

## Reducción fonética de la secuencia consonántica /-st-/ en el español de Bogotá\*

### Phonetic reduction of the consonant sequence /-st-/ in Bogotá Spanish

JOSÉ ALEJANDRO CORREA  
LINDA CAROLINA RODRÍGUEZ

Instituto Caro y Cuervo, Laboratorio de Fonética Experimental, Colombia.  
Correo electrónico: alejandro.correa@caroycuervo.gov.co; linda.rodriguez@caroycuervo.gov.co

En español, la secuencia /-st-/ es frecuente en palabras de función y en palabras de contenido. En la variedad hablada en Bogotá (Colombia), estas consonantes se pronuncian como consonantes dentoalveolares [-ʃt̪-] o como una sibilante seguida de una fricción con baja amplitud [-ʃt̪̥-]. Este artículo examina la frecuencia de aparición de estas realizaciones en habla narrada y en habla leída, así como los efectos de la tonicidad de la sílaba, el tipo de vocal y el sexo del hablante sobre las características acústicas (duración, centro de gravedad, asimetría y curtosis) de esta secuencia. Los resultados muestran que la reducción fonética de /-st-/ varía de hablante a hablante, mientras que el acento y las vocales medias /e, o/ favorecen la realización canónica [-ʃt̪-]. La consonante /s/ se mantuvo acústicamente estable en las diferentes condiciones, mientras que las características espectrales de /t/ sugieren un alto grado de intervariabilidad e intravariabilidad.

*Palabras claves:* reducción fonética, estilos de habla, variación idiolectal, español de Colombia.

In Spanish, the sequence /-st-/ is highly frequent in functional and content words. In the variety spoken in Bogotá (Colombia) these consonants are pronounced as [-ʃt̪-] as well as [-ʃt̪̥-] (a dentoalveolar fricative followed by a low-amplitude frication noise). This study investigates the occurrence of these phonetic realizations in a sample of unscripted narratives and read speech. We examine how the acoustic properties (duration, center of gravity, skewness and kurtosis) of the sequence are affected by stress, vowel type and sex. The results show that the phonetic reduction of /-st-/ depends on the differences between individual speaker productions. Furthermore, stress and close-mid vowels /e, o/ contribute to the occurrence of the canonical form [-ʃt̪-]. In my data, the acoustic

---

\* Este trabajo fue realizado en el marco del proyecto *Propiedades fonéticas de los estilos de habla en el español bogotano*, financiado por el Instituto Caro y Cuervo, y dirigido por José Alejandro Correa.

characteristics of /s/ remain stable across all the experimental conditions, and spectral properties of /t/ suggest strong inter- and intra-speaker variation.

*Key words:* phonetic reduction, speech styles, idiolectal variation, Colombian Spanish.

## 1. INTRODUCCIÓN

### 1.1. La reducción fonética

La pronunciación de las palabras en habla conectada<sup>1</sup> se aparta enormemente de la pronunciación de palabras aisladas de contexto (*citation form*) que sirven de base para los análisis fonológicos tradicionales. Durante una conversación con un amigo o mientras hacemos una breve narración es normal elidir consonantes y sílabas, producir las vocales con poca precisión, fusionar palabras e, incluso, crear secuencias de sonidos que resulten ininteligibles para los hablantes no nativos de nuestra lengua. Por ejemplo, en inglés, la oración *do you have time?* puede pasar de la forma canónica /du ju hæv t<sup>h</sup>aɪm/ a [tjutəm] (Ernestus y Warner 2011: 254), o palabras como *apparently* llegan a reducirse de cuatro sílabas [əp<sup>h</sup>ɛrəntli] a dos [p<sup>h</sup>ɛrɪ] (Johnson 2004). Si bien estas realizaciones podrían ser consideradas variaciones ocasionales o erráticas, las investigaciones de los últimos años han mostrado que la reducción fonética es un fenómeno universal que ocurre de manera sistemática y reglada (Kohler 1998; Barry y Andreeva 2011).

La reducción fonética y otros fenómenos típicos del habla conectada se han explicado apelando al principio de la economía articulatoria. Una de sus versiones más aceptadas, la teoría de la hiper e hipoarticulación (Lindblom 1990), sostiene que la producción del habla es adaptativa y está “prospectivamente organizada”, de modo que el hablante ajusta su articulación a las demandas de la situación comunicativa. El estímulo acústico resultante debe asegurar la discriminación de la información léxica, por lo cual se dice que hay un balance entre la economía articulatoria y la distintividad de la información. La teoría predice, por ejemplo, que los hablantes producirán los sonidos de su lengua con mayor precisión en situaciones formales, mientras que en habla espontánea se reducirán aquellas partes de la señal que no comprometen la interpretación del mensaje.

Un ejemplo del español es el fenómeno de la reducción fonética de las vocales. En un estudio realizado por Harmegnies y Poch-Olivé (1992), se muestra que en habla espontánea los formantes (F1 y F2) se solapan y son más variables, mientras que en la lectura de palabras presentan menos varianza y se ubican en la periferia del triángulo vocálico. En consecuencia, en habla conversacional las vocales se acercan a una vocal

<sup>1</sup> Se entiende por *habla conectada* el estudio de los fenómenos fonéticos que tienen lugar en la cadena de habla. Este concepto se opone al análisis de palabras aisladas que se utiliza para identificar fonemas y pares mínimos (Jones 1973; Crystal 2008). No debe confundirse con un estilo de habla en particular.

central media o *schwa* [ə], y en la lectura de palabras mantienen su timbre (/i e a o u/). En español, esta variación no compromete la identificación del mensaje porque se trata de una adaptación a la situación comunicativa que pasa desapercibida para los hablantes/ oyentes.

Además del estilo de habla, la tonicidad de la sílaba juega un papel importante en el grado de reducción fonética. En un estudio sobre el sueco, Engstrand (1988) analizó las características articulatorias y acústicas de las vocales /i a u/ en función del acento (tónicas *vs.* átonas) y la velocidad de elocución (habla lenta *vs.* rápida). Encontró que las átonas tienen un primer formante más bajo (menor apertura oral) y las tónicas se producen con una mayor constricción lingual. De Jong (1995), en un trabajo que trata sobre la influencia del acento sobre el movimiento lingual y mandibular en inglés, encontró que a mayor grado de prominencia silábica, mayor amplitud y mayor velocidad del movimiento mandibular: en las sílabas tónicas y nucleares las vocales presentaron mayor apertura y las consonantes alveolares un mayor cierre. Para el autor, las sílabas tónicas o acentuadas se hiperarticulan a nivel local (*localized hyperarticulation*) porque permiten identificar la información léxica y sirven de punto de anclaje para la percepción de la palabra.

Ahora bien, como afirma Simpson (2013), uno de los principales problemas de la teoría de la hiper e hipoarticulación es que no pone límites a la reducción fonética. En particular, el principio de economía no siempre explica cuáles son las razones por las cuales los hablantes de una misma lengua producen diferentes realizaciones fonéticas para una misma unidad léxica (Simpson 2013: 163). Por ejemplo, en un estudio reciente sobre la reducción fonética de las vocales del español de Bogotá, Correa (2017) muestra que /i e a o u/ tienen un primer formante más bajo en las átonas producidas en habla no leída que en las tónicas producidas en una lectura. Sin embargo, los hablantes no presentaron el mismo grado de reducción: algunos produjeron /a/ como una vocal abierta [a], otros como una vocal medio abierta [a̠] o como una *schwa* [ə]. Ante este tipo de datos se puede argumentar, como sostienen Ernestus y Warner (2011: 255), que la reducción fonética no es categórica, sino gradiente y, por tanto, abarca un amplio rango de realizaciones. Por el contrario, se puede argumentar, como Ladefoged (1983), que el lenguaje humano es único, que las diferencias entre las lenguas son impredecibles y no pueden explicarse por medio de principios generales como la hiper e hipoarticulación.

Otro problema fundamental es la noción misma de *reducción fonética* pues, como señala Simpson (2001), en ocasiones el resultado de los procesos de debilitamiento conlleva una mayor complejidad articulatoria. A modo de ilustración, el autor menciona la pronunciación de /t/ en algunos hablantes del inglés RP (*Received Pronunciation*). En este caso, el resultado de la reducción de /t/ es una fricativa, un sonido que requiere un gesto articulatorio lo suficientemente preciso como para permitir que el paso de la corriente de aire genere una turbulencia. Es decir, el resultado de la economía articulatoria es, paradójicamente, un sonido más complejo que el original. Como veremos en la próxima sección, este caso problemático también ocurre en español y su estudio puede ayudarnos a comprender la reducción fonética y sus causas.

