



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
LICENCIATURA EN LENGUA Y LITERATURAS HISPÁNICAS

**SISTEMATIZACIÓN DEL FENÓMENO DE
SILABICACIÓN EN EL *CORPUS DIME* PARA SU
APLICACIÓN EN LAS TECNOLOGÍAS DEL HABLA**

TESIS QUE PRESENTA:

PRECIOSA ALEJANDRA ESPINOZA CRUZ

PARA OBTENER EL TÍTULO DE LICENCIADA EN LENGUA Y
LITERATURAS HISPÁNICAS

Asesores: Dr. Luis Alberto Pineda Cortés
Mtro. Javier Octavio Cuétara Priede

Ciudad de México, 2007

Agradecimientos:

Hacer una tesis es una tarea difícil. Tu tiempo y espacio se ven limitados, por una temporada dejas de convivir con las personas que estimas y la presión se vuelve parte de tu vida diaria; sin embargo, ellos son los que, cuando te falta tan sólo un paso para cruzar la meta, te recuerdan quién eres y el por qué te quieren. Me refiero a mi familia, mis amigos y, por su puesto a mi Universidad.

Mi padre, José Sánchez García es la persona que me hereda en vida un magnífico ejemplo a seguir: constancia fuerza y honestidad. Mis madres: María Antonieta Espinosa Cruz y Guillermina Cruz Vargas me han demostrado que lo bello de esta vida lo descubre quien mantiene los ojos bien abiertos, a ambas les agradezco que me hayan enseñado a vivir con la mirada alerta. Agradezco infinitamente a Dios por haberme permitido renacer en el seno de una maravillosa familia.

A mis amigos les agradezco no el cariño que compartimos, sino la paciencia y amor con los que me tratan. Me han consentido tanto que ya me he malacostumbrado. Ana, Judith, Dianas, Toños y ahora César, los querré siempre por compartir conmigo los momentos felices, divertidos y tristes. Anhé, sé que la amistad no se agradece, se complementa, se comparte, se valora; sin embargo, a ti quiero agradecer que me hayas transmitido la fuerza que ahora me invade.

A mis asesores de tesis: Luis Alberto Pineda Cortés y Javier Octavio Cuétara Priede, no sólo por dirigirme la tesis, sino por brindarme su cariño, su valioso tiempo y la confianza que depositaron en mí. Javier, una vez más corroboro que el pedirte que fueras mi asesor no pudo ser mejor, te quiero mucho, eres mi amigo y mi Maestro.

Dr. Pineda te agradezco que me hayas aceptado en el Proyecto, confiar en mí y demostrarme tu aprecio.

Agradezco al Proyecto NSF/CONACYT 39380-A DIME II: *Diálogos Inteligentes Multimodales en Español* y al Proyecto PAPIIT IN 121206-3: *Diálogos Inteligentes Multimodales en Español*. Ambos bajo la coordinación del Dr. Luis Alberto Pineda Cortés, por proporcionarme los recursos económicos, lingüísticos y humanos para realizar este trabajo de tesis. En general agradezco a todo el equipo DIME (Paty).

Como mi carrera, mi tesis y el haber estado enamorada durante todo este proceso: XPG

Soy una universitaria orgullosa de llevar el escudo puma en el corazón.

ÍNDICE

1. Introducción	6
1.1. <i>El problema</i>	7
1.2. <i>Objetivos</i>	8
1.3. <i>Metodología</i>	11
1.4. <i>Corpus de análisis</i>	12
1.5. <i>División de tesis por capítulos</i>	13
2. Fonética y tecnologías del habla	15
2.1. <i>La fonética y la computación</i>	16
2.2. <i>Tecnologías del habla</i>	18
2.2.1. <i>Aplicaciones</i>	24
2.3. <i>Proyecto DIME</i>	24
2.3.1. <i>Herramientas de transcripción del</i> <i>Corpus DIME: Mexbet y Speech View</i>	26
3. La sílaba española y su estructura	30
3.1. <i>El concepto de sílaba y su estructura general</i>	30
3.1.1. <i>Estructura silábica</i>	32
3.2. <i>Diversas definiciones de sílaba</i>	34
3.3. <i>Estructura y características de la sílaba española</i>	38
4. <i>Comportamiento silábico del Corpus dime</i>	43
4.1 <i>Comportamiento general del corpus</i>	43
4.2. <i>Formación de sílaba abierta</i>	46
4.2.1. <i>Sílaba abierta con fonema /r /</i>	47
4.2.2. <i>Sílaba abierta con fonema /s/</i>	52
4.2.3. <i>Sílaba abierta con el fonema /n /</i>	55
4.3. <i>Reducción de segmentos idénticos</i>	57
4.3.1. <i>Reducción de segmentos consonánticos</i>	61
4.4. <i>Sinalefa</i>	63
4.5. <i>Propuestas de reglas de silabificación fonéticas</i> <i>para las tecnologías del habla</i>	70
4.6. <i>Aplicación de las reglas en las tecnologías del habla</i>	71
5. Conclusiones	77
6. Referencias bibliográficas	80
7. Apéndice 1	83
8. Apéndice 2	94
9. Apéndice 3	103

ÍNDICE DE FIGURAS

FIGURA 1	Ejemplo de silabificación de la oración <i>tienes algunas sugerencia....</i>	10
FIGURA 2	Niveles de segmentación.....	12
FIGURA 3	Modelo de conversor de texto.....	21
FIGURA 4	Modelo de un reconocedor de voz.....	23
FIGURA 5	Segmento de diálogo DIME.....	25
FIGURA 6	Interfase del Corpus DIME.....	26
FIGURA 7	Alfabeto fonético <i>Mexbet</i>	27
FIGURA 8	Imagen del <i>Speech View</i>	28
FIGURA 9	Seis grados de abertura propuestos por Saussure.....	35
FIGURA 10	Sílaba <i>a.ga</i> y sílaba <i>ag.a</i> (Quilis: 1979:366).....	37
FIGURA 11	Frecuencia silábica en el poema <i>MioCid</i>	42
FIGURA 12	Cuadro de frecuencias silábicas del español.....	43
FIGURA 13	Imagen espectrográfica en donde se muestra la silabificación fonética de /r e/.....	49
Figura 14	Imagen espectrográfica de sílaba abierta “r a”.....	50
FIGURA 15	Imagen espectrográfica de sílaba abierta “r o”: girar otra vez-gi.ra.ro.tr.a.ves. Diálogo 11; oración 40.....	51
FIGURA 16	Imagen espectrográfica de sílaba abierta “r u”: rotar un-ro.ta.run. Diálogo:2; oración 10.....	52
FIGURA 17	Imagen espectrográfica de sílaba abierta “s a”: tienes algun-ti.ne.sal.gun. Diálogo 11; oración 15.....	54
FIGURA 18	Imagen espectrográfica de sílaba abierta “s e”: las estufas-la ses tu fas. Diálogo 20; oración 4.....	55
FIGURA 19	Imagen espectrográfica de homologación vocálica “e e”: que estes-kes.tes.tan.te. Diálogo 11; oración 92.....	58
FIGURA 20	Imagen espectrográfica de homologación vocálica “a a”: esta alacena-es.ta.la.se.na. Diálogo 7; oración 37.....	61
FIGURA 21	Imagen espectrográfica de homologación consonántica “s s”: estossen-	63

es.to.son. Diálogo 11; oración 90.....	
FIGURA 22 Imagen espectrográfica de sinalefa “a e”: la estufa-laes.tu.fa. Diálogo 11; oración 10.....	66
FIGURA 23 Imagen espectrográfica de sinalefa “e a”: mi idea-mi.dea. Diálogo 2; oración 12.....	68
FIGURA 24 Imagen espectrográfica de sinalefa “o a”.....	69
FIGURA 25 Tabla de ocurrencia de los fenómenos de silabificación del <i>Corpus DIME</i>	70
FIGURA 26 Cuadro de reglas de silabificación fonética para las tecnologías del habla.....	71
FIGURA 27 Cuadro comparativo de porcentaje de sílaba abierta.....	75
Figura 28 Figura 27. Tabla que muestran las tres nuevas entradas de las palabras <i>tienes, algunas</i> y <i>alacenas-tiene algunasalacenas</i>	76

1. Introducción

El presente es un trabajo lingüístico cuyos resultados podrán ser aplicados en las tecnologías del habla, por lo que es indispensable enfatizar la labor interdisciplinaria que se realizó, así como también el trabajo empírico-fonético. Este último fue el que marcó la línea principal de investigación con base en las diferencias entre la lengua escrita y la lengua oral, específicamente las existentes en una unidad de análisis, que ha sido eje central de esta investigación: la sílaba. “El estudio de la sílaba tiene un largo y variado historial que incluye el cuestionarse la existencia de la misma a nivel fonético y al aceptársela ya por razones psicológicas, ya por pura constatación física. Hoy por hoy, se continúa en el sondeo de esta entidad, se penetra en su constitución a nivel fonológico y se la empalma en su realización concreta mediante una serie de principios universales y reglas particulares de lengua”. (Nuñez-Cedeño 2000:455).

El análisis de la sílaba que haremos es fonético, ya que cuando se habla de silabificación se hace referencia primaria a la lengua oral; es decir, es un estudio basado en la forma real que el hablante adopta al emitir o formar grupos silábicos en la cadena hablada; así, se tomará en cuenta el habla como un continuo que no aísla segmentos, sino que entrelaza unidades. “es un hecho comprobado en todas las lenguas del mundo que los sonidos tienden a agruparse, dentro de la cadena, en unidades mayores, dotadas de un entidad propia y más fácilmente aislables que los propios segmentos: son las sílabas” (Gil Hernández, 1999:117).

La interacción e integración que se puede llegar a presentar entre las distintas disciplinas humanas ha ido en aumento, así pues, la que concierne a la lingüística y la computación es sustancial. El análisis lingüístico en las tecnologías del habla es primario, ya que, como define LListerri (2003) se trata justamente de conocer de la forma más detallada posible cuál es nuestra manera de “hablar” e implementarla en las computadoras para hacerlas, así, más adaptadas a nuestra forma de actuar. Las tecnologías del habla: síntesis, interpretación y reconocimiento son una serie de disciplinas cuyo objetivo principal es hacer más natural la interacción entre el humano y la computadora por medio del lenguaje oral, para así poder llevar a cabo distintas tareas.

1.1. EL PROBLEMA

Actualmente el mayor de los retos al que se enfrentan las tecnologías del habla es lograr que éstas sean capaces de reproducir el habla lo más humanamente posible y por consiguiente, tratar de reconocerla tal y como un ser humano la emite en situaciones convencionales. La observación, el estudio y análisis de los distintos niveles de una lengua (gramatical, sintáctico, semántico, fonético, etc.) permiten, cuando éstos son sistematizados y formalizados en reglas, un mejoramiento en dichas tecnologías. Cuanto más detallado y específico sea el estudio de cada nivel de la lengua permitirá una mejor modelación de los fenómenos que dentro de dichos niveles ocurren. El estudio fonético de la silabificación proporciona datos reales de la lengua en uso para que éstos a su vez puedan ser implementados tanto en un

sintetizador como en un reconocedor de habla. La importancia de manejar datos lingüísticos “orales” proporciona a dicho objetivo un mejor y más eficiente estudio del habla que aquel que proporciona el lenguaje escrito; es decir, se hacen evidentes las ventajas que brinda el manejo de modelos de representación de lenguaje oral: corpus orales. Al respecto, Llisterri (1995:49) menciona: “en el campo de la fonética, los corpus constituidos por grabaciones de laboratorio son herramientas indispensables para el estudio experimental del habla [...] En el ámbito de las tecnologías del habla, las bases de datos orales proporcionan datos importantes para el modelamiento de los fenómenos segmentales y suprasegmentales en la conversión de texto a habla y son esenciales para el entrenamiento y la validación de los sistemas de reconocimientos y de diálogo en entornos de comunicación persona máquina, cuyas aplicaciones se extienden desde la oferta de servicios telefónicos automatizados hasta las ayudas para personas con discapacidades”.

1.2. OBJETIVOS

Con base en las diferencias observables entre la división silábica tradicional y la división fonética y, tomando en cuenta las evidentes ventajas que para la modelación del habla proporciona el estudio acústico fonético, el presente trabajo tiene como objetivos principales:

- 1) El análisis empírico de fenómenos propios de la silabificación fonética a nivel de frase¹: formación de sílaba abierta, homologación de segmentos idénticos vocálicos y consonánticos, y sinalefa.
- 2) La comprobación de datos en cuanto a qué tanto difiere la silabificación tradicional de la que realmente ocurre en el habla espontánea.
- 3) Con base en las estadísticas que dicho análisis arrojó, la creación de reglas que sistematicen el fenómeno de silabificación para su aplicación a las tecnologías del habla.

La Figura número 1 muestra la imagen espectrográfica de un fragmento de diálogo del corpus que se analizó, en donde la oración es una interrogativa: *¿Tienes alguna sugerencia?*, la silabificación tradicional nos indica silabear de la siguiente manera: /tié.nes.al.gú.na.su.ge.rén.sia/ y lo que fonéticamente sucede es: [tjé.ne.sal.gú.na.su.ge.rén.sja].

¹ Es decir, no tomar la palabra aislada como parámetro de análisis, sino la cadena hablada, misma que lógicamente no respeta límite entre éstas.

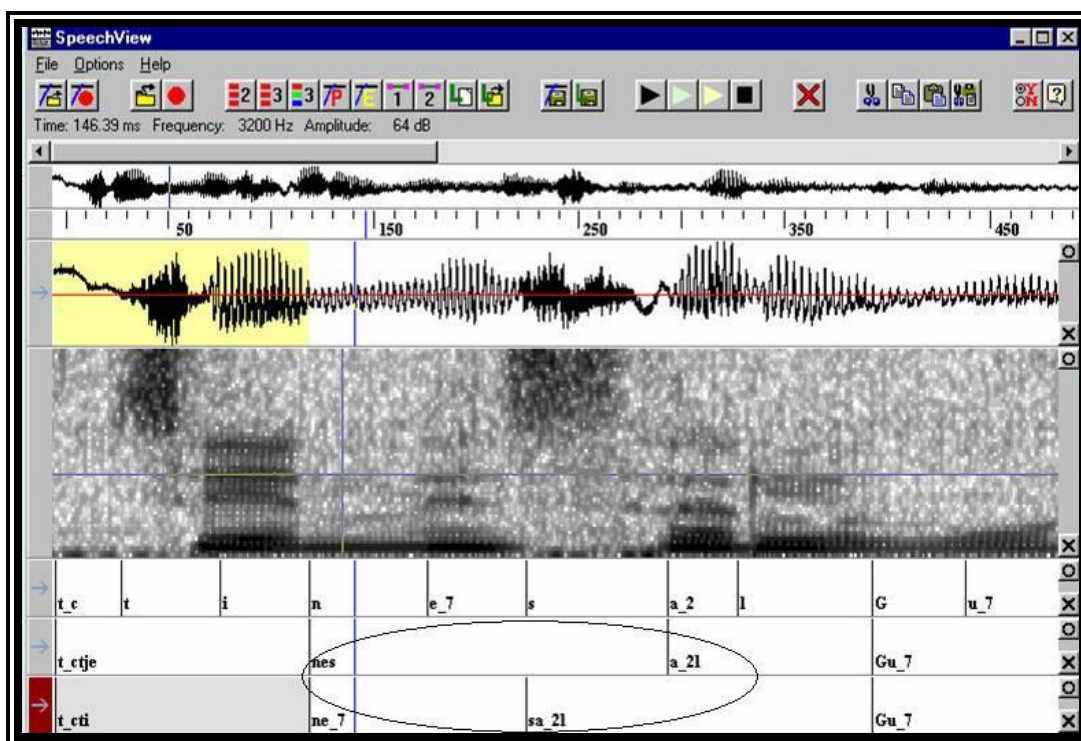


Figura 1. Ejemplo de silabificación de la oración *tienes alguna sugerencia*. Diálogo 7.

El primer nivel de segmentación es alofónico, cuya simbología se extrajo del alfabeto fonético *Mexbet* (Cuétara, 2004); el segundo nivel es silábico, donde se silabifica de manera tradicional o fonológica, y el tercer nivel es el de la silabificación fonética; es decir, donde se presentó un resilabeo entre los límites de las palabras *tienes* y *algún*. La marca ovalada indica la diferencia entre la silabificación tradicional y la silabificación fonética.

1.3. METODOLOGÍA

Los recursos lingüísticos que se utilizaron para poder llevar a cabo el análisis empírico de las muestras del habla fueron extraídos del *Corpus DIME* (Diálogos Inteligentes Multimodales en Español).

Para este efecto se siguieron las siguientes etapas:

1. Asignación de cinco diálogos de este corpus: d02, d07, d11, d18 y d20
2. Transcripción fonética y alineación temporal de los cinco diálogos con el alfabeto fonético: *Mexbet* en los siguientes niveles:
 - a) **Nivel alofónico t54.** Se transcriben en este apartado los alófonos que el hablante realizó en sus emisiones; es decir, que no se transcriben fonemas, sino las realizaciones particulares que de ellos hizo cada hablante.
 - b) **Nivel ortográfico-palabra Tp.** Se transcriben las palabras ortográficamente; es decir, se escribe lo que el hablante “quiso decir” aunque no la haya realizado “correctamente” debido a la velocidad de la elocución, al ruido excesivo o silencios largos.
 - c) **Nivel silábico-fonético Ts.** En este nivel se transcribe el grupo alofónico que el hablante produjo en una sola emisión de voz sin respetar el límite entre palabras, lo que permite que este nivel pueda o no coincidir con la sílaba canónica.

La Figura número 2 muestra los tres niveles de segmentación. El primero, en orden descendente, corresponde al alofónico, el segundo es el nivel silábico fonético y el tercero es el nivel de palabra ortográfica.

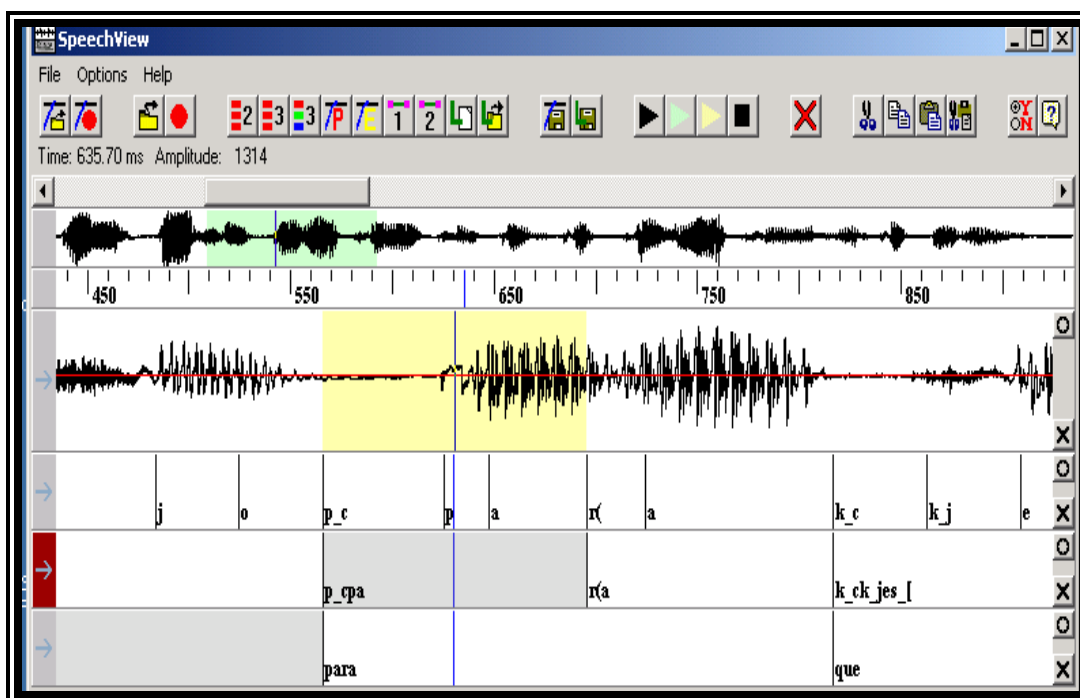


Figura 2. Niveles de segmentación para los diálogos del *Corpus DIME*.

1.4. CORPUS DE ANÁLISIS

Durante la segmentación y alineación temporal de los tres niveles en los cinco diálogos, gráfica y auditivamente, se hicieron claras las diferencias entre la manera tradicional de silabear y la que realmente sucedió, de tal suerte el corpus de trabajo fue depurándose paulatinamente. Así, de las 977 oraciones totales se descartaron 577 por ser monosílabos de respuesta (*sí-no*), archivos de audio con excesiva contaminación por ruido que impedía la clara audición de los mismos, y también se descartaron aquellas grabaciones que presentaron fenómenos de pérdida de segmentos ya que dicho estudio forma parte de otro trabajo de tesis en curso (Ceballos, Ana) 2007. *Fenómenos de pérdida en el Corpus DIME para su inclusión en un reconocedor de habla*).

Por lo tanto, el corpus base de esta investigación consta de 400 oraciones emitidas por cinco hablantes distintos, tres hombres y dos mujeres, todos universitarios que viven en la Ciudad de México, transcritas y alineadas temporalmente en los tres niveles antes mencionados.

1.5. DIVISIÓN DE TESIS POR CAPÍTULOS

Este trabajo de tesis además de la presente introducción contiene cuatro partes: 2) Fonética y Tecnologías del habla, 3) La sílaba española y su estructura, 4) Comportamiento silábico del *Corpus DIME* y 5) Conclusiones.

El objetivo principal del primer capítulo es enmarcar este trabajo lingüístico dentro de un ámbito bidisciplinario, el cual está conformado propiamente por la fonética y las tecnologías del habla; es decir, describe ambas y define el puente que las conecta, éste es la fonética experimental y la acústica. Una vez establecida dicha conexión se hace referencia al tratamiento y las aplicaciones que se les ha dado a las tecnologías del habla, puesto que es el campo de impacto al que se aplicarán los resultados de este trabajo de tesis. Posteriormente, se presenta tanto el Proyecto DIME de cuyo corpus se tomó la información lingüística para realizar el análisis de silabificación.

En el segundo capítulo se aborda la teoría general acerca de la sílaba (conceptos, característica y estructuras). Con base en lo anterior, se hizo hincapié en el hecho de que mucha de la teoría al respecto se centra en lenguaje escrito, por lo que es

pertinente y útil la definición de reglas fonéticas de silabificación extraídas del análisis empírico de una muestra significativa de habla espontánea. El principal objetivo de este capítulo y el de sus apartados es abordar los distintos conceptos y enfoques que se tienen acerca de la sílaba, evidenciando que ésta es una unidad que a lo largo del tiempo ha sido difícil definir, no sólo para el español.

El tercer capítulo es la base medular del trabajo, ya que propondrá la silabificación fonética y la sistematización de reglas con base en el análisis empírico que se realizó. Se describe el comportamiento silábico del Corpus respecto de la tradición, se analizan detalladamente los fenómenos encontrados: sílaba abierta, homologación de segmentos idénticos vocálicos y consonánticos y sinalefa, se muestran ejemplos de los fenómenos. Finalmente se presenta la propuesta de reglas de silabificación fonética.

La parte final de esta tesis corresponde a las conclusiones. Se hace un resumen global de la tesis; es decir, se abordan los problemas principales que el análisis permitió visualizar para, posteriormente, reiterar la importancia de la propuesta de las reglas fonéticas analizadas de una muestra de habla espontánea y no del lenguaje escrito. Se describe la utilidad para las tecnologías del habla: síntesis y reconocimiento y se describe brevemente los respectivos ejercicios que se realizaron en un sintetizador. Finalmente, se habla de las líneas futuras de investigación que pueden tomar como base este trabajo de tesis.

2. FONÉTICA Y TECNOLOGÍAS DEL HABLA

El presente capítulo tiene como objetivo principal dar una visión acerca de la importancia de la información lingüística para las tecnologías del habla, así como también la descripción de los antecedentes de los aparatos de registro y medición que la fonética acústica e instrumental ha utilizado. También se describen las tecnologías del habla y el Proyecto DIME del cual se extrajeron los recursos lingüísticos para el análisis de esta tesis (*Corpus DIME*). Finalmente, se define el *software* que se utilizó para la transcripción y el alfabeto empleado para dicha tarea.

2.1. LA FONÉTICA Y LA COMPUTACIÓN

Hoy en día, el conocimiento humano ya no es exclusivo de una sola disciplina. Ahora éste es compartido e incluso enriquecido por otras materias. El trabajo interdisciplinario que se ha venido desarrollando, no sólo se enfoca hacia el mejoramiento de las técnicas para el beneficio físico y de salud del ser humano, sino que, también, ha llegado a los ámbitos más cotidianos como el de la comunicación por medio del lenguaje. La imagen visionaria del hombre acerca de la comunicación verbal con las computadoras ha dejado de ser, paulatinamente, ficticia. El avanzado grado de especialización de la informática y, por su puesto, de la lingüística, han permitido conjuntamente una interesante y fructífera combinación.

La fonética, en específico la experimental y la instrumental, las cuales toman como base el uso de aparatos de medición y registro de los sonidos, son las que han tenido una exitosa síntesis con la informática, la ingeniería y la computación:

“la fonética instrumental tiene algunos puntos en común con la fonética y la fonología experimental. En primer lugar es producto de la influencia de las ciencias naturales sobre el desarrollo de la lingüística [...] En segundo lugar, la fonética y la fonología experimental también se sirven de instrumentos de registro y medición para la observación objetiva de fenómenos lingüísticos [...] En tercer lugar, los estudios fisiológicos y acústicos de los fonetistas del siglo XIX y su desarrollo de métodos y técnicas instrumentales de análisis han sido una base para la metodología experimental” (Solé, 1985).

Muchas de las herramientas y aparatos de medición que la fonética experimental y acústica han utilizado para el análisis se remontan a finales del siglo XVIII. Tal es el caso de los antecedentes del espectrógrafo, que como define Quilis (1981) tiene la misión de la descomposición automática de la onda sonora compleja en cada uno de sus componentes integrantes y suministrarlos, de este modo, todos los datos que nos interesa conocer.

Métodos como lo son el palatográfico y el plastográfico buscaron en su momento encontrar una manera de registrar los movimientos articulatorios que se llevaban a

cabo en las cavidades supraglóticas durante la producción del habla. Gil Fernández (1999:49) nos da una noción del funcionamiento de un palatógrafo:

Este sistema consiste en introducir en la boca un paladar artificial recubierto con una sustancia especial, de tal modo, que al entrar la lengua en contacto con determinadas zonas del mismo durante la pronunciación de los diferentes sonidos, estas partes de su superficie quedan limpias y claramente delimitadas.

Posteriormente, la plastografía, al mandó de Meyer sustituyó la sustancia que recubría el paladar por pequeños hilos de estaño que se doblaban adaptando la forma que la posición de la lengua los hacía tomar con base en la pronunciación emitida. Sin embargo, no fue hasta el año de 1930 que en los laboratorios americanos *Bell Telephone Co.* fueron implementado el espectrógrafo o fonógrafo, cuyo principal objetivo fue encontrar un método para volver visible el habla, de forma que los sordos pudieran leerla aumentando así sus posibilidades de comunicación (Gil Fernández 1999). Con el paso de los años dicho aparato fue adquiriendo nuevas tareas, tales como identificación e interpretación de mensajes y de emisores.

2.2. TECNOLOGÍAS DEL HABLA

Debido a la creciente necesidad de acceder a sistemas informáticos de manera más rápida, menos complicada y, por su puesto, más natural se ha generado el desarrollo de las tecnologías del habla, las cuales son reconocimiento, síntesis y sistemas de diálogo. El reconocimiento consiste en la decodificación de la onda sonora; la síntesis, en la conversión de texto en habla, y los sistemas de diálogo en un grado mucho más complejo de interacción en donde se involucran tanto la síntesis como el reconocimiento (Llisterri 2002). Su principal objetivo consiste en se logre un avanzado nivel de interacción entre el hombre y las computadoras por medio de la voz. A demás de alcanzar la interacción con los sistemas informáticos mediante la voz, con el fin de que las personas podamos hacer un uso más eficiente de éstos, eliminando las restricciones que, en muchas ocasiones, nos impone las herramientas actuales (Llisterri 2003).

