada(s)	126	22,4	16	4,0
-ido(s)	70	12,2	155	39,1
-ida(s)	32	5,7	30	7,6
TOTAL	563		396	

TABLA 6: Frecuencias absolutas y relativas de los tipos de morfemas según la frecuencia de las palabras

Queda una última cuestión por resolver en el análisis del mantenimiento y elisión de /d/. Bybee 2002, 2003, pp. 148-153, argumenta que cuando las palabras que contienen /d/ morfológica se vuelven más frecuentes los hablantes tienden a tratarlas como lexemas, borrando así

el límite morfemático de su estructura. De ese modo, una palabra como *sentado* (< *sentar* + *ado*) podría ser tratada de la misma manera que una palabra como *abogado* en el caso de volverse muy frecuente. Para verificar esta tesis se llevaron a cabo análisis cruzados del mantenimiento y elisión de /d/ en función del carácter morfemático o lexemático de cada una de las terminaciones: *-ado(s)*, *-ada(s)*, *-ido(s)* e *-ida(s)*. Los resultados (tabla 7) revelan que en los contextos *-ado(s)* e *-ida(s)* la elisión de /d/ no registra diferencias significativas en los morfemas y en los lexemas ($p > 0,05$): 18,5% vs. 17,3% en *-ado(s)* y 6,5% vs. 5,8% en *-ida(s)*. Esto quiere decir que, en ambos casos, los hablantes tratan del mismo modo estas dos categorías de palabras. En las terminaciones *-ada(s)* e *-ido(s)* la elisión es más baja en los morfemas que en los lexemas: 3,5% vs. 9,3% y 2,7% vs. 10,2%, unas diferencias que son significativas $\chi^2 = 4,794$, $p = 0,029$ en *-ada(s)* y $\chi^2 = 5,932$, $p = 0,015$ en *-ido(s)*. Los resultados indican que en estos dos casos morfemas y lexemas son tratados de modo diferente. A diferencia de los resultados encontrados por Bybee 2003 en Nuevo México, en ningún caso las elisiones en los morfemas fueron más altas que en los lexemas.

Estos resultados confirman, con ciertos matices, la tesis de Bybee de que en las palabras con morfemas más frecuentes se borra cualquier estructura interna. Por un lado, *-ado(s)*, el morfema más frecuente en el corpus y el que más aparece en el grupo de palabras de frecuencia alta, registra un porcentaje de elisiones próximo al de los lexemas que tienen la misma terminación. Sin embargo, un comportamiento semejante se observa en los morfemas en *-ida(s)*, que figuran entre los menos frecuentes del corpus y que, además, tienen una escasa presencia dentro de las palabras más frecuentes. Podría plantearse la hipótesis (defendida por Bybee) de que, en virtud de un proceso analógico, el comportamiento del morfema más frecuente, *-ado(s)*, ha podido arrastrar a que otros morfemas tengan su mismo comportamiento. No obstante, de ser cierta esta tesis, lo más lógico hubiera sido que dicho proceso hubiera alcanzado antes a *-ada(s)* que a *-ida(s)*, ya que el primero es también un morfema de la primera conjugación, como *-ado(s)*, y, además, el contexto fónico en el que se encuentra /d/ se asemeja más a este que el contexto fónico de la /d/ en *-ida(s)*. Sin embargo, como ya se ha visto, el porcentaje de elisiones en el morfema *-ida(s)* es más bien bajo (6,5%). Para salvar la hipótesis de Bybee podría argumentarse que la circunstancia de que las elisiones registradas en la terminación *-ida(s)* sean casi las mismas en los morfemas y en los lexemas puede deberse a que estén operando otros factores no directamente relacionados con la frecuencia y que no han sido tenidos en cuenta en esta investigación.

Hay una última cuestión que convendría plantear. Resulta plausible afirmar que la estructura *raíz léxica + morfema* puede ser tratada como

un simple lexema cuando dicha estructura se vuelve muy frecuente. La evolución de las lenguas parece apoyar esta idea. Ahora bien, con el fin de evitar interpretaciones subjetivas acerca de cuándo se produce ese momento, tal vez haya que investigar más acerca de cuál es el umbral de frecuencia que permita decidir cuándo estructuras como las analizadas pierden la organización interna.