### 1.2. La secuencia /-st-/ en el español de Bogotá

La secuencia<sup>2</sup> /-st-/ ocurre en español cuando hay dos sílabas adyacentes en las que /s/ está en posición de coda y /t/ está en posición de ataque. Esta secuencia tiene una alta frecuencia de aparición porque se encuentra al interior de palabras de función (por ejemplo, el verbo *estar*) y de un gran número de palabras de contenido<sup>3</sup>. Como ilustra la figura 1, en el español de Bogotá la pronunciación normativa de la secuencia es [-st̚-] e involucra cuatro fases articulatorias-acústicas: una fricativa sibilante, el cierre u oclusión, la explosión y el *voice onset time* (VOT), que es el intervalo entre la explosión y el inicio de la fonación de la vocal siguiente. Existe una segunda variante en la que el hablante no logra completar el cierre y produce una realización asibilada de /t/. En esta realización —que transcribimos [-st̪̚-] siguiendo a Mitterer y Ernestus (2006)— la fricción se inicia en [s̪] y se mantiene hasta el inicio de la vocal, aunque la amplitud de [t̪] es considerablemente menor a la sibilante precedente (figura 2).

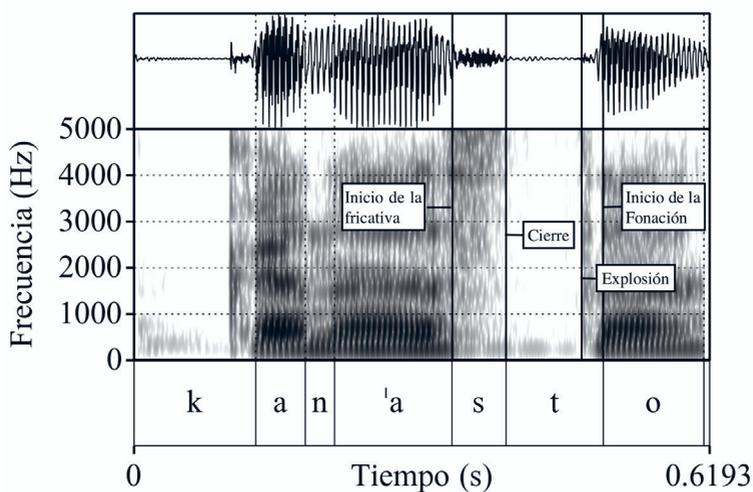


Figura 1. Forma de onda, espectrograma y segmentación de la variante T1, la cual involucra la fricativa [s̪], un cierre, la disolución del cierre o explosión y el VOT.

<sup>2</sup> Siguiendo a Pulgram (1965: 76), en este artículo, *secuencia* es un conjunto de consonantes que aparecen en diferentes sílabas, mientras que en un *grupo consonántico* (*cluster*) se presentan en la misma sílaba. De esta manera /-st-/ es una secuencia en español, mientras que en inglés es un grupo consonántico que aparece en palabras como *stop*, *style* o *stuff*.

<sup>3</sup> El buscador de patrones fonológicos BuFón (Alves *et al.* 2011) arroja, por ejemplo, alrededor de 35 000 palabras de contenido.

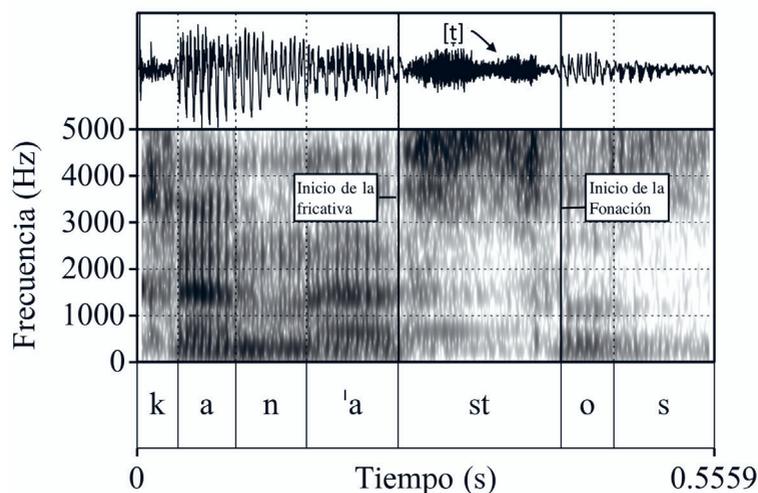


Figura 2. Forma de onda, espectrograma y segmentación de la variante T2, la cual involucra una fricativa sibilante y una realización asibilada de /t/.

La reducción que se observa en la figura 2 no la definimos aquí como el resultado de la economía articulatoria, sino como una simplificación de las fases de producción que se refleja claramente en la señal acústica. Es decir, /t/ no involucra periodo de cierre, explosión y VOT, sino que se reemplazan por una fricción que, como dijimos, requiere mayor esfuerzo y precisión articulatoria. En este artículo analizamos las dos consonantes en conjunto, puesto que también nos interesa saber si el espectro de /s/ se modifica cuando hay asibilación de /t/.

El trabajo de Mitterer y Ernestus (2006) sobre el debilitamiento de /t/ en neerlandés nos da algunas pistas sobre las causas del fenómeno que nos interesa aquí. Los autores analizaron /t/ en un corpus con diferentes estilos de habla y como factor fonético tuvieron en cuenta la consonante precedente. Sus resultados muestran que la /t/ del neerlandés se elide con mayor frecuencia después de /s/ que de otras consonantes como /n/ o /x/. En concordancia con la teoría de Lindblom (1990), la reducción fue más frecuente en el habla conversacional y en narraciones, mientras que en las frases leídas las apariciones fueron escasas. En el mismo estudio se realizaron varios experimentos de percepción que mostraron que la presencia de /s/ permite a los oyentes reconstruir la oclusión que los hablantes eliden, y este efecto fue más sobresaliente cuando el estímulo era una palabra de la lengua. Los autores concluyen que en esta lengua el conocimiento léxico y el entorno consonántico son los factores que permiten compensar perceptivamente la reducción /t/.

Fernández Planas (1999) presenta datos articulatorios sobre /-st-/ en función de la velocidad de elocución en español. En este trabajo se analiza la duración de los sonidos y del tiempo de activación de los electrodos en la zona dentoalveolar durante la producción de la

secuencia *asta*. El análisis electropalatográfico sugiere que la magnitud del gesto lingual de /t/ y la duración de /-st-/ disminuyen cuando aumenta la velocidad, mientras que /s/ se mantiene estable en las diferentes condiciones. Una de las observaciones más interesantes de este trabajo es que, si bien no hay reducción de /t/, se reportan diferencias individuales importantes en el grado de constricción. Esto comprueba, como han demostrado Stevens y Harrington (2011), que las preferencias individuales juegan un papel importante en la variación fonética.

Los trabajos de Mitterer y Ernestus (2006) y Fernández Planas (1999) no exploran otros factores que pueden ser determinantes para explicar esta secuencia, por ejemplo, la tonicidad de la sílaba y el tipo de vocal tras /-st-/. De acuerdo con de Jong (1995), podemos suponer que la tonicidad de la sílaba a la que pertenece /t/ puede inducir a una hiperarticulación de la oclusiva, de manera que la asibilación debería ser más frecuente en palabras como *esta* y menos probable en palabras como *está*. En cuanto a la influencia de la vocal, es probable que la vocal abierta /a/ favorezca un cierre incompleto de /t/, puesto que la lengua y la mandíbula deben realizar movimientos de cierre y apertura en un periodo breve de tiempo. Por el contrario, en la pronunciación de una palabra como *estilo*, las consonantes y la vocal tan solo requieren un cierre progresivo en la parte anterior del tracto vocal.