El sistema especializado de comunicación con el que el ser humano cuenta idealmente sería el más natural, cómodo y, por supuesto, no presupondría interfaces gráficas, para la interacción entre el ser humano y las computadoras. Por ende, es indispensable enfatizar la importancia de los recursos lingüísticos para el mejoramiento de las tecnologías del habla, es decir, la ausencia de dicha información hacen impensable a éstas últimas.

Vamos a otorgar una especial importancia a los conocimientos lingüísticos que son necesarios para desarrollarlas [...] lo que ocurre es que no debemos olvidar que las tecnologías del habla requieren un conocimiento minucioso del sistema de

comunicación privilegiado de la especie humana para, de esta manera, poder construir modelos computacionales.” (Llisterri 2003:249).

La conversión de texto en habla, la síntesis, tiene como principal objetivo la transformación de cualquier entrada gráfica textual en su correspondiente realización sonora, con la cual se busca que asemeje claramente la continuidad y claridad de la voz humana por lo que se debe de apoyar en toda la información lingüística posible, no sólo la fonética segmental, sino la prosódica, la léxica, la gramatical, la sintáctica, etc. Para poder llevar a cabo satisfactoriamente dicho objetivo, se deben atender e incorporar propuestas de información fonética; esto es: reglas. Si a un conversor de texto en habla se le introduce dicha información a manera de reglas, proporcionaría un grado mucho más alto de naturalidad en la síntesis, ya que se trataría al habla como un continuo que agrupa unidades, no que corta y aísla. “La síntesis de habla consiste en la generación de una señal vocal, es decir, creada por el propio sistema informático a partir de un conjunto de datos o regla introducidas de forma previa” (Llisteri, 2003:256). Como menciona Llisterri los datos lingüísticos-fonéticos son introducidos al sistema en una fase previa de aprendizaje en donde es mucho más conveniente, ingresarlos en forma de reglas. Existen tres tipos de síntesis: de formantes, síntesis articulatoria y por reglas, precisamente esta última es la que permite el análisis y manejo de unidades básicas como la sílaba; es decir se establece un conjunto de reglas que agrupan las características de cada unidad y la forma de concatenarlas lo más real y semejante a la que haría un hablante.

Un conversor de texto en habla consta de una entrada ortográfica de unidades, las cuales en un principio fueron solamente palabras aisladas y, posteriormente, unidades más complejas como las oraciones. Dicha entrada será transformada en su respectiva correspondencia sonora; es decir, lo que se busca es que la salida de esa entrada sea una pronunciación lo más parecida a la humana.

Para que se lleve a acabo lo anterior, es necesario que se haga, entre la señal de entrada y la de salida, un análisis lingüístico de las unidades, una transcripción fonética en donde cada símbolo del alfabeto que se utilizó represente exactamente la mayoría de las características articulatorias y acústicas de cada fonema o agrupación de los mismos.

La Figura número 3 muestra un esquema del funcionamiento de un conversor de texto en habla:

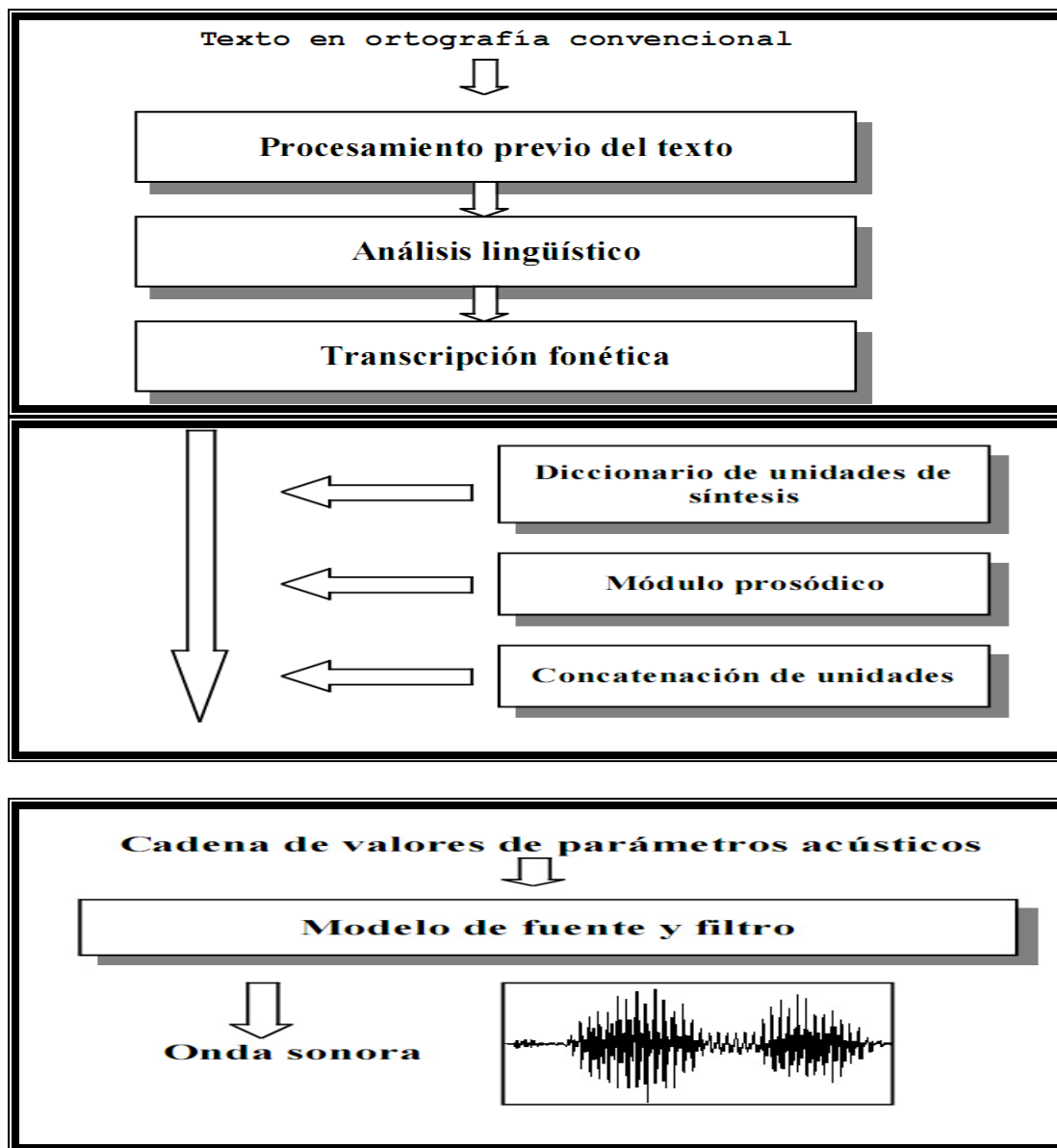


Figura 3. Arquitectura de un conversor de texto en habla (Llisterri, 2003:14).

El reconocedor de voz capta una señal sonora y la convierte de analógica a digital, así extrae sus características y las convierte en su correspondiente representación fonética que a su vez es contrastada con las entradas de los modelos acústicos y los diccionarios de pronunciación. Los elementos básicos de un reconocedor son:

- 1) Diccionario de pronunciación, el cual proporciona los listados de las realizaciones alofónicas reales que un hablante puede pronunciar.

- 2) Modelos de lenguaje, que proporcionan la forma pertinente que cada lengua tiene para hacer combinaciones entre palabras; es decir, el grado de probabilidad de que a determinada palabra le anteceda o lo preceda otra.

Tapias Merino (2002: 192) nos explica el funcionamiento completo de un reconocedor:

La señal de voz entra por el micrófono y se convierte en una señal eléctrica analógica que es posteriormente digitalizada (convertida en secuencias de unos y cero). Esta señal pasa al detector de extremos, que es el encargado de detectar la presencia de voz y de pasar dicha voz al siguiente bloque del reconocedor. El extractor de características calcula una serie de parámetros de la señal de voz que tienen información relevante para el proceso de reconocimiento. Estos parámetros se pasan al descodificador, el cual se apoya en los modelos acústicos, los modelos del lenguaje, las predicciones y el diccionario para generar la frase reconocida.

A continuación se presenta un modelo de un reconocedor de habla.

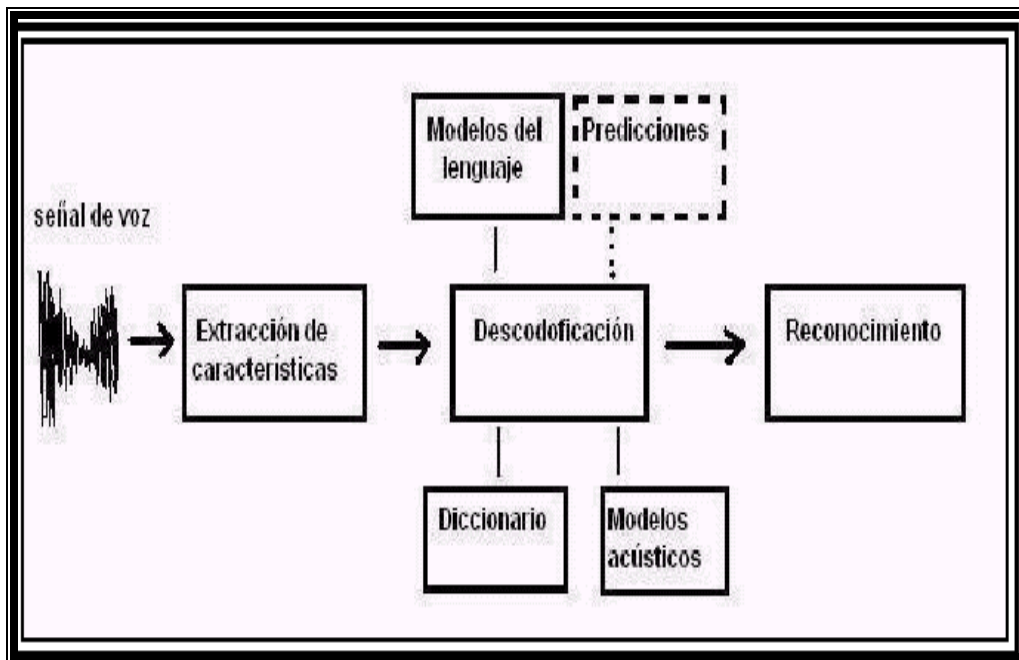


Figura 4. Modelo de un reconocedor de habla. (Llisterri, 2003:8).

Las tecnologías mencionadas, para poder llevar a cabo sus respectivos análisis fonéticos, necesitan del diseño y adecuación de un alfabeto fonético, éste tiene que aportar simbólicamente la mayor cantidad de información y características articulatorias y acústicas para la transcripción de sus corpus. “Así pues, la transcripción fonética en particular, es la representación escrita de los sonidos que pronunciamos, concebida de forma tal que pueden quedar bien reflejados en ella todos los matices que diferencian a estos sonidos entre sí” (Gil Fernández,1999:109).

2.2.1. APLICACIONES

Por el momento, la interacción comunicativa entre la computadora y el hombre es satisfactoria, aún no es posible que se sostenga una interacción oral en la que no se diferencie quien es la máquina y quien el hombre. Si bien, se han logrado interesantes avances comerciales, tales como transacciones bancarias telefónicas, compra y venta telefónica de boletos, dictados automáticos, navegación por Internet, la lectura en voz alta de un archivo de texto.

2.3. PROYECTO Y *CORPUS DIME*

El Proyecto DIME está conformado por un equipo multidisciplinario de ingenieros y lingüistas cuyo principal objetivo es desarrollar investigaciones enfocadas hacia las tecnologías del habla para nuestro país. Comenzó a desarrollarse en 1999 en el IIMAS (Instituto de Investigaciones Matemáticas Aplicadas y en Sistemas) de la UNAM. La primera fase del proyecto fue conformada por tres objetivos encaminados a la construcción de un sistema conversacional (<http://leibniz.iimas.unam.mx/>):

- La constitución y transcripción de un corpus en el diseño de cocinas
- La construcción de un programa de reconocimiento de habla para el español de México
- La definición de una gramática para el español de México

El Corpus desarrollado para el primer objetivo se llama *Corpus DIME*, el cual consta de 31 diálogos multimodales² entre persona/persona³, el cual tiene un dominio específico del diseño de cocinas, lo que permitía, entre otras ventajas, la ocurrencia de referencias orales espaciales. El total de horas de grabación es 17:10. Éstas fueron compiladas digitalmente por un programa de la compañía *Creative* llamado *Wave Studio* (<http://leibniz.iimas.unam.mx/~luis/DIME.html>). Cada diálogo dura 14 minutos en promedio y fueron grabados interrumpidamente para, de esta manera, contar con la conversación espontánea de los participantes⁴. El Corpus consta de 27, 459 palabras (886 por diálogo) y 5779 enunciados (1985 en promedio por diálogo) (Pineda *et al.* 2001).

La siguiente figura es una tabla en donde se muestra un segmento de diálogo:

1	Sistema:	¿Quieres que desplace o traiga algún objeto a la cocina?
2	Usuario:	No.
3	Usuario:	¿Puedes mover la estufa hacia la izquierda?
4	Sistema:	¿Hacia dónde?
5	Usuario:	Hacia... ¡hacia la derecha!
6	Sistema:	Hacia la derecha.
7	Sistema:	Okey.

Figura 5. Segmento de un diálogo del *Corpus DIME* (<http://leibniz.iimas.unam.mx/~luis/DIME.html>).

Los locutores participantes se sentaban en frente de una interfase gráfica de una computadora donde podían pedirle a la máquina (simulada por otra persona) la manera en la que querían diseñar su propia cocina, los espacios determinados en donde querían colocar el mobiliario, lo cual se volvía una negociación

² Consta de archivos de audio y video.

³ Los locutores sabían que a la que le estaban hablando no era una máquina sino otra persona.

⁴ Cabe aclarar que este corpus inicial sufrió una fase de edición donde se retiraron los silencios excesivamente largos al inicio y final de los archivos y la contaminación auditiva.

conversacional, donde el usuario u hablante tenía que quedar completamente satisfecho con el diseño de su cocina. Los propios usuarios podían elegir el mobiliario deseado, desde una ventana inferior donde se desplegaba un catálogo de muebles.

La Figura número 6 muestra una imagen del dispositivo que los locutores que participaron podían ver para que su tarea (el diseño de su propia cocina) fuera llevada a cabo.

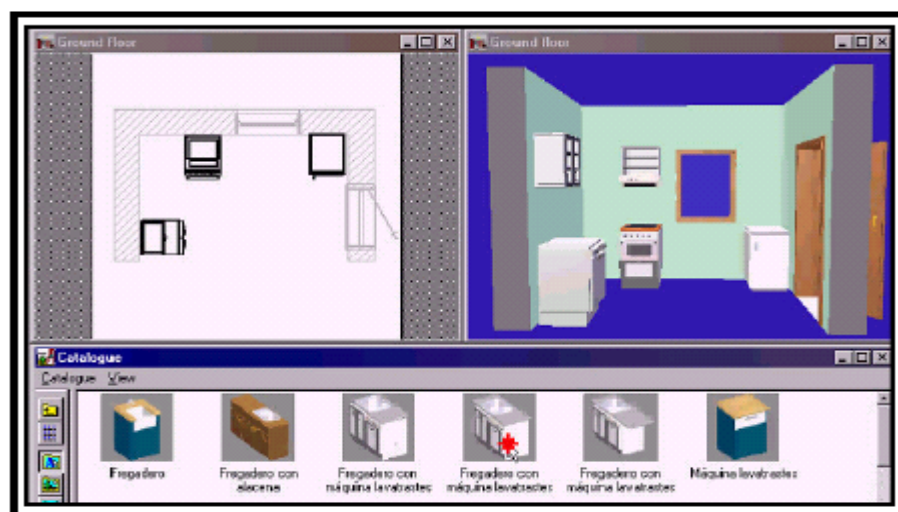


Figura 6. Interfase visual del *Corpus DIME* para el diseño de una cocina.

2.3.1. HERRAMIENTAS DE TRANSCRIPCIÓN DEL *CORPUS DIME*: *MEXBET* Y *SPEECH VIEW*

Como se mencionó en la introducción (§ 1.2.), el alfabeto que se utilizó para la transcripción del *Corpus DIME* fue *Mexbet* (Cuétara, 2004) el cual es un alfabeto computacional que proporciona la información suficiente y pertinente de los

fonemas y alófonos del español de México. La Figura número 7 presenta una tabla del alfabeto *Mexbet*:

Consonantes	Mexbet
Bilabial oclusiva sorda	p
Dental oclusiva sorda	t
Velar oclusiva sorda	k
Velar oclusiva palatalizada	k j
Labial oclusiva sonora	b
Dental oclusiva sonora	d
Velar oclusiva sonora	g
Palatal africada sorda	tS
Palatal africada sonora	dZ
Labiodental fricativa sorda	f
Alveolar fricativa sorda	s
Alveolar fricativa dentalizada	s [
Velar fricativa sorda	x
Alveolar fricativa sonora	z
Palatal fricativa sonora	Z
Bilabial fricativa sonora	V
Dental fricativa sonora	D
Velar fricativa sonora	G
Bilabial nasal	m
Nasal dentalizada	n [
Alveolar nasal	n
Palatal nasal	n~
Nasal velarizada	N
Alveolar lateral	l
Vibrante simple	r(
Vibrante múltiple	r

Vocales	Mexbet
Semivocal palatal anterior	j
Palatal anterior	i
Palatal media	e
Palatal media abierta	E
Central palatalizada	a j
Central	a
Central velarizada	a 2
Velar media abierta	O
Velar media	o
Velar posterior	u
Semivocal velar	w

Figura 7. Alfabeto fonético *Mexbet*.

Una vez asignado el alfabeto fonético, en este caso *Mexbet*, se utilizó para la transcripción el *software Speech View*, el cual es parte de un paquete llamado CSLU Toolkit que pertenece al *Center for Spoken Language Understanding* del *Oregon Graduate Institute of Science and Technology*. Esta herramienta permite la segmentación y alineación temporal de la onda sonora de los archivos de audio, cuyo análisis se vuelve accesible debido a dos oscilogramas (uno fijo y otro movable, el último permite un acercamiento detallado de la onda). También tiene tres espectrogramas, uno de banda ancha y otro de banda estrecha, en escala de grises y a colores, que permiten la identificación gráfica de los formantes y los componentes armónicos.

En la Figura 8 se presenta una imagen de la herramienta *Speech View* y de sus componentes.

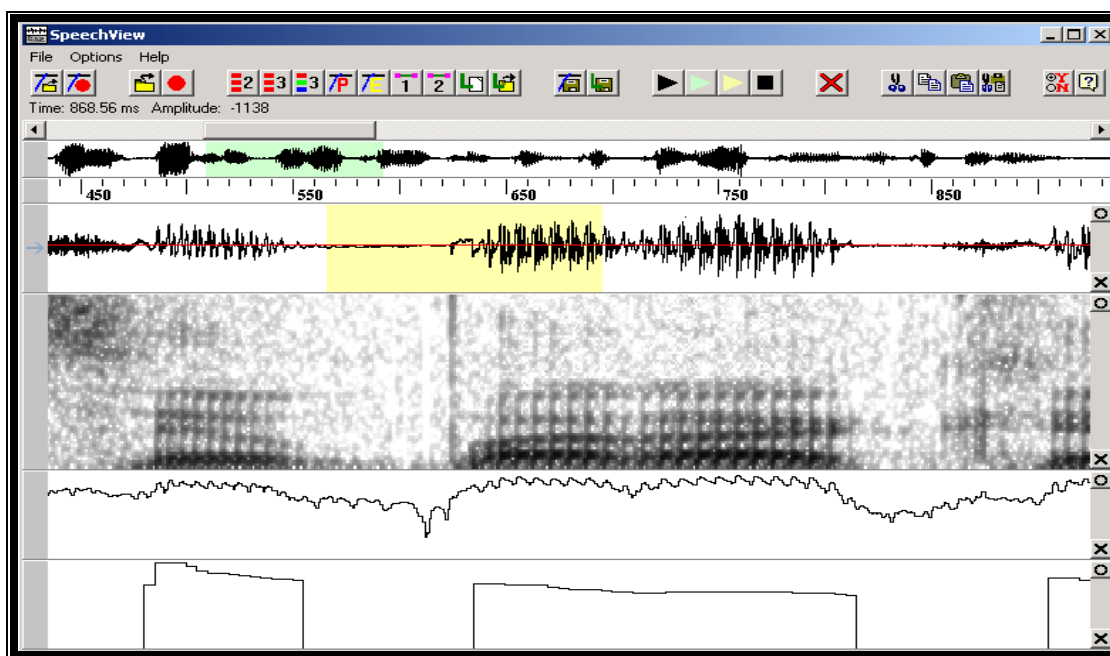


Figura 8. Imagen de Speech View.

En este capítulo pudimos percatarnos de la fructífera síntesis que se ha logrado entre ambas disciplinas (lingüística y computación): desde la fonética instrumental, hasta la lingüística computacional y la ingeniería lingüística.

3. LA SÍLABA ESPAÑOLA Y SU ESTRUCTURA

El objetivo principal de este capítulo es dar una visión teórica acerca del concepto de sílaba y de su estructura general, así como revisar distintos aspectos que a lo largo de la historia de la lingüística se han presentado, tales como el articulatorio, el espiratorio, el acústico, el fisiológico y el generativo. Posteriormente, se abordarán las características específicas de la sílaba española, las pautas pertinentes del español acerca de la estructura de la misma y las reglas fonológicas que el español tiene tradicionalmente para la formación de sílabas. La formación de sílaba a nivel ortográfico difiere de la manera real en que los hablantes tienen para hacer grupos silábicos; por ende, es importante poder tener la base canónica de la formación silábica para, posteriormente, compararla con la manera fonética que el hablante realiza. El análisis comparativo y el estudio empírico de la silabificación real son importantes para la elaboración de las reglas de silabificación, las cuales son la base medular de este trabajo empírico de tesis.

3.1. EL CONCEPTO DE SÍLABA Y ESTRUCTURA GENERAL

Los seres humanos cuando nos comunicamos producimos y escuchamos grupos fónicos secuenciales de sonidos; no escuchamos fonemas aislados, ya que éstos no son unidades con sentido comunicativo en sí; es decir, lo que realmente intercambiamos oral y auditivamente son cadenas fonéticas y entonativas, las cuales a su vez conforman la base primaria fonética de los mensajes que deseamos emitir.

Abundante es la controversia que a lo largo del estudio lingüístico se ha generado en torno a la definición de la sílaba; se han atendido para su conceptualización distintos enfoques, tales como el articulatorio, espiratorio, acústico, generativo y fisiológico; sin embargo, lo que es necesario advertir acerca de esta unidad agrupacional de fonemas es que ésta es una entidad que pareciera ser inherente a cada hablante en particular; es decir, que ésta no sólo atiende a rutinas privativas y convenciones gramaticales pertinentes de cada lengua, sino a capacidades fisiológicas y culturales individuales: “Entre los argumentos aducidos en su favor destacamos el que hace referencia a la capacidad demostrada por los individuos sin conocimientos lingüísticos para silabear las palabras, o aquel otro que resalta el hecho de que los hablantes afectados por afasia motriz subcortical realizan tantos movimientos espiratorios como sílabas contiene el mensaje que intentan expresar” (Gil Fernández, 1999:117). Lo anterior significa que la sílaba, mucho antes de ser explicada en términos lingüísticos, presentó una explicación intuitiva propia de cada hablante según la región geográfica en que le tocó desenvolverse y las capacidades fisiológicas de cada individuo. Quilis (1999) y Alarcos (1981) indican que las dificultades que presenta la definición de sílaba se deben precisamente a que se han adoptado múltiples criterios y que, finalmente, los hablantes tienen una percepción no lingüística de la mismas, como por ejemplo, el hecho de que las palabras cantadas se dividen en sílabas y no en fonemas. Sin embargo, este capítulo se enfocará en los múltiples puntos de vista lingüísticos desde los cuales se ha tratado a la sílaba, no sin antes presentar los elementos estructurales de la misma.

3.1.1. ESTRUCTURA SILÁBICA

La sílaba esta constituida de fonemas vocálicos y consonánticos; los primeros tienen la función de núcleo y los segundos de márgenes anteriores o posteriores; sin embargo, en la mayoría de las lenguas existen sílabas en las que la unidad vocálica forma por sí misma sílaba. El inventario de fonemas de cada lengua precisa la cantidad de sonidos consonánticos que pueden funcionar como codas y los vocálicos como núcleos: “Según las lenguas, el soporte o núcleo silábico puede estar constituido por una vocal, una combinación polifonemática de vocales, una consonante (sólo las sonantes, nasales y líquidas) o una combinación de vocal más consonante” (Hidalgo, 2004:208).

Para que un fonema funcione como núcleo silábico, debe caracterizarse por los siguientes aspectos: 1) duración; es decir, debe ser un fonema que pueda ser prolongado indefinidamente según la capacidad espiratoria de cada emisor; 2) presentar un máximo de abertura en las cavidades supraglóticas (faringea, nasal y bucal), que éstas se encuentren al pronunciar las vocales lo menos obstruidas o cerradas para que el aire salga libremente, sin dificultad y, 3) un máximo de sonoridad y perceptividad; auditivamente los núcleos tienen que ser fáciles de distinguir “lo que hace el receptor es casi “escuchar” directamente el núcleo silábico cuyos componentes transitorios definen a las consonantes contiguas” (D’Introno, 1995:99). Canellada (1987) también describe al núcleo silábico como el fonema que mayor perceptibilidad presenta con respecto a los márgenes o codas silábicos antecedentes y precedentes.

Las consonantes (exceptuando, en algunos casos, las líquidas /l/ o sonorantes y fricativas /s/) forman los márgenes pre y post nucleares, ya que sus características articulatorias, espiratorias, acústicas, etc. no les permiten funcionar como soportes silábicos, cada lengua tiene su inventario cerrado de fonemas consonánticos. Según Hala (1973:26), desde el punto de vista articulatorio y acústico, son no sonoras y en su función en la sílaba, no sonánticas. De hecho, el reconocimiento en la sílaba del sonido consonántico depende de las transiciones formánticas de las vocales, ya que la simple audición de éstas de manera aislada no proporcionarían una certeza de la identidad de la consonante en cuestión. Como bien menciona Malberg con respecto de la concepción tradicional de la estructura de la sílaba, “estaba constituida por una vocal formando soporte o nudo, rodeada de consonantes (de donde se toma el nombre de “consonante” que suena con alguna cosa, que no puede sonar sola.” (1962:75).

En todas las lenguas existe una manera universal de formar sílabas; es decir, existe una forma general de constituir las, ésta es: CV o sílaba abierta, la cual está conformada por un soporte silábico o núcleo vocálico antecedido por un margen anterior consonántico. Autores como Donald Burquest (1993:150), Núñez Cedeño (2000:458) y Martínez Celdrán (1984:362), entre otros, coinciden con la concepción de que este tipo de sílaba es pertinente para todas las lenguas; es un prototipo de sílaba universal: “One fact that is easily understood is that the CV syllable type is considered to be universal. That is, every language has CV syllables, and in fact

there are some languages that have only this syllable type”⁵ (Burquest, 1993:150).

Esta es una de las afirmaciones más acertadas pero no “la explicación” del porqué este tipo de sílaba subyace a las lenguas; podría ser que cómo el núcleo silábico coincide generalmente con el que recibe el acento, las consonantes después del núcleo tienen muy poca tensión y en muchas lenguas incluso llegan a perderse, y la solución más extendida es que formen sílaba con la vocal siguiente.

Una vez mencionados los componentes básicos de la sílaba, así como la estructura silábica universal, conviene hacer una remembranza de distintos aspectos lingüísticos que se han atendido para la definición de la misma.

