



UNIVERSIDAD NACIONAL AUTÓNOMA DE
MÉXICO

FACULTAD DE FILOSOFÍA Y LETRAS
COLEGIO DE LETRAS HISPÁNICAS

EL *TEMPO* EN LLAMADAS DE EXTORSIÓN

T E S I S

QUE PARA OBTENER EL GRADO DE:
LICENCIADA EN LENGUA Y LITERATURAS HISPÁNICAS

PRESENTA:

ROSA MARÍA GARCÍA HERNÁNDEZ

DIRECTORA DE TESIS:

DRA. MARGARITA PALACIOS SIERRA



CIUDAD UNIVERSITARIA, CDMX, 2019



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

*A la mujer que me inspiró
con el sonido de su risa,
a Carolina, cuya vida
el corazón le arrebató.*

Agradecimientos

Agradezco a todas y a cada una de las personas que hicieron posible la realización de este proyecto.

A la Dra. Margarita Palacios, quien me animó a realizar todos mis experimentos, y al Seminario Universitario del Discurso Forense por prestarme el corpus con el que realicé esta investigación.

Al Dr. David Alberto Chávez Rivadeneira por la pronta lectura de esta tesis y por sus palabras de aliento.

A la Mtra. Ana Isabel Tsutsumi Hernández por guiarme hacia la culminación de mis ideas.

A la Dra. Georgina Barraza Carbajal por su paciencia y sus comentarios que me ayudaron a mejorar este estudio.

A la Mtra. Teresita Adriana Reyes Careaga por su compromiso con el seguimiento de este trabajo.

A todos mis compañeros del seminario de investigación, porque sus comentarios me ayudaron a no desviarme del camino.

Doy las gracias también a aquellos que me dieron las herramientas que me facilitaron el procesamiento de los datos y su posterior análisis, es decir, a mis profesores tanto de los cursos curriculares como de los extracurriculares.

Y, finalmente, agradezco a mi madre, a mi hermana y a mis compañeros de la vida, Iván y Desireé, porque estuvieron a mi lado durante todo el proceso brindándome, incondicionalmente, su apoyo.

Índice

1	Introducción.....	6
1.1	JUSTIFICACIÓN	6
1.2	ÁREA PROBLEMÁTICA	7
1.3	PREGUNTA DE INVESTIGACIÓN	8
1.4	OBJETIVOS	9
1.5	ARGUMENTO CENTRAL.....	10
1.6	CAPITULADO	10
2	Estado del arte	12
2.1	VELOCIDAD DE HABLA	12
2.1.1	<i>Constituyentes de la velocidad de habla</i>	14
2.1.1.1	Duración.....	15
2.1.1.2	Pausas.....	16
2.1.2	<i>La velocidad de habla como herramienta para estudiar la oralidad</i>	18
2.2	LA VELOCIDAD DE HABLA EN LA CONVERSACIÓN	19
2.2.1	<i>Los turnos de habla</i>	21
2.2.2	<i>La intención</i>	22
2.3	LAS LLAMADAS DE EXTORSIÓN	23
3	Corpus y metodología.....	25
3.1	CORPUS.....	25
3.2	PROCEDIMIENTO.....	26
3.2.1	<i>Anotación</i>	26
3.2.2	<i>Extracción de datos</i>	28
3.2.3	<i>Determinación de duración segmental</i>	30
3.2.4	<i>Transcripción automática</i>	32
3.2.5	<i>Cálculo de velocidad de habla</i>	34
3.2.6	<i>Herramientas</i>	34
4	Resultados	37
4.1	DISTRIBUCIÓN TEMPORAL INTERNA AL TURNO DE HABLA.....	37
4.1.1	<i>Duración promedio por grabación y sujeto</i>	38
4.1.1.1	Grabación 1 sujeto 1.....	39
4.1.1.2	Grabación 1 sujeto 2.....	41
4.1.1.3	Grabación 2 sujeto 1.....	42
4.1.1.4	Grabación 2 sujeto 2.....	44
4.1.1.5	Distribución temporal general por grabación y sujeto	45
4.1.2	<i>Duración promedio por grabación</i>	47
4.1.2.1	Grabación 1	47
4.1.2.2	Grabación 2	49

4.1.2.3	Distribución temporal general por grabación.....	51
4.1.3	<i>Duración promedio por sujeto</i>	53
4.1.3.1	Sujeto 1.....	54
4.1.3.2	Sujeto 2.....	55
4.1.3.3	Distribución temporal general por sujeto	56
4.1.4	<i>Duración promedio total</i>	58
4.2	VELOCIDAD DE ELOCUCIÓN EN LOS TURNOS DE HABLA.....	60
4.2.1	<i>Grabación 1</i>	61
4.2.1.1	Tempo.....	62
4.2.1.2	Valores de relevancia adquiridos	66
4.2.1.2.1	Conflicto	66
4.2.1.2.2	Acuerdo	70
4.2.2	<i>Grabación 2</i>	73
4.2.2.1	Tempo.....	74
4.2.2.2	Valores de relevancia adquiridos	80
4.2.2.2.1	Conflicto	80
4.2.2.2.2	Acuerdo	82
4.2.2.2.3	Casos excepcionales	86
5	Conclusiones.....	87
5.1	<i>TIME IS MONEY</i>	87
5.1.1	<i>Valores de relevancia adquiridos</i>	88
5.1.2	<i>Uso contrastivo del tempo</i>	89
5.2	PERTINENCIA DE LA VELOCIDAD DE HABLA.....	89
5.3	DISTRIBUCIÓN TEMPORAL	90
6	Referencias	92

1 Introducción

1.1 Justificación

México pasa por una ola de violencia que se ha filtrado en nuestras casas y escuelas, la Universidad Nacional Autónoma de México se ha vuelto otro foco más de inseguridad. Durante mi estadía en la Facultad de Filosofía y Letras ocurrieron grandes injusticias dentro y fuera de la universidad, quizá el caso más escandaloso fue el de Ayotzinapa, que ocurrió un par de meses después de mi ingreso a la licenciatura y que, hasta la fecha, sigue sin resolverse. Pero como he dicho, la violencia está en todas partes, en estos últimos años, estrangularon a una mujer en el Instituto de Ingeniería de la UNAM, mataron a una alumna de la Facultad de Estudios Superiores Iztacala después de un asalto, abusaron sexualmente de una alumna de la Escuela Nacional de Trabajo Social dentro del plantel. Hace apenas unos meses, violaron a una alumna de la Facultad de Estudios Superiores Acatlán a plena luz del día, secuestraron y mataron a una alumna del Colegio de Ciencias y Humanidades Plantel Oriente, y el día 3 de septiembre de 2018 ocurrió un atentado llevado a cabo por grupos porriles en contra de una manifestación pacífica de estudiantes en la que hubo varios heridos.

El corpus de esta tesis es la memoria de un acto violento ocurrido en el centro de México, que cumplió diez años el 6 de septiembre de 2018 y que, como tal, debe servirnos para tratar de mejorar nuestra sociedad. El secuestro es común en un país violento que alarma a todos sus habitantes, pero afortunadamente, según las estadísticas del sitio oficial de la asociación Alto al secuestro (2018) se ha registrado disminución en la incidencia de este delito, tanto en el Estado de México como en la Ciudad de México, desde 2013 hasta 2017. En el Estado de México se denunciaron 562 casos en 2013 y 307 en 2017, mientras que en la Ciudad de México hubo 189 casos en 2013

y 102 en 2017. Cabe señalar que el año 2014 registró un alza en el índice de secuestros en ambas entidades, pero fue disminuyendo con el paso de los años.

Si me preguntaran por qué me decidí a hacer esta tesis, tendría que decir que vivimos en una situación terrible y que cualquier cosa que pueda ayudar a una víctima de algún delito, lo que sea, debe por lo menos intentarse. Lo que trato de hacer es apoyar a las víctimas de llamadas de extorsión, falsas o reales. Mi corpus, de extensión muy limitada, se compone de grabaciones reales, cuyos datos de identidad son anónimos, que obtuve gracias al apoyo del Seminario Universitario de Estudios del Discurso Forense (SUEDIF) que coordina la Dra. Margarita Palacios Sierra. Deseo que, al final, mi investigación aporte algunos conocimientos sobre la velocidad de habla que puedan ser útiles en los casos de extorsión.

1.2 Área problemática

El *tempo*, también conocido como *velocidad de elocución* o *velocidad de habla*, es un rasgo prosódico de la lengua española que matiza los mensajes de diversas maneras, por lo cual los hablantes tienen la posibilidad de aumentar o disminuir la rapidez en infinitas estructuras fónicas que ponen de relieve los comportamientos intencionales y los actos de reflejo emocional. Sin embargo, la información respecto a las características acústicas de la velocidad de habla y sus implicaciones pragmáticas es relativamente poca, ya que se ve reducida a estudios muy específicos, debido en parte a la falta de una metodología estandarizada y a la ausencia de variedad y cantidad de *corpora* orales. Por lo tanto, esta tesis representa una contribución al conocimiento sobre las propiedades que se añaden al discurso con la modificación de la velocidad de elocución de un hablante y sobre el impacto que tiene en la respuesta del oyente.

Específicamente, en el contexto de una llamada telefónica hay varios factores que condicionan

la forma en que se desarrolla la conversación. En primer lugar, se trata de un acto comunicativo que no implica la presencia de los dos hablantes en el mismo espacio, por lo que depende únicamente de la oralidad y no tiene posibilidad alguna de intercambiar gestualidades. Además, la situación misma dicta las pautas de comportamiento entre ambos locutores, pues participan una negociación en la que se pretende intercambiar la vida de un ser humano por dinero. Por ello, los hablantes necesitan utilizar diferentes estrategias fónicas que les permitan mantener la conversación hasta obtener la respuesta que desean de su interlocutor. Por ejemplo, se puede resaltar tanto la información dicha como la implicada disminuyendo la velocidad de elocución, así como diluir un mensaje o aumentar su impacto produciendo una gran cantidad de sílabas por segundo. Si bien, no hay correspondencia entre rapidez y efecto, es posible determinar cuál es la intención con la que se produjo tal o cual variación, ya que los enunciados tienen límites de interpretación debido a su contenido semántico.

1.3 Pregunta de investigación

Las dos grabaciones del corpus utilizado para esta tesis se originan de manera espontánea en un par de conversaciones telefónicas, lo cual permite apreciar la fluctuación de la velocidad de elocución que se establece entre dos hablantes de manera natural. Además, es notable la diferencia entre las duraciones totales de los audios, pues el segundo supera por más de un minuto al primero, lo que implica que no se utiliza la misma dinámica de intercambio en ambos casos. Entonces, surge la pregunta ¿cómo influye la velocidad de elocución en el proceso para mantener abierto el canal de comunicación en situaciones de secuestro? Porque, considerando que el contenido de los mensajes es idéntico en todo el corpus, una o varias características orales tendrían que justificar la ganancia de tiempo que se presenta en la segunda grabación. Por consiguiente, centrarse en un

rasgo suprasegmental, como la velocidad de habla, muestra contrastes marcados a lo largo de los actos comunicativos que explican el cambio en la estructura de la conversación, al menos de manera parcial, porque es imposible que con un solo rasgo se pueda estudiar la totalidad del discurso.

1.4 Objetivos

Esta tesis tiene como objetivo general describir el impacto que tienen las modificaciones de la velocidad de elocución de los mensajes emitidos por un hablante sobre la respuesta de su interlocutor, para verificar si la variación en la rapidez entre turnos de habla contiguos es una estrategia que se utiliza para transmitir información al receptor de manera eficaz y, al mismo tiempo, condicionar las cualidades de la réplica del receptor.

Del mismo modo, se establecieron tres objetivos particulares; el primero es comprobar la existencia de una forma característica de velocidad de elocución para cada hablante, es decir, comparar los datos de ambos hablantes para constatar si existe o no una diferencia significativa entre ambos locutores a partir de su rapidez promedio, el rango que abarcan las medidas de tiempo de todos sus turnos de habla, la cantidad de tiempo utilizado en sus intervenciones y la cantidad de segmentos que realiza. El segundo objetivo particular es establecer un modelo tipográfico en el que se pueda observar de manera gráfica la distribución del tiempo dentro de los turnos de habla, para identificar los momentos de la conversación en los que se enfoca o desenfoca la información mediante la aceleración o la desaceleración del discurso. El último objetivo particular es formular una serie de recomendaciones que puedan apoyar a las víctimas de una llamada de extorsión, según las características de los puntos de «conflicto» y «acuerdo» observados en el corpus.

1.5 Argumento central

Según lo observado en el corpus, la velocidad de elocución de un turno de habla condiciona la rapidez de la siguiente intervención, de manera que toda enunciación es más rápida o más lenta que la inmediatamente anterior. Por ello, propongo que es posible que los hablantes tracen el curso de la conversación a partir del contraste del tiempo entre turnos de habla contiguos, lo cual permite extender o cortar el tiempo del intercambio de información. Además, considero que la diferencia en las medidas de rapidez determina el nivel de conformidad o acuerdo mutuo que existe entre los participantes, es decir, entre mayor sea la distancia entre una medida y otra, mayor será la posibilidad de que surja un conflicto.