De acuerdo con lo anterior, el presente artículo tiene tres objetivos. Primero, busca determinar cuál es la frecuencia de aparición de la realización normativa [-ʃt̪-] (en adelante, *variante T1*) y de la realización asibilada [-ʃt̪̺-] (en adelante, *variante T2*). Segundo, busca obtener datos sobre las características acústicas de las fases involucradas en la producción de cada variante (figuras 1 y 2). Tercero, busca determinar si la tonicidad de la sílaba, el tipo de vocal y las preferencias individuales determinan la aparición de las variantes y/o afectan las características acústicas de la secuencia. Para este fin, en la segunda sección analizamos una muestra de narraciones que nos permitirá obtener datos característicos de una situación no controlada. En la tercera sección analizamos una muestra de habla leída en la que estudiamos la influencia de la tonicidad de las sílabas, el tipo de vocal que sigue después de /-st-/ y el sexo de los hablantes. Este último factor es importante porque los estudios muestran que los hombres tienen preferencia por la reducción fonética de las consonantes y vocales, mientras que las mujeres prefieren una articulación cuidadosa o hiperarticulada (Byrd 1994; Whiteside 1996; Simpson 2009).

## 2. ESTUDIO 1: ANÁLISIS DE HABLA NARRADA

### 2.1. Método

#### 2.1.1. Materiales

Los datos provienen del análisis de las muestras de habla producidas por seis hombres con edades entre los 27 y 36 años, nacidos y criados en la ciudad de Bogotá (Colombia). En el momento de la grabación, los participantes eran estudiantes de posgrado

o investigadores de una institución educativa. Ninguno reportó problemas auditivos y la participación en el experimento fue voluntaria. Para estudiar la realización de la secuencia en habla conectada se usó, como tarea de recolección de datos, la narración de una historia: cada participante observó el video *La historia de las peras* (Wallace 1980) en un ordenador y luego lo narró detalladamente al entrevistador, en una velocidad de habla normal. Esta tarea de recolección de datos permitió limitar el inventario de palabras que los hablantes usaron en la descripción y obtener un número proporcional de casos (tabla 1). Las grabaciones se realizaron en la cabina insonorizada del Laboratorio de Fonética Experimental del Instituto Caro y Cuervo con una tasa de muestreo de 44 100 Hz/16 Bits. Luego se etiquetaron y analizaron con Praat (Boersma y Weenink, 2016) siguiendo el procedimiento que se explica a continuación.

### 2.1.2. Mediciones

Las realizaciones fonéticas de la secuencia /-st-/ presentan una enorme variabilidad, de manera que se agruparon en dos tipos para cuantificarlas. En un primer grupo están las secuencias que involucran una fricativa sibilante /s/ seguida de una oclusiva dentoalveolar /t/. En este caso, la existencia de la oclusiva /t/ se determinó por la presencia de un cierre y la explosión previa al inicio de la sonoridad (figura 1). En el segundo grupo se incluyeron todos aquellos casos en los que el hablante no completa el cierre de /t/ y, en cambio, produce una fricativa sibilante de amplitud y distribución espectral variables (figura 2).

Para superar la dicotomía planteada arriba y capturar la naturaleza gradiente de la realización fonética de /-st-/, medimos la duración de las fases de producción de cada variante y sus características acústicas. En las secuencias con consonante oclusiva se midieron la duración de la fricativa y la duración de /t/, entendida esta última como la duración del cierre consonántico más el VOT. En las secuencias compuestas por dos zonas de fricción, medimos la duración total ante la dificultad de establecer límites precisos entre fricativa y oclusiva asibilada (sobre este punto véase Turk *et al.* 2006: 15). Sin embargo, para medir las características espectrales de la asibilación, en el etiquetado se diferenciaron las dos zonas de fricción teniendo en cuenta la amplitud de la onda y el espectrograma de banda ancha.

Para obtener información acústica de /s/ y de la fricción resultante del cierre incompleto de /t/, se midieron tres momentos espectrales: el centro de gravedad (COG), la asimetría (*skewness*) y la curtosis (*kurtosis*) (Forrest *et al.* 1998; Harrington 2010). Los momentos espectrales son medidas de tendencia central y medidas de dispersión de los componentes frecuenciales del espectro. El centro de gravedad es el promedio de las frecuencias con mayor amplitud. La asimetría aporta información sobre la forma del espectro con referencia al centro de gravedad: en cuanto más se acerquen los valores a cero más se asemejarán a la distribución normal; los valores negativos indican el predominio de frecuencias altas y los valores positivos una concentración de la energía espectral en las frecuencias bajas (Harrington 2010: 300); (Forrest *et al.* 1998). Finalmente, la interpretación generalmente aceptada considera que la curtosis positiva permite determinar

si el espectro tiene picos de frecuencia definidos y una curtosis negativa indica un espectro plano (Harrington 2010: 300); (Forrest *et al.* 1998). Sin embargo, como ha mostrado Westfall (2014), la curtosis no es una medición útil para determinar si una distribución de datos tiene picos, sino que indica la presencia de valores atípicos (*outliers*).

Los momentos espectrales se calcularon de manera automática con un *script* de Praat (Boersma y Weenink, 2016), teniendo en cuenta todo el intervalo del sonido, y se aplicó un filtro hasta los 1000 Hz para evitar la influencia de la sonoridad. Todos los análisis estadísticos se realizaron con el entorno de programación R (R Development Core Team 2012). La regresión logística binomial que se presenta en la tercera sección se realizó con el paquete *lrm* (Harrell 2017) y los modelos de regresión de efectos mixtos con el paquete *nlme* (Pinheiro *et al.* 2017).

## 2.2. Resultados

### 2.2.1. Apariciones de las variantes en habla narrada

La secuencia /-st-/ se encontró en palabras de contenido (53 %) y en palabras de función (47 %). Entre estas últimas, el 37.7 % corresponde a diferentes conjugaciones del verbo *estar* y las palabras restantes a determinantes (5.76 %) y preposiciones (2.56 %). La mayoría de las palabras de contenido analizadas fueron usadas por todos los hablantes debido a su relación con las escenas del video utilizado para obtener los datos. Comparamos, por ejemplo, la palabra *canasta* (22.3 %), su variante *cesta* (7.69 %) o el adverbio *justamente* (6.4 %); solo el 16.7 % de palabras de contenido tuvo una o dos apariciones.

El 49 % de las secuencias se produjo con la variante T1 y el 51 % restante se produjo con la variante T2. El análisis de los casos mostró que las palabras más frecuentes se pronunciaron con las dos variantes consideradas (incluso por el mismo hablante). Las conjugaciones del verbo *estar*, por ejemplo, se pronunciaron el 18.4 % de las veces con T1 y el 19 % con la variante asibilada. A nuestro juicio, esto descarta de momento la influencia de la categoría léxica sobre la reducción de /-st-/.

Como se ve en la figura 3 y en la tabla 1, las dos realizaciones fonéticas ocurren en cada idiolecto, pero las proporciones de aparición varían de hablante a hablante. Realizamos una prueba de ji-cuadrado para comprobar si las realizaciones fonéticas son independientes del hablante, de manera que encontramos una dependencia significativa ( $\chi^2(5) = 49.7$ ,  $p < .001$ ). El diagrama de barras muestra, por ejemplo, que algunos participantes prefieren una pronunciación conservadora con tres fases acústicas (DC09 y DC11), otros prefieren una pronunciación asibilada o T2 (DC01, DC03 y DC07), mientras que un hablante (DC05) acude a las dos pronunciaciones. Para determinar el aporte de cada individuo en la obtención de este resultado, se calcularon los *residuos estandarizados de Pearson* (Levshina 2015: 220-221). De acuerdo con los resultados, todos los participantes contribuyen de manera significativa, pero los identificados como DC07 y DC11 presentan mayor diferencia entre los valores esperados y los valores observados y, por tanto, contribuyen especialmente

al nivel de significatividad de  $p = .001$ . En la figura 3 podemos comprobar que DC07 y DC11 tienen distribuciones opuestas y algo extremas de las pronunciaci3nes estudiadas.