3.2. DIVERSAS DEFINICIONES DE SÍLABA

Uno de los primeros puntos de vista que durante muchos años fue la definición de sílaba por excelencia fue el articulatorio; éste, cuyo principal expositor fue Fernandine de Saussure (1914), quien, para proponer el concepto de ésta, inicialmente categorizó los sonidos en explosivos e implosivos o abrientes y cerrantes, de acuerdo al grado de apertura que realizan los órganos articulatorios al producir los sonidos. Las características articulatorias que presentaran los fonemas son las que determinan su funcionalidad como núcleo o margen silábico; de esta manera, los fonemas que se encuentran antes del núcleo silábico presentan una paulatina abertura de los órganos articulatorios que va aumentando hasta llegar al

⁵ Un hecho fácil de entender es que el tipo de sílaba CV es considerada universal. Todas las lenguas tienen sílabas CV, y de hecho hay algunas lenguas que sólo tienen este tipo de sílaba.

soporte silábico; posteriormente, los sonidos que se localizan después del núcleo silábico van de un máximo de abertura, hasta un mínimo. Saussure establece la siguiente escala de abertura para los fonemas:

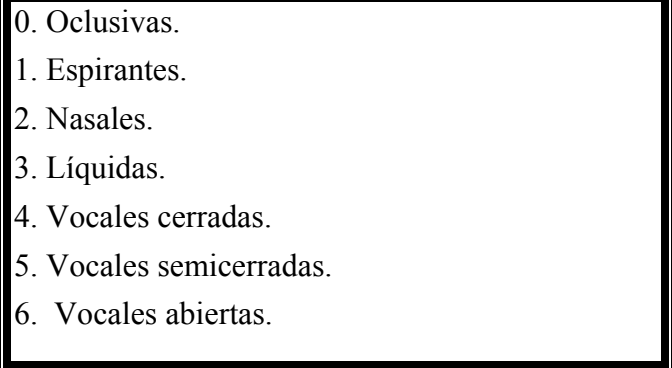
- 
0. Oclusivas.
 1. Espirantes.
 2. Nasales.
 3. Líquidas.
 4. Vocales cerradas.
 5. Vocales semicerradas.
 6. Vocales abiertas.

Figura 9. Los seis grados de abertura propuestos por Saussure.

Una de las explicaciones de sílaba que va de la mano con la definición articulatoria es la referente a la tensión que ejercen los músculos al pronunciar estas unidades; es decir, los músculos faciales e intercostales experimentan un mínimo de presión muscular durante la producción de los fonemas prenucleares y un máximo desde el núcleo hasta un mínimo en los fonemas postnucleares, por ende, los márgenes o ataques prenucleares, por lo general se conservan con mucha mayor facilidad que las codas, ya que estas últimas son más susceptibles a importantes modificaciones (por su débil posición).

Desde el punto de vista acústico la sílaba es la unidad que cumple con tres rasgos primarios: intensidad, sonoridad y perceptibilidad. Los fonemas anteriores al núcleo se caracterizan por presentar un aumento de intensidad desde un mínimo hasta un máximo que se localiza en el núcleo; los fonemas que preceden al núcleo van desde el punto máximo del núcleo hasta el mínimo que recae en el último fonema

marginal. El mismo proceso se cumple con los otros dos rasgos. Por lo tanto, autores como Gil Fernández (1999) definen a la sílaba como una unidad acústica determinada por el grado de fusión y de influencia que ejercen acústicamente entre sí vocales y consonantes. Así, acústicamente, la sílaba es la agrupación de fonemas que el oyente escucha desde un mínimo hasta un máximo de perceptibilidad, según el emisor haya organizado dicho agrupamiento fonémico.

La explicación espiratoria que se tiene acerca de la sílaba consiste en la expulsión máxima y mínima de aire que el hablante despidió durante la emisión de dichas cadenas. Esta definición permite entender por qué los hablantes con problemas de afasia emiten tantos movimientos espiratorios como sílabas contenidas en el mensaje que quieren producir. Lass (1985: Año) indica que “all languages have an inherent rhythmic organization, based on the emission of timed initiator powerburst, each burst having a single peak”⁶.

A finales del siglo XIX y principios del XX comenzaron a buscarse pruebas más asequibles en cuanto la existencia de esta unidad, por lo que comenzaron a desarrollarse aparatos capaces de registrar el movimiento de la lengua y la intensidad del aire en la cavidad bucal y torácica. El quimógrafo, el palatógrafo y el plastógrafo fueron algunos de éstos. El primero consistía en introducir en la boca un paladar artificial recubierto de una sustancia especial, la cual permitía, al ponerse la lengua contacto con dicho aparato recubierto y pronunciar distintos sonidos, registrar los puntos que ésta tocaba dentro de la cavidad bucal, “estas partes quedan limpias y

⁶ Todas las lenguas tienen una organización rítmica inherente, basada en emisión temporal de movimientos espiratorios, cada movimiento posee un solo

claramente delimitadas” (Gil Fernández, 1999). Posteriormente, dentro del área de la plastografía, Meyer fabricó un paladar similar para registrar los movimientos de la lengua; sin embargo, ya no se utilizó la sustancia especial, sino que se introdujeron junto con el paladar hilos de estaño los cuales, por su maleabilidad, permitían reconocer los movimientos “Con posterioridad, la plastografía supuso un importante avance dentro del conjunto de técnicas instrumentales de carácter fisiológico o, si se prefiere acústico” (Gil Fernández, 1999:49).

En 1955, Bertil Malberg proporcionó a la investigación lingüística y, en particular a la fonética instrumental, mediante el uso de aparatos sintetizadores, la tan buscada confirmación que permitiera la corroboración gráfica-visual de la existencia de la estructura silábica. Malberg encontró que era posible ver que la dirección que toma el formante de la vocal está determinada por la consonante de la misma sílaba y no por la consonante de otra sílaba. No obstante, para Canellada esta transición de formantes no es tan clara en el habla espontánea como en la sintetizada. En la Figura 10 se presenta una imagen en la que se muestran las secuencias *a.ga* y *ag.a*.

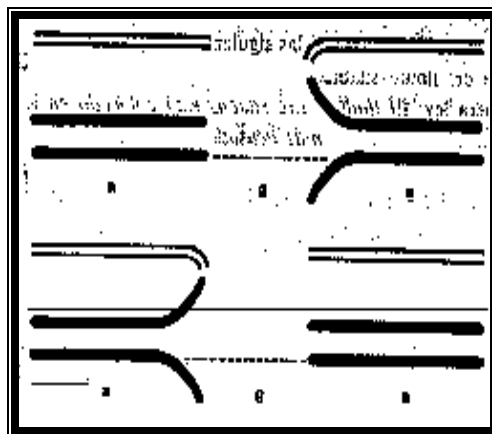


Figura 10. Sílaba *a.ga* y sílaba *ag.a* (Quilis: 1979:366).

Un enfoque muy particular de la sílaba postulado por Soporta y Contreras (1962) durante las décadas de los 60's y 70's), el cual con base en la tesis doctoral de Contreras en 1958, definían que la estructura silábica se organiza de acuerdo con un tipo de plantilla ahormacional que consta de un núcleo (N) obligatorio bordeado de cabeza (Ca) y coda (Co) opcionales, por lo que cualquier palabra o cadena hablada atendía a esta configuración básica; sin embargo, lo único que se reitera con este enfoque es la configuración universal primaria: CV, CVC, VC o V.

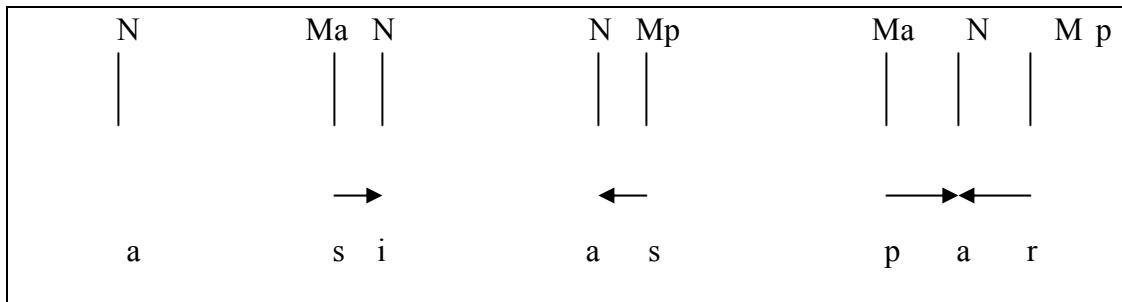
3.3. ESTRUCTURA Y CARACTERÍSTICAS DE LA SÍLABA ESPAÑOLA

La estructura silábica española atiende igualmente a la máxima universal CV; es decir, tiene una clara tendencia a la formación de sílaba abierta CV. Dicho tema se retomará a profundidad en (§ 4.2), donde se comprobará empírica y fonéticamente la siguiente máxima: “Toda secuencia se construye con sílabas; son las divisiones fundamentales de toda secuencia, y en todos los idiomas siguen un modelo de construcción definido que consta de un núcleo y márgenes. Las vocales funcionan en los idiomas como las transportadoras únicas, o por lo menos las más comunes, de los núcleos silábicos, mientras que los márgenes de las sílabas están ocupados principal y únicamente por consonantes” (Jakobson, 1987:67); sin embargo, dicha estructura no es la única que aparece en el español⁷, por consiguiente, en este mismo capítulo se tratarán el resto de las formaciones.

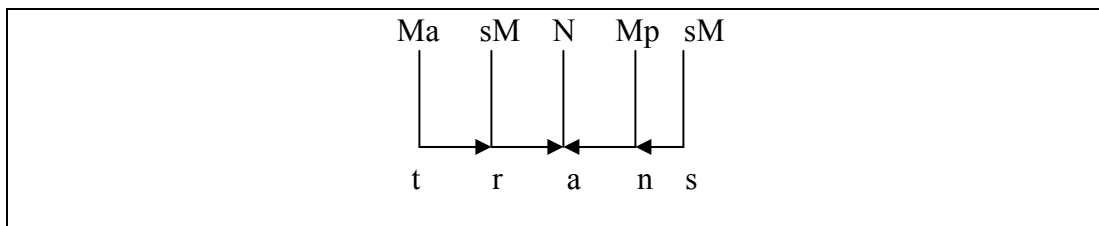
⁷ La lengua aranta australiana tan solo cuenta con esta estructuración silábica.

La sílaba española no puede ser definida sin que su núcleo sea sonido vocálico, partiendo de la premisa anterior tenemos que: 1) Dicho núcleo puede tener márgenes (sólo consonánticos); éstos a su vez son anteriores o posteriores; 2) ambos márgenes dependen directamente del núcleo aunque entre ellos son completamente independientes, por lo que, en español, podemos tener; 3) sílabas formadas tan sólo del núcleo y 4) tenemos sílabas con margen anterior, sin posterior o con posterior sin anterior.

El siguiente es un esquema arbóreo que ejemplifica las cuatro posibilidades sencillas de formación de sílaba en español de las palabras *a*, *si*, *as* y *par* (N: núcleo, Ma- margen anterior y Mp: margen posterior y sM: submargen)⁸.

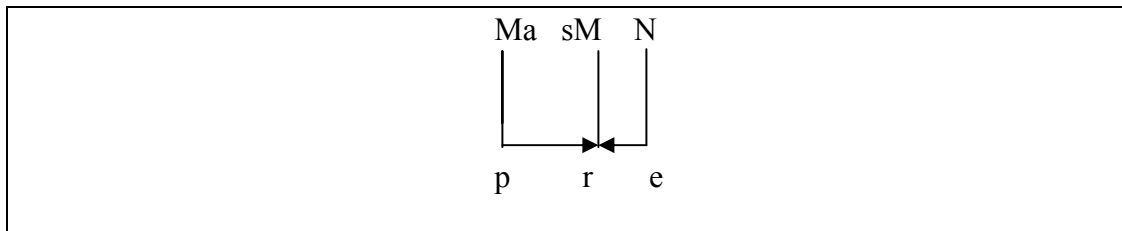


Ahora bien, los márgenes pueden, a su vez estar compuestos por más de un solo fonema, como, por ejemplo, la primera sílaba de la palabra *transformar*: /trans/. A continuación se presenta un esquema arbóreo de dicha sílaba:

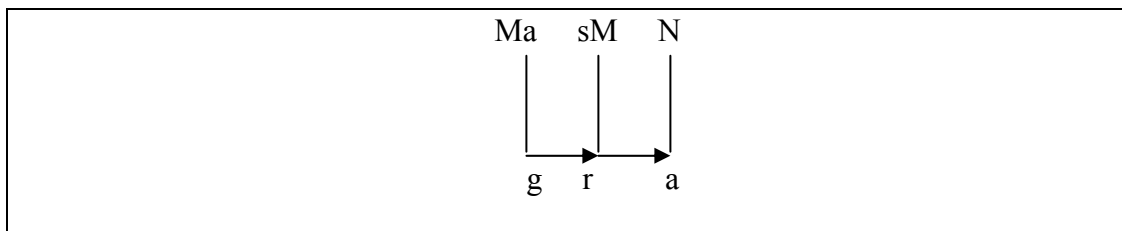


⁸ El esquema arbóreo se utiliza sólo por la facilidad gráfica que representa al leerse.

Como se mencionó, ambos márgenes pueden estar compuestos por más de un solo fonema. En español, por otra parte, son inseparables los grupos formados por consonantes bilabiales o labiodentales acompañados de una líquida: /pr br pl bl fr fl/ A continuación se muestra un esquema arbóreo de la sílaba /pre/ de la palabra *precio*; dicha sílaba presenta un margen anterior compuesto por el grupo /pr/ inseparable en el idioma español:

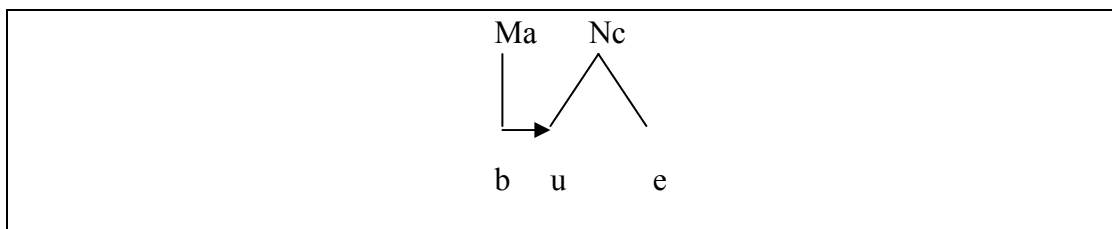


Existen otros dos grupos de fonemas inseparables en el español. Éstos son los formados por consonantes linguovelares más líquidas: /gr gl kr kl/ y el conformado por linguodentales más vibrantes: /dr tr/. A continuación se muestra un esquema arbóreo de la sílaba /gra/ de la palabra *grano*. Dicha sílaba presenta un margen anterior compuesto por el grupo /gr/, inseparable en el idioma español:

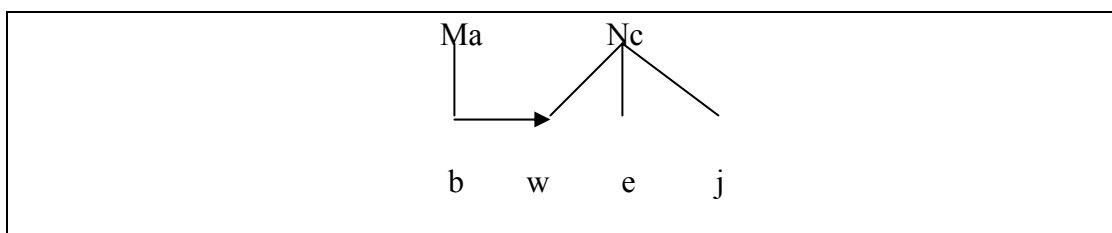


Al igual que los márgenes, el núcleo puede ser compuesto; es decir, conformado por más de un sonido vocálico. Teóricamente en español forman diptongo una vocal baja o media y otra alta o viceversa; sin embargo, como se verá en (§ 4.2.1.1.) la

formación de diptongos no sólo se da entre dichas vocales, ya que la pronunciación ordinaria y espontánea permiten la agrupación en diptongo de otras vocales, que teóricamente tendrían que formar hiatos. A continuación se presenta un esquema del núcleo silábico de diptongo del español, la sílaba es /bue/ de la palabra *bueno*:



También en español es posible la formación de un núcleo silábico formado por tres elementos vocálicos; triptongo. A continuación se presenta un esquema arbóreo del núcleo silábico compuesto-triptongo: [wej] de la palabra *buey*:



Una vez mostradas las combinaciones posibles de la formación silábica en español es pertinente preguntarse cuál es la estructura con mayor frecuencia en esta lengua. La respuesta es la sílaba libre o abierta (CV). Perissinotto (1975) se refirió a ésta como un hecho bastante conocido: la predilección del español a la formación de sílaba abierta. De hecho, Núñez Cedeño (2000) se refiere a esta forma de silabear como “automático”, no sólo del español, sino de muchas de las lenguas del mundo. Varios han sido los estudios estadísticos que comprueban dicha sentencia: Navarro

Tomás (1966:204) Granada (1966), Rafael Guerra (1983), entre otros. Es necesario advertir que el estudio de Navarro Tomás fue hecho con base en la literatura y el de Granada se cimentó sobre el mismo estudio. El análisis de Guerra está basado en lengua hablada. También cabe mencionar la importancia de hacer análisis lingüísticos fundados en lengua oral, ya que estos presuponen un acercamiento real al objeto de estudio de esta tesis. La producción oral propicia, debido a la velocidad de emisión (entre otros aspectos), la aparición de fenómenos fonéticos de silabificación, los cuales se tratarán en (§ 4.1.).

Navarro Tomas (1966:204) presenta un recuento estadístico de la frecuencia de la sílaba española basado en el *Poema de Mio Cid*; es decir, de la lengua escrita, y sus resultados fueron:

CV:	48.50%
CVC:	28.14%
V:	11.99%
VC:	6.88%
CCV:	3.35%
CCVC:	0.88%
CVCC:	0.13%
CCVCC:	0.12%

Figura 11. Frecuencia silábica en el poema *Mio Cid*. Navarro Tomas (1966:204) (C: consonante y V: vocal).

Núñez Cedeño (2000:467) enlista de forma descendente las estructuras silábicas más comunes del español y la primera casilla es ocupada por CV, “el prototipo CV es el núcleo silábico por excelencia de la lengua española siguiéndole los prototipos de CVC y VC que suman un 30%. Lo que en realidad tenemos es un arquetipo trabado por una consonante sencilla, para reiterar una verdad obvia”.

En la Figura número 12 se presenta el cuadro de frecuencias de tipos de estructura silábica del español que realizó Rafael Guerra (Guerra, 1983, *Apud.* Quilis)

CV:	55,81%
CVC:	21,61%
V:	9,91%
VC:	8,39%
CCV	3,14%
CVCC:	0,02%
CCVC:	0,98%
CCVCC:	0,01%
VCC:	0,13%

Figura 12. Cuadro de frecuencias silábicas del español Guerra, 1983, *Apud.* Quilis. Estudio realizado sobre la lengua hablada).

Los estudios estadísticos realizados por Guerra y por Navarro Tomás serán tomados como base teórica comparativa respecto de los datos porcentuales obtenidos en este estudio empírico en el próximo capítulo.

4. COMPORTAMIENTO SILÁBICO DEL *CORPUS DIME*

El presente capítulo es la base medular de este trabajo de tesis. El objetivo principal consiste en la propuesta de un conjunto de reglas de silabificación fonética, las cuales fueron obtenidas con base en el análisis empírico porcentual de los siguientes fenómenos de silabificación: 1) formación de sílaba abierta; 2) homologación de segmentos idénticos vocálicos y consonánticos y 3) sinalefa. Cada fenómeno cuenta con una explicación teórica, posteriormente se presentan las características y explicaciones fonéticas de cada fenómeno, así como también figuras que muestran imágenes, tablas porcentuales y ejemplos representativos de cada uno. Finalmente se presenta un análisis comparativo entre los datos obtenidos y algunos estudios fonéticos representativos. El contar con las siguientes reglas permiten la extracción real de la silabificación que los hablantes realizan en sus producciones, por lo que el empleo de éstas en las tecnologías del habla proporcionan una base de datos que posibilitan un incremento sustancial en el índice de reconocimiento y, por su puesto, proporcionarían a la síntesis una mejoría en la naturalidad de la voz como se explicó (§ 1 y § 2.2.).

4.1. COMPORTAMIENTO GENERAL DEL CORPUS

Como se mencionó (§ 1), el *Corpus DIME* es una muestra significativa de habla espontánea que proporcionó el análisis empírico de fenómenos propios del campo de estudio de la fonética. La producción de habla en su carácter natural presenta ciertas

peculiaridades; por ejemplo, la rapidez con la que se emiten las elocuciones cuando existe un alto grado de familiaridad con el receptor y el tema o tarea específica, propicia la aparición de fenómenos como el de la formación de sílaba abierta entre palabras; es decir, sin que se respete el límite de palabra, la reducción de segmentos idénticos a uno solo y el agrupamiento de fonemas en una sola emisión silábica (sinalefa). Halá (1966:100) menciona al respecto: “Estos agrupamientos de consonantes han dado lugar a diversos cambios fonéticos, de suerte que actualmente no siempre es fácil descubrir el límite silábico con precisión; además, el hombre ha aprendido a hablar como consecuencia de su desarrollo mental y del aumento de las necesidades de comunicación, con mayor rapidez y con más economía, y esto también ha modificado de alguna manera los grupos consonánticos” (1966:100). Malberg (1962) y Company (2001) explican en términos de economía articulatoria el agrupamiento y combinación de fonemas en una sola emisión silábica; es decir, el emisor busca el ahorro en los movimientos articulatorios que no sean indispensables para lograr el efecto acústico deseado. La formación de la sílaba atiende principalmente a razones articulatorias; los movimientos de nuestra lengua y labios podrían decirse que son un tanto perezosos y los sonidos se contagian unos con otros con facilidad y nuestros movimientos articulatorios no llegan a “contemplar la configuración del resonador que se corresponde con los timbres vocálicos normales y es bastante común que se hagan realizaciones híbridas” (D’Introno, 1995:99). El análisis del *Corpus DIME*, con respecto a la teoría tradicional, demostró un comportamiento característico de la lengua en uso; es decir, sus mayores quiebres o discontinuidades fueron específicamente con la manera tradicional de silabear las

palabras aisladas o representadas ortográficamente. Se encontró un alto índice de consistencia con la teoría consultada con respecto a la sílaba abierta.

La conducta de los datos analizados permitieron la simplificación de dichos fenómenos encontrados en reglas específicas de comportamiento silábico de la lengua oral, las cuales son el resultado de un análisis empírico fonético exhaustivo. A continuación se presenta el análisis correspondiente al fenómeno de sílaba abierta.

4.2. FORMACIÓN DE SÍLBA ABIERTA

La sílaba abierta, como se mencionó (§ 3.1.1.), es la tendencia general primaria de todas las lenguas de dejar al núcleo silábico (vocálico en español) sin un margen u apoyo posterior, el cual se convertirá en el apoyo anterior de la siguiente sílaba, si es que ésta comienza con vocal. En las muestras de habla analizadas se encontró que para el español se forman con mayor sistematización los siguientes apartados de sílaba abierta:

1. Vibrante simple con las vocales / e a u o /; /r e/, /r a/, /r u/ y /r o/.⁹
2. Fricativa alveolar con las vocales / a e/; /s a/ y /s e/.
3. Nasal alveolar con vocal / e /; / n e/.

⁹ Es indispensable mencionar que la formación de sílaba abierta con vibrante simple fue la que mayor índice de ocurrencia presentó debido a la alta frecuencia de verbos en infinitivo con la que el dominio del corpus cuenta.

A continuación se hará una descripción de cada fonema con el que se formaron las sílabas abiertas y se mostrarán los datos analizados con tablas y gráficas porcentuales de cada categoría y subcategoría. El total de las fichas, imágenes y tablas se localizan en el Apéndice número 1 al final de este trabajo. Fue necesario hacer dicho apartado para que la lectura de este capítulo no se viera interrumpida por una gran cantidad de información visual.

4.2.1 SÍLABA ABIERTA CON FONEMA /r /

Fonema alveolar que de acuerdo a su posición en la sílaba (prenuclear o postnuclear) se pronuncia de distintas maneras.

- a) Vibrante múltiple a comienzo de palabra o en interior de palabra, comienzo de sílaba cuando va precedida de vocal. Ejemplo: *ratón*
- b) Vibrante simple a comienzo de sílaba precedida de vocal o en grupo consonántico prenuclear C /r/ (en estos casos C es una consonante /p, t, k, b, d, g/). Ejemplo: *dragón*

Se presentó en su modalidad de vibrante simple en las 34 fichas que cuentan con la combinación /r e/; en el 86 % de dichos casos, esta vibrante simple, pertenecía al infinitivo de verbos propios del corpus, tales como: *mover, poner, llevar, deslizar, colocar, eliminar, girar, revisar, centrar, acomodar y seleccionar*. El 14% restante corresponde a la preposición *por* y el sustantivo *color*. Todos los casos presentan a /r/ en posición postnuclear; sin embargo, no se presenta ningún proceso de debilitamiento o trueque, lo que ocurrió fue que la siguiente vocal, en este caso /e/, habilitó a la vibrante para que formara el margen inicial o ataque de la siguiente

sílaba inicial de palabra. Cabe mencionar que la pronunciación de esta última sílaba se presenta con la modalidad del alófono alveolar vibrante simple, es decir, que no se realiza como teóricamente tendría que cumplirse, no ocurre una realización de vibrante múltiple; aunque sea inicio de sílaba y de palabra precedida de vocal. De esta manera, tenemos la siguiente la realización del fonema /r/ en un fragmento de dialogo. La primera columna del lado izquierdo muestra el número de ficha, la segunda el segmento que presenta el fenómeno transcrito alofónicamente en el nivel T54 y la silabificación, la tercera corresponde a la transcripción ortográfica de la emisión, la cuarta al número de diálogo del *Corpus DIME* al que corresponde, la quinta al número de oración y la sexta nos indica si dicha silabificación fue emitida por el emisor que tenía el rol de usuario o por el que tenía el papel de sistema.¹⁰

Ficha	Segmento	Transcripción	Diálogo	Oración	Emisor
1	mo Ve_7 r(e les	puedes mover el estante	11	107	u

La pronunciación del segmento resilabeado corresponde a [mo bé re les]. En la Figura número 13 se presenta una imagen espectrográfica del fenómeno en cuestión, donde se aprecia cómo la vibrante simple forma sílaba con la siguiente palabra dejando libre la sílaba a la que inicialmente pertenecía. El primer nivel de celdas que se encuentra debajo del oscilograma muestra la segmentación alofónica; el segundo es donde se localiza la silabificación tal y como la produjo el hablante: (la silabificación fonética.) y el tercero presenta la silabificación tradicional.

¹⁰ Las fichas completas de todos los fenómenos se localizan en los apéndices al final de este trabajo, tan sólo se muestra un fragmento a modo de ejemplo.