1.6 Capitulado

En la introducción se presenta de manera general la línea que sigue esta investigación, desde la justificación hasta la argumentación. En el primer capítulo aparece un recorrido a través de textos teóricos, manuales de fonética y artículos de investigaciones que tocan los temas de velocidad de habla y de conversación. Primero se muestra la terminología que rodea al fenómeno del tiempo, después aparecen los elementos que modifican la velocidad de elocución y sus aplicaciones en distintas áreas. Finalmente, se mencionan los efectos que produce el cambio de velocidad dentro de una conversación.

En el segundo capítulo se muestra la descripción del corpus de este trabajo: las características de los archivos, los rasgos particulares de los hablantes desde un enfoque diatópico y las circunstancias que individualizan la situación comunicativa llevada a cabo en las dos grabaciones que lo constituyen. También se describe, paso a paso, el procedimiento metodológico que se siguió para llegar a las conclusiones de esta investigación.

En el tercer capítulo se exponen los resultados y el análisis de los datos obtenidos a partir un procesamiento riguroso. Este apartado está constituido por dos partes, la primera contiene la distribución temporal interna de los turnos de habla y la división del corpus que genera las transcripciones. En la segunda parte del capítulo aparece, para cada grabación, una descripción de la interacción entre los dos locutores del corpus y los valores de relevancia que adquieren en el continuum del tiempo, así como una comparación de las características generales de la velocidad de elocución de ambos hablantes.

Finalmente, en las conclusiones se hace una revisión de la interacción entre los dos hablantes del corpus, los valores de relevancia que adquieren sus intervenciones y se formulan algunas recomendaciones a partir de los puntos de «conflicto» y «acuerdo» presentes en las conversaciones. Posteriormente, se comenta sobre la pertinencia del tempo como rasgo identitario. Finalmente, se habla sobre los métodos y la información aportada por las transcripciones de la distribución temporal interna de los turnos de habla.

2 Estado del arte

2.1 Velocidad de habla

La *velocidad de habla* o *tempo* es la rapidez con la que un hablante pronuncia uno o más enunciados, que se calcula dividiendo el tiempo total de la elocución, incluido el de las pausas, entre el número de segmentos realizados (Gil Fernández, 2007, p. 308). En esta tesis se utiliza la medida de segmentos por segundo (s./s.) porque utilizar el término «sílabas» en vez de «segmento» invisibiliza la intervención de las pausas en el discurso. Además, a lo largo del texto, se utiliza indistintamente los términos *velocidad de elocución*, *velocidad de habla* y *tempo*, ya que funcionan como sinónimos, lo cual evita, hasta cierto punto, la repetición constante de uno solo de ellos.

Sin embargo, entre los antecedentes del tema hay inconsistencia en el uso de todos los términos asociados a este fenómeno, el más extendido es *velocidad de elocución* (Asuaje et al., 2005; Battaner et al., 2003; Blondet et al., 2003; Cederborg y Chávez Aguilar, 1992; Gil Fernández, 2007; Llisterri, 2016; Martínez Sánchez et al., 2016; Machuca, 2009; Machuca y De la Mota, 2006; Schwab, 2015), en especial en los estudios más modernos, aunque también se utiliza *velocidad de habla* (Contreras, 2013; Gil Fernández, 2007; Hidalgo Navarro y Quilis Merín, 2004), *tempo* (Gil Fernández, 2007; Gili Gaya, 1975; Hidalgo Navarro y Quilis Merín, 2004; Machuca, 2009; Navarro Tomás, 2004) y *rapidez ordinaria* (Navarro Tomás, 2004, p. 197) con la misma acepción. Además, existe un término similar, *velocidad de articulación*, que se utiliza para nombrar la velocidad con la que se producen los enunciados sin tomar en cuenta el tiempo de las pausas, este término está generalizado en los textos, no obstante, existe un caso inusual en el que se usa *velocidad de elocución* (Contreras, 2013, p. 239) para esta acepción. Ahora bien, la elección de *tempo* para el título de esta tesis se fundamenta en que “esta denominación alude más bien a la

sensación perceptiva que produce” (Gil Fernández, 2007, p. 308) en el receptor el cambio en la velocidad de habla en los mensajes, es decir, la velocidad de un enunciado condiciona la respuesta dependiendo de la «sensación» que produce en el oyente la rapidez o la lentitud.

En español aumentar (*accelerando*) o disminuir (*ritardando*) la velocidad de uno o varios enunciados no cambia el contenido semántico de ningún componente de estos, por lo que se puede afirmar que este rasgo carece de pertinencia fonológica, lo cual no niega que dichas variaciones puedan adquirir significaciones de otro tipo, pues representan un medio expresivo fundamental del lenguaje humano (Hidalgo Navarro y Quilis Merín, 2004, p. 46; Gili Gaya, 1975, p. 41). Además, la velocidad de elocución, gracias a su carácter de rasgo suprasegmental de la lengua española, permite a los hablantes “[caracterizar] trozos determinados de la cadena hablada no coincidentes siempre con un fonema” (Alarcos Llorach, 1974, p. 88), es decir, puede incidir en unidades de sentido mayores, como los enunciados, para modificar el matiz pragmático de los mensajes emitidos, ya sea por razones involuntarias o intencionales. Tales variaciones en el discurso son posibles ya que, a diferencia de lo que ocurriría con algún rasgo distintivo, el tempo actúa como elemento contrastivo, pues la distinción entre sus diferentes realizaciones no es de alternancia, sino que se contrastan entre sí debido a su concurrencia y sucesión durante el tiempo del discurso (D'Introno, Teso, y Weston, 1995, p. 124), es decir, la dicción rápida o lenta solo se puede detectar si se compara con distintas realizaciones del mismo hablante o de otro.

Se supone que la velocidad de elocución «normal» de un hablante está determinada por la rapidez que el individuo tenga como hábito articulatorio (Machuca, 2009, p. 110); sin embargo, a partir del punto medio que se establezca, en determinados momentos de la enunciación, la dicción podría parecer rápida o lenta según se aparte hacia un extremo o al otro de lo que se considera común para un individuo (Navarro Tomás, 2004, p. 197-198). En otras palabras, el tempo puede

modificarse, de manera voluntaria o involuntaria, a causa de diferentes razones, ya sea por la expresión accidental de estados de ánimo o asuntos semejantes (Gili Gaya, 1975, p. 40), por circunstancias relacionadas al contenido del mensaje como el tipo de información y la relevancia, o por razones externas al mensaje que dependen de la situación en la que se desarrolla la comunicación (Gil Fernández, 2007, p. 309).

Ahora bien, lo que marca la diferencia entre un acto voluntario (forma de comportamiento) y uno involuntario (acto reflejo) es la intención con que se emite el mensaje, puesto que las expresiones accidentales forman parte del compendio de información que se infiere entre los participantes de un acto comunicativo y, al no ser declaradas intencionalmente, no pueden considerarse como información comunicada (Escandell Vidal, 2014, pp. 23-24). Por ejemplo, cuando un hablante se encuentra sumido en determinado estado emocional se expresará de forma anormal, hablará lento cuando esté triste o rápido cuando esté molesto. Por otra parte, “los hablantes adaptan su velocidad de elocución a las necesidades de sus oyentes” (Gil Fernández, 2007, p. 309), por lo que es posible manipular la velocidad para lograr eficacia en la transmisión de un mensaje, dependiendo del contexto particular de la enunciación. De modo que, los hablantes pueden lograr efectos comunicativos concretos, como la insistencia o la burla; no obstante, no existe correspondencia entre una modificación producida y el sentido que adquiere, por lo que ha de establecerse dentro del contexto de cada muestra de comunicación particular (Escandell Vidal, 2014, p. 66) a partir del contenido semántico.

2.1.1 Constituyentes de la velocidad de habla

Las estrategias fónicas que se siguen para modificar la velocidad son el número de pausas y la duración de los segmentos o sílabas, todo ello dentro de una estructura de distribución temporal

específica que resulta inversamente proporcional a la velocidad de elocución empleada en la producción de los enunciados. De modo que, entre más sílabas y pausas se produzcan en la misma unidad de tiempo, menor será el tiempo que le corresponda a cada una y mayor será la velocidad del mensaje; por el contrario, si los segmentos son largos, emplearán más tiempo y volverán lenta la producción. Para cambiar la rapidez de un mismo discurso es necesario reestructurarlo, ya que no se trata de un hecho mecánico, sino de un acto comunicativo, por lo que no es suficiente alargar o acortar sílabas, o agregar o quitar pausas, hay que reorganizar y distribuir el tiempo de las enunciaciones (Blondet et al., 2003, p. 2), pues en los enunciados siempre aparecen segmentos contiguos, cuya duración depende, entre otras cosas, de la posición que tengan dentro de la cadena hablada.

2.1.1.1 Duración

La *cantidad* es una cualidad física del sonido que se refiere a la duración de los sonidos y de las sílabas (Hidalgo Navarro y Quilis Merín, 2004, p. 46), cuya dimensión psicológica es la persistencia del sonido (Quilis, 1981, p. 75), más aún, es una característica que destaca fácilmente para los oyentes de cualquier mensaje porque “el tiempo es [lo] que mejor se percibe a través de la audición” (Quilis, 1981, p. 81). Este rasgo “obedece en español a razones meramente fonéticas [ya que] las modificaciones de la cantidad española afectan [...] a la forma y fisionomía de las palabras, pero no alteran la significación de estas” (Navarro Tomás, 2004, p. 197), mientras que en otras lenguas, como el latín o el finés, se distingue entre sonidos largos y breves para cambiar el significado de las palabras, desde un punto de vista fonológico, se podría interpretar este fenómeno como la correlación de la homologación de vocales o la geminación de consonantes (Alarcos Llorach, 1974, pp. 91-93). Sin embargo, es preferible hablar de cantidad, porque no

necesariamente hay dos sonidos gemelos, sino uno alargado (Martínez Celdrán, p. 248), y, para que puedan aparecer sonidos gemelos, es necesario que haya también un quiebre silábico que los separe.

La *cantidad* se encuentra en el grupo de rasgos suprasegmentales junto con la velocidad de elocución, es decir, no es necesariamente un componente del tiempo, sino que ambos rasgos coexisten en la oralidad y se pueden relacionar porque la dimensión que los caracteriza es el tiempo (Llisterri, 2016, pp. 14-16). Hay que mencionar, además, que las distintas duraciones que presentan las sílabas modifican significativamente la medida que se obtiene al calcular la velocidad de habla, ya que la distribución de los segmentos fónicos concurrentes en una misma unidad de tiempo es irregular.

Si bien “[casi] cualquier sonido o grupo de sonidos puede ser alargado cuanto lo permite el aire acumulado en los pulmones” (Martínez Celdrán, 1983, p. 244), existe la creencia de que hay una correspondencia entre cantidad y acento que se mantiene más o menos constante (D’Introno, Teso, y Weston, 1995, p. 129), pues se supone que “la posición acentuada favorece la duración y, entre las tónicas, la posición aguda es la más larga” (Martínez Celdrán, p. 246). De ahí que se considere que las sílabas acentuadas son siempre largas y las inacentuadas siempre breves; sin embargo, en español, todas las sílabas resultan aproximadamente iguales (Navarro Tomás, 2004, pp. 205-206) y su duración depende solo del acto concreto de habla, ya que el contexto y el estado físico y emocional de los hablantes varían de una situación a otra.

2.1.1.2 Pausas

La pausa se puede definir como “la interrupción temporal del proceso de producción de los sonidos, es decir, de la fonación y de la articulación” (Gil Fernández, 2007, p. 298), en otras palabras, es el

silencio que ocurre entre los diversos grupos de sonidos, que sirve para delimitar el grupo fónico, y que puede ocurrir por razones fisiológicas, como respirar, o puede tener alguna función lingüística, ya sea demarcativa o estilística (Llisterri, 2016, p. 14). Aunque se escapa parcialmente a esa acepción el concepto de pausa llena, la cual tiene contenido sonoro, o sea, vocalizaciones que varían entre las diferentes lenguas, e incluso de un dialecto a otro (Gil Fernández, 2007, p. 299), entre las que se pueden incluir alargamientos vocálicos o elementos vocales, como [e:] [m:] en español (Llisterri, 2016, p. 15), o [s:] [o:] en coreano. Sin embargo, para esta investigación solo se consideran como pausas aquellas que estén vacías¹, por lo que los elementos sonoros se incluyen dentro del compendio de sílabas.