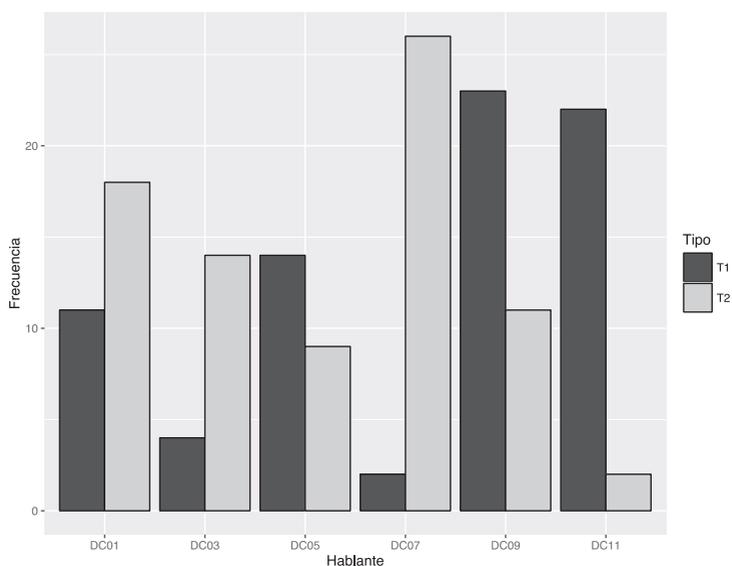


Figura 3. Número de apariciones de las variantes fonéticas de la secuencia /-st-/ según el hablante.

Tabla 1. Proporciones y número de apariciones (paréntesis) de las variantes fonéticas de la secuencia /-st-/ por hablante

| Hablante | T1        | T2        | Total     |
|----------|-----------|-----------|-----------|
| DC01     | 0.07 (11) | 0.11 (18) | 0.18 (29) |
| DC03     | 0.02 (4)  | 0.09 (14) | 0.11 (18) |
| DC05     | 0.09 (14) | 0.06 (9)  | 0.15 (23) |
| DC07     | 0.01 (2)  | 0.16 (26) | 0.17 (28) |
| DC09     | 0.15 (23) | 0.07 (11) | 0.22 (34) |
| DC11     | 0.14 (22) | 0.01 (2)  | 0.15 (24) |
| Total    | 0.49 (76) | 0.51 (80) | 1 (159)   |

### 2.2.2. Duración de /-st-/ en habla narrada

Como se dijo anteriormente, la duración de las fases acústicas de la secuencia /-st-/ se calculó con base en la variante T1 ante la imposibilidad de delimitar entre la fricativa sibilante y la fricción correspondiente a /t/ en la variante T2. La fricativa /s/ tuvo una duración de  $67.4 \pm 24.5$  ms<sup>4</sup>, y la oclusiva /t/  $77.6 \pm 14.5$  ms (cierre =  $46.7 \pm 14.2$  ms VOT=  $31 \pm 6.4$  ms). La duración de T1 fue de  $143.8 \pm 32$  ms y presentó un descenso de  $-20$  ms en la variante T2 ( $123.2 \pm 31$  ms). Realizamos la prueba no paramétrica de Wilcoxon y encontramos que la diferencia entre la duración total de la variante no asibilada y la variante asibilada es altamente significativa ( $W = 4081$ ,  $p < .001$ ).

De acuerdo con los resultados presentados, la variante [-s̥t̥-] se distingue de [-st̥-] por tener una menor duración. Sin embargo, la duración de [-s̥t̥-] está por debajo de los valores presentados por Fernández Planas (1999) para la secuencia *asta* pronunciada en velocidad de habla normal ( $218.4 \pm 32.4$  ms) y, en cambio, se asemeja a los valores reportados por la autora para el habla rápida ( $166.3 \pm 27.9$  ms) y muy rápida ( $141.2 \pm 12.8$  ms). Dado que en dicho trabajo se analiza la secuencia *asta* pronunciada de manera aislada, consideramos que las diferencias están asociadas principalmente a la tarea de recolección de datos usada en cada estudio, es decir, la producción consciente de una secuencia aislada de contexto en condiciones de laboratorio conlleva necesariamente a una mayor duración de los sonidos que cuando la secuencia se produce en palabras reales dentro de frases conectadas.

### 2.2.3. Momentos espectrales en habla narrada

La fricativa /s/ tiene un centro de gravedad de  $4220 \pm 1602$  Hz en T1 y desciende a  $3925 \pm 1817$  Hz en T2. Esta diferencia es más sobresaliente en algunos hablantes, pero no es estadísticamente significativa ( $t_{(153,16)} = 1.0769$ ,  $p = .2$ ). De acuerdo con los valores de la asimetría, la energía espectral se acerca a una distribución normal (T1=  $0.5 \pm 0.8$ ,) aunque, en T2, /s/ tiene la desviación típica más alta (T2=  $0.7 \pm 1.2$ ). Al comparar ambos grupos de datos no encontramos diferencias significativas con respecto a la asimetría ( $W = 2866$ ,  $p = .5$ ). Finalmente, la curtosis de /s/ indica que la distribución del espectro tiene pocos valores atípicos (T1=  $-0.1$ , T2=  $1.2$ ) y, como en el caso del centro de gravedad y la asimetría, las diferencias entre variantes no son significativas ( $W = 2516$ ,  $p = 0.06$ ). Estos datos indican que las características acústicas de /s/ se mantuvieron estables en las dos variantes estudiadas, de manera que la reducción de /t/ no modifica la fricativa precedente.

La tabla 2 ilustra los momentos espectrales de /s/ y /t/ en la variante T2 por cada hablante. Hemos excluido al hablante DC11, pues produjo tan solo dos realizaciones asibiladas. Nótese que el centro de gravedad de /t/ tiene un valor medio de  $1698 \pm 1454$  Hz, es decir, cerca de la mitad del centro de gravedad de /s/. El hecho más sobresaliente, sin embargo, es que hay amplias diferencias entre hablantes: algunos tienen una realización

<sup>4</sup> En lo que sigue, así reportamos la media  $\pm$  la desviación típica de los datos.

de /t/ con un centro de gravedad alto (DC03, DC05 y en menor grado DC01) mientras que otros producen una fricativa con un centro de gravedad ubicado en la parte baja del espectro (DC07, DC09). En la misma línea, la asimetría ( $4.1 \pm 3.5$ ) indica que los datos se alejan de la distribución normal y los valores altos de la curtosis ( $34.2 \pm 65.7$ ) muestran que el espectro de la variante asibilada de /t/ contiene un gran número de valores extremos. A partir de estos datos se observa que los hablantes producen la oclusiva dentoalveolar con diferentes grados de constricción y, por tanto, el grado de asibilación presenta un alto grado de intervariabilidad.

Un punto que es necesario resaltar es la interpretación de los valores de curtosis. Como se dijo anteriormente, la interpretación difundida entre fonetistas y estadísticos es que los valores positivos y superiores a 9 indican un espectro con picos claramente definidos, mientras que los valores bajos y negativos muestran un espectro plano y difuso (Jongman *et al.* 2000); (Newell y Hancock 1984). Nótese que todos los valores correspondientes a /t/ son positivos y bastante altos. Por ejemplo, los hablantes DC07 y DC09 tienen un centro de gravedad menor a 1000 Hz y, de acuerdo con la asimetría, un espectro que se aleja de la distribución normal. Si la interpretación generalmente aceptada de la curtosis fuera cierta, para estos hablantes el valor medio debería ser negativo y no excesivamente alto, como muestra la tabla 2. Nos suscribimos, por tanto, al planteamiento de Westfall (2014), quien afirma que esta medición permite determinar si los datos tienen valores atípicos. Así las cosas, la curtosis baja indica que el hablante es consistente en la manera como asibila /t/ (DC03 y DC05), mientras la curtosis alta sugiere que el hablante asibila la oclusiva dentoalveolar con diferentes grados de constricción (DC01, DC07, DC09).