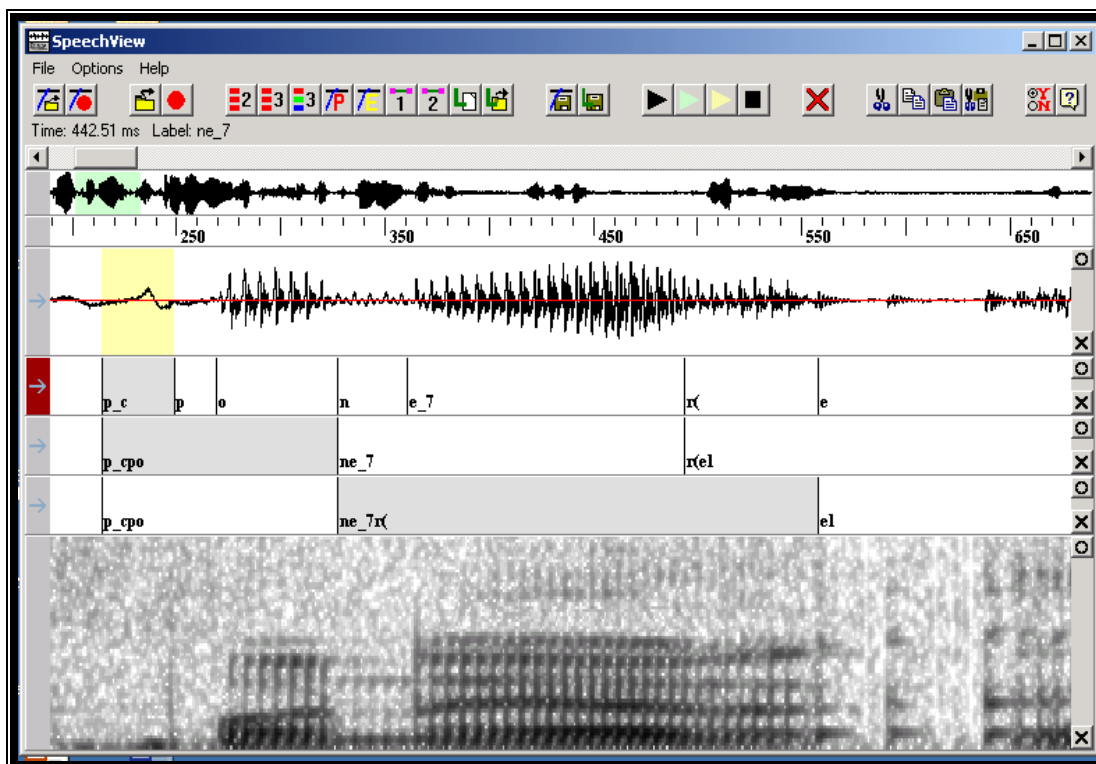


Figura 13. Imagen espectrográfica de sílaba abierta “r e”: en la frase *poner el* [poné rel]. Diálogo 3; oración 5.

En el caso de la formación de sílaba abierta /r a/ se obtuvo un 69% de aparición. El 100% de las fichas analizadas con el fenómeno se caracterizaron, como en el caso anterior, por presentar a la vibrante en su modalidad simple. En el 90% de los casos la vibrante pertenecía al sustantivo *extractor*. En el 100% de los casos la vibrante fue postnuclear con núcleo tónico, por lo que se puede anticipar debido a que dicho soporte tónico recayó en esa sílaba, ya no fue necesario el margen posterior, en este caso la vibrante simple; por ende, y porque la siguiente era una vocal, pasó a formar el margen inicial de la siguiente sílaba, dejando a la sílaba abierta; es decir, después de pronunciar el núcleo silábico tónico, los sonidos subsecuentes en dicha palabra

pierden fuerza articulatoria, espiratoria y muscular. En la siguiente tabla se presenta una ficha con este fenómeno.

Ficha	Segmento	Transcripción	Diálogo	Oración	Emisor
5	ne _7 r(a k _ck _jj	poner aquí	2	42	u

La silabificación de dicho segmento es [po ne ra kí] y lo que fonológicamente tendría que ser /po nér a kí/. En la Figura número 14 se presenta la imagen espectrográfica del fenómeno donde se puede apreciar silabificación y formación de silababa abierta en el segundo nivel de segmentación.

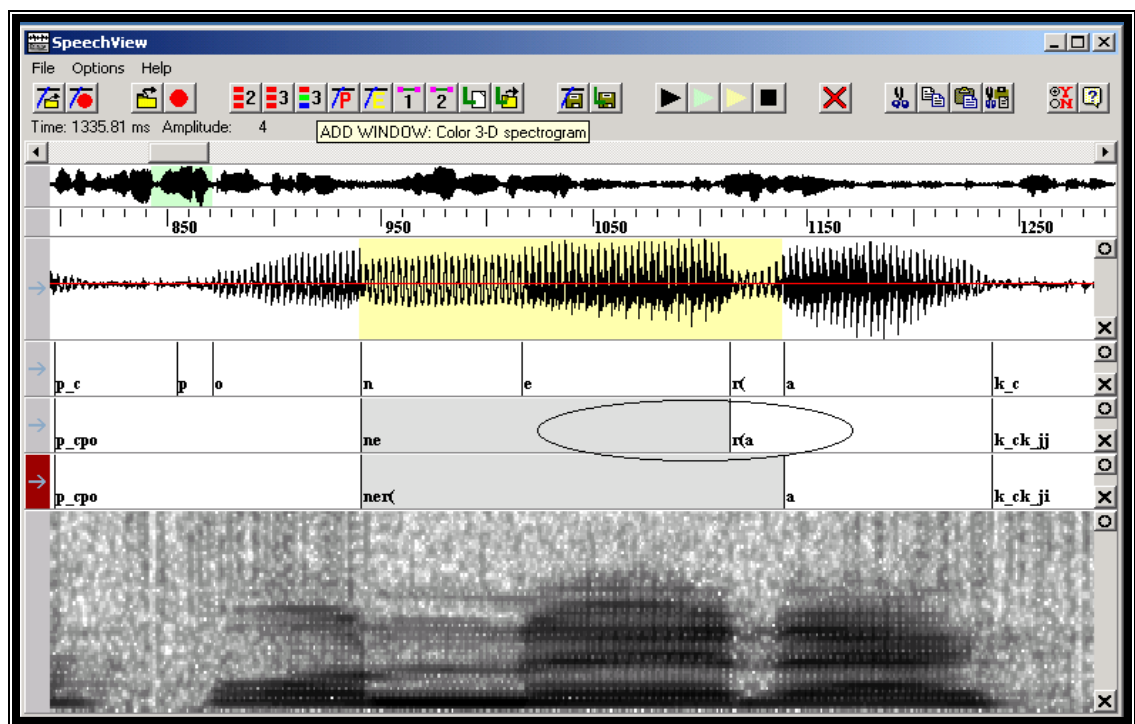


Figura 14. Imagen espectrográfica de sílaba abierta “r a”: en la frase *poner aquí* [po ne ra kí]. Diálogo 7; oración 6.

En el caso de la formación de sílaba abierta con la vibrante múltiple y la vocal /o/ el 90% de las fichas corresponden a verbos en infinitivo y el 10% restante corresponde al sustantivo *refrigerador* del corpus seleccionado. El 100% de las fichas analizadas al que pertenecía la vibrante simple también era núcleo silábico, por lo que, probablemente, el margen silábico posterior /r/ perdió fuerza de pronunciación y en vez de perderse prefirió ser ataque o margen anterior de la siguiente vocal. En la siguiente ficha se presenta un ejemplo de ficha con este fenómeno.

Ficha	Segmento	Transcripción	Diálogo	Oración	Emisor
3	xi r(a_7 r(o t_ctr(a Ves	girar otra vez	11	40	u

En la Figura número 15 se muestra gráficamente dicho fenómeno

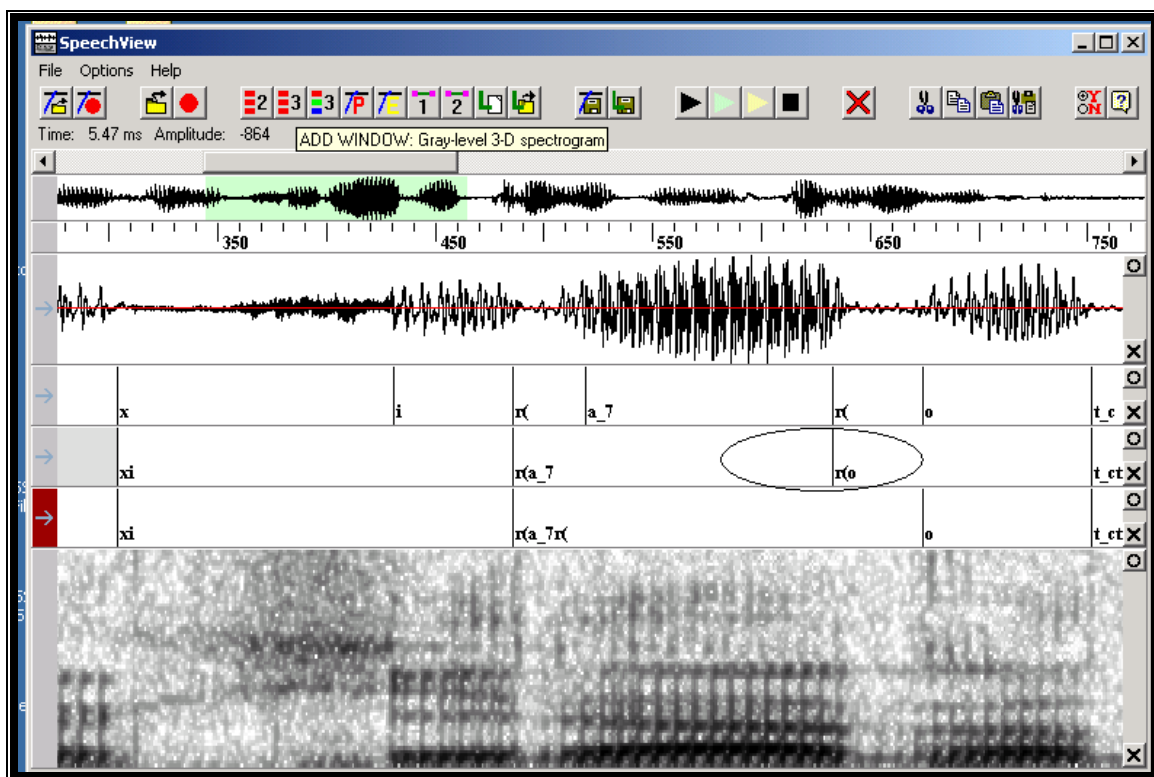


Figura 15. Imagen espectrográfica de sílaba abierta “r o”: en la frase *girar otra vez* [gi.rá.ro.tra.bes.] Diálogo 11; oración 40.

El último caso de formación de sílaba abierta fue con la vocal /u/; en el corpus de trabajo se presentó un 63% de aparición del fenómeno. El 100% de los casos la vibrante pertenece a verbos en infinitivo y la siguiente palabra es el artículo indeterminado singular masculino *un* o al indeterminado femenino singular *una*. El 80% de los casos con silabificación o resilabeo, el núcleo silábico a la que pertenece la vibrante es el tónico. A continuación se muestra un ejemplo de ficha y su respectiva Figura número 16.

Ficha	Segmento	Transcripción	Diálogo	Oración	Emisor
1	r(a ro_7 t_ cta r(u_7n	rotar un poco	2	10	u

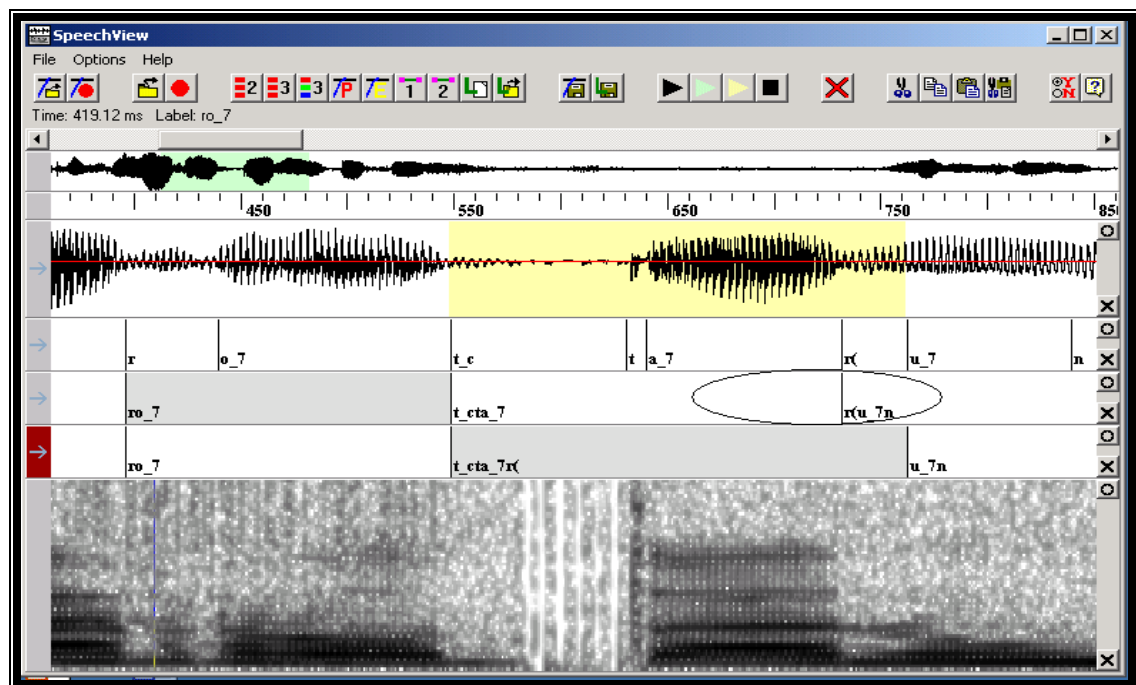


Figura 16. Imagen espectrográfica de sílaba abierta “r u” en la oración *rotar un* [ro.tá.run.] Diálogo:2; oración 10.

4.2.2. SÍLABA ABIERTA CON FONEMA /s/

El fonema alveolar /s/ que se realiza en el español de México es de una articulación predorsal a comienzo de sílaba; es decir, con el ápice de la lengua mucho más relajado al contacto con los dientes incisivos superiores. En posición final de sílaba, la articulación de /s/ varía mucho más de acuerdo a cada dialecto. Generalmente el fonema /s/ sonoriza y disminuye ligeramente su tensión cuando precede a una consonante sonora.

Las fichas detectadas en el corpus analizado corresponden a la articulación predorsal a final de sílaba y presentaron una clara tendencia a conservar el fonema mucho mejor como inicial de sílaba que como final; es decir, el último fonema de la última sílaba de la palabra era /s/ y el fonema inicial de la siguiente palabra era vocal /a/, por lo que la /s/ final de sílaba dejaba libre a la vocal anterior y formaba sílaba con la siguiente vocal, aunque esta correspondiera a otra palabra. El corpus presentó un 70% de sílaba abierta con fonema /s/ y vocal /a/; dicho fonema corresponde al margen silábico posterior de verbos conjugados en presente con la segunda persona singular y la primera persona plural: *tienes*, *movemos*, *podemos*, etc. También este sonido corresponde al artículo indeterminado plural *unos* y al definido plural femenino *las*. A continuación se muestra un ejemplo de sílaba abierta /s a/ el cual muestra fonéticamente la silabificación o resilabeo de un fragmento del diálogo 11 emitido por el usuario. La silabificación tradicional o teórica no sucede; el segmento de oración es *tienes algún*. La silabificación canónica sería: /tié nes al gún/ y lo que realmente sucede es [te ne sal gún].

Ficha	Segmento	Transcripción	Diálogo	oración	emisor
1	t_cti ne sa_2l Gu_7	tienes algún	11	143	u

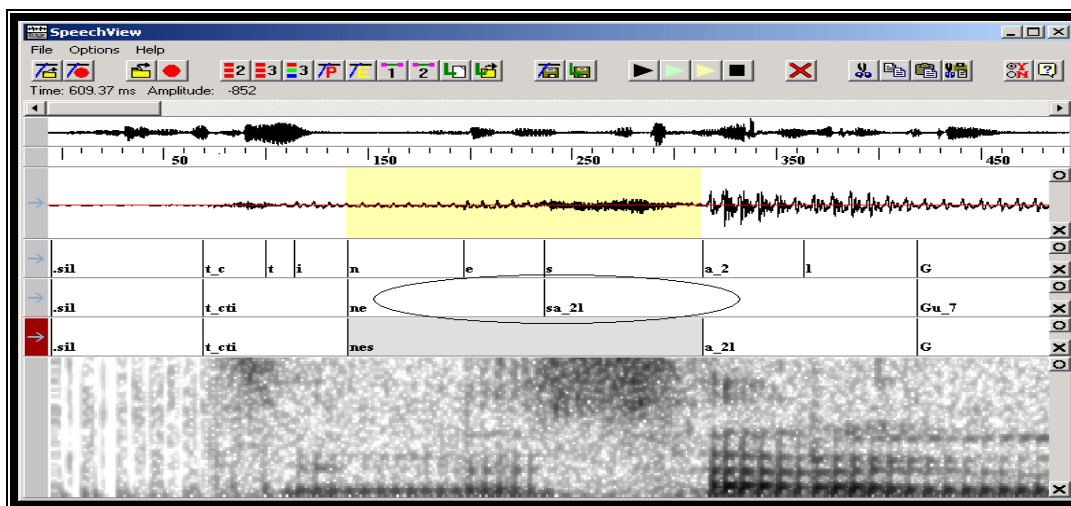


Figura 17. Imagen espectrográfica de sílaba abierta “s a”: *tienes algún* [ti.ne.sal.gun]. Diálogo 11; oración 15.

La Figura número 17 muestra cómo el sonido /s/ de la sílaba final /nes/ forma sílaba con la vocal siguiente dejando libre a la sílaba anterior. Existe una posible explicación fonética de este fenómeno con el fonema alveolar y no sólo por la tendencia del español de formar sílaba abierta, ésta tiene que ver con la naturaleza de cantidad con la que cuenta el fonema /s/; es decir, para el hablante que esta emitiendo una cadena continua de habla le resulta poco cómodo hacer un corte entre el último sonido de la palabra y el siguiente, por ende, resulta más fácil adherir esta consonante como ataque de la siguiente sílaba que en este caso es la vocal /a/ que conservarla y pronunciarla largamente en el lugar donde estaba.

La formación de sílaba abierta /s e/ reportó un 56% de aparición en el corpus analizado. La naturaleza gramatical de las palabras en las que el fonema /s/ funcionó como coda en la última sílaba fue muy variada: verbos, sustantivos, adjetivos demostrativos, artículos, sustantivos propios del corpus, etc. A continuación se

presenta una ficha con un ejemplo de la transcripción de este fenómeno y la Figura número 18 es la imagen espectrográfica de dicho fenómeno.

Ficha	Segmento	Transcripción	Diálogo	Oración	Emisor
16	.sil la ses_ [t_ctu	las estufas	20	4	s

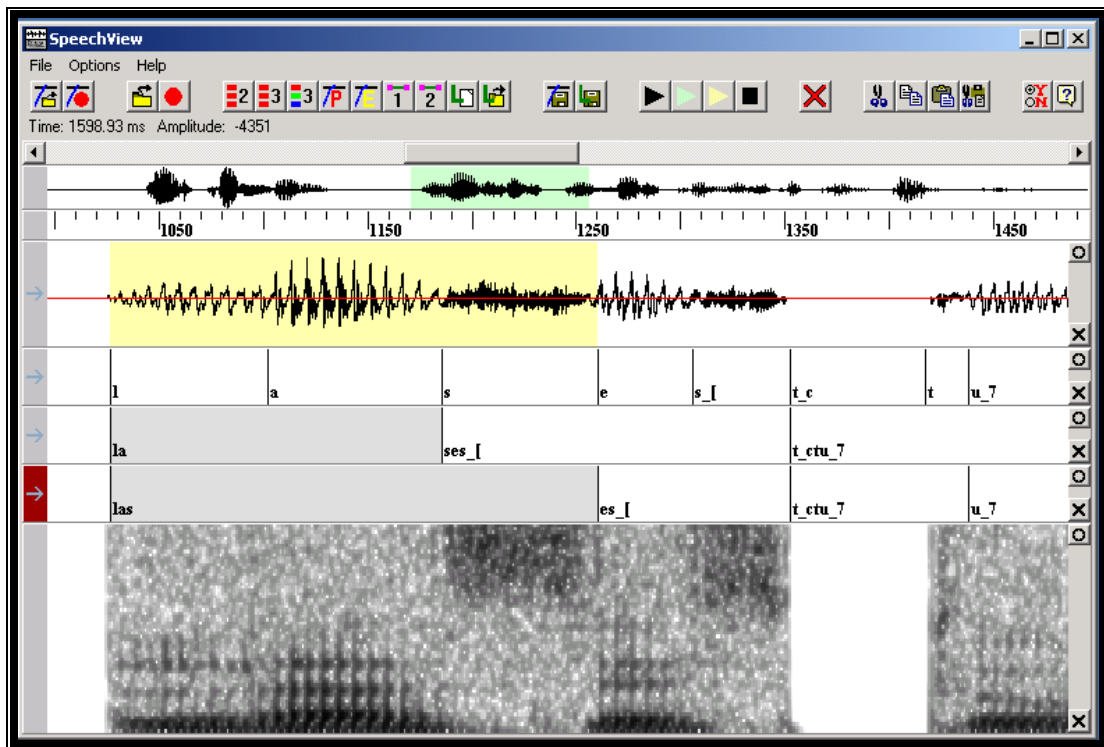


Figura 18. Imagen espectrográfica de sílaba abierta “s e “: las estufas-la ses tu fas. Diálogo 20; oración 4.

4.2.3. SÍLABA ABIERTA CON EL FONEMA /n/

El último registro de sílaba abierta fue con /n/, uno de los fonemas con mayor aparición en el español de México (Cuétara, 2004). De esta manera se explica el porque en todo el corpus se tuvo un alto índice de aparición las sílabas que presentaron este fonema como coda silábica: exactamente 89 casos, de los cuales 61 presentaron la formación de sílaba abierta; es decir, la formación de sílaba abierta /n/

e/ fue del 68% en el corpus de trabajo. El fonema /n/ se articula de distintas maneras.

D’Introno (1995:309) considera que si ordenamos las realizaciones más comunes de /n/ a final de sílaba (incluyendo posición posnuclear) sobre una escala de debilitamiento se obtiene:

1. Realización alveolar (variante más fuerte y también se da a comienzo de sílaba). Ejemplo: /na da/
2. Realización coarticulatoria /nm/ que resulta de la “anticipación” del punto de articulación de la consonante bilabial siguiente, una articulación secundaria que se sobrepone a la alveolar siguiente de /n/.Ejemplo: /kon mu ta dór/
3. Nasal asimilada en punto de articulación a la consonante siguiente. /pa te ón/
4. Variante debilitada o articulación velar del fonema. /mán go/
5. Realización no nasal, antes con “restos” de ésta en la forma de nasalización de la vocal precedente. /fun da ciŃ/
6. Variante más débil sin nasalización de la vocal. /kos ti tu sjón/

Cabe mencionar que el comportamiento de esta consonante en el *Corpus DIME* fue consistente con la categoría número uno expuesta por D’Introno. Dicho fonema originalmente era la coda silábica de una palabra, esta posición, como sabemos, propicia el debilitamiento del fonema, por lo que dicho sonido prefirió ser ataque silábico y formar sílaba con la siguiente palabra y dejar libre a la anterior sílaba. En el corpus de trabajo aparecieron 61 casos exitosos de los cuales en 32 ocasiones el fonema /n/ formaba parte de la preposición *en* seguido de los adjetivos demostrativos *este, esta y estas*.

4.3 REDUCCIÓN DE SEGMENTOS IDÉNTICOS

Cuando entran en contacto dos sonidos idénticos, la solución más generalizada y habitual que el hablante prefiere utilizar es la reducción de éstos en un solo sonido; dicha reducción se debe normalmente a la velocidad de emisión; sin embargo, es bien sabido que cuando la realización es enfática o cuidada el hablante produce dichos fonemas de las siguientes maneras:

1. Pronunciación pausada y definida de los dos segmentos idénticos
2. Compensación de la vocal en el aspecto de duración; es decir, se emite un solo fonema de duración más prolongada sin que suene como dos.

La homologación vocálica dentro del corpus DIME en el nivel de frase, entre la última sílaba de la palabra y la primera de la siguiente, origina que se modifique y que se formen sílabas entre límites de palabra. El fenómeno de homologación vocálica en el corpus de análisis se dio con la vocal media anterior /e/ y la vocal abierta /a/, fonemas que tienen la mayor frecuencia de aparición en el español.

El *Corpus DIME* presentó 169 fichas donde estaban en contacto dos palabras cuya sílaba final de la primera terminaba en -e y la sílaba inicial de la siguiente palabra iniciara en e- también; sin embargo, sólo 48 fichas redujeron el segmento a uno solo; es decir, el fenómeno contó con un porcentaje de aparición de 29% debido que el resto de las palabras estaban separadas por silencio o ruido intermedio por lo que no estaban directamente concatenadas. A continuación se muestra un ejemplo donde la

reducción de /e/ se transcribe alofónicamente y se puede observar un proceso de silabificación por homologación de segmentos.

Ficha	Segmento	Transcripción	Diálogo	Oración	emisor
16	t_ctes t_cta_7n_	este estante	11	92	u

La producción canónica indica que este segmento tendría que silabearse /és te es tán te/; sin embargo, lo que realmente ocurre es [es tes tán te]. Las dos vocales medias anteriores se redujeron a una sola y la duración de producción también corresponde a una sola vocal. En la Figura 19 se muestra en la imagen espectrográfica cómo no existe ninguna compensación de tiempo en la producción de la vocal.

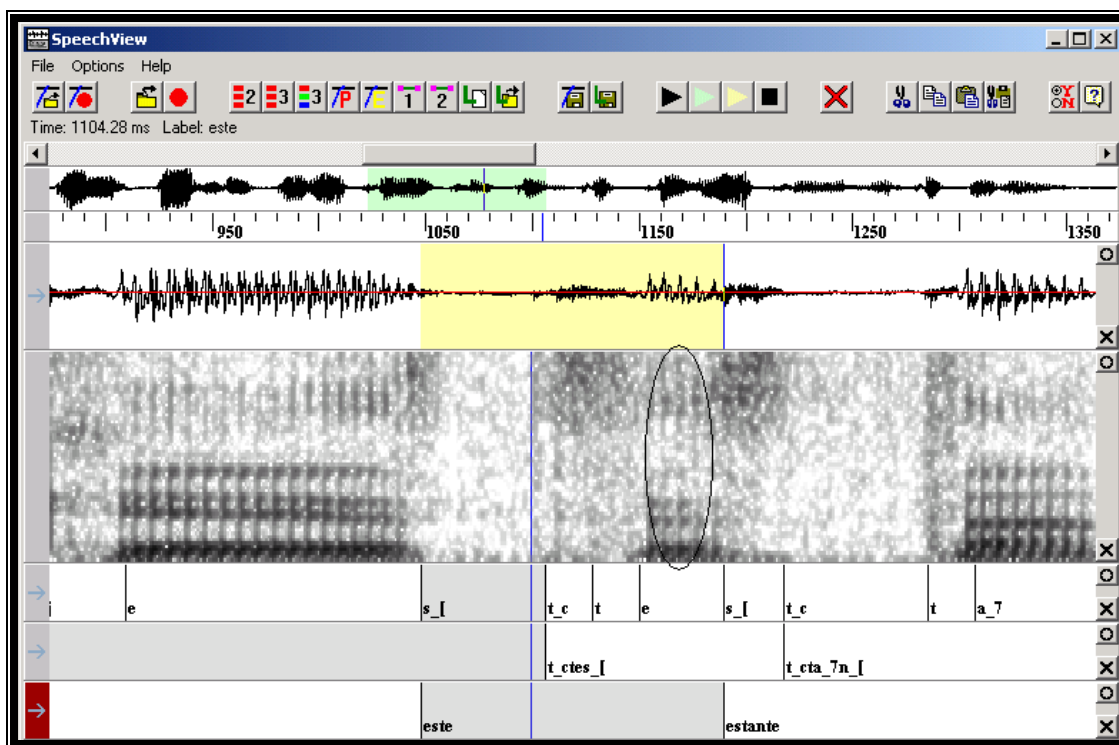


Figura 19. Imagen espectrográfica de homologación vocálica “e e”: *que este estante* [kes.tes.tan.te.] Diálogo 11; oración 92.

Quilis (1999) indica que cuando dos vocales que se encuentran en contacto son átonas o inacentuadas, lo que el hablante produce es una vocal normal átona y que lo mismo ocurre cuando una de las dos vocales pertenece a una palabra inacentuada (artículo determinado, preposición, conjunción, etc). El 31% de las fichas analizadas con el fenómeno presentan la característica que describe Quilis, puesto que la vocal /e/ pertenece, en dichos casos al adjetivo demostrativo *este* y el siguiente segmento homólogo vocálico es de sustantivos inacentuados en dichos segmento; es decir, en palabras que no tienen acento en la primera sílaba, (por ejemplo, *estante*, *extractor* y *espacio*). En un 25% de las fichas analizadas se tiene que la vocal homóloga corresponde a la conjunción *que*, la cual va seguida de los verbos *ser* y *estar* conjugados tanto en modo indicativo como subjuntivo: *es*, *está*, *estoy*, *esté* y *estaba*. Lo anterior comprueba que ninguno de los dos elementos vocálicos es átono, por ende, la solución que prefirieron emplear los hablantes del *Corpus DIME* fue un solo segmento, que es producido y percibido con la duración de uno solo. Lo mismo ocurre con un 16% de las fichas cuyos elementos vocálicos son átonos; siendo el primero de éstos parte de la preposición *de* seguida de adjetivos demostrativos *esta* y *este*.