	Morfemas		Lexemas	
	N/T	%	N/T	%
-ado(s)	98/530	18,5	23/133	17,3
-ada(s)	5/142	3,5	31/332	9,3
-ido(s)	6/225	2,7	5/49	10,2
-ida(s)	4/62	6,5	5/86	5,8

TABLA 7: Frecuencias absolutas y relativas de elisión de /d/ en cuatro tipos de morfemas y lexemas

5. CONCLUSIONES

El análisis precedente permite alcanzar algunas conclusiones interesantes tanto desde un punto de vista teórico como metodológico. Desde un punto de vista teórico los resultados demuestran que la frecuencia de uso tiene un efecto sobre la variación de /d/ y /tʃ/ en Santa Cruz de Tenerife, pero este efecto es más bien moderado. Es cierto que las palabras de frecuencia alta favorecen tanto la elisión de /d/ como la tensión de /tʃ/, y en ambos casos las diferencias fueron estadísticamente significativas. Sin embargo, como ha revelado el análisis individual, existe un número muy importante de casos que quedan sin explicar por la teoría. Por tanto, es probable que otros factores (como el papel del acento, que suele retrasar los procesos de debilitamiento y elisión, o, en el caso de /d/, aspectos de tipo psicosocial, como las actitudes) estén interfiriendo en la función reductora o reforzadora de la frecuencia.

Los resultados relativos a /tʃ/ permiten concluir, además, que la frecuencia no solo desempeña una función reductora sobre el significante de las palabras, propiciando la aparición de las realizaciones que tienen un menor coste articulatorio, sino también una función de refuerzo fonético. El hecho de que las variantes tensas de /tʃ/ se produzcan sobre todo en las palabras más frecuentes y no en las menos frecuentes,

como se había presupuesto, puede tener que ver con el tipo de proceso fonético que supone la tensión. Por un lado, los procesos fónicos que implican debilitamiento articulatorio (tales como sonorización de oclusivas sordas, debilitamiento de consonantes implosivas, etc.) están muy arraigados en el dialecto que se ha analizado, por lo que la tensión de la articulación de la /tʃ/ tradicional puede considerarse una excepción. Por otro lado, el cambio /tʃ/ laxa > tensa supone la sustitución de una variante vernácula por otra variante con la que los canarios no están tan familiarizados. Todo ello puede provocar en los individuos una cierta inseguridad lingüística, que es la que finalmente los puede haber conducido a emplear la variante innovadora en aquellas palabras con las que se sienten más seguros: las palabras que más usan.

El papel relativamente moderado que juega la frecuencia se observa no solo cuando se analizan los porcentajes globales de palabras de frecuencia baja y alta, sino también cuando, como en el caso de /d/, se tiene en cuenta el tipo de palabra en función de su categoría léxica (morfemas vs. lexemas). Los resultados mostraron que el tipo de morfema más frecuente, *-ado(s)*, es también el que registra el porcentaje más alto de elisión. Esta correlación puede estar reforzada por el hecho de que el morfema masculino de la primera conjugación es también el más frecuente en el corpus de Santa Cruz de Tenerife. Sin embargo, los análisis realizados plantean también una serie de interrogantes acerca de en qué momento los hablantes tienden a procesar las palabras constituidas por una raíz más un morfema como si se tratara de lexemas (es decir, sin tener en cuenta la estructura interna).

Desde el punto de vista metodológico hay dos aportes que destacar. En primer lugar, la conveniencia de realizar análisis no solo de grupos de palabras (lemas) en función de la frecuencia, sino también de cada unidad léxica del corpus. Este análisis puede servir para determinar la interacción de otras variables sobre la frecuencia. En segundo lugar, esta misma idea puede extenderse al análisis de los tipos de morfemas y lexemas: parece preferible analizar las palabras en función de cada una de las terminaciones *-ado(s)*, *-ada(s)*, *-ido(s)* e *-ida(s)* que analizarlas estableciendo grupos, *-ado(s) + -ada(s)* por un lado e *-ido(s) + -ida(s)* por el otro. Es cierto que una descripción lingüística que emplee categorías más generales como «morfemas de la primera conjugación» es preferible a otra que emplee categorías más particulares como «morfemas en *-ado(s)*», «morfemas en *-ada(s)*», etc. Sin embargo, el establecimiento de una determinada categoría tiene que estar justificada tanto por la teoría como por los datos, y en el caso que se estudia los resultados obtenidos en el análisis morfológico no parecen justificar el uso de categorías generales.