El número de pausas dentro de los enunciados modifica la velocidad de habla porque aumenta la duración de la unidad de tiempo y modifica la distribución de la cantidad de los segmentos. Es una estrategia utilizada para cambiar la rapidez de las diferentes modalidades discursivas (Contreras, 2013, p. 240), ya que regularmente los contextos formales requieren un habla más pausada que la de una conversación coloquial. Por otra parte, “combinadas con las variaciones en la velocidad de articulación y con otras claves rítmicas, las pausas, finalmente, ayudan a regular los turnos de palabra conversatorios y pueden condicionar la recepción positiva o negativa del mensaje por parte del oyente” (Gil Fernández, 2007, p. 305), aunque no siempre aparezcan como límites de turno de habla, pues dentro de estos pueden aparecer separando un grupo fónico de otro. También es importante decir que, al igual que el rasgo de cantidad, las pausas no están subordinadas de ninguna manera a la velocidad de elocución, ya que ni siquiera coaparecen en todas las ocasiones.

¹ Las pausas vacías son aquellas que separan grupos fónicos con silencio.

2.1.2 La velocidad de habla como herramienta para estudiar la oralidad

Como se ha mostrado, la velocidad de habla varía según diversos factores, ya sea por el temperamento, el carácter o las emociones que invadan al individuo en el momento de la producción, y, aunque no hay suficientes datos comparativos entre distintas lenguas (Hidalgo Navarro y Quilis Merín, 2004, p. 46), se cree que existen diferencias cuantitativas en la velocidad que distinguen a lenguas y dialectos. Por ejemplo, se ha encontrado que la velocidad de elocución promedio del español es menor que la del francés (Manuel Jerez, 2003, p. 58), mientras que, dentro del mundo hispanohablante, más concretamente entre San José, Costa Rica y Barcelona, España, se determinó que no existe diferencia significativa en la velocidad que pueda diferenciar una variante de otra (Schwab, 2015, pp. 136-137).

Por otra parte, en el mismo estudio comparativo (Schwab, 2015, pp. 136-137) se encontró un correlato de edad en el que, a mayor cantidad de años, disminuye la velocidad de elocución, y que la disminución no se debe a ningún tipo de déficit cognitivo o visual. Al mismo tiempo, dicho correlato de edad permite hacer investigaciones sobre la velocidad de habla en grupos de sujetos que padecen algún tipo de enfermedad o deterioro cognitivo; por ejemplo, en la Universidad de Salamanca, se ha estudiado los casos de las enfermedades de Alzheimer y Parkinson. En el primer caso, se intentó comprobar si la velocidad de elocución funciona para predecir el desarrollo de la enfermedad; sin embargo, aunque sí se presenta deterioro y disminución de la rapidez, los resultados no permiten concluir que sea un método de alta exactitud para el diagnóstico temprano del Alzheimer (Martínez Sánchez et al., 2013, p. 328-239). En el segundo caso, se encontró que la disminución del tempo contrasta ampliamente con la velocidad de un envejecimiento normal, debido al deterioro de los músculos necesarios para la articulación de sonido (Martínez Sánchez et al., 2016, p. 467-469).

Otros estudios al respecto han registrado velocidades diferentes que dependen de la situación comunicativa; por ejemplo, una investigación sobre la variación acústica inter e intra locutor (Battaner et al., 2003, p.2) encontró que la rapidez de los hablantes se mantiene más o menos estable a pesar de que se les haya grabado en sesiones separadas temporalmente. En cuanto a la velocidad de habla en lectura, se ha llegado a la conclusión de que la narración de noticias y de prosa son muy similares, pero que estas dos son más rápidas que la lectura de poesía (Contreras, 2013, p. 239).

En áreas más específicas, se ha determinado que “la velocidad de elocución con que se lee [un] anuncio puede incrementarse de forma súbita al interpelar al destinatario” (Machuca y De la Mota, 2006, p. 1137); y que los profesores de español como segunda lengua tienden a disminuir la rapidez al introducir tópicos, dar consignas y durante explicaciones gramaticales, mientras que suelen aumentar la velocidad cuando se trata de alguna explicación sociocultural o de algún tema que no se trata en clase en ese momento (Cederborg y Chávez Aguilar, 1992, p. 296).

Por último, se sabe que las investigaciones fonéticas con la finalidad de desarrollar programas de síntesis de voz proponen que es posible generar distintas velocidades de habla a partir de las alteraciones en la duración silábica o difonemática, todo ello sin discriminar entre las diferentes posiciones de sílaba ni entre la duración concreta que cada una presenta (Blondet et al., 2003, p. 2).

2.2 La velocidad de habla en la conversación

Para visualizar la influencia de la velocidad de habla en la conversación es necesario partir de una definición de «conversación». Entiendo este concepto de manera general como una interacción hablada, más específicamente como un acto comunicativo sin planificación previa y de finalidad

interpersonal (Hidalgo Downing, 2003, p. 84-85) que se organiza “según un sistema de cambio de turnos por el cual se da una alternancia constante entre hablantes y en la que los papeles de hablante y oyente se encuentran en continuo intercambio” (Hidalgo Downing, 2003, p. 87). Si bien, un acto espontáneo carece de planificación, los hablantes son capaces de adaptarse a cada situación nueva porque no solo atienden a las propiedades únicas e irrepetibles que caracterizan a cada una de ellas, sino que siguen la tendencia natural de conceptualizarlas como un caso particular de algún tipo de contexto genérico más conocido (Escandell Vidal, 2014, p. 57). Además,

de alguna manera, [explícita o tácita], se reconoce por ambas partes que, permaneciendo constantes todas las condiciones restantes, la transacción habría de proseguir con un estilo adecuado, a menos que unos y otros se pongan de acuerdo en darle fin. No se trata ni de irse sin más ni de comenzar de improviso a hacer otra cosa (Grice, 2012, p.528).

Durante el transcurso de una conversación, la velocidad de elocución de los participantes varía considerablemente de un mensaje a otro, incluso dentro del mismo enunciado de un sujeto, tales variaciones tienen la función de orientar al oyente en la interpretación del discurso (Asuaje et al., 2005, p. 454). Las modificaciones se pueden deber a varios factores, tanto internos como externos al mensaje, en el caso del habla espontánea puede ocurrir que la información más relevante se produzca de manera más lenta, mientras que la información de carácter más incidental presente una velocidad mayor (Asuaje et al., 2005, p. 453). También es notoria la diferencia entre los registros formal e informal, ya que en el primero se suele articular menos sílabas por segundo que en el otro (Contreras, 2013, p. 240), debido a que en el registro formal se cuida mucho más la producción de sonidos y pausas. Además, en el habla cuidada es posible la concurrencia de dos fonemas homólogos que dan origen a una prolongación cuantitativa del sonido resultante, que en una conversación normal bien podría aparecer como un solo sonido (Quilis, 1999, p. 376), aunque este último caso, al menos en esta tesis, no se considera como un alargamiento porque entre ambos

sonidos existe un quiebre silábico.

2.2.1 Los turnos de habla

La conversación está conformada por un conjunto de turnos de habla sucesivos, es decir, por una serie de espacios interactivos ocupados por un solo hablante al que su o sus interlocutores le prestan atención simultánea y manifiesta (Musselman Shank, 2006, p. 159), ya que “las contribuciones de los participantes han de encajar unas con otras, siendo mutuamente independientes” (Grice, 2012, p. 528). Los turnos pueden ser largos o breves, y estar formados por un enunciado, varios o ninguno (Hidalgo Downing, 2003, p. 88), en otras palabras, puede estar constituido por varios grupos fónicos, uno solo, o incluso una interjección basta para completar un turno de habla.

Para que uno de los participantes mantenga el turno de palabra es necesario que no siga ninguna estrategia de distribución del turno, es decir, que nunca deje explícita alguna señal para que otro tome la palabra, mientras que, quien retenga el turno ha de utilizar algún marcador, que podría ser cualquier indicio para que el oyente interprete la dirección adoptada, ya sea una pausa breve, un grupo fónico trunco, un alargamiento, un tonema suspensivo (Musselman Shank, 2006, pp. 161-162), o cualquier otro marcador discursivo. En este caso solo se mencionan los marcadores fónicos ya que esta investigación se ubica más en el área de la forma que en la del contenido.

Una vez que un hablante ha terminado su turno, alguien más debe tomarlo o terminar la conversación, pero esto requiere que, además de un cambio de hablantes, haya reconocimiento y aceptación del cambio por parte del o los interlocutores (Musselman Shank, 2006, p. 159). Para marcar la consecución del turno se utilizan estrategias ligeramente diferentes a las que se usan para mantenerlo; por ejemplo, una pausa larga o alguna otra clave entonativa, no verbal, gestual, el uso de ciertos marcadores o incluso un breve resumen final de la intervención pueden ser indicación

de que el turno ha finalizado, aunque, aun sin ninguna señal, los hablantes pueden anticipar la terminación de la intervención de su interlocutor (Hidalgo Downing, 2003, p. 87-88). Además de todo ello, cada turno de habla presenta una velocidad de elocución distinta, que depende de la rapidez utilizada en el mensaje inmediatamente anterior.

2.2.2 La intención

Se dice, desde la perspectiva de la producción del mensaje, que hay comunicación cuando hay una intención manifiesta por parte de un emisor de producir un gesto visible para que, a partir de ese gesto, el receptor pueda inferir los contenidos que pretendía transmitirle (Escandell Vidal, 2014, p. 22). Por lo tanto, los participantes de una llamada telefónica de extorsión se valen de la modificación de la velocidad de elocución, entre otras estrategias, para intentar transmitir, de manera implícita, un mensaje matizado para presionar o disuadir a su interlocutor, y que ello los conduzca a lograr su propio objetivo. De la misma manera, quien recibe el mensaje no solo decodifica la señal recibida, sino que trata de identificar el propósito comunicativo con el que aquella fue emitida (Escandell Vidal, 2014, p. 22), es decir, ambos participantes de la conversación reconocen que el mensaje llega cargado de las intenciones del otro, y es su deber comprenderlo para responder de manera adecuada, ya sea para cumplir los deseos propios o los del otro hablante.

El propósito comunicativo se puede entender como

la relación dinámica entre el emisor y el aspecto de su entorno sobre el que este quiere actuar, bien para producir cambios, bien para evitar que estos se produzcan: los cambios pueden afectar a la situación externa, o al destinatario, o a ambos (Escandell Vidal, 2014, p. 24),

por lo tanto, se relaciona con la intención en cuanto a que este es el fin buscado por aquella. Además, la finalidad funge un papel muy importante para el desarrollo de la conversación, ya que es el eje que regula el comportamiento de los hablantes y los conduce a través de una determinada

serie de estrategias que les permitan alcanzar sus objetivos (Escandell Vidal, 2014, p. 24) Sin propósito no habría necesidad de comunicarse.

El poder es otro factor que condiciona la forma de hablar de los hablantes,

lo que me interesa en el problema del discurso es el hecho de que alguien ha dicho algo en determinado momento. No es el sentido lo que pretendo poner en evidencia, sino la función que se puede asignar al hecho de que no haya sido dicho en este momento (Foucault, 2010, p. 741).

Cuando se entabla una conversación en la que uno de los participantes puede ejercer control sobre el comportamiento de la otra persona, existe una relación vertical asimétrica en la que destacan las diferencias de ambos sujetos (Musselman Shank, 2006, p. 171), lo cual implica que ambos se encuentran en límites distintos de expresividad. Por ejemplo, en el caso tratado en esta tesis, dos hombres negocian mediante llamadas telefónicas: quien tiene el poder es el extorsionador porque retiene en contra de su voluntad al familiar de su interlocutor; por lo tanto, el extorsionado debe ser cuidadoso al responder a las exigencias del otro para que la persona secuestrada no muera en el acto.

2.3 Las llamadas de extorsión

El medio en el que se desarrolla la comunicación “sea oral o escrito, se pone al servicio de la representación que se tiene de las diferentes situaciones y del tipo de registro requerido en cada una de ellas” (Escandell Vidal, 2014, p. 65); por ejemplo, el contexto del caso que se toca en esta tesis es muy específico pues se trata de un par de llamadas telefónicas de extorsión efectuadas por los mismos dos sujetos en ambas ocasiones, lo cual implica un medio oral de comunicación en el que falta la copresencia de ambos participantes. Por lo tanto, las estrategias fónicas, como la velocidad de habla, la entonación y el volumen, son prácticamente las únicas maneras de ampliar la significación de los enunciados, ya que no se intercambia ningún tipo de información

visualmente. Ahora bien, una conversación telefónica se abre con el par llamada-respuesta y, por lo regular, sigue la identificación de quien efectúa la llamada y el reconocimiento de quien la recibe; sin embargo, el caso de una llamada de extorsión es una transacción con cierta planificación previa, por lo que la apertura se reduce o no aparece explícitamente y la conversación se dirige rápidamente al motivo de la interacción (Hidalgo Downing, 2003, pp. 90-92).