Como se mencionó, la homologación vocálica también se presentó en el corpus con la vocal abierta /a/. El fenómeno contó con un porcentaje de incidencia del 32%. La reducción del segmento idéntico se comportó de manera similar a las fichas con vocal /e/; el 99% de las vocales homologadas fueron átonas y no existió ninguna compensación de duración en su producción. El 31% de las vocales pertenecen a adjetivo demostrativo *esta* seguido del sustantivo *alacena*. A continuación se

presenta un segmento de una ficha con el fenómeno. La silabificación tradicional nos indica que el enunciado *esta alacena* debe de ser silabificado /és ta a la sé na/; sin embargo, lo que realmente sucede en el habla común es [es ta la sé na]. Ninguna de las dos vocales son núcleos silábicos, por lo que son aún más susceptibles de reducirse y de formar una sola sílaba. En § 4.4. (sinalefa) se tratarán las fichas cuyas vocales distintas son susceptibles a formar un solo grupo silábico aunque sean vocales distintas.

Ficha	Segmento	Transcripción	Diálogo	Oración	Emisor
1	t_ cta la se_7 na	esta alacena	7	37	u

En la Figura número 20 se presenta la imagen espectrográfica en la que se puede apreciar la sílaba que se produjo debido a la reducción de los segmentos idénticos, así como también el tamaño de la vocal, la cual es del mismo tamaño de las simples. La vocal encerrada en el óvalo continuo pertenece a la vocal reducida a una sola en la misma sílaba y la señalada en un rcuadro pertenece a una vocal sin reducción.

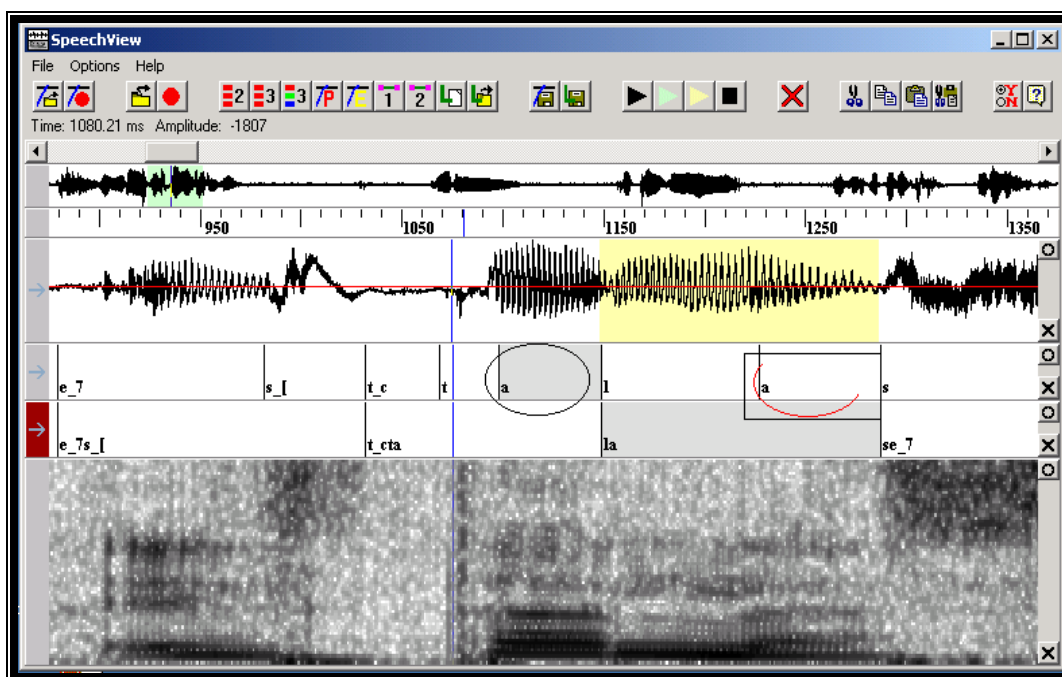


Figura 20. Imagen espectrográfica de homologación vocálica “a a”: esta alacena-es.ta.la.se.na. Diálogo 7; oración 37.

4.3.1 REDUCCIÓN DE SEGMENTOS CONSONÁNTICOS

La consonante que mayor índice de homologación fue /s/. El corpus de análisis presentó 25 casos en los que el fonema se puso en contacto directo; es decir, no fueron separadas por silencio o ruido las palabras a las que pertenecía el fonema. En el 100% de las fichas ocurrió la reducción de la consonante; la duración del segmento corresponde al de una sola consonante. El 48% de los fonemas consonánticos pertenecen a los adjetivos demostrativos *estos* y *esos*, los cuales van seguidos del verbo conjugado en la tercera persona plural *son*. El resto de los fonemas corresponden al artículo determinado plural femenino *las* seguido del sustantivo *sillas*.

Se comprobó empíricamente lo que dice Quilis (1999:375): “Cuando dos consonantes linguoalveolares sordas se encuentran en contacto, la solución es una sola [s] cuya duración viene a ser aproximadamente la misma que la de cualquier consonante [s] intervocálica”; cabe resaltar que la /s/ resultante no se sonorizó en ninguno de los casos, a pesar de ser intervocálica. A continuación se muestra un ejemplo de la reducción consonántica en una ficha de trabajo. La tradición nos indica que la silabificación del segmento tendría que haber sido /és tos són/ y lo que el hablante produjo fue [es to són].

Núm. de ficha	Segmento	Transcripción	Diálogo	Oración	Emisor
4	es_ [t_cto so	estos son los estantes	11	90	s

En la Figura número 21 se muestra la imagen espectrográfica de dicho fenómeno donde se señala el espacio donde se presenta la reducción consonántica y la nueva silabificación que dio origen dicha homologación.

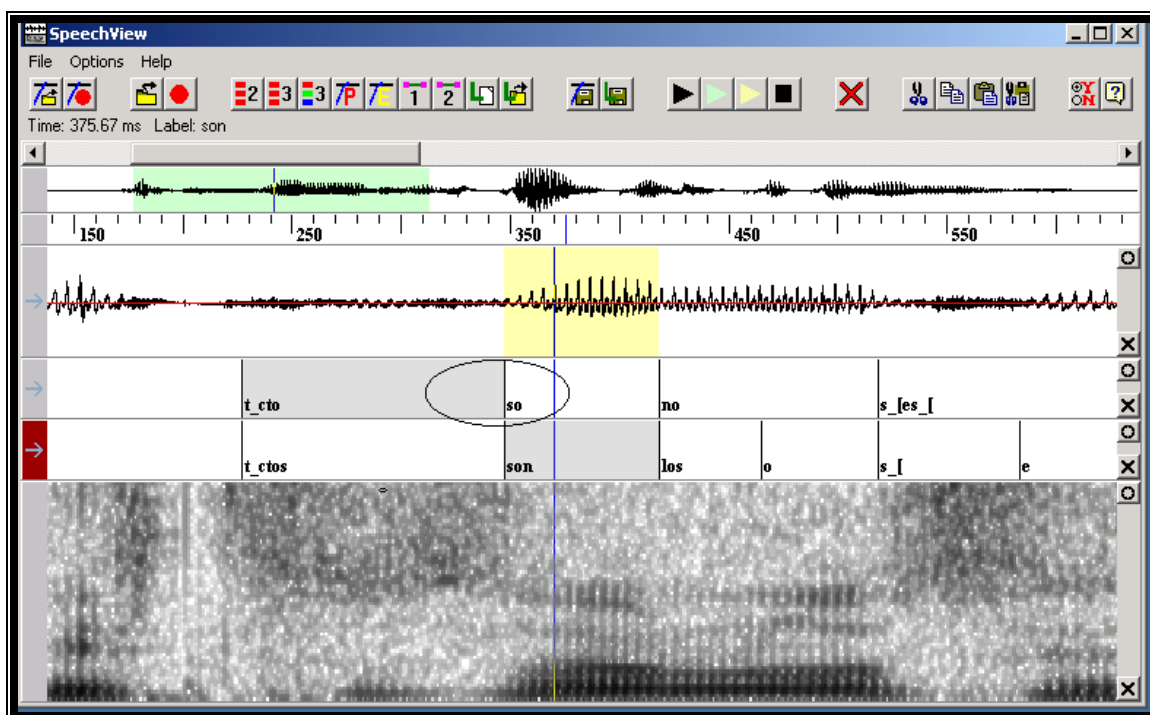


Figura 21. Imagen espectrográfica de homologación consonántica “s s”: *estos son* [és.to.són.] Diálogo 11; oración 90.

4.4. SINALEFA

Es conveniente comenzar este apartado con una cita de Navarro (2004:150) que define perfectamente el concepto de homologación vocálica: “Fonéticamente, dos vocales, cualesquiera que sean, son siempre susceptibles de reducción a una sola sílaba, aunque en determinados casos por razones gramaticales o de otro tipo, deje de cumplirse dicha reducción”. Cuando dos o más vocales están en contacto, éstas se vuelven muy breves y relajadas (excepto la más abierta), por ende, se logra un efecto auditivo en el que dichas vocales forman parte de una misma sílaba. Matluck (1951) define como una característica del valle de México la reducción de dos sílabas a una, sin importar que éstas pertenezcan a dos palabras distintas. Dicho fenómeno es conocido como *sinalefa* y el español tiene una clara tendencia a la producción de la

misma (Quilis, 1979). La *sinalefa* es un fenómeno de silabificación fonético que consiste en la agrupación de vocales en una sola sílaba que pertenecen a la última sílaba de una palabra y la primera vocal de la sílaba de la siguiente palabra, siempre y, cuando no exista pausa entre dichas palabras. Moreno de Alba (1994:51) explica que las vocales en sinalefa ofrecen una múltiple y variada combinación y como la sinalefa depende de la abertura de las vocales, tan sólo se pueden concebir las siguientes formaciones:

- 1) Progresión de abertura de más cerrada a más abierta (*mea le gro*).
- 2) De más abierta a más cerrada (*laes cue la*).
- 3) Presencia de la abertura más grande en el centro (*po coe fec to*).

En el *Corpus DIME* la sinalefa fue un fenómeno recurrente; sin embargo, es necesario advertir que no fueron tan abundantes los casos en los que el contexto ideal para la aparición de la sinalefa se presentara. Esto significa que se registró un alto índice de segmentos en los que la última sílaba de una palabra fuera vocal y la siguiente palabra comenzara en vocal; sin embargo, en una cantidad considerable de casos dichas palabras estaban separadas por algún tipo de ruido o de pausa que volvían improbable la formación de sinalefas. Las categorías que sí formaron sinalefas en el corpus analizado fueron:

- 1) /ae/ (de abierta a cerrada).
- 2) /ea/ (de cerrada a abierta).
- 3) /oa/ (de cerrada a abierta).

Navarro (1908) dice que la pronunciación española se caracteriza por la preferencia a convertir, siempre que sea posible, todo conjunto de vocales en un grupo monosilábico, pero que por distintas circunstancias históricas, analógicas o eruditas se suele oponer en muchos casos dicha tendencia.

El primer caso de sinalefa registró un porcentaje de aparición exitosa (que sí se cumplieran las características contextuales fonéticas) de 12%. En su mayoría, las vocales pertenecen al artículo determinado femenino *la* seguido del sustantivo *estufa* y la preposición *a* seguida del adjetivo demostrativo femenino *esta*. A continuación se muestra una ficha donde ocurre el agrupamiento de dichas vocales en un solo grupo silábico. Para el hablante es mucho más fácil emitir en un solo grupo fónico la mayor cantidad de sonidos; articulatoriamente, la vocal /a/ no está muy distante de la vocal /e/, por lo que el hablante entrelaza las vocales, aun cuando éstas pertenezcan a palabras distintas, sin corte ni separación de la sonoridad. Por ende, pasa suave y gradualmente de un sonido vocálico a otro en la misma emisión silábica.

Ficha	Segmento	Transcripción	Diálogo	Oración	Emisor
1	<i>laes_ [t _ctu _7</i> fa	la estufa	11	10	u

El segmento en cursivas señala la sinalefa que va de un segmento abierto a uno más cerrado en un solo grupo silábico. Para Butragueño (2002) las dos vocales no se pronuncian completamente en el mismo grupo silábico, sino que la más abierta es la que predomina y la /e/ es sometida a un proceso de debilitamiento en el que es absorbida por la abierta.

En la Figura número 22 se presenta la imagen espectrográfica de la sinalefa [laes tú fa]. En esta imagen podemos apreciar cómo los formantes de las vocales en sinalefa no presentan ningún tipo de corte o pausa entre estas; de hecho es difícil apreciar visualmente el final de la /a/ y el inicio de la /e/. Lo anterior está señalado en un óvalo para su apreciación.

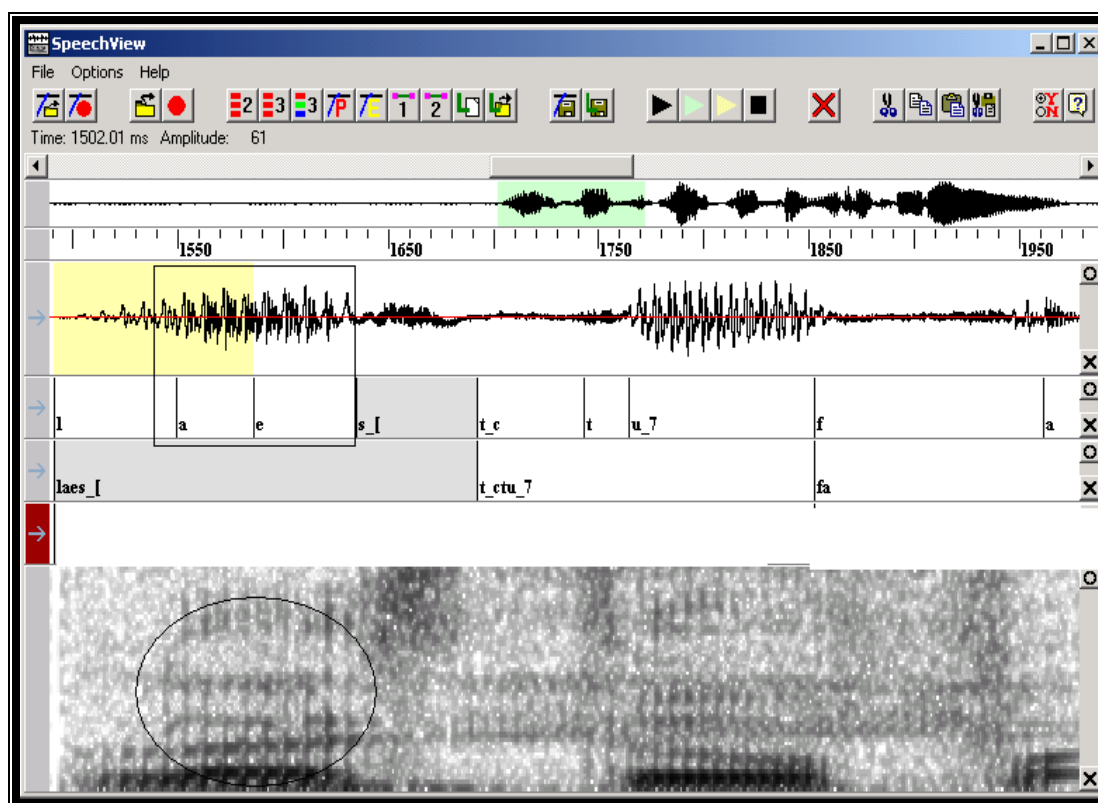


Figura 22.Imagen espectrográfica de sinalefa “a e”: *la estufa* [laes.tú.fa]. Diálogo 11; oración 10.

El segundo caso de *sinalefa* más frecuente en el *Corpus DIME* es el que va del cierre a la abertura máxima /ea/. Cuenta con un índice de aparición de 21%. El 100% de las vocales de las fichas analizadas son átonas, por lo que presentan una clara facilidad de emisión en un solo grupo silábico; es decir, la producción de estas vocales es

continua, sin ningún tipo de corte. A continuación se presenta un ejemplo de ficha de sinalefa del tipo de cerrado a abierto /ea/.

Ficha	Segmento	Transcripción	Diálogo	Oración	Emisor
2	mea sen	me hacen falta	2	134	u

En la Figura número 23 se muestra la agrupación de las vocales en sinalefa. La silabificación tradicional prescribe que dicho segmento debe ser silabeado de la siguiente forma: /me á sen/; sin embargo, lo que el hablante produjo es [meá sen]. En un solo grupo silábico se juntan dos vocales que teóricamente forman hiato; es decir, el hiato se pierde y se forma una sílaba cuyas vocales son fuertes. Autores como Quilis (1979) y Matluck (1951) definen al español como una lengua que tiende a la formación de sinalefa; es decir, que se caracteriza por tener una postura antihiática. Gili Gaya (1975:121) menciona cómo el espectrograma facilita la observación del enlace de las vocales que forman un diptongo. Se puede apreciar cómo las vocales forman una sola sílaba y cómo el entrelazamiento de la diptongación es evidente.

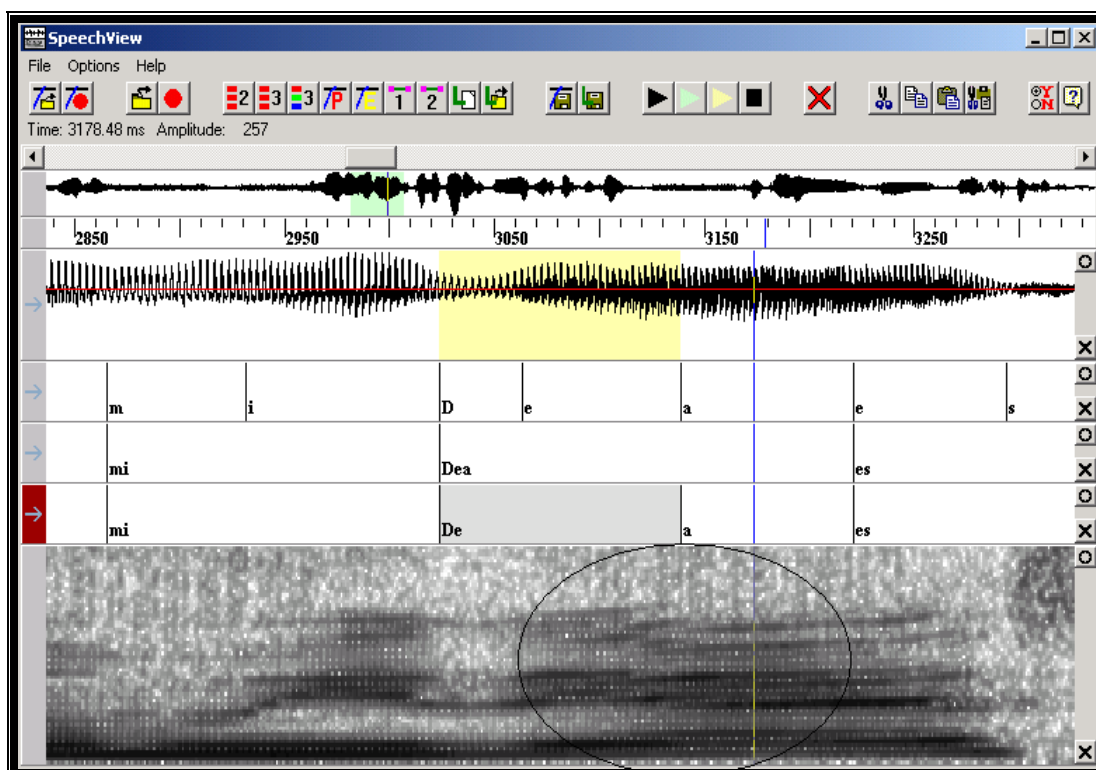


Figura 23. Imagen espectrográfica de sinalefa “e a”: *mi idea* [mi.dea.] Diálogo 2; oración 12.

El último fenómeno analizado de sinalefa es el conformado por la diptongación de las vocales /oa/; es una sinalefa del tipo de cerrada a abierta. En el *Corpus DIME* se encontró un índice de aparición de 30%. El 100% de las vocales de las fichas analizadas que presentaron dicho fenómeno fueron vocales átonas, por lo que su entrelazamiento monosilábico se realizó fácilmente. Gili Gaya (1975:120) menciona que a esta clase de vocales, que en vez de formar sinalefa se agrupan en diptongo en un solo grupo silábico, se les conoce como diptongos no consolidados, a los cuales la gramática tradicional da el nombre de impropios, por no haber en ellos semiconsonante o semivocal. A continuación se presenta una ficha donde el fenómeno de sinalefa se resalta. La gramática tradicional nos indica que la

silabificación de este segmento tiene que ser /ko lo kár lo a kí/; lo que el emisor produjo es [ko lo kar loa kí].

Ficha	Segmento	Transcripción	Diálogo	Oración	Emisor
6	loa k_ck_ji_7	colocar lo aquí	2	177	u

La Figura número 24 muestra la sinalefa que se forma con las vocales fuertes; también se aprecia el entrelazamiento de dichas vocales en la banda ancha del espectrograma.

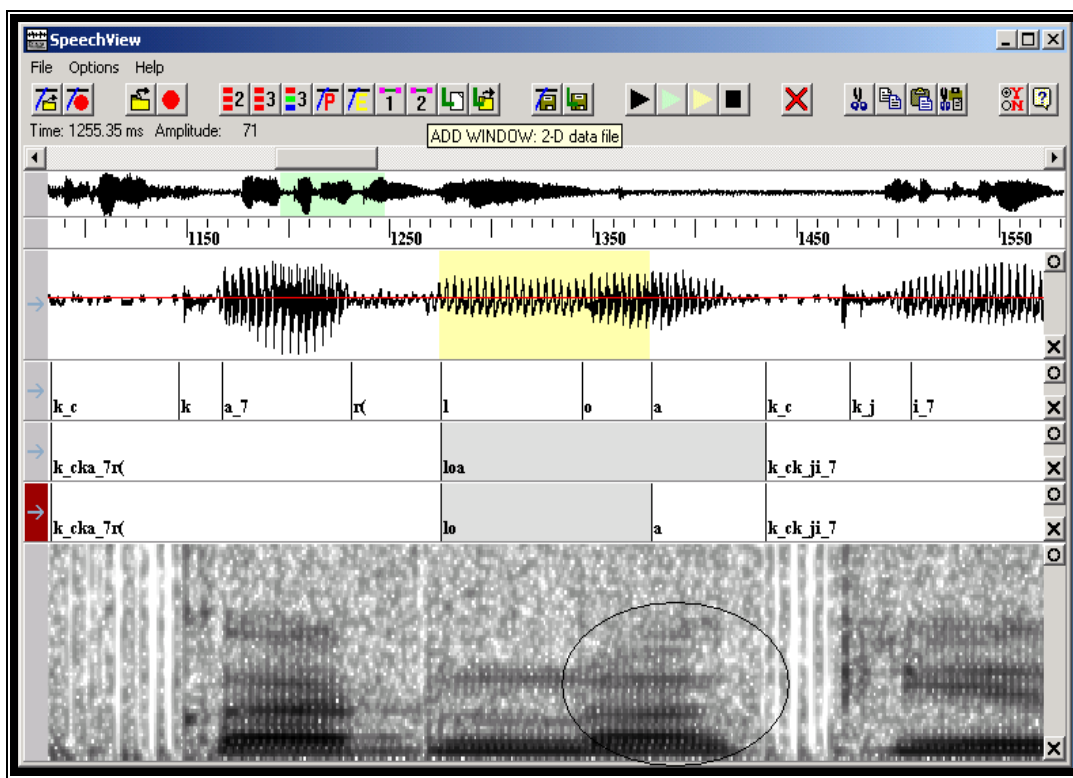


Figura 24. Imagen espectrográfica de sinalefa “o a” de la frase *colocar lo aquí* [ko lo kar loa kí]. Diálogo 2; oración 177.

La Figura número 25 muestra los resultados globales de la ocurrencia total de los fenómenos analizados.

Fenómeno	Núm. total de apariciones exitosas en el Corpus DIME	Núm. de ocurrencias	Porcentaje de ocurrencia
Sílaba abierta: "r e"	55	34	61%
Sílaba abierta: "r a"	13	9	69%
Sílaba abierta "r u"	11	7	63%
Sílaba abierta "r o"	8	5	62%
Sílaba abierta "s a"	42	30	71%
Sílaba abierta "s e"	41	23	56%
Sílaba abierta "n e"	89	61	68%
Homologación vocálica "e e"	169	49	28%
Homologación vocálica "a a"	50	16	32%
Homologación consonántica "s s"	25	25	100%
Sinalefa "a e"	169	21	12%
Sinalefe "e a"	32	12	37%

Figura 25. Tabla de ocurrencia de los fenómenos de silabificación del *Corpus DIME*.

4.5. PROPUESTA DE REGLAS DE SILABICACIÓN FONÉTICAS PARA LAS TECNOLOGÍAS DEL HABLA

Con base en el análisis empírico del corpus y del estudio de las fichas que se clasificaron de acuerdo al fenómeno que presentaron en el *Corpus DIME* se obtuvieron las siguientes reglas:

Tipo de silabificación	Regla	Descripción
Formación de sílaba abierta	$VCV \rightarrow V.CV$	Cuando una C se encuentra entre dos V, la C se agrupa o forma sílaba con la V siguiente (no se respeta límite de palabra).
Reducción de segmentos idénticos vocálicos	$V_{\alpha} \# \# V_{\alpha} \rightarrow V_{\alpha}$	Cuando la última sílaba de una palabra termina en V y la primera sílaba de la siguiente también es la misma V, éstas se reducen a una sola V y la V resultante forma sílaba con la primera palabra.
Reducción de segmentos consonánticos	$C_{\alpha} \# \# C_{\alpha} \rightarrow C_{\alpha}$	Cuando la última sílaba de una palabra termina con C y la primer sílaba de la siguiente también es la misma C, éstas se reducen a una sola C.
Sinalefa	$V_{\alpha} \# \# V_{\beta} \rightarrow \$V_{\alpha}V_{\beta}\$$	Cuando una sílaba termina en V y la siguiente sílaba comienza en otra V, éstas se agrupan para formar una sola sílaba con la primera.

Figura 26. Cuadro de reglas de silabificación fonética para las tecnologías del habla.

4.6. APLICACIÓN DE LAS REGLAS EN LAS TECNOLOGÍAS DE HABLA.

Como se mencionó (§ 1.2.) uno de los objetivos futuros del trabajo es poner en práctica las reglas obtenidas a las tecnologías del habla, por lo que probaron dichos resultados a manera de ejercicios en el sintetizador del Proyecto DIME¹¹. Se introdujeron dos entradas textuales; es decir, dos frases por cada fenómeno de silabificación analizado, la primera se escribió de manera tradicional y la segunda se

¹¹ Es un sintetizador de voz con acento venezolano, el de acento mexicano que se encontraba gratuitamente en la red no contaba con un nivel aceptable de naturalidad en la producción de voz.

modificó de tal manera que ejemplificara los fenómenos de silabificación. El sintetizador reprodujo las frases en orden consecutivo para comparar si el grado de naturalidad aumentaba entre la forma canónica y la forma silabificada.