Tabla 2. Momentos espectrales (centro de gravedad, asimetría y curtosis) de /s/ de la realización asibilada de /t/ (derecha)

| Hablante | COG /s/     | Asimetría /s/ | Curtosis /s/ | COG /t/     | Asimetría /t/ | Curtosis /t/ |
|----------|-------------|---------------|--------------|-------------|---------------|--------------|
| DC01     | 3320 (1204) | 1.0 (0.6)     | 0.5 (1.9)    | 1697 (1155) | 3.1 (2.1)     | 15.3 (21.9)  |
| DC03     | 5037 (1438) | -0.001 (0.6)  | -0.08 (0.8)  | 3350 (1592) | 1.1 (1.1)     | 1.2 (3.7)    |
| DC05     | 4855 (1307) | 0.3 (0.29)    | -0.7 (0.2)   | 2612 (1380) | 2.0 (1.5)     | 5.7 (10.4)   |
| DC07     | 4344 (1870) | 0.3 (1.0)     | -0.07 (3.1)  | 955 (959)   | 5.8 (3.5)     | 55.3 (79)    |
| DC09     | 2310 (1541) | 1.6 (1.17)    | 2.7 (4.5)    | 760 (407)   | 6.7 (4.4)     | 80.4 (100)   |
| Media    | 3925 (1817) | 0.7 (1.2)     | 1.2 (7.2)    | 1698 (1454) | 4.1 (3.5)     | 34.2 (65.7)  |

Como se ve, la aparición de las variantes T1 y T2 y sus características acústicas dependen de las preferencias individuales: unos hablantes favorecen una pronunciación

conservadora en la que se mantiene la oclusión [-st̚-], otros asibilan la oclusión [-st̚h-] la mayoría de las veces, mientras que otros usan ambas realizaciones. La intervariabilidad también se refleja en el grado de asibilación de /t/, pues algunos de los participantes producen /t/ con un centro de gravedad semejante a los valores de /s/, y otros producen la asibilación con un espectro difuso y con una gran cantidad de valores atípicos. Esto significa que el grado de asibilación es gradiente y no categórico, como lo podría suponer la división del fenómeno en dos variantes. La fricativa /s/ presentó estabilidad en la duración y en las características espectrales, de manera que, como se dijo atrás, podemos afirmar que la reducción de /t/ no modifica las características acústicas desde la consonante precedente.

### 3. ESTUDIO 2: ANÁLISIS DE HABLA LEÍDA

Los resultados presentados anteriormente sugieren que la reducción fonética no solo se puede explicar como un fenómeno asociado al habla conectada, sino que el individuo puede limitar o propiciar la reducción de los sonidos de una lengua. No obstante, para conocer la importancia de las variaciones individuales debemos analizar otros factores como el estilo, la tonicidad, la vocal tras /-st-/ y el sexo del hablante. A continuación, presentamos los resultados de un segundo estudio en el que controlamos estos factores en una tarea de lectura. Si las preferencias individuales determinan la reducción de la secuencia /st/, entonces debemos descartar la influencia de estas variables sobre la aparición, duración y características espectrales de estos sonidos.

#### 3.1. Método

##### 3.1.1. Materiales

Como se dijo en la segunda sección, este experimento tiene como objetivo establecer si el acento, el estilo de habla (habla leída), el tipo de vocal y el sexo del hablante imponen restricciones sobre la reducción fonética de las fases articulatorias y determinan las características acústicas de la secuencia /-st-/. Para este fin, se elaboraron tres textos en los que se incluyeron 50 palabras de contenido con una longitud de tres sílabas. Las palabras se eligieron con el buscador de patrones fonológicos BuFón (Alves *et al.* 2011) teniendo en cuenta las siguientes condiciones: la presencia de una vocal central media átona /e/ delante de la consonante sibilante /s/, y cada una de las vocales del español /i e a o u/ tras la oclusiva /t/. Por cada uno de los 8 hablantes se obtuvieron 5 ejemplos de las 5 vocales y en 2 condiciones de acento ( $8 \times 5 \times 5 \times 2 = 400$  casos).

En este estudio participaron 4 mujeres y 4 hombres nacidos en Bogotá con edades entre los 27 y 45 años, quienes leyeron los textos en una velocidad de habla normal. Todos los participantes eran profesionales de diferentes áreas y, cuando se recolectaron los datos, trabajaban como empleados en una institución pública. Las grabaciones se realizaron en

una cabina insonorizada y se analizaron con Praat (Boersma y Weenink 2016) siguiendo el procedimiento descrito en la sección 2.1.2.

### 3.2. Resultados

#### 3.2.1. Apariciones de las variantes en habla leída

La figura 4 y la tabla 3 presentan las apariciones y las proporciones de las realizaciones estudiadas en habla leída: el 75 % corresponde a la variante T1 y el 25 % a T2, lo cual muestra que esta última tiene un 26 % de apariciones menos que en las narraciones. A nivel individual, los participantes produjeron [-st̪-] entre el 7 y el 10 % de las veces, y [-st̪̥-] entre el 2 y el 5.5 %. Algunas mujeres (identificadas en el diagrama con números pares) presentan un mayor número de secuencias asibiladas.

A pesar de que el número de apariciones presenta una distribución equilibrada, la prueba de ji-cuadrado indica que la aparición de las variantes analizadas es dependiente de los hablantes ( $\chi^2(7) = 19.4$ ,  $p = .006$ ). El análisis de los residuos estandarizados de Pearson indica que, en efecto, todos los participantes de este estudio contribuyeron al resultado. Después de examinar las respuestas de cada hablante, concluimos que esto ocurre porque los hablantes produjeron las variantes T1 y T2 en diferentes ítems léxicos.

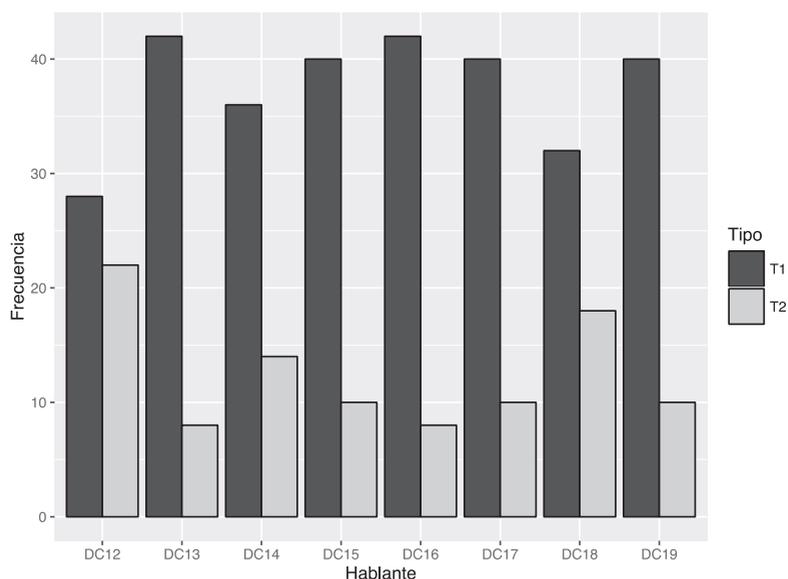


Figura 4. Número de apariciones de las variantes fonéticas de la secuencia /-st-/ por hablante. Las mujeres se codifican con números pares y los hombres con números impares.

Tabla 3. Proporciones y número de apariciones (paréntesis) de las variantes fonéticas de la secuencia /-st-/ por hablante

| Hablante | T1         | T2         |
|----------|------------|------------|
| DC12     | 0.070 (28) | 0.055 (22) |
| DC13     | 0.105 (42) | 0.020 (8)  |
| DC14     | 0.090 (36) | 0.035 (14) |
| DC15     | 0.100 (40) | 0.025 (10) |
| DC16     | 0.105 (42) | 0.020 (8)  |
| DC17     | 0.100 (40) | 0.025 (10) |
| DC18     | 0.080 (32) | 0.045 (18) |
| DC19     | 0.100 (40) | 0.025 (10) |
| Total    | 0.75 (300) | 0.25 (100) |

Para conocer la probabilidad de aparición de T1 y T2 según la tonicidad de la sílaba, el tipo de vocal y el sexo del hablante, realizamos una regresión logística binomial sobre 400 observaciones. Con el fin de evitar la multicolinealidad<sup>5</sup>, usamos la técnica de penalización descrita por Baayen (2008: 205-207) para calcular los coeficientes. La significatividad de los factores se determinó con el procedimiento de incorporación de variables (*stepwise selection*) y luego se verificaron los intervalos de confianza. Encontramos un efecto significativo del tipo de vocal ( $\chi^2(4) = 34.9$ ,  $p < .001$ ), la tonicidad de la sílaba ( $\chi^2(1) = 19.7$ ,  $p < .0001$ ) y el sexo ( $\chi^2(1) = 8.82$ ,  $p = .002$ ). El índice de concordancia del modelo ( $C = 0.74$ ) sugiere que la discriminación es aceptable.