Aunque es cierto que participar en una llamada de extorsión no es algo común, casi todas las personas han participado en un acto de compraventa y han realizado alguna llamada telefónica en su vida. De modo que, con esos dos antecedentes, es posible conversar en este nuevo contexto porque el modo en que los hablantes se comportan en cada situación está condicionado por las expectativas que se asocian a los diferentes tipos de intercambio (Escandell Vidal, 2014, p. 59). Por consiguiente, cualquier persona que reciba una llamada de extorsión conoce lo que ocurre generalmente en una llamada telefónica y espera que suceda lo mismo que en cualquier otra ocasión, igual que con el intercambio de bienes. Por otro lado, quien realiza la llamada ha aprendido los métodos específicos para extorsionar por teléfono, por lo que podría decirse que esa persona ya ha institucionalizado el intercambio, y que por eso la forma y el contenido de sus mensajes están regulados (Escandell Vidal, 2014, p. 60-61) con el fin de obtener un objetivo específico.

3 Corpus y metodología

De entre todas las posibilidades, se eligió el *tempo* o *velocidad de habla* y su influencia en la conversación como tema de esta tesis, pues gracias a la naturaleza del corpus, es posible estudiar aspectos que pertenecen exclusivamente a la oralidad del discurso. Sin embargo, hay inconsistencia terminológica en los pocos trabajos teóricos sobre el tema (véase 2.1), por lo tanto, para el cálculo de la velocidad, he tenido que implementar métodos y parámetros basados en los datos que pude recopilar tanto de la teoría como de otras investigaciones. De tal manera que el corpus está dividido en los turnos de habla de cada uno de los hablantes en las dos grabaciones, y dentro de cada turno se hizo una subdivisión en segmentos, es decir, sílabas y pausas, ya que a partir de esa división es posible calcular la velocidad de habla presente en cada turno, y observar la estructura de distribución temporal que ocurre dentro de ellos.

3.1 Corpus

El corpus oral alrededor del que se desarrolla esta investigación proviene de una fuente fortuita de origen jurídicamente privado, recuperado en el SUEDIF/UNAM. consta de dos grabaciones de conversaciones telefónicas en formato WAV, ambas fechadas el 6 de septiembre de 2008, y son parte de una negociación por motivo de secuestro entre dos locutores varones dentro del rango de edad de 20 a 35 años, habitantes del Altiplano Central mexicano, cuya habla se presenta en registro informal. La *grabación 1* tiene una duración de 2 minutos con 24 segundos, de los cuales aproximadamente 1 minuto con 9 segundos están ocupados por el *sujeto 1* (víctima de la extorsión), casi 54 segundos por el *sujeto 2* (extorsionador), y el resto del tiempo por pausas y ruido. Mientras que, la *grabación 2* cuenta con una duración de 3 minutos con 41 segundos, de los cuales

aproximadamente 1 minuto con 37 segundos están ocupados el *sujeto 1*, alrededor de 1 minuto con 44 segundos por el *sujeto 2*, y el resto por ruido y pausas.

3.2 Procedimiento

3.2.1 Anotación

Tabla 1. Alfabetos fonéticos para el español de México

AS	RFE	AFI
p	p	p
t	t	t
k	k	k
b	b	b
d	d	d
g	g	g
y	ÿ	j
c	ĉ	ç
f	f	f
s	s	s
x	x	x
m	m	m
n	n	n
ñ	ɲ	ɲ
l	l	l
r	r/ṛ	r/r
a	a	a
e	e	e
i	i	i
o	o	o
u	u	u

Antes de comenzar con la anotación del corpus, se elaboró un alfabeto fonético simplificado para el español de México, en el cual solo se utilizan signos del alfabeto latino, ya que utilizar los signos fonéticos puede representar un riesgo de pérdida de información durante la conversión de archivos. Por lo tanto, se sacrificó especificidad en pro de practicidad, lo que dio como resultado un alfabeto de 21 signos correspondientes al sistema fonológico del español de México, ya que no se hace distinción entre la vibrante simple y la múltiple dentro del alfabeto, solo es posible determinar la presencia de una u otra en los contextos específicos donde se presenten.

Cabe resaltar que la elección de los caracteres que representan a cada fonema se basa en el formato del alfabeto de la Revista de Filología Española (RFE), la transcripción, por su parte, tiene la forma utilizada en RFE en cuanto a la ausencia de quiebres silábicos, pero mantiene la acentuación usada en el AFI para resaltar las sílabas tónicas con el espacio creado debajo del apóstrofe, mientras que las pausas se marcan con el signo «#»,

solo si su duración promedio se considera breve, en caso contrario la marca es doble «##». En la Tabla 1 se presenta el alfabeto simplificado (AS) con sus signos correspondientes en RFE y AFI.

A continuación, anoté de manera separada las dos grabaciones con del programa Praat, la información de los sonidos segmentados se almacenó en dos archivos, uno para cada grabación, con terminación TextGrid, creados igualmente dentro del programa. Ya que las medidas necesarias para calcular la velocidad de habla solo son la cantidad de segmentos por turno de habla y la duración de los turnos de habla, se hizo una anotación en cuatro niveles numerados del 1 al 4 como se aprecia en la parte inferior de la figura 1.

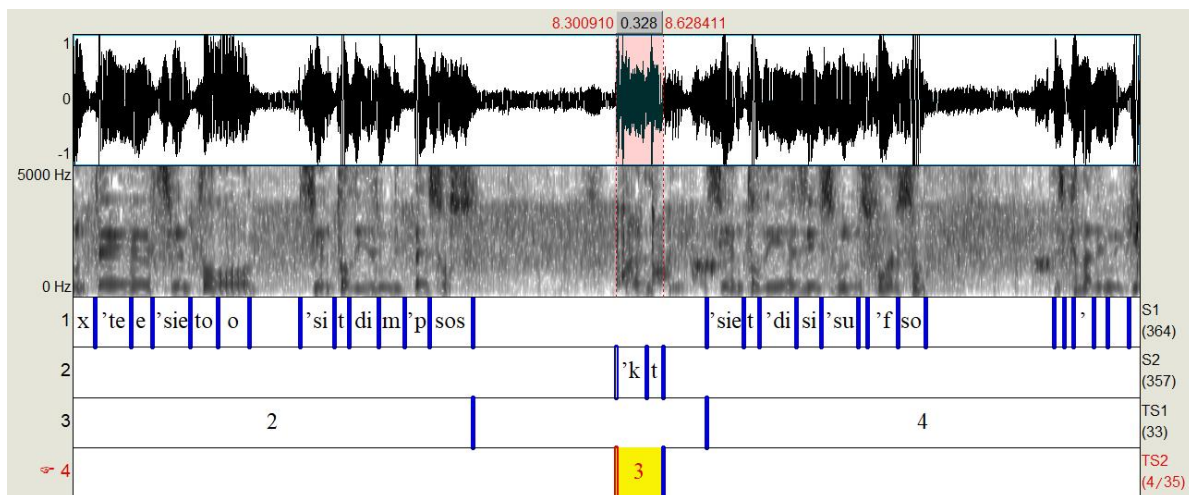


Figura 1. Niveles de anotación en Praat

Las dos primeras líneas, marcadas como S1 y S2, corresponden a los segmentos realizados por cada locutor, mientras que en la tercera (TS1) y la cuarta (TS2) se presentan numerados los espacios ocupados por cada turno de habla completo, cuyo límite suele ser una pausa seguida de la intervención de su interlocutor, aunque dichas pausas no se toman en cuenta para el procesamiento de datos, pues no aparecen dentro del tiempo utilizado durante las intervenciones de los hablantes. No obstante, los cambios de turno no siempre respetan la norma de intercambio y se presentan solapamientos como el que aparece en la figura 2 donde el turno 17 del *sujeto 2*

interrumpe el turno 16 del *sujeto 1*, o como en el caso de la figura 3 donde el turno 25 ocurre dentro del mismo lapso que el turno 24, por lo que en estos casos pueden aparecer turnos con números contiguos realizados por un mismo locutor.

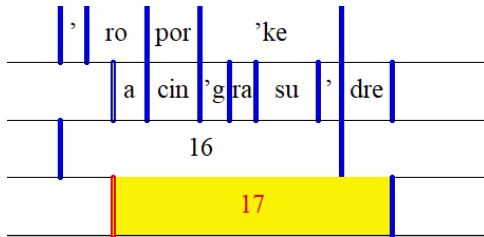


Figura 2. Solapamiento

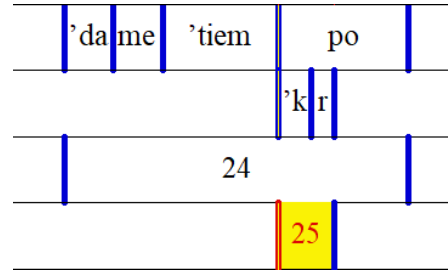


Figura 3. Solapamiento

3.2.2 Extracción de datos

Después de anotar todo el corpus, se utilizó un código script de Praat para acelerar la extracción de datos y crear un documento de texto que contuviera la duración de cada segmento presente en todos los niveles de anotación de las grabaciones (ver figura 4).

```
# Duración:
start = Get starting point... tier interval
end = Get end point... tier interval
duration = end-start
midpoint = (start + end) / 2

# Guardar el resultado en un archivo de texto:
resultline$ = "'soundname$' 'tier' 'label$' 'duration' 'start' 'end' 'newline$'"
fileappend "'resultsfile$'" 'resultline$'

# Selecciona el TextGrid para iterar el siguiente intervalo:
select TextGrid 'soundname$'
```

Figura 4. Script

Para ejecutar el script, como se puede observar en la figura 5, es necesario introducir. Se hizo lo mismo con los cuatro niveles de cada una de las grabaciones y toda la información se guardó en el mismo documento de texto (figura 6).

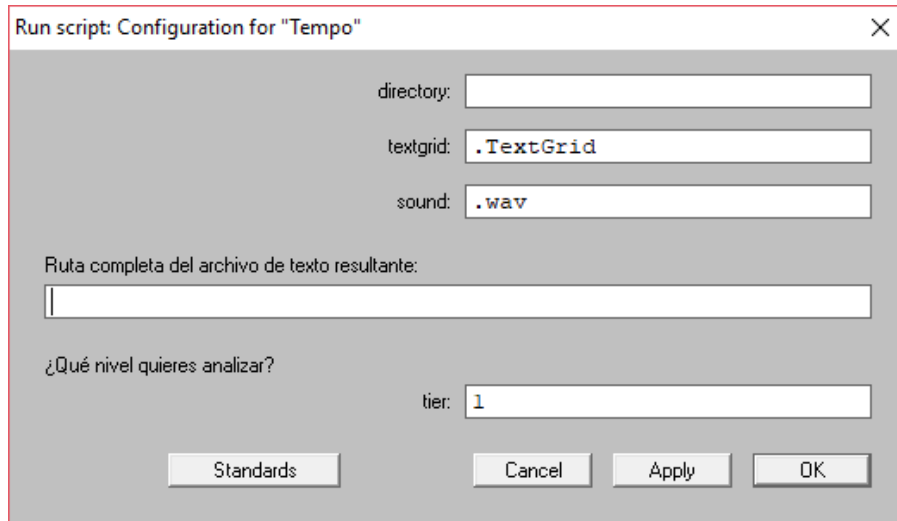


Figura 5. Datos para ejecución del script

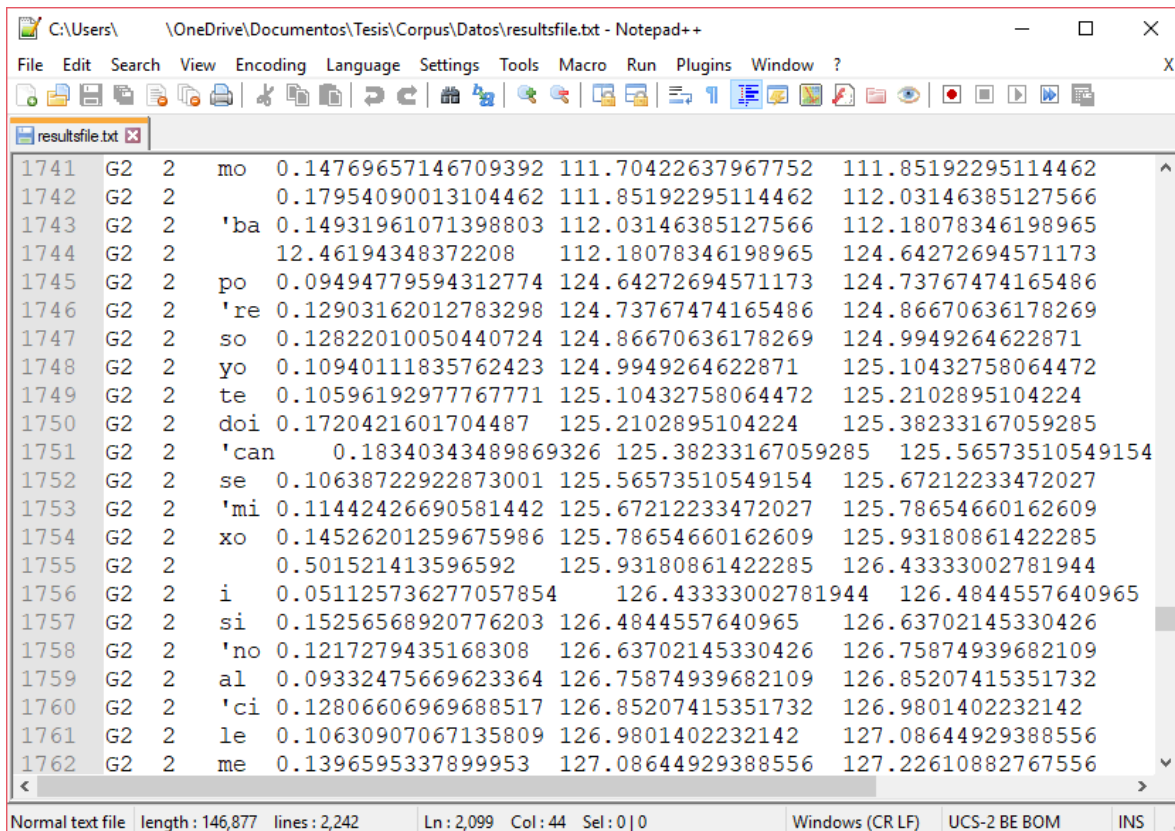


Figura 6. Archivo de resultados

Posteriormente, para facilitar el cálculo y la extracción de los datos de turnos y segmentos mediante un procesamiento digital, se trasladaron los datos a un archivo CSV (documento

delimitado por comas), en el que la columna «sujeto» corresponde a los niveles 1 y 2 de la anotación en Praat, y la columna «turno» a los niveles 3 y 4 (tabla 2). Durante el proceso de anotación, hubo un segmento² de la grabación 2 donde resultó complicado determinar los límites de las sílabas realizadas por el *sujeto* 2, por lo tanto, se etiquetó con el símbolo «-» y se mantuvo como parte del corpus únicamente para no alterar la duración del turno al que pertenecía, pero no se tomó en cuenta para el cálculo de la velocidad de habla.