Con base en dichas prácticas se observó lo siguiente:

1. Con respecto a sílaba abierta:

- a) “r e”. La primer entrada fue *Puedes mover el estante*, la segunda frase se escribió de la siguiente manera *Puedes move relestante*. El sintetizador produjo la palabra *relestante* con vibrante múltiple; las muestras del *Corpus DIME* pronuncian dicha palabra con vibrante simple debido a la continuidad de la conversación; sin embargo las reglas fonológicas con las que cuenta el sintetizador le indica que el fonema /r/ inicial de palabra debe de ser vibrante múltiple. Lo mismo ocurre con la las combinaciones: “r a”, “r e” y “r u”.
- b) “s a”. La primer entrada textual fue *tiene alguna que sea*, la segunda entrada fue *tiene salguna que sea*. La segunda emisión contó con mucha más naturalidad de producción. La pronunciación del sintetizador se vio mejorada.
- c) “s e”. La primer entrada fue *en ese espacio*, la segunda fue *e nese espacio*. Ambas pronunciaciones suenan de la misma manera. No aumento el índice naturalidad pero tampoco se vio modificado.
- d) “n e”. La primer entrada fue *las campanas son estas*, la segunda fue *las campanas so nestas*. En la segunda emisión el sintetizador no pronunció la /s/ del verbo *son*, por lo que se escuchó: *las campanas*

honestas; sin embargo, en esta prueba se agregó una tercera entrada: *las campanas sonestas*. La producción de ésta última fue la que mayor naturalidad presentó, esto debido en gran medida por la desaparición de espacios entre palabras, lo cual permitió una mayor velocidad de elocución.

2. Con respecto a homologación de segmentos idénticos:

- a) “e e”. La primer entrada fue *este estante*, la segunda fue *estestante*.

La segunda emisión superó a la primera en naturalidad de elocución y en rapidez, no se distorsionó la claridad de la emisión. De hecho, esta es la regla que mayor éxito tuvo en las pruebas.

- b) “a a”. La primer entrada fue *esta alacena*, la segunda fue *esta lacena*.

Una vez más la primer elocución fue superada por la segunda en cuanto a naturalidad y rapidez.

- c) “s s”. La pronunciación de ambos enunciados resultó muy similar, no hubo ningún mejoramiento en la segunda, a pesar de ser el fenómeno que mayor éxito tuvo en las grabaciones del *Corpus DIME*.

3. Con respecto de sinalefa

- a) “a e”. La primer entrada fue *aesta esquina*, la segunda fue *aestasquina*. La segunda elocución no fue mejor que la primera; sin embargo la velocidad en la frase escrita con el fenómeno fue emitida con mayor velocidad.

- b) “e a”. No hubo ninguna diferencia entre los enunciados, es decir, el segundo no aportó mayor naturalidad.

c) “o a”.La primer frase fue *fregadero aquí*, la segunda *fregaderoa qui*.

Lo que produjo el sintetizador en la segunda ocasión sonó completamente afectado puesto que acentuó a /a/ y perdió completamente la tonicidad de /i/ correspondiente a *aquí*.

Es importante advertir que los ejercicios anteriores ponen de manifiesto los posibles beneficios que dichas reglas proporcionarían a la síntesis en futuras pruebas con otros sintetizadores.

Con base en la información lingüística obtenida en este trabajo, se pueden explicar las aportaciones de la misma para que el índice de reconocimiento aumente. Una vez que los modelos acústicos hayan transformado la señal sonora en la representación fonética y ésta a su vez corresponda con su respectivo equivalente léxico en el diccionario de pronunciación éste se vería afectado por la silabificación, por ende, tendrían que agregarse muchas mas entradas léxicas que contengan los fenómenos. La silabificación por homologación de segmentos idénticos y la sílaba abierta aportarían una nueva realización a la palabra. Por ejemplo, en el caso del fenómeno de sílaba abierta, a las pronunciaciones canónicas de los verbos *mover*, *llevar* y *poner* se les agregarían las realizaciones donde se omita la consonante final de la palabra; es decir, *move*, *lleva* y *pone*, otro tipo de palabras frecuentes en el corpus que empiezan con vocal como lo son los adjetivos demostrativos *este*, *esta* y *estas*, los artículos determinados e indeterminados *el*, *un*, *una*, los deícticos espaciales *aquí* y *ahí* y, por su puesto, los sustantivos propios del dominio del corpus *alacena*, *estante* y *extractor*, se verían modificadas en el ataque de su primer sílaba, es decir,

ganarían las consonantes que dichos verbos cederían por la tendencia de la sílaba libre o abierta del español. Las reglas de silabificación proporcionarían al diccionario de pronunciación el reconocimiento de las palabras *tene salguna salacenas*, las cuales corresponden a la frase *¿Tienes algunas alacenas?* La figura número 27 muestra una tabla de dicho ejemplo.

Entrada canónica	Entrada modificada por el fenómeno de silabificación
<i>Tines</i>	<i>tiene</i>
<i>Algunas</i>	<i>salguna</i>
<i>Alacenas</i>	<i>salacenas</i>

Figura 27. Tabla que muestran las tres nuevas entradas de las palabras *tienes*, *algunas* y *alacenas-tiene salguna salacenas*.¹²

Cabe mencionar el impacto que estas nuevas entradas proporcionarían a los reconocedores, ya que no se tiene el conocimiento de que este tipo de información aplicada actualmente en el reconocimiento.

Durante el análisis del corpus se encontró que el fenómeno de sílaba abierta es un rasgo característico de la silabificación; es decir, fue un aspecto recurrente de los hablantes y del español. La Figura número 28 muestra los porcentajes obtenidos de sílaba abierta entre dos estudios de lengua oral y uno de lengua escrita. Los resultados del *Corpus DIME* son consistentes con los que presente Guerra en 1983.

¹² Las cuales pueden ser agregadas al diccionario de pronunciación debido a la aplicación de las reglas obtenidas en este estudio.

Navarro (1918)	Guerra (1983)	<i>Corpus DIME</i> (2007)
<i>lengua escrita</i>	<i>lengua hablada</i>	<i>lengua hablada</i>
58%	68.05%	67%

Figura 28. Cuadro comparativo de porcentaje de sílaba abierta.

5. CONCLUSIONES

El propósito de este trabajo de tesis es sistematizar el fenómeno de silabificación fonética en el *Corpus DIME* como una muestra del habla del español de México para poder crear un conjunto de reglas fonéticas que pueden ser aplicadas en las tecnologías del habla para el mejoramiento de las mismas. Fue necesario partir de las evidentes diferencias existentes entre la manera convencional de formar sílaba y la que realmente produce cualquier hablante; es decir, tomando el habla como un continuo que agrupa unidades. En primera instancia, se necesitaba tener un panorama teórico y amplio de la sílaba; estructura, comportamiento, características esenciales y, por su puesto, la definición de la misma en español. Una vez sentada dicha base se presentaron los resultados del análisis, los cuales permitieron la sistematización de tres fenómenos:

- a) Sílaba abierta.
- b) Reducción de segmentos idénticos vocálicos y consonánticos.
- c) Sinalefa.

Con la finalidad de entender el impacto que dichas reglas tendrían en las tecnologías del habla fue de primordial importancia saber qué son las tecnologías, así como también tener claro en qué pueden auxiliar para su mejoramiento las reglas obtenidas, por lo que el primer capítulo de esta tesis consistió en la explicación y concepción de las tecnologías del habla y el papel que juega la lingüística en las mismas.

El segundo capítulo está dedicado a la explicación estructural de la sílaba, a los distintos enfoques que a lo largo de la historia de la lingüística se ha atendido para su explicación y, finalmente, a la estructura convencional que tiene el español para formar sílabas. La importancia de esta sección radica en que es la base teórica lingüística de este trabajo de tesis; es decir, una vez que se atendió a la manera convencional de esta unidad se podían contrastar con ésta los resultados obtenidos empíricamente y poder analizar y corroborar las diferencias existentes entre la silabificación tradicional y la oral.

Finalmente, el tercer capítulo plasma el análisis empírico fonético de la muestra de habla espontánea tomada del Corpus DIME; en el estudio de cinco diálogos se pudo encontrar un comportamiento peculiar en la manera de formar sílabas con respecto de la convencional y fue posible sistematizarlas.

- a) Con respecto de sílaba abierta se puede concluir que efectivamente sí es una pauta pertinente preferencial del español de México, no nada más porque es una tendencia primaria universal para formar sílabas en todas las lenguas, sino por el comportamiento general del español; es decir, el español es una lengua que fácilmente entrelaza unidades en distintos niveles como el silábico a nivel de frase y el fonosintáctico: sirremas. La formación de sílaba abierta con las consonantes analizadas seguidas de vocal comprueba que la tendencia a dejar a la sílaba como no trabada tiene que ver con la tonicidad de la vocal precedente y con la comodidad articulatoria que presupone la emisión de dicha consonante como ataque de sílaba. Vimos que todos los contextos presentan más del 50% de

ocurrencia del fenómeno; los casos en los que no ocurrió el fenómeno están determinados por circunstancias diversas, principalmente por silencios que separan las cadenas fónicas.

- b) En cuanto a la homologación, se concluye que la mayoría de las veces en que dos segmentos idénticos están en contacto se pronuncian como uno solo, sin características especiales; es decir, no ocurren sonorizaciones, alargamientos ni ningún otro tipo de compensación. El caso con mayor porcentaje de reducción es el de dos /s/, que como vimos, se dio en 100% de las ocasiones.
- c) Por lo que respecta a sinalefa, se concluye que el entrelazamiento de vocales en un solo grupo silábico es una constante del español del altiplano; pues como vimos en los resultados los contextos /ea, oa, ae/ muestran que estos segmentos tienden a agruparse en sinalefa. Sin embargo, sólo los últimos dos muestran además de sistematización, una frecuencia mayor de 30%, por lo que serán éstos los que son considerados para la formulación de reglas.

Como conclusión global, se puede decir que la tonicidad de las vocales es un factor determinante en la aparición de estos fenómenos; de la misma forma en que la velocidad de elocución propicia la mayoría de los fenómenos.

6. REFERENCIAS BIBLIOGRÁFICAS

- ALARCOS LORACH, EMILIO. 1950/1991. *Fonología española*, Madrid: Gredos.
- BYBEE, JOAN. 2001. *Phonology and Language use*, Cambridge: Cambridge University Press.
- BERNAL BERMÚDEZ, JESÚS, BOBADILLA SANCHO, JESÚS Y PEDRO GÓMEZ VILDA. 2000. *Reconocimiento de voz y fonética acústica*, Madrid: RA- Ma.
- BURQUEST, DONALD. 1993. "Natural Phonological Processes: conditionain by syllable structure" en *Phonological Analysis. A Funtional approach*, Summer Institute of Linguistics.
- BUTRAGUEÑO, PEDRO. 2002. *Variación lingüística y teoría fonológica*, México: El Colegio de México.
- CANELLADA, MARÍA JOSEFA y JOHN KUHLMANN MADSEN. 1987. *Pronunciación del español. Lengua hablada y literaria*, Madrid: Castalia.
- CUÉTARA PRIEDE, JAVIER. 2004 (Ms.). *Fonética y fonología del habla espontánea de la ciudad de México. Su aplicación e las tecnologías del habla*, tesis de maestría inédita, México: UNAM.
- COMPANY COMPANY, CONCEPCIÓN y MARCELA FLORES CERVANTES. 2001. *Manual de Filología Hispánica*. México: UNAM.
- D'INTORNO, FRANCESCO. 1995. *Fonética y fonología actual del español*, Madrid: Cátedra.
- GIL FERNÁNDEZ, JUANA. 1999. *Los sonidos del lenguaje. Textos de apoyo*, Madrid: Síntesis.
- GILI GAYA, SAMUEL. 1975. *Elementos de fonética general*, Madrid: Gredos.
- GRANADA, G. 1966. *La estructura silábica*, Madrid: C.S.I.C.
- GUERRA, RAFAEL. 1983. "Estudio estadístico de la sílaba en español" en *Estudios de fonética*.
- HALA, BOHUSLAV. 1996. *La sílaba. Su naturaleza, su origen y sus transformaciones*, Madrid: C.S.I.C.
- HERRERA Z, ESTHER. 2001. *Temas de fonética instrumental*, México: El Colegio de México.
- HIDALGO, ANTONIO y MERCEDES QUILIS. 2004. *Fonética y fonología españolas*, Valencia: Tirant lo blanch.
- JAKOBSON, ROMAN. 1979/1987. *La forma sonora de la lengua*, México: Fondo de Cultura Económica.
- LLISTERRI BOIX, JOAQUIM, 1991. *Introducción a la fonética: El método experimental* Barcelona: Anthropos.
- _____. 2003. "Las tecnologías del habla", en *Tecnologías del lenguaje*, M. A. Martí (Coord.), Barcelona: UOC, pp. 249-281.
- _____. 2003-A "Lingüística y tecnologías del lenguaje", *Lynx. Panorámica de Estudios Lingüísticos* (Departament de Teoria dels Llenguatges, Universitat de València) 2: 9-71.
- LLISTERRI BOIX, JOAQUIM, LOURDES AGUILAR, JUAN M. GARRIDO, MARÍA JESÚS MACHUCA, RAFAEL MARÍN, CARME DE LA MOTA y ANTONIO RÍOS. 1999. "Fonética y tecnologías del habla", en *Filología e Informática. Nuevas*

- tecnologías en los estudios filológicos*, J.M. Blecua, G. Clavería, C. Sánchez y J. Torruella (Eds.), Barcelona: Milenio i Universitat Autònoma de Barcelona, pp. 449-479.
- LLISTERRI BOIX, JOAQUIM y MARÍA ANTONIA MARTÍ. 2002. "Las tecnologías lingüísticas en la sociedad de la información", en *Tratamiento del lenguaje natural*, M. A. Martí y J. Llisterri (Eds.), Barcelona: Fundación Duques de Soria y Edicions Universitat de Barcelona, pp. 13-28.
- LLISTERRI BOIX, JOAQUIM, CARMEN CARBÓ, MARÍA JESÚS MACHUCA, CARMEN DE LA MOTA, MONTSERRAT RIERA y ANTONIO RÍOS. 2003 (Ms.). "El papel de la lingüística en el desarrollo de las tecnologías del habla", en *Séptimas Jornadas de Lingüística*, Cádiz: Servicio de Publicaciones de la Universidad de Cádiz.
- LASS, ROGER. 1984/1985. *Phonology. An introduction to basic concepts*, Cambridge: Cambridge University Press.
- MACHUCA, RAFAEL MARÍN, CARMEN DE LA MOTA y ANTONIO RÍOS. 1999. "Fonética y tecnologías del habla", en *Filología e Informática. Nuevas tecnologías en los estudios filológicos*, J. M. Blecua, G. Clavería, C. Sánchez y J. Torruella (Eds.), Barcelona: Milenio i Universitat Atonóma de Barcelona.
- MALBERG, BERTIL. 1954/1981. *La fonética*, Buenos Aires: Editorial Universitaria de Buenos Aires.
- MARTÍNEZ CELDRÁN, EUGENIO. 1984. *Fonética, (con especial referencia a la lengua castellana)*, Barcelona: Teide.
- _____. 1998. *Análisis espectrográfico de los sonidos del habla*, Barcelona: Ariel.
- MATLUCK, JOSEPH. 1951. *La pronunciación en el español del Valle de México*, México: Edición de autor.
- MORENO DE ALBA, JOSÉ G. 1994. *La pronunciación del español en México*, México: El Colegio de México.
- NAVARRO TOMÁS, TOMÁS. 1918/1982. *Manual de pronunciación española*, Madrid: RAYCAR.
- _____. 1944/1966. *Manual de entonación española*, México: Colección Málaga.
- NÚÑEZ CEDEÑO, RAFAEL. 2000. "Teoría de la organización silábica" en *Panorama de la fonología actual*, Madrid: Arco libros.
- PÉREZ PAVÓN, ELIA PATRICIA. 2006. *Construcción de un reconocedor de voz utilizando Sphinx y el Corpus DIME*, tesis de Licenciatura inédita, México: UNAM
- PERISSINOTTO, GIORGIO SABINO ANTONIO. 1975. *Fonología del español hablado en la Ciudad de México. Ensayo de un método sociolingüístico*, México: El Colegio de México.
- PINEDA, LUIS A., ANTONIO MASSÉ, IVÁN MEZA, MIGUEL SALAS, ERIK SCHWARZ, ESMERALDA URAGA y LUIS VILLASEÑOR. 2001 (Ms.). "El Proyecto DIME", México: Universidad Nacional Autónoma de México.
- QUILIS, ANTONIO. 1981/1988. *Fonética acústica de la lengua española*, Madrid: Gredos.
- _____. 1993/1999. *Tratado de fonología y fonética españolas*, Madrid: Gredos.

- SOLÉ SABATER. 1985. "La experimentación en fonética y fonología", en *Estudios de fonética experimental*, E. Martínez Celdrán y M. J. Solé Sabater (Eds.), Barcelona: Promociones Publicaciones Universitarias.
- SAPORTA Y CONTRERAS, H. (1962). *A Phonological Grammar of Spanish* Seattle, University of Washington Press.
- SAUSSURE, FERNANDINE DE. *Curso de Lingüística general*, México: Nuevo Mar
- TAPIAS MERINO, DANIEL. 2002. "Interfaces de voz con lenguaje natural", en *Tratamiento del lenguaje natural*, María Antonia Martí, Joaquim Llisterri Boix (Eds.), Barcelona: Edicions Universitat de Barcelona
- VILLASEÑOR, LUIS, ANTONIO MASSÉ y LUIS A. PINEDA. 2001 (Ms.). "The DIME Corpus", en *Tercer Encuentro Internacional de Ciencias de la Computación ENC-01*, Aguascalientes: SMCC-INEGI, pp. 591-600.

7. Apéndice 1

1. TABLA DE SÍLABA ABIERTA (r e)

Fenómeno: Sílabas abiertas “r e”

Regla: VCV → V.CV

Fichas	Segmento	Transcripción	Núm. diálogo	Oración	emisor
1	mo Ve_7 r(e les	puedes mover el estante	11	107	u
2	mo Ve r(el	mover el	18	152	u
3	mo Ve_7 r(es_ [t_ cte	mover este	18	184	u
5	p_ cpo ne r(e les	puedes poner el estante	11	129	u
6	p_ cpo ne r(es_ [t_ cte	poner este fregadero	11	198	u
7	p_ cpo ne r(el fr(e Ga De_7	poner el fregadero	11	71	u
8	ne_7 r(e la Va	poner el lavatrastos	2	111	u
9	Ze Va_7 r(es_ [	puedes llevar este otro	11	163	u
10	Ze Va_7 r(es_ [t_ cte fr(e	llevar este fregadero	11	57	u
11	Ze Va_7 r(es_ [t_ cta	llevar esta	11	20	u
12	Ze Va_7 r(es_ [t_ cte	llevar este	11	215	u
13	Ze Va_7 r(es_ [t_ cte fr(e	llevar este fregadero	11	217	u
14	Ze Va_7 r(es_ [t_ cta	llevar ésta	11	232	u
15	Ze Va_7 r(es_ [t_ cte	llevar este	11	258	u
16	d_ cdes li sa r(el	deslizar el lavatrastes	11	187	u
17	t_ ctr(a r(el	mostrar el catálogo	18	3	u
18	mos_ [t_ ctr(a r(el me un_7	mostrar el menú	11	203	u
19	se li mi na r(es t_ cte	eliminar este fregadero	11	209	u
20	k_ cka_7 r(es_ [	colocar esta	2	12	u
21	k_ cka_7 r(es_ [	colocar este	18	9	u
22	k_ ck_7 r(es	colocar este	18	12	u
23	k_ cka r(es_ [[	colocar esta	2	27	u
24	sen_ [t_ ctr(a r(es_ [t_ cte	centrar este el fregadero	11	227	u
25	xi r(a r(el k_ ckwar(t_ cto	girar el cuarto	11	25	u
26	xi r(a r(el k_ ckwar(t_ cto	girar el cuarto	11	273	u
27	d_ cde k_ ck_ je k_ cko lo r(es	de qué color es	7	75	u
28	d_ cde Ge k_ cko lo r(es	de qué color es	7	78	u
29	De le xi_7 r(el fr(e Ga de_7 r(o	favor de elegir el fregadero	20	36	s

30	p_cpo r(el mo	por el momento	2	150	u
31	p_cpo r(es_[t_cte	por este fregadero	2	84	s
32	sjo na r(el	seleccionar el	18	111	u
33	Vi sa r(el	revisar el	18	206	u
34	k_cko Da_7 r(es_[	acomodar estos	18	87	s

2. TABLA DE SÍLABA ABIERTA (r a)

Fenómeno: Sílabas abiertas “r a”

Regla: VCV → V.CV

Fichas	Segmento	Transcripción	Diálogo	Oración	Emisor
1	p_cpe_7 Ga r(a les	puedes pegar a la esquina	11	155	u
2	Vol Ve_7 r(a mos_[	volver a mostrar	11	203	u
3	mo Ve r(a seal fr(en_[	mover hacia el frente	11	63	u
4	mo Ve r(a sja	mover hacia acá	11	66	u
5	ne r(a k_ck_jj	poner aquí	2	42	u
6	Ze Va_7 r(a	llevar hacia	18	21	u
7	re Gr(e sa_7 r(a el	regresar al cuarto	11	238	u
8	fri Go Va r(a k_ck_ji	frigover aquí	11	198	u
9	es_[t_ctr(ak_ck t_cto r(aj	extractor ahí	11	250	u

3. TABLA DE SÍLABA ABIERTA (r u)

Fenómeno: Sílabas abiertas “r u”

Regla: VCV → V.CV

Fichas	Segmento	Transcripción	Diálogo	Oración	Emisor
1	r(a ro_7 t_cta r(u_7n	rotar un poco	2	10	u
2	De_7 mo a se r(u na	podemos hacer una rotación	20	54	u
3	a se_7 r(u	hacer un éste	2	110	u
4	Ve_r(un	mover un poco	2	36	u
5	k_cka r(u na	colocar una	2	48	u
6	k_cka_7 r(un fr(e	colocar un fregadero	2	64	u
7	k_ck_je se r(u no	que ser uno	18	105	s

4. TABLA DE SÍLABA ABIERTA (r o)

Fenómeno: Sílabas abiertas “r o”

Regla: VCV → V.CV

Fichas	Segmento	Transcripción	Diálogo	Oración	Emisor
1	re fri xe r(a Do_7 r(oa_2l	refrigerador o algo así	11	194	u
2	b_cba xa_7 r(o	puedes bajar o	11	108	u
3	xi r(a_7 r(o t_ctr(a Ves	girar otra vez	11	40	u
4	p_cpo De_7 moz Ve_7 r(o t_ctr(os	podemos ver otros muebles	20	143	s
5	p_cpo ne_7 r(o	poner otro	18	98	s

5. TABLA DE SÍLABA ABIERTA (s a)

Fenómeno: Sílabas abiertas “s a”

Regla: VCV → V.CV

Fichas	Segmento	Transcripción	Diálogo	Oración	Emisor
1	t_cti ne sa_2l Gu_7	tienes algún	11	143	u
2	tje ne sa_2l Gu_7 na	tienes alguna alacena	11	149	u
3	t_cti ne sa_2l Gu_7nak_ck	tienes alguna que sea	11	15	u
4	t_cti ne sa_2l Gu na	tienes alguna	11	161	u
5	p_po_7 Go ma_7 sas	un poco más a esta esquina	11	187	u
6	t_ctje ne sa_2l Gu_7n	tienes algún menú	11	194	u
7	t_ctje ne sa_2l Gu_7n	tienes algún extractor	11	244	u
8	t_cti ne sa_2l Gu_7 na su xe r(en	tienes alguna sugerencia	7	69	s
9	k_ck_je t_cte ne mo sa ri Va	que tenemos arriba	7	78	u
10	k_cko lo k_cka_7 mo sa_2l se_7n [	y la colocamos al centro	20	120	u
11	la mo Ve_7 mo sa .sil a e_7s [	la movemos a este lugar	20	137	u
12	Ve_7r(la sa la se_7 nas	deseo ver las alacenas	20	145	u
13	moz Ve_7r(la sa la se_7 nas me	podemos ver las alacenas mejor	20	49	u
14	la sa la se_7 nas	las alacenas	20	50	u
15	t_cte ne mo sa la se nas	tememos alacenas	20	53	s
16	la sa la se_7 nas	son todas las alacenas	20	68	u
17	u no sar(ma_7	unos armarios	2	158	s
18	b_cba_7 mo sa	vamos a ver	2	110	u
19	mo sa k_ck_ji	pusimos aquí	2	191	u

20	ne sa k_cka_7	tienes acá	2	22	s
21	ne sa k_ck_ji_7	tienes aquí	2	18	u
22	t_cte la sa la se_7	las alacenas	2	49	u
23	ma_7 sa k_ckja Va_7	una más aquí	2	7	s
24	la sa la se_7 nas	las alacenas	2	51	s
25	Ga_7 Do sa	pegados hacia	18	175	u
26	t_ctr(e la sa la	las alacenas	18	202	s
27	la sa la se_7 nas	las alacenas	18	204	s
28	na sa r(o	alacenas a los lados	18	60	s
29	ma_7 san tS_c tSa	más ancha	18	63	s
30	t_cti ne sa_2l Gu_7	tienes algún	11	161	u

6. TABLA DE SÍLABA ABIERTA (s e)

Fenómeno: Sílabas abiertas “s e”

Regla: VCV → V.CV

Ficha	Segmento	Transcripción	Diálogo	Oración	Emisor
1	k_ckwa le sel ma_7s	cuál es el más	11	206	u
2	p_pwe_7 se li mi na_7 r(es	puedes eliminar este fregadero	11	209	u
3	t_ctoz Do_7 ses [t_ctan [t_ctes	entre estos dos estantes	11	258	u
4	Gr(a Do sen la mi_7s mae	noventa grados en la misma	11	48	U
5	k_cko lo k_cka_7 mo sen si_7 ma	y lo colocamos encima de la estufa	20	108	u
6	e IE xi_7 mo se les [t_ctr(ak_ck	elegimos el extractor	20	105	u
7	mo Ve_7 mo ses [t_cta	movemos ésta la alacena que está	20	136	u
8	es [t_cta ses [t_ctu_7 fas	contamos con estas estufas	20	45	s
9	k_cka mo se ne_7s	colocamos en	20	83	u
10	e_7 se su_7N	ese es un gabinete	20	88	s
11	ne_7 ses	en ese espacio	2	109	u
12	t_cto se_7n	lavatrastos en	2	113	u
13	ma_7 sel	nada más el	2	183	u
14	e ses [t_cta	es esta	2	23	s
15	e ses [t_cta	es esta	2	25	s
16	.sil la ses [t_ctu	las estufas	20	4	s
17	t_ctr(a_7 sel	muestras el	18	5	u
18	e se_7s {t_cte	es éste	18	7	s
19	mwe Vle sen	muebles encima	18	28	s
20	la ses [t_ctu_7 fas	las estufas	18	60	s
21	ne sen	tienes en	18	87	s
22	noe sel	no es el	18	94	u

23	t_ctr(a_7 mo se_7 se	encontramos ése	18	98	s
----	----------------------	-----------------	----	----	---

7. TABLA DE SÍLABA ABIERTA (n e)

Fenómeno: Sílabas abiertas “n e”

Regla: VCV → V.CV

Fichas	Segmento	Transcripción	Diálogo	Oración	Emisor
1	e nes_7s_ [t_ cta	en esta	11	130	s
2	e nes_ [t_ ctes k_ ck_ ji_ na	en esta esquina	11	132	s
3	e nes_ [t_ cta p_ cpa r(e_7	en esta pared	11	138	u
4	u nes_ [t_ ctr(ak_ ck t_ to_7	tienes algún extractor o algún menú	11	161	u
5	a k_ ck_ ji_7 e nes_ [t_ ctes	aquí en esta esquina	11	163	u
6	.sil e nes_ [t_ ctes k_ ck_ ji_ na	en esta esquina	11	179	s
7	r(a k_ ck_ jje nes_ [t_ cta	aquí en esta pared	11	198	u
8	Dea k_ ck_ je nes_ [t_ cta_ cpa r(e_7	quede aquí en esta pared	11	92	u
9	e nes_ [t_ cta p_ cpa r(e_ D	en esta pared pegado a la esquina	7	37	u
10	e su nes_ [t_ ctak_ ck t_ ctor(	este es un extractor	7	42	u
11	e nes_ [t_ cte reN g_ cglo	en este reglón	7	81	u
12	so ne_7s_ [t_ ctas	las campanas son estas	20	104	s
13	k_ cko lo k_ cka mos e ne_7s_ [t_ cte	y la colocamos en este sitio	20	156	u
14	e nes_ [t_ cte si t_ ctjo	en este sitio	20	161	u
15	k_ cko nes_ [t_ cte	contamos con este frigobar	20	17	s
16	e nes_ [t_ ctes k_ ck_ ji_7 na	lo ponemos en esta esquina aquí	20	20	u
17	e nes_ [t_ ctes k_ ck_ ji_7 na	en esta esquina	11	57	u
18	k_ cko nes_ [t_ ctos.sil k_ ckwa t_ ct	contamos con estos cuatro	20	35	s
19	k_ ck_ je De_7 nen	queden en frente	20	57	u
20	e ne_7s_ [t_ cte	en esta pared	20	76	u
21	e nes_ [ta	en esta	20	87	u
22	e ne_7s_ [t_ cte	en esta esquina	20	9	u