La tabla 4 resume los coeficientes y estadísticos de la regresión logística binomial. Los coeficientes positivos significan que hay una mayor probabilidad de aparición de la variante T2 y los coeficientes negativos indican una mayor probabilidad de la variante T1. Como se ve, las vocales medias ocurren preferentemente en variantes T1 y este efecto es significativo tanto para /e/ (logit: -1.1, DT=0.40,  $p = .005$ ) como para /o/ (logit: -1.3, DT= 0.42,  $p = .001$ ). Las vocales /i/ y /u/ ocurrieron principalmente en la variante T2, pero, como muestra la tabla 4, la probabilidad no fue estadísticamente significativa. Las secuencias /-st-/ seguidas por una vocal tónica se produjeron con la variante T1 (logit: -1.0, DT= 0.25,  $p < .0001$ ). Como ilustra la figura 4 y los coeficientes, los hombres presentan una menor aparición de la variante T2 y este efecto es estadísticamente significativo (logit: -0.7, DT= 0.24,  $p = .004$ ).

<sup>5</sup> La multicolinealidad ocurre cuando tenemos errores típicos y valores de  $p$  elevados (véase también Levshina 2015: 272).

Tabla 4. Coeficientes y estadísticos de la regresión logística binomial. El intercepto representa el *logit* de la probabilidad para vocal para las secuencias /-st-/ seguidas por la vocal /a/ átona de los hablantes de sexo femenino

|                 | Coeficiente | DT   | Wald Z | P       |
|-----------------|-------------|------|--------|---------|
| Intercepto      | 0.001       | 0.29 | 0.01   | .9958   |
| Vocal: /e/      | -1.1        | 0.40 | -2.80  | .0052   |
| Vocal: /i/      | 0.4         | 0.34 | 1.32   | .1873   |
| Vocal: /o/      | -1.3        | 0.42 | -3.19  | .0014   |
| Vocal: /u/      | 0.05        | 0.34 | 0.17   | .8629   |
| Acento: tónica  | -1.0        | 0.25 | -4.23  | < .0001 |
| Sexo: masculino | -0.7        | 0.24 | -2.87  | .0041   |

En resumen, en habla leída y en los datos de los hombres, se reportaron menos casos de la variante [-st̚-], aunque las dos variantes coexisten en la pronunciación de todos los hablantes. En cuanto a la influencia de los factores de nuestro interés, de acuerdo con nuestras expectativas, cuando /t/ pertenece a una sílaba tónica hay mayor probabilidad de que se pronuncie [-st̚-]. Las vocales medias /e o/ también favorecieron la aparición de esta variante, lo cual no es un resultado sorprendente dado que no requieren posiciones linguales o movimientos mandibulares tan precisos como los implicados en la producción de las vocales periféricas /i a u/.

### 3.2.2. Duración en habla leída

En la variante T1, la fricativa /s/ tuvo una duración de  $71.3 \pm 22$  ms y la oclusiva /t/  $80.7 \pm 25$  ms (el cierre =  $53.1 \pm 20.5$  ms y VOT=  $27.61 \pm 17.1$  ms). En cuanto a la duración total, la variante T1 se produjo en  $149 \pm 30$  ms, mientras que la variante T2 tuvo una duración media de  $126 \pm 22$  ms. La prueba no paramétrica de Wilcoxon indica que la diferencia es estadísticamente significativa ( $W= 20989$ ,  $p < 0.001$ ).

Como notará el lector, estos valores son casi idénticos a los resultados obtenidos en el análisis presentado en el apartado 2.2.2. Por esta razón, realizamos un modelo de regresión de efectos mixtos para obtener más información sobre la duración del cierre de la oclusión de la variante T1 y la duración total de las realizaciones de /-st-/<sup>6</sup>. Los participantes en el estudio y las palabras analizadas se consideraron variables aleatorias. Las variables independientes fueron: 1) la vocal que sigue a la secuencia, 2) la tonicidad de la vocal y 3) el sexo del hablante.

<sup>6</sup> No realizamos un modelo para la fricativa /s/ porque los parámetros estudiados se mantuvieron estables en las condiciones estudiadas.

La duración del cierre de la oclusión se afectó principalmente por las vocales medias ( $\chi^2(9) = 20.8$ ,  $p < .001$ ): /o/ produjo un descenso de  $-13.2 \pm 5.3$  ms y la vocal anterior media /e/ produjo un descenso de  $-10 \pm 5.3$  ms. En las vocales /i, a, u/, la tonicidad de la vocal ( $\chi^2(5) = 0.007$ ,  $p = 0.9$ ) y el sexo ( $\chi^2(4) = 0.11$ ,  $p = 0.7$ ) no afectaron la duración del cierre de la oclusión. Finalmente, la duración total de la secuencia /-st-/ presentó un incremento de  $14.8 \pm 2.3$  ms cuando la vocal es tónica ( $\chi^2(5) = 54.2$ ,  $p < .0001$ ).

Aunque en este apartado analizamos un grupo de participantes y un estilo de habla diferente, la duración de las variantes no tuvo modificaciones con respecto al habla narrada. De otro lado, en la sección anterior vimos que [-ʃt-] aparece tras vocales medias /e o/, pero los datos presentados aquí nos muestran que con estas vocales la oclusión de /t/ es breve. La duración total de la secuencia incrementa cuando la vocal tras /-st-/ es tónica, lo cual comprueba una vez más que la prominencia de la sílaba determina el grado de reducción fonética.

### 3.2.3. Momentos espectrales en habla leída

El centro de gravedad de la fricativa /s/ fue de  $4094 \pm 1537$  Hz en T1 y  $3873 \pm 1636$  Hz en T2, sin embargo, esta diferencia fue marginalmente significativa ( $W = 1679$ ,  $p = .07$ ). La asimetría alcanzó un valor de  $-0.2 \pm 1.1$  en la variante T1 y de  $0.09 \pm 1.3$  en la variante T2, diferencia que resultó apenas significativa ( $W = 12682$ ,  $p = .02$ ). Estos datos indican que el espectro de /s/ tiene una distribución normal; los valores no se acercaron completamente a 0 porque, como muestra la tabla 5, el hablante DC15 presentó un centro de gravedad más bajo que los demás hablantes. Finalmente, la curtosis se mantuvo baja en ambos casos (T1=  $-0.2 \pm 2.4$ , T2=  $0.1 \pm 4.0$ ), lo cual quiere decir que, a pesar de las diferencias individuales, los datos correspondientes a /s/ presentan pocos valores atípicos. Como se ve, al igual que en habla narrada, las propiedades espectrales de /s/ se mantuvieron estables en las dos realizaciones estudiadas.

La tabla 5 muestra los datos del centro de gravedad para /s/ y para la realización asibilada de /t/. El valor del centro de gravedad de /t/ es de  $2082 \pm 1745$  Hz, lo cual significa que la fricción es el 50 % más baja que en /s/. Como puede verse, el grado de fricción de /t/ varía de hablante a hablante y es gradiente: mientras que algunos participantes tienen un centro de gravedad bajo, otros tienen una pronunciación con valores cercanos a su pronunciación de /s/ (DC13, DC15, DC18). En consecuencia, la asimetría muestra que los datos se alejan de la distribución normal ( $2.3 \pm 4.5$ ) y los valores elevados de la curtosis ( $29.9 \pm 2082$ ) indican que los datos correspondientes a la asibilación de /t/ tienen una gran cantidad de valores atípicos.

Tabla 5. Momentos espectrales de /s/ (izquierda) y /t/ (derecha) en la variante T2

| Hablante | COG /s/     | Asimetría /s/ | Curtosis /s/ | COG /t/     | Asimetría /t/ | Curtosis /t/ |
|----------|-------------|---------------|--------------|-------------|---------------|--------------|
| DC12     | 4197 (1161) | -0.63 (0.88)  | -0.32 (2.14) | 2225 (1139) | 1.04 (0.99)   | 0.35 (2.67)  |
| DC13     | 4061 (719)  | -0.39 (0.49)  | -1.40 (0.59) | 3149 (1487) | 0.28 (1.21)   | -0.39 (2.25) |
| DC14     | 3671 (1314) | 0.11 (0.99)   | -0.94 (2.0)  | 1069 (979)  | 3.27 (1.67)   | 12.6 (11.11) |
| DC15     | 1091 (667)  | 2.51 (1.49)   | 7.11 (9.08)  | 1454 (884)  | 2.04 (1.93)   | 6.8 (13.6)   |
| DC16     | 3287 (1800) | 0.36 (1.24)   | -0.37 (1.31) | 796 (404)   | 3.67 (1.37)   | 14.6 (10.7)  |
| DC17     | 4043 (830)  | -0.27 (0.51)  | -1.45 (0.31) | 1988 (933)  | 1.56 (1.97)   | 4.4 (15.4)   |
| DC18     | 4611 (2047) | 0.36 (1.13)   | -0.17 (2.06) | 3752 (2659) | 4.0 (9.9)     | 136 (528)    |
| DC19     | 4150 (1210) | -0.68 (1.07)  | 0.07 (2.51)  | 1076 (973)  | 2.9 (1.8)     | 11.6 (13.9)  |
| Media    | 3783 (1636) | 0.09 (1.34)   | 0.19 (4.01)  | 2082 (1745) | 2.38 (4.5)    | 29.9 (224)   |