Grabación	Sujeto	Turno	Segmento	Duración	Inicio	Fin
1	1		–	2.76	0.00	2.76
1	1	2	ps	0.37	2.76	3.13
1	1	2	ya	0.28	3.13	3.42
1	1	2	'yo	0.15	3.42	3.57
1	1	2	ya	0.08	3.57	3.66
1	1	2	te	0.20	3.66	3.85

3.2.3 Determinación de duración segmental

Con la ayuda de la aplicación Jupyter Notebook, se obtuvieron las estructuras fónicas de cada uno de los turnos de habla a partir de la duración promedio de los segmentos en cuatro diferentes versiones: promedio por grabación y sujeto (PGS), promedio por grabación (PG), promedio por sujeto (PS) y promedio total (PT), ya que establecerse en una sola de esas medidas podría sesgar los resultados. Dentro de cada lista se modificó el espacio en blanco que marcaba tipográficamente las pausas vacías por el símbolo '#'. Para la versión PGS se trasladaron los datos del documento CSV con los datos a cuatro listas de diccionarios nombradas G1S1, G2S1, G1S2 y G2S2, donde la S corresponde a «sujeto» y la G a «grabación». Del mismo modo, para la versión PG se crearon las listas G1 y G2, para la versión PS fueron S1 y S2, y para la versión PT solo se necesitó una

² <arriesgando la vida de tu hijo>, G2S2T13 [00:01:14 – 00:01:15].

lista.

Después se obtuvo una lista de segmentos sin repetir para cada una de las sublistas, con ellas se realizó el conteo de la frecuencia y se calculó el promedio de cada segmento, sin contar aquellos que separan los turnos de un mismo sujeto, ya que son fragmentos imprecisos que no tienen correspondencia en ambos locutores dentro de la línea temporal (ver segmentos vacíos en figuras 1, 2 y 3). Posteriormente, los datos «segmento», «promedio» y «frecuencia», se guardaron en nuevas listas diccionarios para confrontarlos con los registros completos de los segmentos, si la duración del segmento es mayor al promedio se le considera larga y se marca tipográficamente con mayúsculas, si es menor o igual al promedio se le considera breve y se mantiene en minúsculas.

Los datos de los promedios se guardaron en hojas de cálculo, donde además se hizo el conteo de los segmentos totales y los segmentos sin repetir, y se calculó la duración promedio para un segmento en cada lista. Sin embargo, este método se desechó casi en su totalidad debido a la falta de muestras de todos los segmentos, ya que había algunos que solo aparecían una vez, y, a pesar de ser muy largos, se les clasificaba como breves (ver tabla 3). Por lo tanto, se mantuvieron las etiquetas PGS, PG, PS y PT y se calculó un solo promedio de duración de sílaba para cada lista o sublista en lugar de un promedio para cada segmento. A partir de ese punto medio se determinó la cantidad, larga o breve, de los segmentos, como se muestra en la tabla 4, donde presentan los mismos segmentos de la tabla 3 para hacer una comparación.

Tabla 3. Dos segmentos del *sujeto 2* en la *grabación 1* que presentan irregularidad en el etiquetado

Turno	Segmento	Duración (c.s.)	Promedio (c.s.)	Frecuencia	Cantidad
11	nas	39	39	1	Breve
19	dre	30	21	8	Larga

Tabla 4. Dos segmentos del *sujeto 2* en la *grabación 1* etiquetados a partir del mismo promedio

Turno	Segmento	Duración (c.s.)	Promedio (c.s.)	Cantidad
11	nas	39	16	Larga
19	dre	30	16	Larga

En el caso particular de las pausas se mantuvo el promedio del cálculo de duración para cada segmento, ya que su duración es mucho más variable y requirió una medición individual. Además, se realizó un procedimiento aparte para apreciar el rango de variabilidad existente alrededor de la media, lo cual fue posible ya que el número de muestras presente en el corpus es el más elevado. Para ello se separaron los registros de la lista original donde se guardaron los datos de todos los segmentos y se crearon cuatro listas para el conjunto GS (grabación y sujeto), dos para G (grabación), dos para S (sujeto) y una para el total de pausas (T). En cada una de dichas listas cada elemento según su cantidad, se marcó la pausa más larga y la más breve en cada una de ellas, asimismo se marcaron todos los elementos que presentaron diferencias de cantidad de una lista a otra, causadas por los distintos promedios de duración. La tabla 5 muestra algunas pausas que presentan discrepancias en la información de cantidad, pues a partir del promedio GS y G se consideran breves, pero con el promedio S y T se etiquetan como largas.

Tabla 5. Pausas con etiquetas irregulares del sujeto 1 en la grabación 1

Grabación	Sujeto	Turno	PGS (s.)	PG (s.)	PS (s.)	PT (s.)	Duración (s.)	Inicio	Fin
1	1	29	0.65	0.59	0.53	0.46	0.54	127.83	128.37
1	1	6	0.65	0.59	0.53	0.46	0.55	33.62	34.16
1	1	4	0.65	0.59	0.53	0.46	0.56	13.19	13.76

3.2.4 Transcripción automática

Una vez que se clasificaron los segmentos por cantidad, se procedió a crear una transcripción automática de cada turno de habla en las cuatro versiones de estructura fónica. Las versiones PGS,

PS y PT adquirieron la misma división por grabación que mantenía PG para que los turnos tuvieran consecución lógica y no se repitiera o eliminara alguno (figura 7). A continuación, se creó un documento en Excel para cada grabación donde se introdujeron los datos de sujeto, número de turno, duración del turno, cantidad de segmentos, velocidad de elocución, inicio, final y transcripción en los cuatro formatos para comparar las estructuras de cantidad obtenidas, observar la distribución del tiempo y comprobar si había concordancia entre ellas.

PGS	9	17	te'diXEke'NOke'riketeKO'liera'nadieno'puto
PGS	14	19	per'noMES'toipa'SANDode#de'MADA##no'esno'en
PGS	27	1	da
PGS	10	24	es'toiaSIEN'tolopo'sible'nadiemes'tako'LIANDO#'tulostas'BIENdo'nadIE
PGS	25	2	'kiero
PGS	28	7	aCIN'garatu'maDRE
PGS	15	12	acin'garas#aCIN'garasu'maDRE
PGS	23	5	'soilaMU'ñeKA
PGS	26	12	'kierosa'kara'LINDIO##iasu'XENTE
PGS	3	2	'KUANTo
PGS	7	20	po'REso'PUTOsea'basapo'neREN'riesgola'bidade'TUIxo

Figura 7. Transcripción automática según PGS

En las tablas 6 y 7 se presenta la información del *turno 29* de la *grabación 1*, en las que se puede observar irregularidad en las diferentes transcripciones, como ya se había mostrado en la tabla 5 de pausas. Los segmentos que resultaron diferentes de una transcripción a otra se señalaron con negritas. Las sílabas que de manera conjunta constituyen una unidad de sentido en la que la velocidad del discurso se acelera o disminuye también se señalaron tipográficamente según su tipo, las sílabas breves (aceleración) se marcan con cursivas y las sílabas largas (desaceleración) se marcan con subrayado. No se marcan enteros todos los fragmentos con alargamientos o acortamientos porque, aunque el sonido es un continuum, no es posible interpretar palabras desmembradas, por lo mismo, tampoco es posible hacer distinciones de velocidad con monosílabos, ya que ello implicaría seccionar cada turno de habla en pares de sonidos sin significado.

Tabla 6. Datos del turno 29 de la grabación 1

Sujeto	Turno	Duración (s.)	Segmentos	Pausas	Velocidad (s./s.)	Inicio (s.)	Fin (s.)
S1	T29	4.84	25	1	5.16	124.05	128.90

Tabla 7. Transcripciones del turno 29 de la grabación 1

Promedio	Transcripción
PGS	'dame' <u>TIEMPO</u> 'dame'daMEUNra'tito'MASmMIfa'miliame'batra'ER#'yolos'toi
PG	'dame' <u>TIEMPO</u> 'dame'daMEUNra'tito'MASmMIfa'miliame'baTRA'ER#'YOLOS'toi
PS	'dame' <u>TIEMPO</u> 'dame'daMEUNra'tito'MASmMIfa'miliame'batra'ER##'YOLOS'toi
PT	'dame' <u>TIEMPO</u> 'dame'daMEUNra'tito'MASMMIfa'miliame'baTRA'ER##'YOLOS'toi

3.2.5 Cálculo de velocidad de habla

Después, se calculó, para cada locutor, la cantidad de turnos en la grabación, la duración total de la intervención, el promedio de duración de los turnos de habla y la velocidad promedio de las elocuciones, datos con los que se realizaron distintas gráficas de líneas, en las que además se determinó la tendencia de velocidad de cada grabación. Finalmente, con todo ello se pudieron comparar las fluctuaciones en la velocidad de habla en el curso de la conversación, es decir en el contraste existente entre turnos de habla adyacentes tomando en cuenta la velocidad y el sentido del mensaje. Además, dentro de las realizaciones de un mismo locutor también se puede apreciar cómo esos cambios influyen en la duración de ambos audios y en la información que se resalta o se oculta dentro de la línea temporal.

3.2.6 Herramientas

❖ Praat

Software de uso libre desarrollado por Paul Boersma y David Weenink del Instituto de Ciencias

Fonéticas de la Universidad de Ámsterdam. Permite analizar el habla con un oscilograma y un espectrograma, donde se tiene la opción de observar también los datos de tono, formantes e intensidad, además de poder anotar segmentos del audio en varios niveles. También tiene usos más especializados como el aprendizaje por redes neuronales, funciones estadísticas, creación de imágenes y programación, ya sea por medio de algún *script* que contenga instrucciones ejecutables específicamente para Praat, o por la adición de los módulos para escribir en C o C++ (Boersma & Weenink, 2017).

❖ Jupyter Notebook

Según su página web, el Jupyter Notebook es una aplicación para explorador de uso libre que permite crear y compartir, códigos, ecuaciones, visualizaciones y textos narrativos. Entre sus varios usos incluye: limpieza y transformación de datos, simulación numérica, modelación estadística, visualización de datos, aprendizaje de máquinas, entre otros. Soporta alrededor de 40 lenguajes de programación, incluyendo Python, que es el que yo utilicé, R, Julia y Scala. Su interfaz es mucho más amable que la programación tradicional, ya que es posible ejecutar partes del código de manera individual, lo cual facilita la corrección de errores (Project Jupyter, 2018).

❖ Excel

Software de hojas de cálculo que forma parte de la suite Microsoft Office capaz de aprender patrones y organizar datos, además de crear gráficos y tablas para presentar la información, y crear pronósticos a partir de las tendencias (Microsoft, 2018).