23	e ne_7 se	en ese	2	109	u
24	e ne_7	en ese espacio	2	109	u
25	k_cke De ne	queden en	2	117	u
26	e nes_[]	en este	2	12	
27	e nes_[]	en este	2	135	u
28	e ne_7s_[]	en este	2	135	u
29	e nes	en este	2	139	u
30	e nes_[]	en esta	2	139	u
31	e nes_[]	en esta	2	153	u
32	e nes_[] t_cta	en esta	2	153	u
33	e nes_[] t_cta	en esta	2	153	u
34	e nes_[]	en este	2	161	s
35	e nes_[] t_cta	en esta	2	27	u
36	e nes_[] t_ctas	en esta esquina	2	27	u
37	w nes_[] t_ctr(ak_ck t_ctor(	un extractor	2	31	u
38	e nes_[]	en este espacio	2	48	u
39	e nes_[]	en este espacio			
40	e nes_[]	en esta	2	64	u
41	e nes_[]	en esta	2	64	u
42	k_cko nes_[]	con esta	2	74	u
43	e nes_[]	en estas	2	69	s
44	pae ne ses	quepa en ese	2	81	u
45	k_cko nel	con el	18	103	s
46	p_cpo_7 nel k_ck_je	pon el que	18	107	s
47	e ne la_7	en el lado	18	116	u
48	k_cko nel	con el	18	117	s
49	t_ctae nel	está en el	18	137	u
50	k_cko nel	con el	18	146	s
51	enes_[] t_cta	en esta	18	184	u
52	u nes_[] t_ctr(aor(	un extractor	18	195	u
53	e nel	en el	18	35	u
54	e nel	en el	18	35	u
55	e nel	en el	18	43	u
56	e nel	en el	18	75	u
57	e nel	en el	18	75	u
58	so_7 nes_[]	son éstos	18	96	s
59	k_cko nel	con el	18	97	u
60	e nel	en el	11	4	u
61	a_2l Gu_7 nes_[]	tienes algún este menú este de	11	143	u

		cajones			
--	--	---------	--	--	--

8. TABLA DE HOMOLOGACIÓN VOCÁLICA (e e)

Fenómeno: Homologación vocálica “e e”

Regla: $V_{\alpha} \# \# V_{\alpha} \rightarrow V_{\alpha}$

Fichas	Segmento	Transcripción	Diálogo	Oración	Emisor
1	d_cdes_[t_cte a la p_c	de este a la pared	11	119	u
2	Go d_cdes_[t_ctu_7 fas	de estufas	11	11	u
3	los_[d_cdes_[t_ctu_7 fas	de estufas	11	17	s
4	Des_[t_cte es	de este	11	23	u
5	Des_[t_ctr(aG	de extractores	11	246	u
6	Des_[t_cte la_7 Do	de este lado	11	263	u
7	r(e Des_[t_cte lao	de este lado	11	119	u
8	r(e_7 Des_[t_cte	de este lado	11	263	u
10	t_ctoe_7s_[t_ctes_[	este estante	18	207	u
11	es_[t_ctes_[t_ctan_[t_cte	este estante	11	132	s
12	es_[t_ctes_[t_cta_7n_[	este estante	11	172	u
13	es_[t_ctes_[t_cta_7n_[t_cte	este estante	11	173	s
14	es_[t_ctes_[t_cta_7n_[	este estante	11	179	s
15	es_[t_ctes_[t_cta_7n_[	este estante	11	185	s
16	t_ctes t_cta_7n_[	este estante	11	92	u
17	t_ctes_[t_can_[	este estante	11	99	s
18	t_cte sun	este es un	18	83	s
19	t_ctes_[t_ctaG	este extractor	11	252	u
20	t_ctes_[p_cpa sjo	este espacio	7	12	u
21	t_ctes_[t_cte	este este	2	161	u
22	ro_7 t_ctes_[t_cte	rote este	18	165	s
23	t_cten_[t_cttjen_[	no te entiendo	18	41	s
24	sis_[t_cte na	consiste en	18	87	s
25	t_ctr(es_[t_ctos	entre estos	11	250	u
26	t_ctr(el	entre el	2	152	u
27	en_[t_cr(el	entre el	11	62	u

28	d_cdes_[t_cte	grande este	2	161	u
29	xi_7r(el k_ckwar t_cto	gire el cuarto	11	274	s
30	so_7Vr(es [t_cta	sobre esta	11	217	u
31	e se suN	ese es un gabinete	7	60	u
32	k_ck_k_jes [t_cto_7o	que es todo	20	177	u
33	k_ck_jel k_c_je	que el que	7	22	u
34	k_ck_jes la Va ra	que es la barra	2	135	u
35	k_ck_jes la	que es la estufa	2	159	u
36	k_ck_jes [t_cte_7	que este	2	42	u
37	k_ck_jes [t_ctaj	que está ahí	18	103	s
38	k_ck_jes [t_cta_7	que está	18	164	u
39	k_ck_jes [t_cta_7	que está	18	35	u
40	k_ck_jes [t_cta_7	que está	18	37	u
41	k_ck_jes [t_cta_7 Va	que estaba	18	39	u
42	k_ck_jes `t_cta_7	que esá	18	45	u
43	k_ck_jes [t_ctoj	que estoy	2	168	u
44	mes k_cku tS_c tSa	me escuchas	2	4	s
45	mwe Vle ses [t_cte	mueble es este	18	189	u
46	Ve ne ses p_pa	cabe en ese espacio	2	79	s
47	Ve ne ses p_pa	cabe en ese	18	122	s
48	k_cka_7 Ve ne ses	cabe en ese	2	79	s

9. TABLA DE HOMOLOGACIÓN VOCÁLICA (a a)

Fenómeno: Homologación vocálica “a a”

Regla: $V_{\alpha} \# \# V_{\alpha} \rightarrow V_{\alpha}$

Fichas	Segmento	Transcripción	Diálogo	Oración	Emisor
1	t_cta la se_7 na	esta alacena	20	137	u
2	es [t_cta la se_7 na	esta alacena	20	83	u
3	e_7s [t_cta la se_7 na	esta alacena	20	76	u
4	es [t_cta la se_7 na	esta alacena	7	37	u
5	es [t_cta la se_7 na	esta alacena	2	113	u
6	mwe Va_2l se_7_7n [	mueva al centro	2	44	s
7	mu Va_2l Gu_7 no	mueva alguno	18	123	s
8	w na la se_7 na	una alacena	18	47	s
9	w na la se_7 na	una alacena	18	49	u
10	su na la se_7 na	una alacena	18	84	s
11	su na la se_7 na	una alacena	18	85	s
12	t_ctao t_ctr(a la se_7 na	otra alacena	11	163	u
13	la la se_7 na	la alacena	20	136	u
14	a k_cka Va_7 Go	acá abajo	2	124	u

15	p_cpoN Ga_2l Gu_7	ponga algún	18	61	s
16	t_ctr(a_7e Ga_2l g_cgün	traiga algún	20	1	s

10. TABLA DE HOMOLOGACIÓN CONSONÁNTICA (s s)

Fenómeno: Homologación consonántica “s s”

Regla: $C_\alpha \# \# C_\alpha \rightarrow C_\alpha$

Fichas	Segmento	Transcripción	Diálogo	Oración	Emisor
1	es_ [t_ cta son	estas son las alacenas	11	147	s
2	es_ [t_ cta son	estas son las alacenas	2	51	s
3	es_ [t_ cta son	éstas son las alacenas	18	204	s
4	es_ [t_ cto so	estos son los estantes	11	90	s
5	es_ [t_ cta son	estas son las sillas	20	154	s
6	es_ [t_ cta son	estas son sillas	2	145	s
7	es_ [t_ cta son	estas son las estufas	20	4	s
8	es_ [t_ cta son	estas son las estufas	2	7	s
9	t_ cta son	éstas son las estufas	18	60	s
10	es_ [t_ cto son	estos son cuatro	7	18	s
11	es_ [t_ cto zon	éstos son los cuatro	18	196	s
12	es_ [t_ cto son	estos son cuatro	2	34	s
13	es_ [t_ cta son	estas son las mesas	2	120	s
14	es_ [t_ cto so lo	estos son fregaderos	2	67	s
15	e so so los	esos son los muebles	18	4	s
16	es_ [t_ cta son	éstas son las máquinas	18	68	s
17	ma_7 se p_cpa	màs se parezca	18	107	s
18	p_cpwe_7 se lek_ck	puedes seleccionar	18	111	u
19	mo_7 Vle_7 se	muebles se llaman	7	59	s
20	la si Zas	las sillas	2	145	s
21	la sa la se nas	las alacenas	11	147	s
22	la si Zas	las sillas	20	152	u
23	la s_7 Zas	las sillas	20	154	s
24	d_cdo_7 si Zas	dos sillas	20	171	u
25	a_2l Gu na si_7 Zas	algunas sillas	2	143	u

11. TABLA DE SINALEFA (a e)

Fenómeno: Sinalefa “a e”

Regla: $V_{\alpha} \# \# V_{\beta} \rightarrow \$V_{\alpha} V_{\beta} \$$

Fichas	Segmento	Transcripción	Diálogo	Oración	Emisor
1	laes [t ctu 7 fa	la estufa	11	10	u
2	laes [t ctu 7 fa	la estufa	11	10	u
3	laes [t ctu 7 fa	la estufa	11	119	u
4	aes [t cta	a esta pared	11	119	u
5	aes [t cta	a esta pared	11	119	u
6	t ctaes [t ctu fa	esta estufa	11	123	s
7	aes [t ctas	a esta esquina	11	123	s
8	b cbje naes [t cta	también a esta	11	130	u
9	se 7n [t ctr(aes [	centra esta	11	138	u
10	aes [t cta	a esta	11	152	u
11	aes [t ctes	a esta	11	172	u
12	laes [t ctu 7 fa	la estufa	11	268	u
13	laes [t ctu 7 fa	la estufa	11	32	u
14	laes [t ctu 7 fa	la estufa	11	71	u
15	d cdea k ckwer(d cdo	de acuerdo	20	132	s
16	laes [t ctu 7 fa	la estufa	2	111	u
17	laes [t ctu 7 fa	la estufa	2	117	u
18	laes [t ctu 7 fa	la estufa	2	60	u
19	laes [t ctu 7 fa	la estufa	2	64	u
20	laes [t ctu 7 fa	la estufa	7	3	u
21	laes k ck ji 7 na	la esquina	7	3	u

12. TABLA DE SINALEFA (e a)

Fenómeno: Sinalefa “e a”

Regla: $V_{\alpha} \# \# V_{\beta} \rightarrow \$V_{\alpha} V_{\beta} \$$

Fichas	Segmento	Transcripción	Diálogo	Oración	Emisor
1	Dea es	mi idea es colocar	2	12	u
2	mea sen	me hacen falta	2	134	u
4	Dea o 7r(a	de ahora	2	1	s
5	t ctea mi	frente a mi	11	239	u
6	Dea k ck je	quede aquí en esta	11	92	u
7	se rea la **esta?	cuál sería la	2	134	u
8	se rea k ckom b c**está?	sería conveniente	2	152	s
9	xe r(a Do 7 r(aoa 2l	refrigeradora o al	18	125	u
10	xi r(ael** no sé si entra	girar el cuarto	11	43	u

11	mwe_7 Vlea la	mueble a la cocina	20	1	s
----	---------------	-----------------------	----	---	---

13. TABLA DE SINALEFA (o a)

Fenómeno: Sinalefa “o a”

Regla: $V\alpha\#V\beta \rightarrow \$V\alpha V\beta\$$

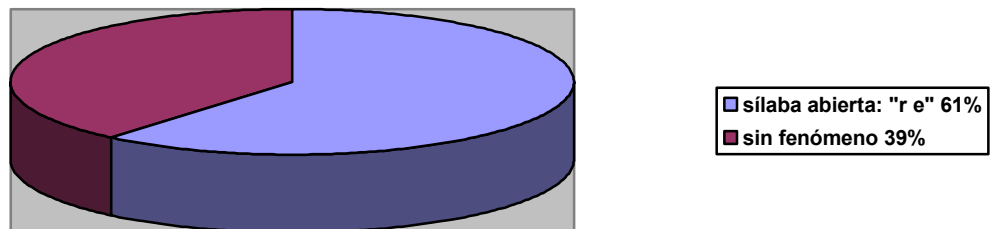
Fichas	Segmento	Transcripción	Diálogo	Oración	Emisor
1	xu_7n_ [toa la	junto a la	11	131	u
2	De r(oa k_ck_ji_7	fregadero aquí	11	138	u
3	r(a Do_7 roa_2l Go**sila a	refrigerador o algo así	11	194	u
	r(oa_2l	fregadero al	11	215	u
4	xu_7n_ [toa_2l	junto al	11	233	u
5	xu_7n_ [toa_2l	junto al	20	76	u
6	loa k_ck_ji_7	colocarlo aquí	2	177	u
7	loa k_ck_ji_7	controlarlo aquí	2	188	u
8	nob_cb xe_7 t_ctoa	objeto a la cocin	18	1	s
9	Ga_7 Doa	pegado hacia	18	125	u
10	Ga_7 Doa	pegado a la estufa	18	132	u
11	k_ckoa sjael	poco haicia	18	19	u

8. APÉNDICE 2. PORCENTAJES Y GRÁFICAS

1. RESULTADOS DE SÍLABA ABIERTA (r e)

FENÓMENO: SÍLABA ABIERTA: “R E”	NÚMERO TOTAL DE APARICIONES DEL FENÓMENO EN EL CORPUS <i>DIME</i>	NÚMERO DE FICHAS EN DONDE EL FENÓMENO OCURRE	PORCENTAJE DE OCURRENCIA DEL FENÓMENO
	55	34	61%

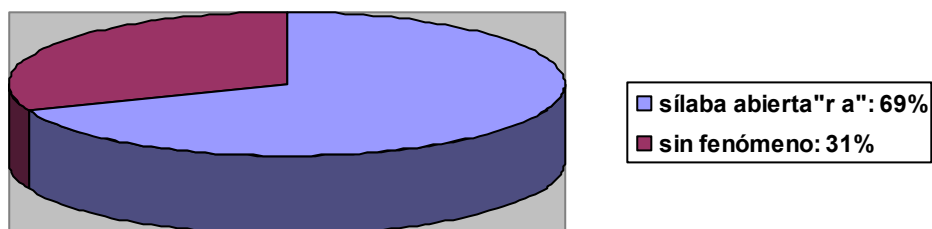
1. GRÁFICA DE RESULTADOS: SÍLABA ABIERTA (r e)



RESULTADOS SE SÍLABA ABIERTA (r a)

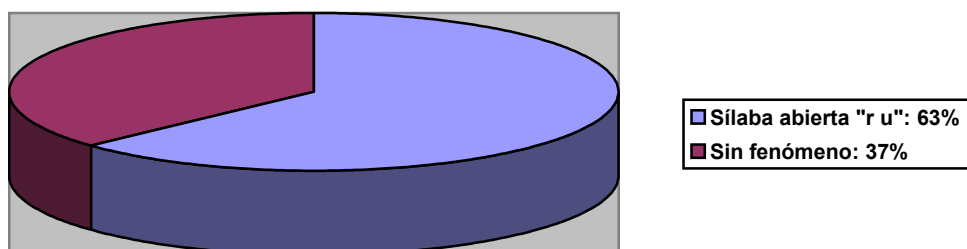
FENÓMENO: SÍLABA ABIERTA “R A”	NÚMERO TOTAL DE APARICIONES DEL FENÓMENO EN EL CORPUS <i>DIME</i>	NÚMERO DE FICHAS EN DONDE EL FENÓMENO OCURRE	PORCENTAJE DE OCURRENCIA DEL FENÓMENO
	13	9	69%

GRÁFICA DE RESULTADOS: SÍLABA ABIERTA (r a)



RESULTADO DE SÍLABA ABIERTA (r u)

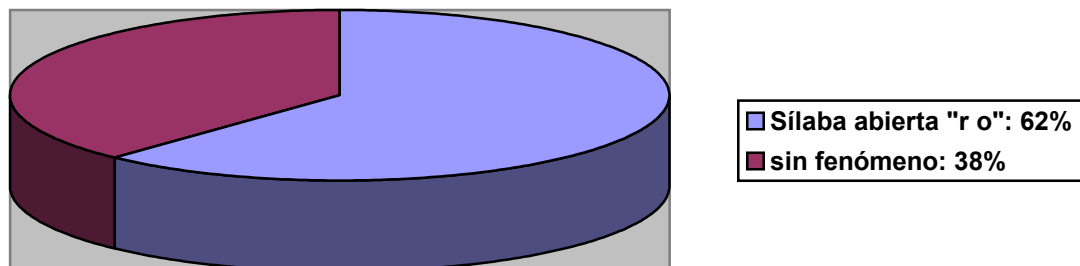
FENÓMENO: SÍLABA ABIERTA "R U"	NÚMERO TOTAL DE APARICIONES DEL FENÓMENO EN EL CORPUS <i>DIME</i>	NÚMERO DE FICHAS EN DONDE EL FENÓMENO OCURRE	PORCENTAJE DE OCURRENCIA DEL FENÓMENO
	11	7	63%



RESULTADO DE SÍLABA ABIERTA (r o)

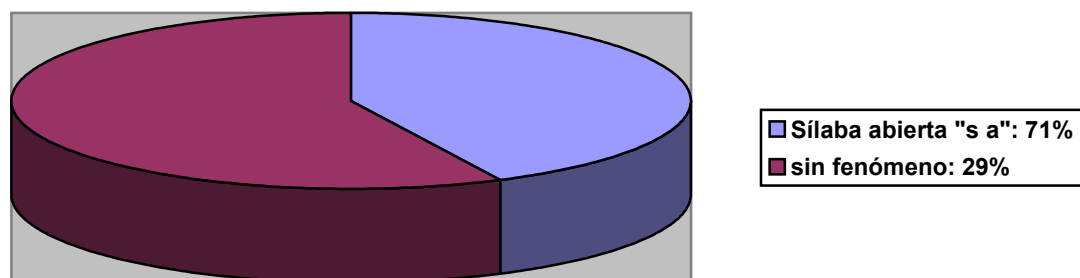
FENÓMENO: SÍLABA ABIERTA "R O"	NÚMERO TOTAL DE APARICIONES DEL FENÓMENO EN EL CORPUS <i>DIME</i>	NÚMERO DE FICHAS EN DONDE EL FENÓMENO OCURRE	PORCENTAJE DE OCURRENCIA DEL FENÓMENO
	8	5	62%

GRÁFICA DE RESULTADOS: SÍLABA ABIERTA (r o)



RESULTADOS DE SÍLABA ABIERTA (s a)

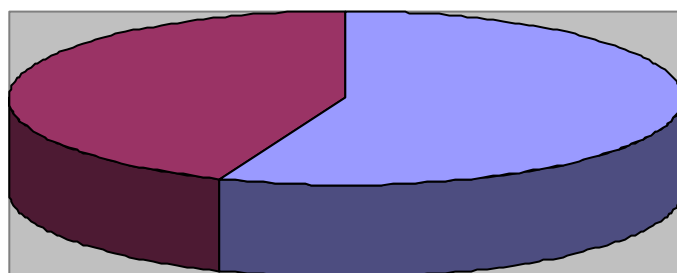
FENÓMENO: SÍLABA ABIERTA "S A"	NÚMERO TOTAL DE APARICIONES DEL FENÓMENO EN EL CORPUS <i>DIME</i>	NÚMERO DE FICHAS EN DONDE EL FENÓMENO OCURRE	PORCENTAJE DE OCURRENCIA DEL FENÓMENO
	42	30	71%



RESULTADOS DE SÍLABA ABIERTA (s e)

FENÓMENO: SÍLABA ABIERTA "S E"	NÚMERO TOTAL DE APARICIONES DEL FENÓMENO EN EL CORPUS <i>DIME</i>	NÚMERO DE FICHAS EN DONDE EL FENÓMENO OCURRE	PORCENTAJE DE OCURRENCIA DEL FENÓMENO
	41	23	56%

GRÁFICA DE RESULTADOS: SÍLABA ABIERTA (s e)

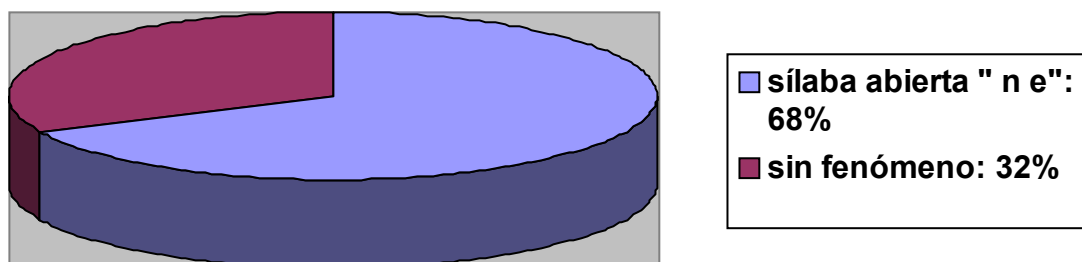


■ Sílabas abiertas "s e": 56%
■ sin fenómeno: 44%

RESULTADOS DE SÍLABA ABIERTA (n e)

FENÓMENO: SÍLABA ABIERTA "S E"	NÚMERO TOTAL DE APARICIONES DEL FENÓMENO EN EL CORPUS <i>DIME</i>	NÚMERO DE FICHAS EN DONDE EL FENÓMENO OCURRE	PORCENTAJE DE OCURRENCIA DEL FENÓMENO
	89	61	68%

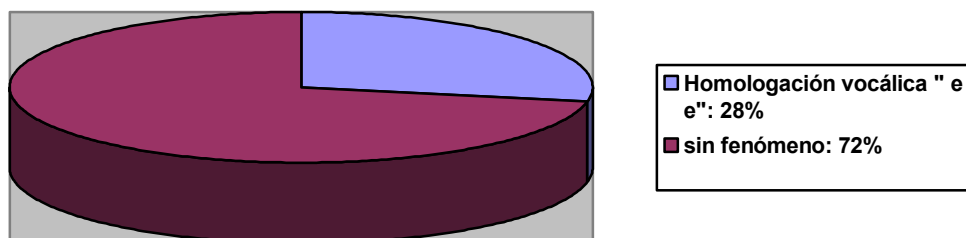
GRÁFICA DE RESULTADOS: SÍLABA ABIERTA (n e)



RESULTADOS DE HOMOLOGACIÓN VOCÁLICA (e e)

FENÓMENO: HOMOLOGACIÓN VOCÁLICA "E E"	NÚMERO TOTAL DE APARICIONES DEL FENÓMENO EN EL CORPUS <i>DIME</i>	NÚMERO DE FICHAS EN DONDE EL FENÓMENO OCURRE	PORCENTAJE DE OCURRENCIA DEL FENÓMENO
	169	49	28%

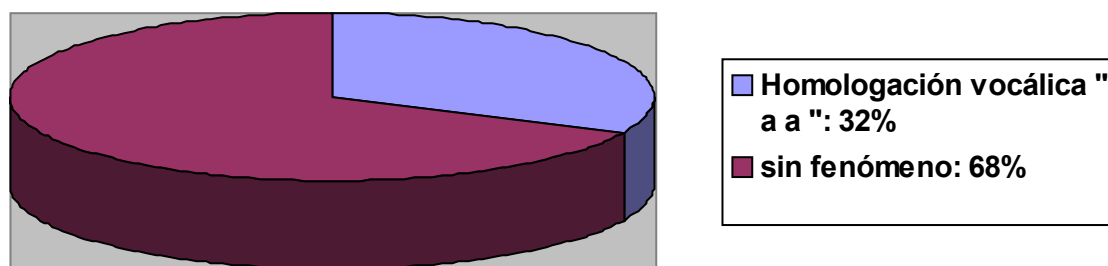
GRÁFICA DE RESULTADOS: HOMOLOGACIÓN (e e)



RESULTADOS DE HOMOLOGACIÓN VOCÁLICA (a a)

FENÓMENO: HOMOLOGACIÓN VOCÁLICA "E E"	NÚMERO TOTAL DE APARICIONES DEL FENÓMENO EN EL CORPUS <i>DIME</i>	NÚMERO DE FICHAS EN DONDE EL FENÓMENO OCURRE	PORCENTAJE DE OCURRENCIA DEL FENÓMENO
	50	16	32%

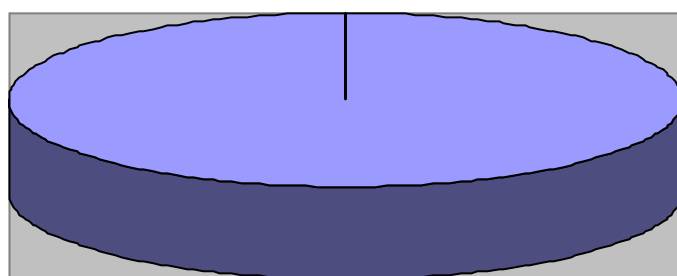
GRÁFICA DE RESULTADOS: HOMOLOGACIÓN (a a)



RESULTADOS DE HOMOLOGACIÓN CONSONÁNTICA (s s)

FENÓMENO: HOMOLOGACIÓN VOCÁLICA “S S”	NÚMERO TOTAL DE APARICIONES DEL FENÓMENO EN EL CORPUS <i>DIME</i>	NÚMERO DE FICHAS EN DONDE EL FENÓMENO OCURRE	PORCENTAJE DE OCURRENCIA DEL FENÓMENO
	25	25	100%

GRÁFICA DE RESULTADOS: HOMOLOGACIÓN (s s)

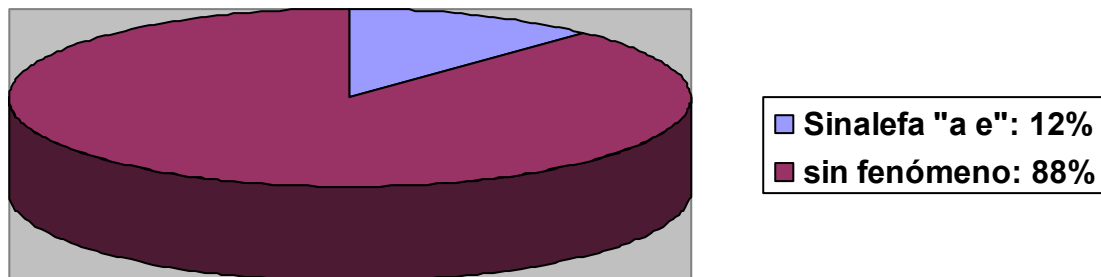


■ Homologación
consonántica "s s":
100%

RESULTADO DE SINALEFA (a e)

FENÓMENO: SINALEFA “A E”	NÚMERO TOTAL DE APARICIONES DEL FENÓMENO EN EL CORPUS <i>DIME</i>	NÚMERO DE FICHAS EN DONDE EL FENÓMENO OCURRE	PORCENTAJE DE OCURRENCIA DEL FENÓMENO
	169	21	12%

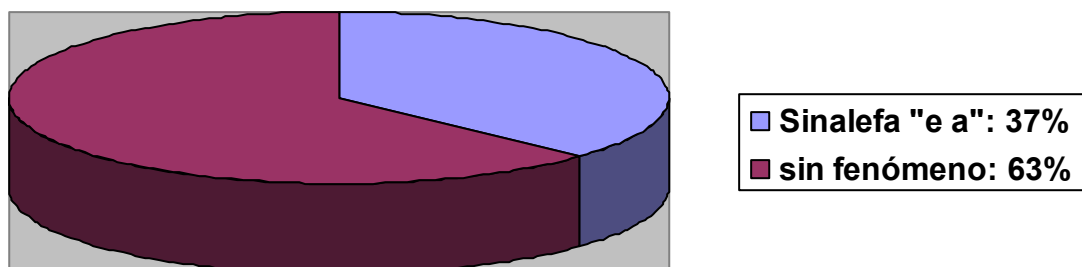
GRÁFICA DE RESULTADOS: SINALEFA (a e)



RESULTADOS DE SINALEFA (e a)

FENÓMENO: SINALEFA "E A"	NÚMERO TOTAL DE APARICIONES DEL FENÓMENO EN EL CORPUS <i>DIME</i>	NÚMERO DE FICHAS EN DONDE EL FENÓMENO OCURRE	PORCENTAJE DE OCURRENCIA DEL FENÓMENO
	32	12	37%

GRÁFICA DE RESULTADOS: SINALEFA (e a)



The screenshot displays the SpeechView software interface, which is used for analyzing speech signals. The interface is divided into several sections:

- Top Window (Waveform):** Shows a black waveform on a white background. A green rectangular highlight is visible on the left side of the waveform. Below the waveform, a time scale is marked with values 250, 350, 450, 550, and 650. The text "Time: 442.51 ms" and "Label: ne_7" is displayed at the bottom left of this window.
- Middle Window (Spectrogram):** Shows a spectrogram with a yellow rectangular highlight on the left side. The x-axis represents time, and the y-axis represents frequency.
- Bottom Window (Phonetic Transcription Table):** Displays a table with three rows and several columns. The first row contains the phonetic segments: "p_c", "p", "o", "n", "e_7", "r()", and "e". The second row contains "p_cpo", "ne_7", and "r(e)". The third row contains "p_cpo", "ne_7r()", and "el".