Realizamos un análisis de regresión de efectos mixtos sobre la variante T2 con el objetivo de determinar si el grado de asibilación está determinado por variables independientes, como el tipo de vocal, el acento o el sexo del hablante. Cada uno de los momentos espectrales (centro de gravedad, asimetría y curtosis) se consideraron como variables dependientes, mientras que el hablante y las palabras se consideraron variables aleatorias. El centro de gravedad, que es el promedio de las frecuencias con mayor amplitud, incrementó  $637 \pm 313$  Hz en las vocales tónicas ( $\chi^2(2) = 4.5$ ,  $p < .05$ ). El tipo de vocal ( $\chi^2(3) = 8.9$ ,  $p = .06$ ) y el sexo del hablante ( $\chi^2(4) = .02$ ,  $p < .08$ ) tuvieron efectos marginalmente significativos. Los modelos correspondientes a la asimetría y a la curtosis no presentaron ningún efecto estadísticamente significativo sobre las variables consideradas.

#### 4. RESUMEN Y DISCUSIÓN

El concepto de reducción fonética resulta útil para describir y explicar procesos de debilitamiento vocálico y consonántico que se alejan de la pronunciación canónica o normativa (*citation form*). Como se discutió al inicio, la reducción fonética se explicó como una adaptación del hablante a la situación comunicativa, la cual mantiene un balance entre la economía articulatoria y la discriminación de la información en un continuo de hiper e hipoarticulación (Lindblom 1990; Kohler 1998; Barry y Andreeva 2001). Los estudios sobre el inglés y el sueco (Engstrand 1988; de Jong 1995) extendieron este principio y muestran que las sílabas tónicas favorecen la hiperarticulación de vocales y consonantes, mientras que las átonas favorecen la reducción fonética o la hipoarticulación.

De acuerdo con Simpson (2001; 2013), esta teoría presenta dos problemas. Primero, no pone límites a la reducción, pues no explica por qué algunos hablantes reducen ciertas unidades léxicas mientras que otros no. Y segundo, el principio de economía no explica los procesos de debilitamiento cuyo resultado son sonidos articulatoriamente más complejos que los originales. Uno de estos ejemplos es la asibilación de las oclusivas, que ha sido reportado en varias lenguas como el inglés (Simpson 2001) y neerlandés (Mitterer y Ernestus 2006). En este trabajo mostramos que en el español hablado en Bogotá (Colombia) este fenómeno ocurre tras /s/ en secuencias intervocálicas /-st-/. Para evitar explicaciones basadas en la economía articuladora, definimos este tipo de reducción fonética como una simplificación de las fases de producción y de sus correlatos acústicos (particularmente en /t/), lo cual nos permitió clasificar el fenómeno en dos variantes [-sʃ-] y [-sʃ̥-].

El primer objetivo de este trabajo fue determinar cuál es la frecuencia de las realizaciones [-sʃ-] y [-sʃ̥-] en dos tipos de habla conectada (habla narrada *vs.* habla leída). Segundo, buscamos aportar datos sobre las características acústicas de /s/ y /t/ en cada una de estas variantes y, por último, determinar si la tonicidad, el tipo de vocal tras la secuencia y las preferencias individuales determinan la aparición de las variantes y sus propiedades acústicas. A continuación, resumimos los principales resultados de este trabajo.

#### 4.1. Frecuencia de aparición de las variantes

En habla narrada, [-sʃ-] y [-sʃ̥-] tienen un número igual de proporciones (49 % y 51 % respectivamente), pero el número de apariciones varía de hablante a hablante, de manera que unos prefieren una variante conservadora [-sʃ-], otros favorecen la variante relajada [-sʃ̥-] y otros producen las dos variantes. En habla leída analizamos otro grupo de hablantes e incluimos participantes de ambos sexos. La pronunciación [-sʃ̥-] tuvo 26 % menos de apariciones en la lectura que en las narraciones y fue más común en el habla femenina. Un análisis de regresión lineal binomial mostró que [-sʃ̥-] es más común tras vocales tónicas y, particularmente, cuando se trata de vocales medias /e o/. La relación de [-sʃ̥-] con la tonicidad se explica, siguiendo a de Jong (1995), porque el grado de prominencia es fundamental para la identificación de la información léxica. Ahora, la aparición de [-sʃ̥-] tras vocales medias se explica porque estas vocales no requieren el movimiento lingual o mandibular de las vocales periféricas, de manera que el cierre de /t/ se puede producir sin anticipar movimientos articulatorios opuestos, como los que requieren las vocales /a u/.

#### 4.2. Características acústicas

La duración de [-sʃ-] se ubica alrededor de 146 ms y [-sʃ̥-] tiene un valor medio de 124 ms. No se reportaron variaciones asociadas al estilo de habla, sin embargo, estos valores corresponden a los valores reportados por Fernández Planas (1999) para la secuencia *asta* pronunciada de manera aislada con velocidad de habla rápida y muy rápida. Nosotros interpretamos estos resultados como una consecuencia de la tarea de recolección de datos

(habla conectada *vs.* palabras leídas), aunque nuestra conclusión es limitada porque en este estudio no consideramos los efectos de la velocidad de elocución.

La probabilidad de aparición de [-ʃt̪-] es más alta ante las vocales medias, pero el modelo de efectos mixtos muestra que en este entorno vocálico el cierre de la oclusiva dentoalveolar es entre 10 y 13 ms más breve. Esto significa que las apariciones de [-ʃt̪-] tienen en general una constricción débil. Ahora, cuando la vocal ubicada después de la secuencia es una tónica, el cierre u oclusión incrementa 14.8 ms, contrarrestando así el efecto de las vocales medias.

La fricativa /s/ tuvo un centro de gravedad alrededor de 4000 Hz, lo cual concuerda con los datos reportados en la bibliografía (Gordon *et al.* 2002). Estos valores presentaron algunas diferencias individuales, pero en general /s/ se mantuvo estable en las dos variantes analizadas, pues el espectro se acercó a la distribución normal (asimetría cercana a cero) y los datos presentaron pocos valores atípicos (curtosis baja). De manera que el análisis de los momentos espectrales nos permite concluir que la fricativa es resistente a la reducción fonética y al estilo de habla en la secuencia intervocálica /-st-/.

El análisis acústico de [t̪] muestra que hay una gran intervariabilidad, de modo que algunos participantes producen una oclusiva asibilada con un centro de gravedad que se acerca a su realización de /s/, mientras que otros producen una fricción difusa y con una concentración de la energía en la parte baja del espectro. Sin embargo, tomados los datos en conjunto, observamos que el espectro de [t̪] tiene un centro de gravedad 50 % más bajo que /s/. Los datos acústicos de [t̪] se alejan de la distribución normal y tienen un gran número de valores atípicos, lo cual es un reflejo de la intervariabilidad e intravariabilidad en el grado de constricción de /t/.

## 5. CONCLUSIÓN

Los datos del presente estudio muestran, en primer lugar, que el estilo de habla no es el único factor determinante de la reducción fonética de la secuencia /-st-/ como sugiere la teoría de la hiper e hipoarticulación, aunque el análisis del habla leída y el habla narrada resultó ser clave para entender la complejidad del fenómeno. Por ejemplo, en las narraciones, la variante [-ʃt̪-] fue más frecuente y las diferencias individuales más acusadas, pero el número de casos no resultó suficiente para establecer la relevancia de factores como el tipo de vocal y la tonicidad. A su vez, el habla leída determinó la influencia de estos factores, pero mostró que [-ʃt̪-] y [-ʃt̪-] ocurren en todos los idiolectos estudiados. En segundo lugar, las tónicas favorecieron la aparición de la variante [-ʃt̪-] y la producción de un cierre dentoalveolar más largo, lo cual comprueba que las sílabas tónicas son resistentes a la reducción (Engstrand 1998; de Jong 1995) y las sílabas átonas tienen tendencia a la reducción vocálica y consonántica (Barry y Andreeva 2011: 54). Tercero, las vocales medias ejercieron la misma influencia sobre la secuencia /-st-/ que la tonicidad, pues favorecieron la aparición de una variante no asibilada e incrementaron la duración del cierre de /t/.