4 Resultados

4.1 Distribución temporal interna al turno de habla

La distribución del tiempo dentro de cada turno de habla depende de la media empleada para determinar la cantidad de cada segmento, por lo tanto, como se ha mencionado anteriormente, el estándar de duración promedio de los segmentos se calculó desde las cuatro versiones en las que se distribuye el corpus: grabación y sujeto (GS), grabación (G), sujeto (S) y total (T), lo cual no quiere decir que el corpus sea la suma de esas listas, sino que cada una de ellas es el corpus entero ordenado de manera distinta. A partir de ellas se calculó la duración promedio de los segmentos, por un lado, los elementos que carecen de articulación y que presentan un amplio rango de variabilidad amplio (pausas), por el otro, todos los grupos de sonidos sin hacer distinciones debido a sus características articulatorias (sílabas), pues no hay suficientes muestras de cada una de ellas.

La división del corpus y la presentación de resultados en cuatro versiones se debe a que las intervenciones de los locutores no se realizan con la misma cantidad de segmentos ni en el mismo período de tiempo, de tal manera que, distribuir el corpus de distintas formas permite observar las diferencias en las estructuras fónicas presentes en cada turno de habla, ya que cada postura se elabora desde una suposición distinta. Si partimos de la creencia de que cada acto de comunicación es diferente y que cada hablante tiene sus propias estrategias para repartir el tiempo, entonces la versión PGS (promedio por grabación y sujeto) del corpus nos mostrará el comportamiento de los sujetos de manera individual y además podremos ver cómo varía ese comportamiento de una grabación a otra y de un hablante a otro. De manera similar, si aceptamos como cierto que la producción de un hablante depende del intercambio de mensajes, es decir, que está condicionada por la respuesta de algún interlocutor, quiere decir que cada acto comunicativo está aislado y que

no es relevante la variación intra hablante sino la variación inter hablante, por lo tanto, la versión PG (promedio por grabación) es la única que podría demostrarlo.

De igual manera, si asumimos que las realizaciones de los hablantes son siempre constantes independientemente del acto comunicativo, la versión PS tendría que demostrar que las estrategias para distribuir el tiempo son las mismas en ambas grabaciones, o por lo menos tendrían que resultar muy similares. Por otra parte, si partimos de la idea de que todos los hablantes de español presentan las mismas características generales de cantidad en todos los segmentos que producen y que todas las variaciones ocurridas a lo largo de los actos de comunicación son mínimas y dependen de causas ajenas a la lengua, ya sea intencionadas o no, se tendría que tomar el corpus completo para realizar cualquier cálculo y observar si los diversos rangos de cantidad se mantienen constantes.

A continuación, se presentan, en el mismo orden presentado en este apartado, las cuatro posturas que se utilizaron para determinar si un segmento es largo o breve, desde su distribución interna y sus características particulares, hasta sus similitudes y diferencias con las demás versiones.

4.1.1 Duración promedio por grabación y sujeto

El promedio por grabación y sujeto (PGS) consta de ocho medidas de duración de segmentos, como se muestra en la tabla 8, cuatro para pausas y cuatro para sílabas, todos medidos en centésimas de segundo (c. s.), cuyas etiquetas son G1S1, G1S2, G2S1 y G2S2, en las

Tabla 8. Duración promedio por grabación y sujeto

Lista	Duración promedio de sílaba (c. s.)	Duración promedio de pausa (c. s.)
G1S1	20	65
G1S2	16	51
G2S1	17	46
G2S2	16	34

que la letra G corresponde a «grabación» y la letra S a «sujeto». La división del corpus se hace de esta manera ya que cada una de las listas se crea con la idea de que la participación de los sujetos en cada grabación es un conjunto aislado que depende únicamente de la situación individual del

hablante.

Por lo tanto, las listas presentan aspectos que las distinguen, con los que es posible establecer relaciones de dependencia y similitud entre ellas, que caracterizan no solo a la dinámica de los actos comunicativos de las dos grabaciones, sino a la variación que existe inter e intra hablante. Se puede apreciar en la tabla 8 que las duraciones promedio del *sujeto 1* se distinguen entre sí mucho más que las del *sujeto 2*, ya que incluso presenta la misma medida para las sílabas de ambas grabaciones. Además, en la primera grabación la diferencia de duraciones promedio entre ambos sujetos es mucho mayor en la *grabación 1*, de 33 turnos de habla, que en la *grabación 2*, de 70 turnos de habla.

4.1.1.1 Grabación 1 sujeto 1

En la lista G1S1 se registraron 347 segmentos, de entre los cuales 23 son pausas, la más larga de ellas es la única que aparece dentro del *turno 8* con una duración de 2.54 segundos y, al ser mayor que el promedio de 65 centésimas de segundo (c. s.), se considera larga y se marca tipográficamente como «##»; mientras que la pausa más breve se encuentra en la sexta y última posición dentro del *turno 18* y dura solamente 21 c. s., por lo que su anotación tipográfica se mantiene como «#» al estar por debajo del promedio. En la tabla 9 se muestra la

Tabla 9. Pausas de G1S1 (s.)

Turno	Pausa	Duración	Inicio	Fin
8	##	2.54	45.29	47.83
18	##	1.42	97.32	98.74
18	##	1.26	100.46	101.72
14	##	0.93	89.58	90.51
4	##	0.90	10.47	11.38
6	##	0.82	29.62	30.44
4	#	0.56	13.19	13.76
18	#	0.55	106.41	106.96
6	#	0.55	33.62	34.16
29	#	0.54	127.83	128.37
18	#	0.45	103.04	103.50
2	#	0.45	3.85	4.30
12	#	0.45	65.58	66.03
12	#	0.42	63.97	64.38
6	#	0.40	28.34	28.74
6	#	0.36	36.33	36.69
2	#	0.36	5.72	6.08
10	#	0.36	52.98	53.33
18	#	0.34	96.03	96.38
4	#	0.34	14.42	14.76
14	#	0.34	88.48	88.82
12	#	0.30	63.29	63.59
18	#	0.21	107.65	107.85

Tabla 10. Sílabas de G1S1 (c. s.)

Turno	Segmento	Duración
18	'ser	49
20	'si	47
33	se	45
6	do	44
4	'toi	44
16	'ke	43
29	'er	43
14	da	43
12	me	43
29	'mas	43
4	d	6
12	s	6
4	al	6
6	'ri	6
18	mi	6
18	do	5
8	a	5
4	a	5
6	a	5
6	p	4

lista completa de pausas de G1S1, acompañadas de su turno, duración y posición en el archivo de audio. Las 324 sílabas de la lista, por su parte, tienen una duración promedio de 18 c. s., de todas ellas la más larga fue la tónica /'ser/ con 49 c. s., que solo se realizó en el *turno 18*, mientras que la sílaba más breve fue la átona /p/ de 4 c. s. del *turno 6*. En la tabla 10 se encuentran las 10 sílabas más largas y las 10 más breves.

De todo el conjunto de intervenciones realizadas por el *sujeto 1* en la *grabación 1*, habría que resaltar la concurrencia de uno de los elementos más largos y uno de los más breves dentro del mismo turno, pues en el *turno 18* de este audio se encuentran tanto la pausa más breve

como la sílaba más larga, además de otras cinco pausas (entre ellas la segunda y la tercera más largas) y dos de las sílabas más breves. Lo cual habla de una gran variación interna en la distribución del tiempo.

Por otro lado, el *turno 6* contiene la sílaba más breve, lo que indica que las enunciaciones de este sujeto también presentan elementos incompletos, ya que la sílaba en cuestión es una forma anómala constituida por un solo sonido consonántico: de la palabra «pues» se eliden los dos sonidos vocálicos y el que debería convertirse en núcleo (/s/) se traslada a la sílaba siguiente, /pwes/ > /p/.

Los segmentos más largos de esta lista se reparten entre las dos mitades del audio, por lo que se podría pensar que la variación temporal es estable. Sin embargo, según las muestras, se observa

que, mientras que la distribución de las pausas es uniforme, las sílabas más breves aparecen al principio de la grabación y las más largas desde la mitad hasta el final, lo cual indica que las intervenciones se hacen más lentas conforme avanzan, pues se asume que, entre mayor sea el tiempo que toma la realización de los segmentos, menor será la velocidad en que se enuncien los mensajes.

4.1.1.2 Grabación 1 sujeto 2

La lista G1S2 contiene 339 segmentos en total, 16 de ellos son pausas, cuya duración promedio es de 51 centésimas de segundo (c. s.), de las cuales la más larga cuenta con una duración de 1.26 segundos (s.) y aparece como la tercera de cinco pausas durante el *turno 31*, mientras que la más breve dura solo 12 c. s. y es la única que se presenta en el *turno 19*. En la tabla 11 se muestran todas las realizaciones de pausas de esta lista en medida de segundos, el turno al que

Turno	Pausa	Duración	Inicio	Fin
31	##	1.26	132.67	133.93
13	##	1.02	74.32	75.34
13	##	1.01	71.50	72.51
13	##	0.81	77.04	77.85
31	##	0.64	136.02	136.66
26	##	0.56	121.22	121.79
13	#	0.49	84.06	84.56
5	#	0.49	22.52	23.01
13	#	0.47	80.00	80.47
5	#	0.47	21.51	21.98
31	#	0.26	138.28	138.54
15	#	0.18	91.38	91.56
13	#	0.16	82.03	82.20
31	#	0.14	129.20	129.34
31	#	0.13	129.88	130.01
19	#	0.12	111.28	111.39

pertenecen y su posición en el audio. Los 323 segmentos restantes son sílabas, cuya duración promedio es de 16 c. s., la más larga de todas es la átona /nas/ que dura 39 c. s. y aparece en el *turno 11*, la más breve es la átona /a/ de 3 c. s. que aparece en el *turno 19*. En la tabla 12 se presentan las 10 sílabas más largas y las 10 más breves.

Según estas muestras, el *sujeto 2* en la *grabación 1* presenta concentración de pausas a la mitad y al final del audio, concretamente en los *turnos 13* y *31*, es decir, las pausas se reparten de manera

equitativa. Mientras que en la distribución de tiempo por sílaba se observa que los segmentos más breves se presentan hacia el final del audio y los más largos aparecen al principio, por lo tanto, hay una marcada tendencia a la aceleración. Además, en el *turno 19* coinciden los dos elementos más breves de la lista G1S2, lo que indica que el *sujeto 2* utiliza menos tiempo para realizar los segmentos de sus respectivas intervenciones a partir de la mitad del audio.

4.1.1.3 Grabación 2 sujeto 1

La lista G2S1 contiene 572 segmentos en total, de los cuales 39 son pausas con una duración promedio de 46

centésimas de segundo (c. s.), la más larga de ellas aparece dentro del *turno 5*, mientras que la más breve dura apenas 7 c. s. y ocurre en el *turno 55*. En la tabla 13 se muestran las 10 realizaciones más largas y las 10 más breves de pausas de esta lista en medida de segundos, el turno al que pertenecen y su posición en el audio. Los 533 elementos restantes son sílabas, las cuales presentan una duración promedio de 16 c. s., de entre ellas la más larga es la tónica /'kres/ que dura 48 c. s. y aparece dentro del *turno 3*, mientras que la más corta es la átona /e/ que aparece en el *turno 34* con duración de 4 c. s. Las 10 sílabas más largas y las 10 más breves se presentan en la tabla 14.

En los *turnos 3* y *5* aparecen los segmentos más largos de toda la lista G2S1, lo que indica que el *sujeto 2* comenzó sus participaciones a una velocidad lenta. Sin embargo, en el *turno 5* se encuentran también dos de las sílabas más breves de toda la lista, por lo que su velocidad podría

Tabla 12. Sílabas de G1S2 (c. s.)

Turno	Segmento	Duración
11	nas	39
19	dre	30
11	ka	29
11	'yo	27
5	'bron	27
7	to	27
11	'mas	26
11	'bron	26
21	'pu	25
1	xo	25
5	'bi	6
25	ro	6
26	ra	6
28	ra	6
13	a	6
13	ra	5
19	te	5
13	a	5
7	a	4
19	a	3

verse acelerada en ese punto, ya que el tiempo utilizado en la articulación es menor que en el *turno* 3. Por otra parte, los *turnos* 34 y 55 contienen los segmentos más breves, es decir, que la participación del *sujeto* 1 en la *grabación* 2 alcanza un pico de rapidez en la parte intermedia del audio.

Tabla 13. Pausas de G2S1 (s.)

Turno	Pausa	Duración	Inicio	Fin
5	##	1.15	35.05	36.20
3	##	1.03	26.04	27.06
14	##	1.00	74.67	75.67
11	##	0.88	62.81	63.69
11	##	0.85	64.92	65.78
30	##	0.79	122.47	123.26
3	##	0.78	27.52	28.30
28	##	0.78	100.04	100.82
32	##	0.75	133.94	134.70
51	##	0.72	178.58	179.30
1	#	0.23	1.80	2.03
5	#	0.20	33.26	33.45
30	#	0.18	116.76	116.94
38	#	0.17	159.33	159.50
5	#	0.14	32.48	32.62
28	#	0.11	102.45	102.56
28	#	0.10	97.44	97.55
38	#	0.10	158.23	158.33
1	#	0.07	0.28	0.35
55	#	0.07	189.98	190.05

Tabla 14. Sílabas de G2S1 (c. s.)