SpeechView

File Options Help

Time: 1335.81 ms Amplitude: 4

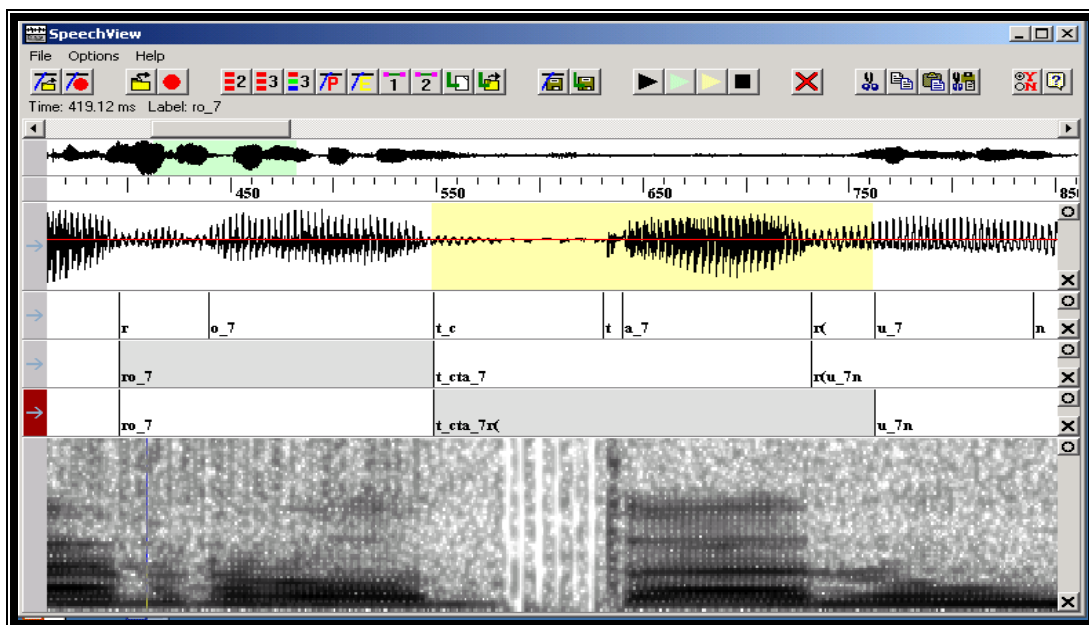
ADD WINDOW: Color 3-D spectrogram

850 950 1050 1150 1250

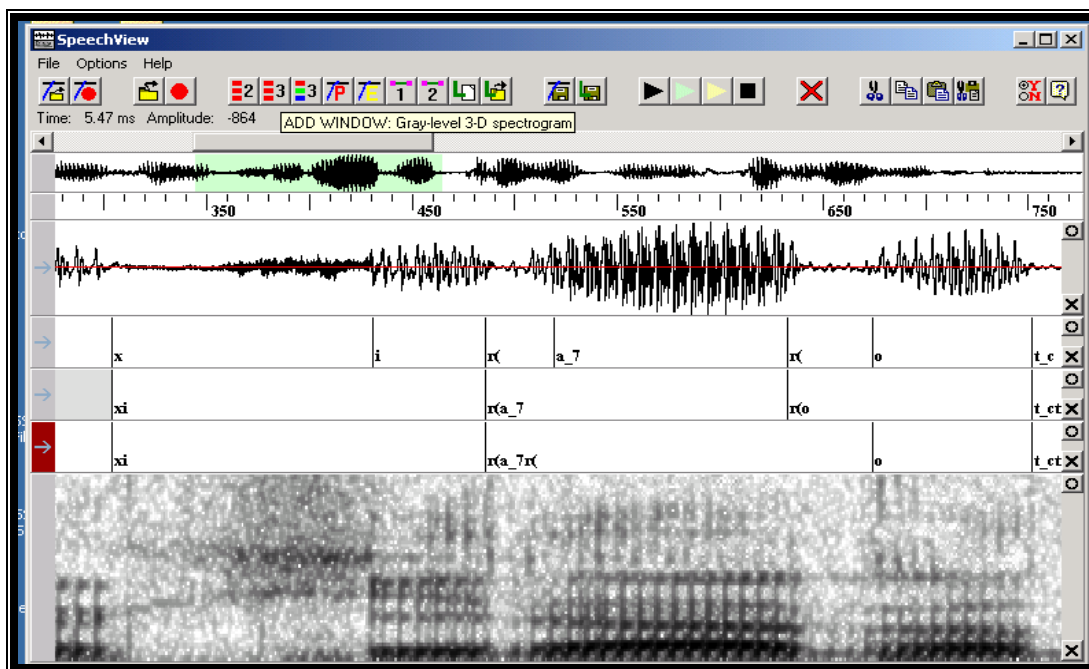
p_c p o n e r(a k_c

p_cpo ne r(a k_ck_jj

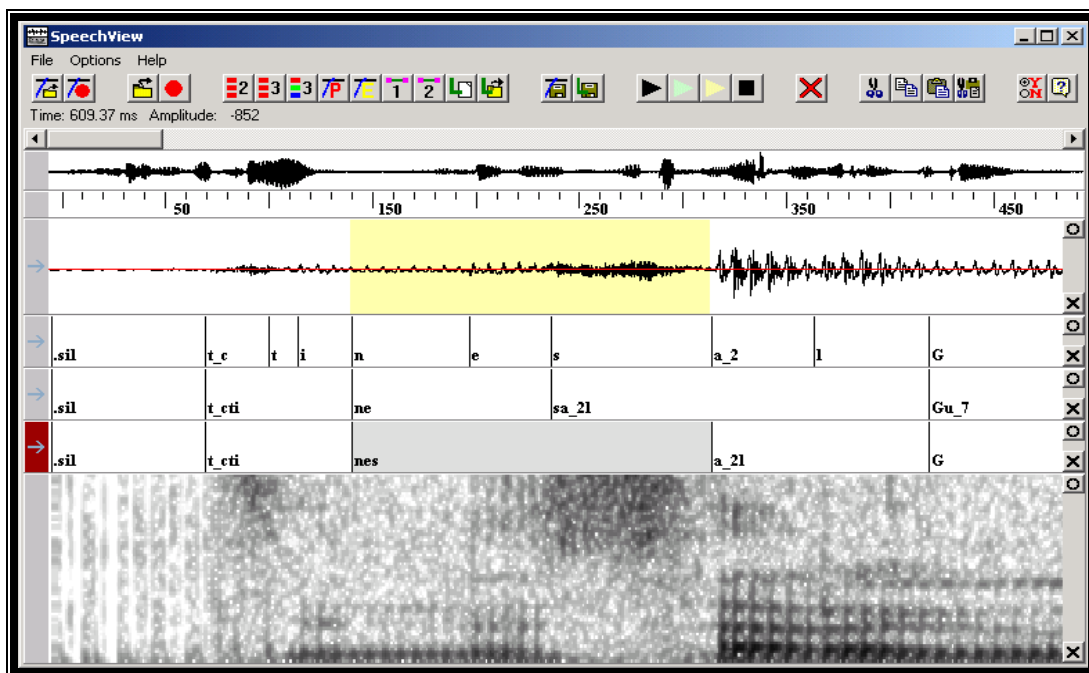
p_cpo ner(a k_ck_ji



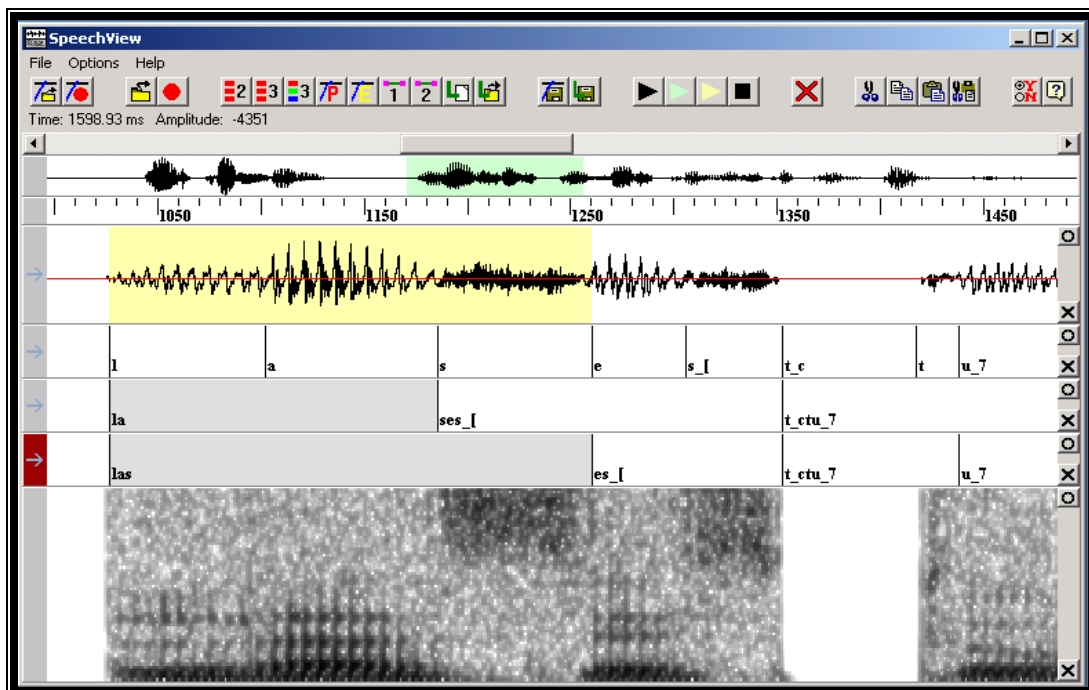
3. Imagen espectrográfica de sílaba abierta “r u”:
rotar un-ro.ta.run. Diálogo:2; oración 10.



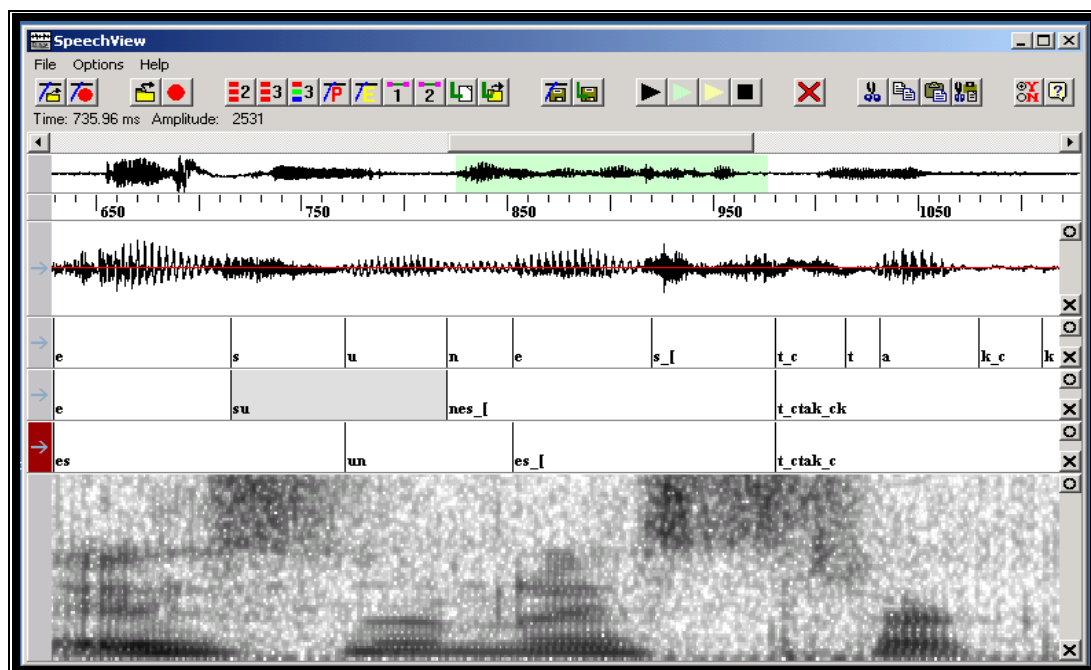
6. Imagen espectrográfica de sílaba abierta “r o”:
girar otra vez-gi.ra.ro.tra.ves. Diálogo 11; oración 40.



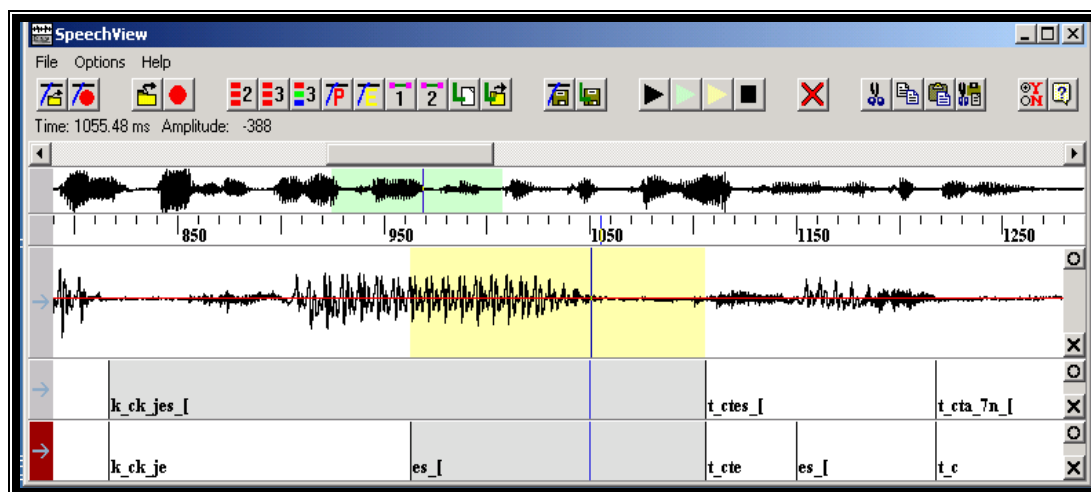
7. Imagen espectrográfica de sílaba abierta “s a”:
tienes algun-ti.ne.sal.gun. Diálogo 11; oración 15.



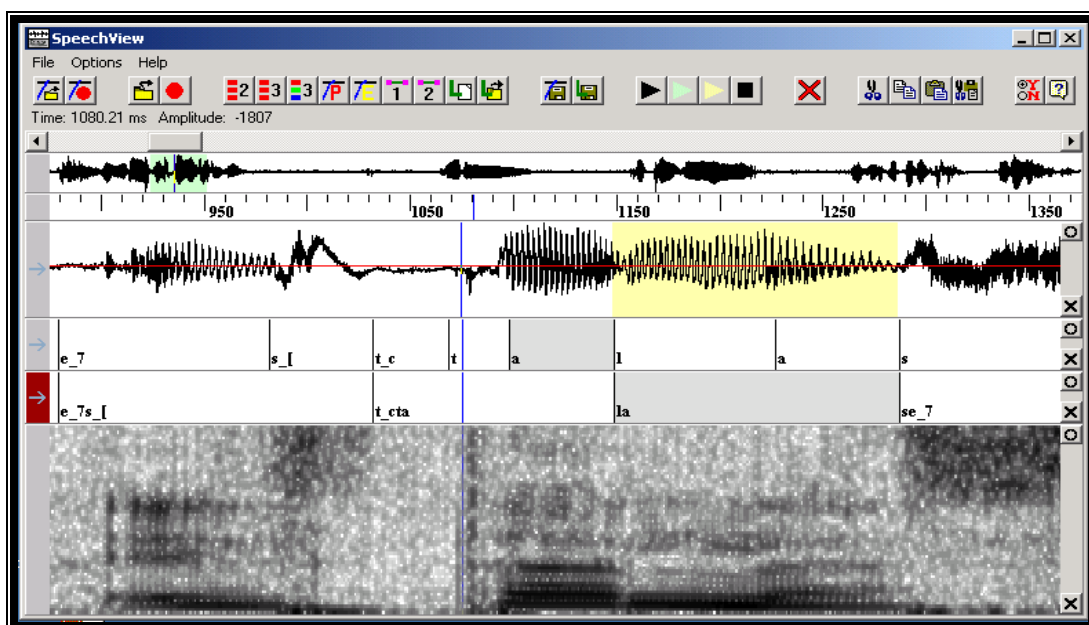
8. Imagen espectrográfica de sílaba abierta “s e”.



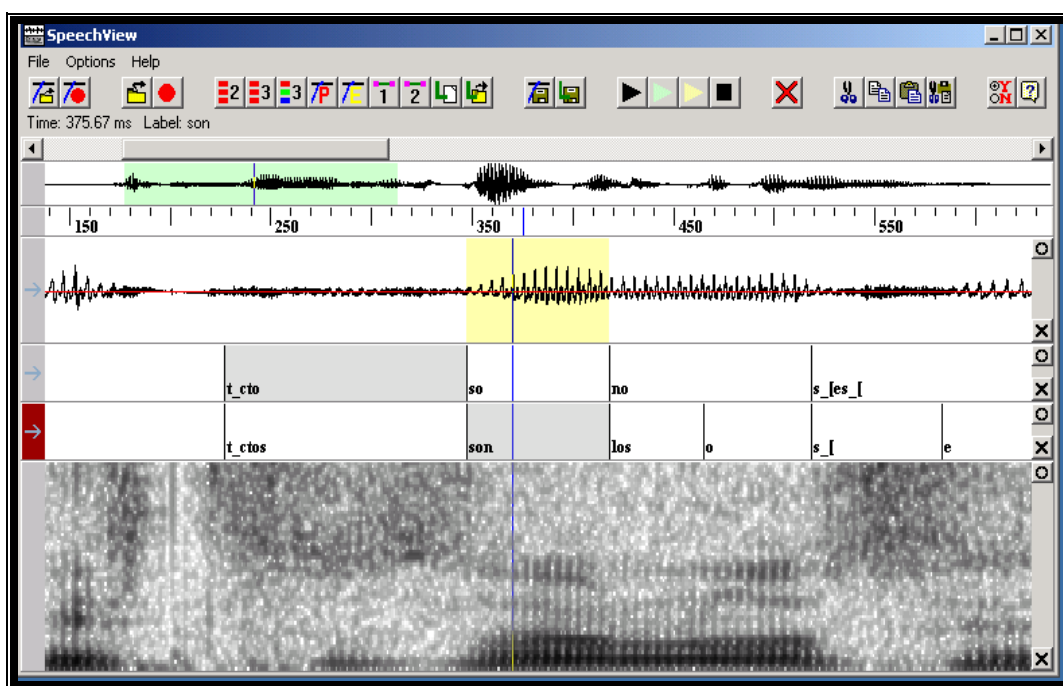
9. Imagen espectrográfica de sílaba abierta “n e”.



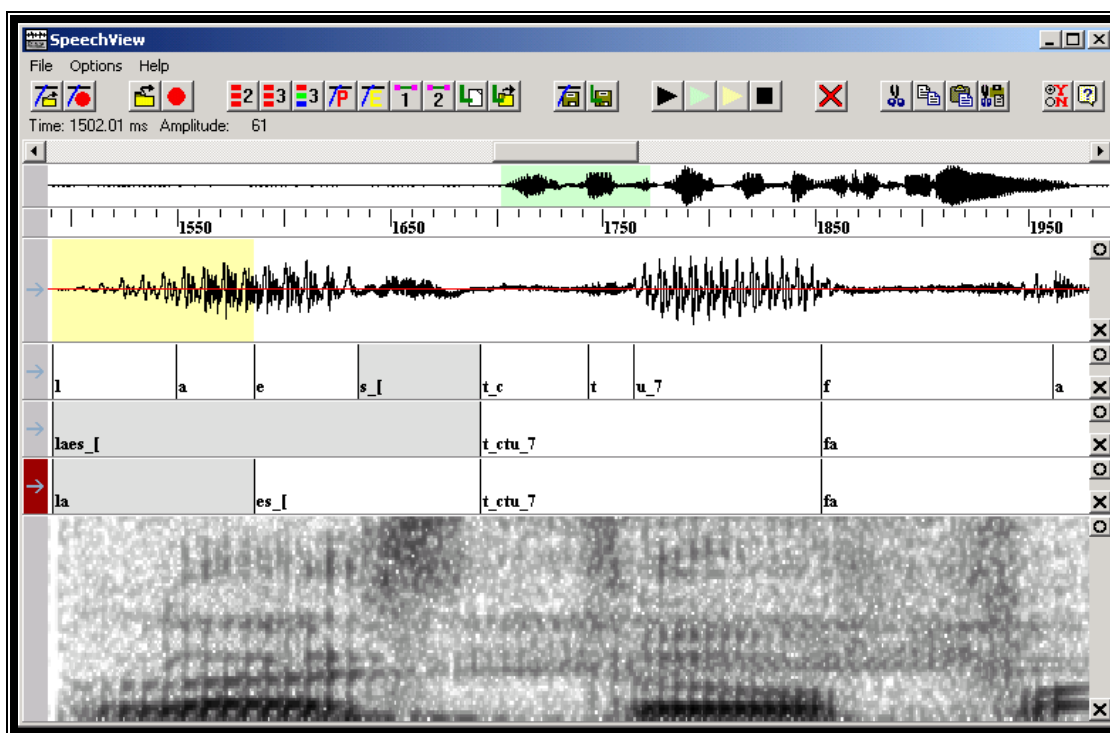
10. Imagen espectrográfica de homologación vocálica “e”: que este-kes.tes.tan.te. Diálogo 11; oración 92.



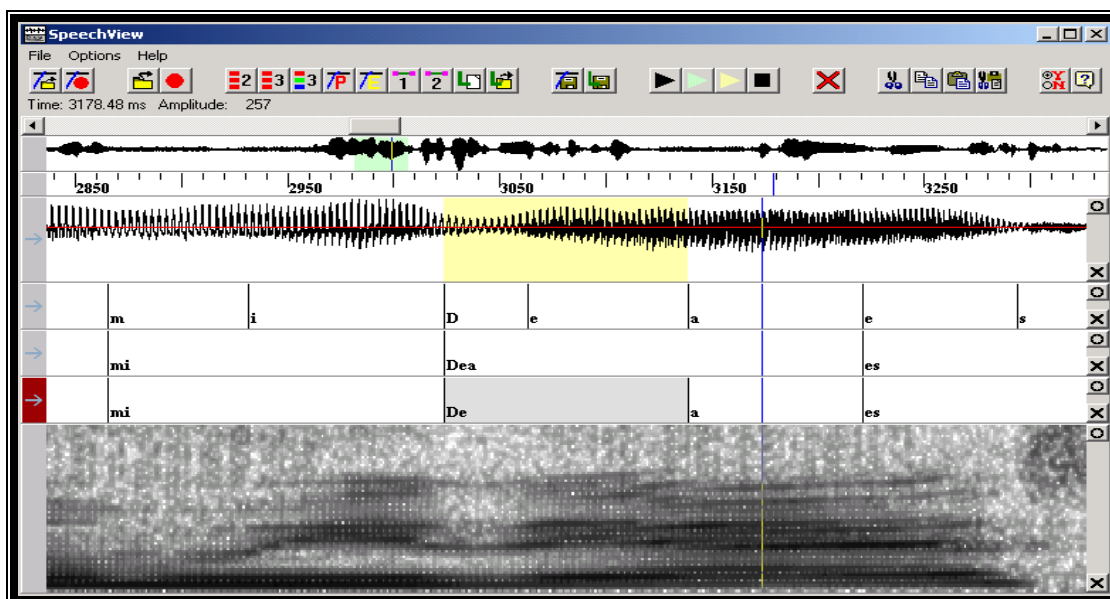
11. Imagen espectrográfica de homologación vocálica “a”: esta alacena-es.ta.la.se.na. Diálogo 7; oración 37.



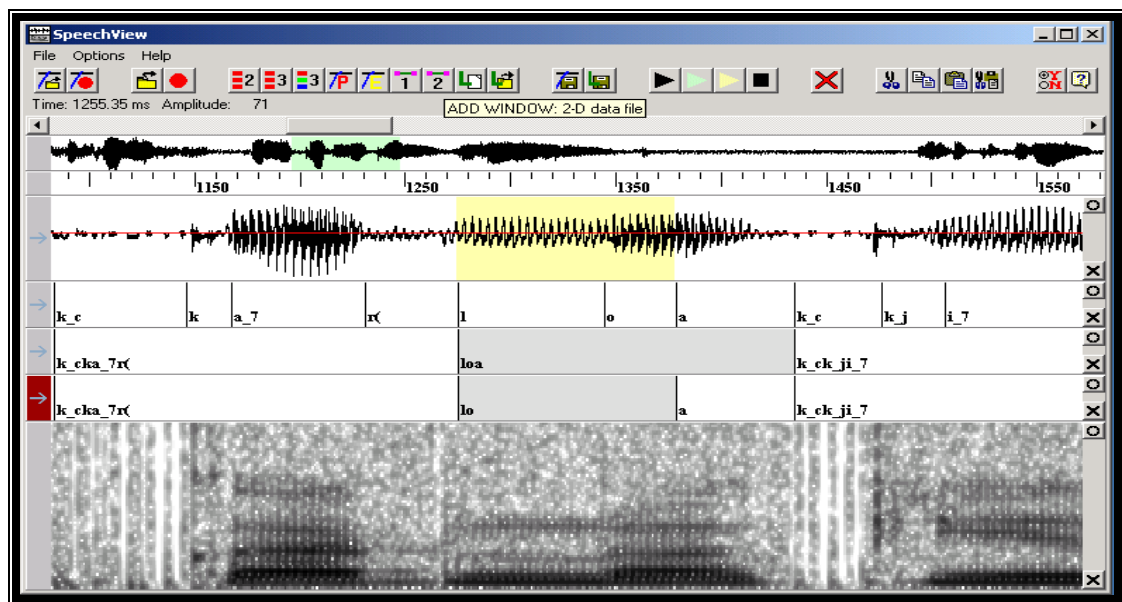
12. Imagen espectrográfica de homologación consonántica “s”: estos son-es.to.son. Diálogo 11; oración 90.



13. Imagen espectrográfica de sinalefa “a e”: la estufa-laes.tu.fa.
Diálogo 11; oración 10.



14. Imagen espectrográfica de sinalefa “e a”: mi idea-mi.dea. Diálogo 2; oración 12.



15.Imagen espectrográfica de sinalefa “o a”.