REFERENCIAS BIBLIOGRÁFICAS

- ALMEIDA, M. (1994-1995): «Sociolinguistic mechanisms of phonetic change: /tʃ/ in Santa Cruz de Tenerife», *Journal of Hispanic Research* 3, pp. 45-56.
- (2003 [1999]): *Sociolingüística*, La Laguna, Universidad de La Laguna.
- (2011): «Restricciones sociolingüísticas en un caso de alternancia (morfo)fonológica: /d/ intervocálica en una comunidad de habla canaria», *Lingüística Española Actual* 33, pp. 29-53.
- ALMELA, R., CANTOS, P., SÁNCHEZ, A., SARMIENTO R. y ALMELA, M. (2005): *Frecuencias del español. Diccionario y estudios léxicos y morfológicos*, Madrid, Universitas.
- BERKENFIELD, C. (2001): «The role of frequency in the realization of English *that*», en Bybee, J. y Hopper, P. (eds.), *Frequency and the emergence of linguistic structure*, Ámsterdam-Philadelphia, John Benjamins, pp. 281-307.
- BERRUTO, G. (2005): «Dialect/standard convergence, mixing, and models of language contact: The case of Italy», en Auer, P., Hinskens, F. y Kerswill, P. (eds.), *Dialect change: Convergence and divergence in European languages*, Cambridge, Cambridge University Press, pp. 81-95.
- BLAS ARROYO, J. L. (2006): «*Hasta aquí hemos llega(d)o*: ¿un caso de variación morfológica? Factores estructurales y estilísticos en el español de una comunidad bilingüe», *Southwest Journal of Linguistics* 25, pp. 1-35.
- BRITAIN, D. (2009): «One foot in the grave? Dialect death, dialect contact, and dialect birth in England», *International Journal of the Sociology of Language* 196-197, pp. 121-155.
- BYBEE, J. (2002): «Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change», *Language Variation and Change* 14, pp. 261-290.
- (2003): *Phonology and language use*, Cambridge, Cambridge University Press.
- (2006): «From usage to grammar: The mind's response to repetition», *Language* 82, pp. 711-733.
- (2007): *Frequency of use and the organization of language*, Oxford-Nueva York, Oxford University Press.
- DÍAZ CAMPOS, M. (2004): «Acquisition of sociolinguistic variables in Spanish: Do children acquire individual lexical forms or variable rules?», en Face, T. L. (ed.), *Laboratory approaches to Spanish phonology*, Berlín, Mouton de Gruyter, pp. 221-236.
- y RUIZ-SÁNCHEZ, C. (2008): «The value of frequency as a linguistic factor: The case of two dialectal regions in the Spanish speaking world», en Westmoreland, M. y Thomas, J. A. (eds.), *Selected proceedings of the 4th workshop on Spanish sociolinguistics*, Somerville, Cascadilla Proceedings Project, pp. 43-53.
- FILE-MURIEL, R. (2009): *Lexical frequency and s-realization in Barranquilla, Colombia*, Puentes del Sarre, VDM.
- GONZÁLEZ, C. (2006): «Efecto de la posición en la oración y la frecuencia léxica en /d/ final en español del País Vasco», en Face, T. L. y Klee, C. A. (eds.), *Selected proceedings of the 8th Hispanic linguistics symposium*, Somerville, Cascadilla Proceedings Project, pp. 89-102.

- HOOPER, J. B. (1976): «Word frequency in lexical diffusion and the source of phonological change», en Christie, W. (ed.), *Current progress in historical linguistics*, Ámsterdam, North Holland, pp. 96-105.
- JULLIAND, A. G. y CHANG-RODRÍGUEZ, E. (1964): *Frequency dictionary of Spanish words*, La Haya, Mouton.
- JURAFSKY, D., BELL, A., GREGORY, G. y RAYMOND, W. D. (2001): «Probabilistic relations between words: Evidence from reduction in lexical production», en Bybee, J. y Hopper, P. (eds.), *Frequency and the emergence of linguistic structure*, Ámsterdam-Philadelphia, John Benjamins, pp. 229-254.
- KIPARSKY, P. (1980): «Historical Linguistics», en Lyons, J. (ed.), *New horizons in linguistics*, Londres, Penguin, pp. 302-315.
- KLATT, H. (1980): «Pauses as indicators of cognitive functioning in aphasia», en Dechert, H. W. y Raupach, M. (eds.), *Temporal variables in speech: Studies in honour Frieda Goldman-Eisler*, La Haya-París-Nueva York, Mouton, pp. 113-120.
- LAVIER, J. (1995): *Principles of phonetics*, Cambridge, Cambridge University Press.
- MENÉNDEZ PIDAL, R. (1980 [1904]): *Manual de gramática histórica española*, Madrid, Espasa-Calpe.
- MORENO FERNÁNDEZ, F. (2008): «Cambios vivos en el plano fónico del español: variación dialectal y sociolingüística», en Cano, R. (coord.), *Historia de la lengua española*, Barcelona, Ariel, pp. 973-1009.
- PHILLIPS, B. S. (1981): «Lexical diffusion and Southern *tune, duke, news*», *American Speech* 56, pp. 72-78.
- (1984): «Word frequency and the actuation of sound change», *Language* 60, pp. 320-342.
- (2001): «Lexical diffusion, lexical frequency, and lexical analysis», en Bybee, J. y Hopper, P. (eds.), *Frequency and the emergence of linguistic structure*, Ámsterdam-Philadelphia, John Benjamins, pp. 123-136.
- ROJO, G. (2011): «Frecuencia de inventario y frecuencia de uso en los elementos gramaticales», *Revista Española de Lingüística* 41, pp. 5-43.
- TRUDGILL, P. (1986): *Dialects in contact*, Oxford, Basil Blackwell.
- ZIPF, G. K. (1965 [1935]): *The Psycho-biology of language: An introduction to dynamic philology*, Cambridge, The MIT Press.