Turno	Segmento	Duración
3	'kres	48
34	no	47
11	'si	47
9	'si	44
11	'gual	41
24	no	39
3	so	38
32	'no	38
26	no	35
3	b	35
1	te	5
5	'ko	5
3	'ko	5
5	'bo	5
18	a	5
34	a	4
34	yo	4
30	o	4
16	ro	4
34	e	4

El comportamiento de *sujeto* en G2S1 es, en general, irregular, no mantiene una tendencia a la rapidez o a la lentitud durante todo el transcurso de sus intervenciones. A pesar de la existencia de dos focos, al principio (segmentos largos) y a la mitad de la conversación (segmentos breves), existen irregularidades, por lo que no se puede hacer una aproximación certera de la velocidad general con estos datos porque no concurren suficientes elementos similares y la distribución de las muestras no abarca toda la grabación. Pese a todo lo anterior, que no exista regularidad indica

que el *sujeto 1* realiza una variación estable en la distribución temporal, es decir, no se presentan cambios abruptos ni inclinación alguna hacia la aceleración o la lentitud.

4.1.1.4 Grabación 2 sujeto 2

La lista G2S2 contiene un total de 666 segmentos, de todos ellos 44 son pausas, cuya duración promedio es de 34 centésimas de segundo (c. s.), la más larga de ellas tiene una duración de 92 c. s. y aparece en la cuarta posición de cinco dentro del *turno 29*, dentro del mismo turno, pero en la segunda posición, se encuentra la pausa más breve con duración de 10 c. s. En la tabla 15 se muestran las 10 realizaciones más largas y las 10 más breves de pausas de esta lista en medida de segundos, el turno al que pertenecen y su posición en el audio. Dentro de los 622 segmentos restantes, hay uno

Tabla 15. Pausas de G2S2 (s.)

Turno	Pausa	Duración	Inicio	Fin
29	##	0.92	110.32	111.24
8	##	0.81	42.54	43.35
23	##	0.70	90.68	91.37
60	##	0.69	199.72	200.41
2	##	0.67	18.52	19.19
2	##	0.61	10.02	10.63
13	##	0.54	66.46	67.00
2	##	0.53	13.93	14.46
8	##	0.52	41.87	42.39
31	##	0.50	125.93	126.43
25	#	0.16	93.90	94.07
15	#	0.15	78.88	79.03
46	#	0.14	171.00	171.14
2	#	0.13	8.18	8.31
23	#	0.13	91.56	91.69
15	#	0.12	78.32	78.44
56	#	0.12	192.52	192.64
8	#	0.11	43.55	43.66
33	#	0.11	135.39	135.50
29	#	0.10	107.06	107.15

que tuvo que descartarse debido a la dificultad para determinar sus límites (ver 3.2.3), por lo que esas sílabas y las 85 c. s. que duran no se tomaron en cuenta para ningún cálculo. Las demás sílabas presentaron una duración promedio de 16 c. s., la más larga es la átona /t̃e/ del *turno 8* que dura 41 c. s., mientras que la más breve es la átona /a/ que aparece también dentro del *turno 8* y dura solamente 3 c. s. Las 10 sílabas más largas y las 10 más breves se muestra en la tabla 16.

En el *turno 8* aparecen tanto la sílaba más larga como la más breve de toda la lista G2S2, y en

Tabla 16. Sílabas de G2S2 (c. s.)

Turno	Segmento	Duración
8	ce	41
8	yo	37
8	treil	36
15	'dios	34
52	ke	32
2	'mi	32
8	dal	32
15	'dios	29
8	ta	29
29	ser	28
62	a	5
2	a	5
31	i	5
23	te	5
21	de	4
37	po	4
29	mo	4
35	a	4
29	a	4
8	a	3

el *turno 29* las pausas más larga y más breve, lo que implica que la variación interna del *sujeto 2* en la *grabación 2* no determina si el turno será rápido o lento, debido a la falta de una marcada tendencia hacia una u otra velocidad. Es posible que sus intervenciones sigan la misma línea desde el principio hacia la mitad, pero no hay indicios de su comportamiento en la segunda mitad de la grabación.

Si se toman en cuenta solo estas listas, se puede apreciar que los elementos más breves aparecen, en su mayoría, en los primeros turnos de habla, y que se alcanza un pico de rapidez hacia la mitad del audio. En general, el *sujeto 2* en la *grabación 2* no muestra cambios significativos a lo largo

de sus participaciones, por lo que la distribución temporal utilizada en G2S2 presenta una variación estable durante toda la grabación.

4.1.1.5 Distribución temporal general por grabación y sujeto

El número de segmentos presenta una distribución más o menos similar entre ambos locutores en los dos audíos, pues las listas donde se realizan más pausas (G1S1 y G2S2) son las que ocupan una mayor

Tabla 17. Segmentos del corpus por grabación y sujeto

Lista	Segmentos	Sílabas	Pausas
G1S1	347	324	23
G1S2	339	323	16
G2S1	572	533	39
G2S2	666	622	44

porción del corpus, en comparación con sus contrapartes de su respectiva grabación (ver tabla 17).

De igual manera, G1S1 y G2S2 tienen la mayor cantidad de sílabas articuladas, aunque la diferencia entre G1S1 y G1S2 es mínima ya que solo se distancian por un segmento, por lo que, como se puede apreciar en la figura 8, su porcentaje de información dentro del corpus es el mismo.

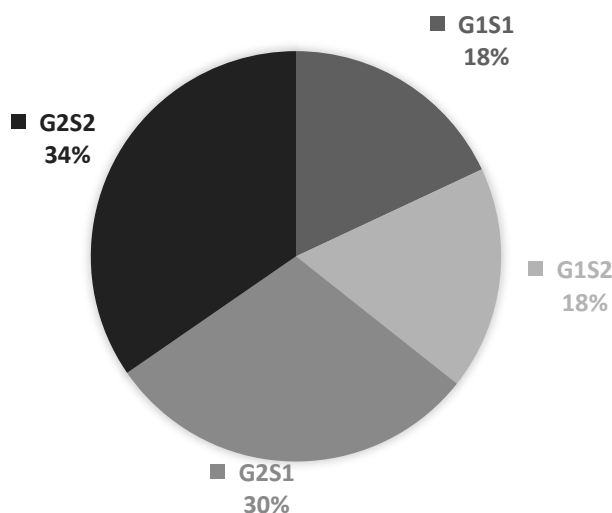


Figura 8. Distribución del corpus por grabación y sujeto

Los rangos de variación en la duración de los segmentos hablan de la tendencia de cada locutor hacia la aceleración o la desaceleración. En general, como se observa en la tabla 18, la duración de los segmentos es mayor para el *sujeto 1* que para el *sujeto 2*, es decir, el primero utiliza más tiempo por segmento para hablar que el segundo, y, por lo tanto, habla más lento.

Sin embargo, la diferencia de cantidad silábica de los dos locutores en la *grabación 1* es mayor que en la *grabación 2*, por lo que puede decirse que hay más estabilidad en el segundo caso y que ambos sujetos presentan aceleración con respecto al primer audio. Además, el rango de variación interna de los segmentos muestra que el *sujeto 1* tiende a los alargamientos, mientras que el *sujeto 2* suele hacer lo contrario.

Tabla 18. Variación de PGS

Lista	Pausa (s.)	Sílaba (c. s.)
G1S1	0.21 a 2.54	4 a 49
G1S2	0.12 a 1.26	3 a 39
G2S1	0.07 a 1.15	4 a 48
G2S2	0.1 a 0.92	3 a 41

El promedio por grabación y sujeto (PGS) implica que los distintos valores de velocidad y la distribución temporal interna de los turnos de habla son características inherentes a cada producción individual, por lo que no tendrían que verse influidos por factores externos al locutor en sí mismo. Empero, tanto la situación comunicativa como la respuesta de uno o más

interlocutores necesariamente afecta la estructura de los mensajes enunciados. Si bien tomar únicamente las cuatro listas (G1S1, G1S2, G2S1 y G2S2) de manera aislada pone de relieve sus cualidades, al hacer este tipo de análisis se obvian los estímulos que provocan dichas cualidades y los resultados se basan en suposiciones.

4.1.2 Duración promedio por grabación

El promedio por grabación (PG) se obtuvo mediante la división del corpus en dos partes, tomando cada grabación como un conjunto aislado en el que no se distingue entre un sujeto y otro. Se asume que las

Tabla 19. Duración promedio por grabación

Lista	Duración promedio de sílaba (c. s.)	Duración promedio de pausa (c. s.)
G1	18	59
G2	16	39

modificaciones temporales están condicionadas por la manera en que se desarrolla el acto comunicativo, por lo tanto, cada intervención depende únicamente del mensaje inmediatamente anterior. Los cuatro promedios de duración correspondientes a las pausas y sílabas de G1 y G2, donde la letra G corresponde a «grabación», se pueden observar en la tabla 19, donde las medidas aparecen en centésimas de segundo (c. s.).

4.1.2.1 Grabación 1

En la lista G1 se registraron en total 686 segmentos, de los cuales 39 son pausas, cuya duración en promedio es de 59 centésimas de segundo (c. s.), la más larga de ellas es la única que aparece en el *turno 8* del *sujeto 1*, que dura 2.54 segundos, mientras que la más breve fue realizada por el *sujeto 2* durante el *turno 19* con duración de 12 c. s. Las 10 realizaciones más largas y las 10 más breves de pausas se muestran en la tabla 20. Los 647 segmentos restantes son sílabas, cuya duración promedio se calculó en 18 c. s., la más larga de todas ellas es la tónica /'ser/ que dura 49

c. s. y fue articulada en el *turno 18* por el *sujeto 1*, la más breve es la átona /a/ con duración de 3 c. s. que fue articulada por el *sujeto 2* en el *turno 19*. Las 10 sílabas más largas y las 10 más breves se presentan en la tabla 21.

Tabla 21. Pausas de G1 (s.)						Tabla 20. Sílabas de G1 (c. s.)			
Sujeto	Turno	Pausa	Duración	Inicio	Fin	Sujeto	Turno	Segmento	Duración
1	8	##	2.54	45.29	47.83	1	18	'ser	49
1	18	##	1.42	97.32	98.74	1	20	'si	47
2	31	##	1.26	132.67	133.93	1	33	se	45
1	18	##	1.26	100.46	101.72	1	6	do	44
2	13	##	1.02	74.32	75.34	1	4	'toi	44
2	13	##	1.01	71.50	72.51	1	16	'ke	43
1	14	##	0.93	89.58	90.51	1	29	'er	43
1	4	##	0.90	10.47	11.38	1	14	da	43
1	6	##	0.82	29.62	30.44	1	12	me	43
2	13	##	0.81	77.04	77.85	1	29	'mas	43
1	4	#	0.34	14.42	14.76	2	13	ra	5
1	14	#	0.34	88.48	88.82	1	18	do	5
1	12	#	0.30	63.29	63.59	2	19	te	5
2	31	#	0.26	138.28	138.54	1	8	a	5
1	18	#	0.21	107.65	107.85	1	4	a	5
2	15	#	0.18	91.38	91.56	2	13	a	5
2	13	#	0.16	82.03	82.20	1	6	a	5
2	31	#	0.14	129.20	129.34	1	6	p	4
2	31	#	0.13	129.88	130.01	2	7	a	4
2	19	#	0.12	111.28	111.39	2	19	a	3

En la lista G1 hay una marcada diferencia en la distribución temporal de cada sujeto, ya que los dos segmentos más largos forman parte de las intervenciones del *sujeto 1*, mientras que los dos más breves son realizados por el *sujeto 2*. Por lo tanto, se asume que el primero tiende a la desaceleración y el segundo a la aceleración. Las posiciones de los turnos donde aparecen los segmentos más largos y más breves muestran que el *sujeto 1* mantiene lentitud constante porque el *turno 8*, donde aparece la pausa más larga, está en la primera mitad de la conversación, y el

turno 18, donde se encuentra la sílaba más larga, está en la segunda mitad. Por otro lado, los segmentos más breves se presentan en el mismo turno de habla por lo que no se puede determinar si hay una tendencia constante por parte del *sujeto 2*.

Desde otra perspectiva, las listas muestran un panorama más amplio con el que se afirma que el *sujeto 2* tiende a hablar mucho más rápido que el *sujeto 1*. Además, según las tablas 20 y 21, la tendencia a la aceleración del *sujeto 2* es tan constante como la tendencia a la desaceleración del *sujeto 1*, ya que ambos presentan elementos muy largos o muy breves respectivamente a lo largo de toda la *grabación 1*. No obstante, habría que resaltar el hecho de que las muestras presentadas en estas listas son mayoritariamente realizaciones del *sujeto 1*, lo cual sugiere que su variación es mucho más grande.