Este resultado está soportado en los estudios sobre coarticulación (Recasens 1999: 81-84) aunque no se suele tener en cuenta en el análisis de la reducción consonántica (véanse los trabajos de van Son y Pols 1996; Shockey y Farnetani 1992).

La cuarta y última conclusión es que el comportamiento individual —incluyendo el sexo del hablante— juega un papel importante en la reducción fonética (para una revisión completa de este factor, véase Stevens y Harrington 2014). En los datos presentados aquí, las variaciones idiosincráticas se reflejan en la frecuencia de aparición de cada variante que, en últimas, es una consecuencia del uso de diferentes estrategias para realizar la constricción de [t̪]. Así, debemos reconocer que las características espectrales de [t̪] son gradientes, como sostienen Ernestus y Warner (2011: 255), pero también son un aspecto impredecible que no se puede explicar con principios como la hiper e hipoarticulación, como argumentaba Ladefoged (1983).

#### OBRAS CITADAS

- Alves, Helena *et al.* 2011. *BuFón. Buscador de patrones fonológicos*. Madrid: Consejo Superior de Investigaciones Científicas (CSIC). [<http://www.estudiosfonicos.cchs.csic.es/fonetica/bufon?p=presentacion>].
- Baayen, Harald. 2008. *Analyzing Linguistic Data. A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Barry, William & Bistra Andreeva. 2001. “Cross-language similarities and differences in spontaneous speech patterns”. *Journal of the International Phonetic Association* 31 (1): 51-66.
- Boersma, Paul & David Weenink. 2016. *Praat: doing phonetics by computer* (versión 6.0.19) [<http://www.praat.org/>].
- Byrd, Dani. 1994. “Relations of sex and dialect to reduction”. *Speech Communication* 15 (1-2): 39-54.
- Correa, José Alejandro. 2017. “Caracterización acústica de la reducción vocálica en el español de Bogotá (Colombia)”. *Estudios de Fonética Experimental* 26: 63-91.
- Crystal, David. 2008. *A Dictionary of Linguistics and Phonetics*. Oxford: Blackwell.
- De Jong, Kenneth. 1995. “The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation”. *Journal Acoustical Society of America* 97 (1): 491-504.
- Engstrand, Olle. 1988. “Articulatory correlates of stress and speaking rate in Swedish VCV utterances”. *Journal Acoustical Society of America* 83 (5): 1863-1975.
- Ernestus, Mirjam & Natasha Warner. 2011. “An introduction to reduced pronunciation variants”. *Journal of Phonetics* 39: 253-260.
- Fernández Planas, Ana María. 1999. “Aproximación al estudio de la influencia de la velocidad de habla en grupos consonánticos intervocálicos –st–”. En: J. Fernández, C. Fernández, M. Sánchez, E. Prieto & S. Río. Eds. *Lingüística para el siglo XX*. Salamanca:

- Ediciones Universidad de Salamanca: 641-651.
- Forrest, Karen *et al.* 1998. "Statistical analysis of word-initial voiceless obstruents: preliminary data". *Journal Acoustical Society of America* 84 (1): 115-123.
- Gordon, Matthew *et al.* 2002. "A cross-linguistic acoustic study of voiceless fricatives". *Journal of the International Phonetic Association* 32 (2): 141- 174.
- Harmegnies Bernard & Dolores Poch Olivé. 1992. "A study of style-induced vowel variability: Laboratory versus spontaneous speech in Spanish". *Speech Communication* 11 (4-5): 29-437.
- Harrell, Frank. 2017. "Regression Model Strategies. Package rms". [R package], version 5.1-1. [<https://cran.r-project.org/web/packages/rms/>]
- Harrington, Jonathan. 2010. *Phonetic Analysis of Speech Corpora*. Oxford: Willey Blackwell.
- Johnson, Keith. 2004. "Massive reduction in conversational American English". En: K. Yoneyama & K. Maekawa. Eds. *Spontaneous speech: data and analysis*. Tokio: The International Institute for Japanese Language: 29-54.
- Jones, Daniel. 1973. "Some thoughts on the phoneme". En: W. E. Jones & J. Laver. Eds. *Phonetics in Linguistics. A book of readings*. London: Longman Group: 168-186.
- Jongman, Allard. *et al.* 2000. "Acoustics characteristics of English fricatives". *Journal Acoustical Society of America* 108 (3): 1252-1263.
- Kohler, Klaus. 1998. "The disappearance of words in connected speech". *ZAS Working Papers in Linguistics* 11: 21-34.
- Ladefoged, Peter. 1983. "The limits of biological explanations in phonetics". *UCLA Working Papers in Phonetics* 57: 1-10.
- Levshina, Natalia. 2015. *How to do Linguistics with R. Data exploration and statistical analysis*. Amsterdam: John Benjamins.
- Lindblom, Bjorn. 1990. "Explaining phonetic variation: a sketch of the H&H theory". En: W. Hardcastle & A. Marchal. Eds. *Speech production and speech modeling*. Dordrecht: Kluwer: 403-439.
- Mitterer, Holger & Mirjam Ernestus. 2006. "Listeners recover /t/s that speakers reduce: Evidence from /t/-lention in Dutch". *Journal of Phonetics* 34: 73-103.
- Newell, Karl & Peter Hancock. 1984. "Forgotten moments: a note on skewness and kurtosis as influential factors in inferences extrapolated from response distributions". *Journal of Motor Behavior* 16 (3): 320-355.
- Pinheiro, Jose *et al.* 2017. "nlme: Linear and Nonlinear Mixed Effects Models". [R package], version 3.1-131. [<https://cran.r-project.org/web/packages/nlme/index.html>]
- Pulgram, Ernst. 1965. "Consonant Cluster, Consonant Sequence, and the Syllable". *Phonetica* 13: 76-81.
- R Core Team. 2013. *R: A language and environment for statistical computing*. Vienna: Austria [<https://www.r-project.org>]
- Recasens, Daniel. 1999. "Lingual coarticulation". En: W. Hardcastle & N. Hewlett. Eds. *Coarticulation: Theory, data and techniques*. Cambridge, Cambridge University Press: 80-104.

- Shockey, Linda & Edda Farnetani. 1992. "Spontaneous speech in English and Italian". *The Second International Conference on Spoken Language Processing*. 767-770.
- Simpson, Adrian. 2001. "Does articulatory reduction miss more patterns than it accounts for". *Journal of the International Phonetic Association* 31 (1): 29-39.
- \_\_\_\_\_. (2009). Phonetic differences between male and female speech. *Language and Linguistics Compass* 3 (2): 621-640.
- \_\_\_\_\_. 2013. "Spontaneous speech". En: M. Jones & R. Knight. Eds. *The Bloomsbury companion to phonetics*. London: Bloomsbury: 155-170.
- Stevens, Mary & Harrington Jonathan. 2014. "The individual and the actuation of sound change". *Loquens* 1 (1): e003.
- Turk, Alice *et al.* 2006. "Acoustic segment durations in prosodic research". En: S. Sudhoff, D. Lenertova, R. Meyer, S. Pappert, P. Augurzky, I. Mleinek, N. Richter, J. Schließer. Eds. *Methods in Empirical Prosody Research*. New York/Berlin: de Gruyter. 1-27.
- Van Son, Rob & Louis Pols. 1996. "An acoustic profile of consonant reduction". *Spoken Language, 1996. ICSLP 96. Proceedings, Fourth International Conference*.
- Wallace, Chafe. 1980. *The Pear Stories: Cognitive, Cultural and Linguistic Aspects of Narrative Production*. Norwood, NJ: Ablex.
- Westfall, Peter. 2014. "Kurtosis as Peakedness, 1905-2014. R.I.P." *The American Statistician* 68 (3): 191-195.
- Whiteside, Sandra. 1996. "Temporal-based acoustic-phonetic patterns in read speech: some evidence for speaker sex differences". *Journal of the International Phonetic Association* 26 (1): 23-40.