La *grabación 1* presenta discordancia y oposición entre los dos locutores que aparecen en ella, por lo que es normal que sus intervenciones se distingan en cuanto a su distribución temporal. A pesar de que el *sujeto 1* aparezca en ambos extremos de las listas de manera regular, el hecho de que el *sujeto 2* se mantenga inclinado hacia las menores cantidades de tiempo indica, por contraste, que el primer sujeto tiene una distribución temporal variada pero con inclinación hacia los alargamientos, ya que el espacio que ocupan las muestras del segundo sujeto en las listas de sílabas y pausas raramente se extiende demasiado hacia arriba de las tablas, por lo que se entiende que sus realizaciones no varían mucho y que se mantiene en la tendencia de los acortamientos.

4.1.2.2 Grabación 2

La lista G2 contiene 1238 segmentos, de los cuales 83 son pausas, cuya duración promedio es de 39 centésimas de segundo (c. s.), la más larga de todas ellas dura 1.15 segundos y fue realizada por el *sujeto 1* en el *turno 5*, la más breve duró solo 7 c. s. y fue realizada también por el *sujeto 1*

pero esta vez durante el *turno 55*. Las 10 realizaciones más largas y las 10 más breves de pausas se muestran en la tabla 22. Los 1155 segmentos restantes son todos sílabas, excepto uno que fue descartado por las razones mencionadas anteriormente (ver 3.2.3). La sílaba más larga de esta lista es la tónica /'kres/ articulada en el *turno 3* por el *sujeto 1* que dura 48 c. s., mientras que la más breve es la átona /a/ con duración de 3 c. s. que fue articulada por el *sujeto 2* en el *turno 8*. La lista con las 10 sílabas más largas y las 10 más breves aparece en la tabla 23.

Tabla 22. Pausas de G2 (s.)						Tabla 23. Sílabas de G2 (c. s.)			
Sujeto	Turno	Pausa	Duración	Inicio	Fin	Sujeto	Turno	Segmento	Duración
1	5	##	1.15	35.05	36.20	1	3	'kres	48
1	3	##	1.03	26.04	27.06	1	34	no	47
1	14	##	1.00	74.67	75.67	1	11	'si	47
2	29	##	0.92	110.32	111.24	1	9	'si	44
1	11	##	0.88	62.81	63.69	2	8	ce	41
1	11	##	0.85	64.92	65.78	1	11	'gual	41
2	8	##	0.81	42.54	43.35	1	24	no	39
1	30	##	0.79	122.47	123.26	1	3	so	38
1	3	##	0.78	27.52	28.30	1	32	'no	38
1	28	##	0.78	100.04	100.82	2	8	yo	37
2	15	#	0.12	78.32	78.44	2	21	de	4
2	56	#	0.12	192.52	192.64	1	34	yo	4
2	8	#	0.11	43.55	43.66	1	30	o	4
1	28	#	0.11	102.45	102.56	2	37	po	4
2	33	#	0.11	135.39	135.50	2	29	mo	4
1	28	#	0.10	97.44	97.55	1	16	ro	4
1	38	#	0.10	158.23	158.33	2	35	a	4
2	29	#	0.10	107.06	107.15	1	34	e	4
1	1	#	0.07	0.28	0.35	2	29	a	4
1	55	#	0.07	189.98	190.05	2	8	a	3

En la *grabación 2* se observa que el *sujeto 1* realiza tres de los cuatro segmentos de medidas más largas o más breves mencionados anteriormente, mientras que el *sujeto 2* solo articula la sílaba más breve de toda la lista. Además, las realizaciones de sílabas y pausas del primer locutor se

distribuyen equitativamente en las tablas, mientras que las del segundo locutor se mantienen en la parte inferior, correspondiente a las medidas más breves. Por lo tanto, se asume que el *sujeto 2* se mantiene en una tendencia hacia la aceleración y que el *sujeto 1* se muestra más flexible a los cambios, pero a la vez es inconsistente.

Sin embargo, hay que resaltar el hecho de que las muestras solo abarcan la primera mitad del audio, por lo que no sería pertinente afirmar que las tendencias de variabilidad y aceleración se mantienen durante toda la conversación. No obstante, se puede suponer que la aceleración en el *sujeto 2* se mantiene en toda la *grabación 2* porque la sílaba más breve aparece en el *turno 55*. Por su parte, el *sujeto 1* presenta mucho contraste en la primera mitad, lo cual quiere decir que en la segunda mitad no hay gran diferencia en la duración de sus segmentos y, por lo tanto, mantiene un discurso más estable.

4.1.2.3 Distribución temporal general por grabación

Lista	Segmentos	Sílabas	Pausas
G1	686	647	39
G2	1238	1155	83

La porción del corpus que ocupa cada lista permite que haya mayor variabilidad en la lista G2, ya que dobla el número de elementos de G1, manteniendo la

proporción de sílabas y pausas presente en el primer audio, como se ve en la tabla 24. Observando los mismos datos en una representación gráfica diferente, se aprecia que la *grabación 2* conforma el 64% del corpus, mientras que la *grabación 1* apenas representa el 36% de las muestras (ver figura 9), por lo que el comportamiento de ambos sujetos en el segundo audio es mucho más diverso. A pesar de que los promedios de duración, tanto de sílabas como de pausas, son más cercanos en G2 que en G1, tener muchas más muestras implica la posible existencia de más

segmentos de la conversación en los que el intercambio de información se realice con dinámicas consecutivas distintas y que quizá el transcurso de la conversación no sea tan estable como parece en primera instancia.

A partir de esta división del corpus se puede destacar la diferencia entre las dinámicas que se

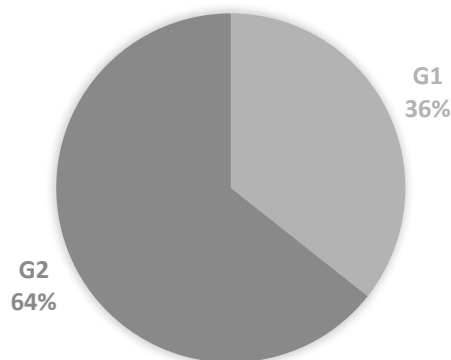


Figura 9. Proporciones del corpus por grabación

desarrollan en ambas grabaciones. Si bien, en los dos audios existe un contraste entre las tendencias a la aceleración (*sujeto 2*) y a la desaceleración (*sujeto 1*), en G1 es mucho más marcado que en G2, ya que en esta última se presenta mucha más variabilidad, al menos para el *sujeto 1*. Por otra parte, en las muestras de segmentos con medidas extremas de la *grabación 1* hay datos de todas las partes de la conversación, mientras que las muestras de la *grabación 2* apenas llegan a la segunda mitad del audio, por lo que se puede intuir que en el segundo caso se pudo llegar a un nivel de acuerdo común que permitió que la conversación se extendiera.

Tabla 25. Variación de PG

Lista	Pausa (s.)	Sílaba (c. s.)	Segmentos
G1	0.12 a 2.54	3 a 49	686
G2	0.07 a 1.15	3 a 48	1238

La variación en la duración de los segmentos para ambas grabaciones resulta casi idéntica en el número de sílabas, sus duraciones van desde las 3 centésimas de

segundo (c. s.) en ambas grabaciones, hasta las 49 c. s. para G1 y 48 c. s. para G2, mientras que las duraciones de pausas entre grabaciones varían un poco más, yendo de 0.12 segundos (s.) a 2.54 s. en G1 y de 0.07 s. a 1.15 s. en G2 (ver tabla 25). Dicho lo anterior, se puede afirmar que, a pesar de que la duración total de la primera llamada telefónica indica que la conversación fue muy rápida, en la segunda llamada no se utilizaron largos espacios de tiempo, ocupados por sonidos o silencios,

para alargar la conversación y, por lo tanto, el acto comunicativo tuvo un ritmo más fluido sin los grandes contrastes presentes en G1.

Según el promedio por grabación (PG), es posible determinar de manera contrastiva las características propias del sujeto presentes en cada acto comunicativo, ya que los cambios de velocidad y distribución temporal varían siempre por razón de la situación comunicativa en la que se encuentre al momento de la enunciación. La velocidad de cada individuo no está sujeta a una característica identitaria sino a la situación comunicativa en que se desenvuelva el hablante, ya que contrastar las características de un hablante sin tomar en cuenta el contexto no resulta significativo. Solo es posible determinar si las realizaciones han sido largas o breves si existen datos para contrastarlos, ya sea de otro hablante o del mismo hablante en un momento distinto, siempre teniendo en cuenta que los estímulos externos modifican el tiempo.

4.1.3 Duración promedio por sujeto

Tabla 26. Duración promedio por sujeto		
Lista	Duración promedio de sílaba (c. s.)	Duración promedio de pausa (c. s.)
S1	18	53
S2	16	38

El promedio por sujeto (PS) se calculó a partir de la división del corpus en dos listas, una para cada locutor, S1 y S2, donde la letra S corresponde a «sujeto». En la tabla 26 se presentan los promedios obtenidos de

duración de pausas y sílabas para cada una de estas listas en centésimas de segundo (c. s.). La distribución se hace de esta manera porque se presume que existe uniformidad en la producción de cada hablante, por lo que las variaciones en la distribución del tiempo solo serían comparables entre distintos hablantes; sin embargo, también es posible que la variación interna se deba a factores externos.

4.1.3.1 Sujeto 1

En la lista S1 hay un total de 919 segmentos, de los cuales 62 son pausas, cuya duración promedio es de 53 centésimas de segundo (c. s.), la más larga de todas pertenece al *turno 8* de la *grabación 1* y presenta una duración de 2.54 segundos, mientras que la más breve dura 7 c. s. y pertenece al *turno 55* de la *grabación 2*. Las 10 realizaciones más largas y las 10 más breves de pausas se presentan en la tabla 27. Los 857 segmentos restantes son sílabas, cuya duración promedio es de 18 c. s., la más larga de todas es la tónica /'ser/ del *turno 18* de la *grabación 1* con una duración de 49 c. s., mientras que la más breve es la átona /e/ que dura 4 c. s. y aparece en el *turno 34* de la *grabación 2*. La lista de las 10 sílabas más largas y las 10 más breves se muestran en la tabla 28.

G	Turno	Pausa	Duración	Inicio	Fin
1	8	##	2.54	45.29	47.83
1	18	##	1.42	97.32	98.74
1	18	##	1.26	100.46	101.72
2	5	##	1.15	35.05	36.20
2	3	##	1.03	26.04	27.06
2	14	##	1.00	74.67	75.67
1	14	##	0.93	89.58	90.51
1	4	##	0.90	10.47	11.38
2	11	##	0.88	62.81	63.69
2	11	##	0.85	64.92	65.78
1	18	#	0.21	107.65	107.85
2	5	#	0.20	33.26	33.45
2	30	#	0.18	116.76	116.94
2	38	#	0.17	159.33	159.50
2	5	#	0.14	32.48	32.62
2	28	#	0.11	102.45	102.56
2	28	#	0.10	97.44	97.55
2	38	#	0.10	158.23	158.33
2	1	#	0.07	0.28	0.35
2	55	#	0.07	189.98	190.05

G	Turno	Segmento	Duración
1	18	'ser	49
2	3	'kres	48
2	34	no	47
1	20	'si	47
2	11	'si	47
1	33	se	45
2	9	'si	44
1	6	do	44
1	4	'toi	44
1	16	'ke	43
2	3	'ko	5
1	6	a	5
2	5	'bo	5
2	18	a	5
2	34	a	4
2	34	yo	4
2	30	o	4
1	6	p	4
2	16	ro	4
2	34	e	4

La distribución de segmentos en las tablas indica que el *sujeto 1* realiza segmentos más largos

en la *grabación 1* y más breves en la *grabación 2*. Según las muestras, el comportamiento tendiente a la desaceleración se presenta a lo largo de toda la *grabación 1*, debido a la escasez de segmentos breves en la parte inferior de las listas. Mientras que las muestras de la *grabación 2* aparecen en ambos extremos de las listas, por lo que se puede concluir que presenta más variación en esta grabación, con una pequeña aceleración en comparación con el primer audio. El comportamiento de este sujeto en la *grabación 1* muestra un uso de alargamientos estable y en la *grabación 2* la duración de sus segmentos no es marcadamente larga ni breve.

4.1.3.2 Sujeto 2

En la lista S2 hay un total de 1005 segmentos, de entre ellos 60 son pausas, cuya duración promedio es de 38 centésimas de segundo (c. s.), la más larga de ellas dura 1.26 segundos y aparece en el *turno 31* de la *grabación 1*, mientras que la más breve aparece en la *grabación 2* durante el *turno 29* con una duración de 10 c. s. Las 10 realizaciones más largas y las 10 más breves se presentan en la tabla 29. Los 945 segmentos restantes, excepto uno (ver 3.2.3), son sílabas, cuya duración promedio es de 16 c. s., la más larga de ellas es la átona /tʃe/ (/ce/) que dura 41 c. s. y que aparece en el *turno 8* de la *grabación 2*, mientras que la más breve solo dura 3 c. s. y aparece durante el *turno 19* de la *grabación 1*. La lista con las 10 sílabas más largas y las 10 más breves se muestra en la tabla 30.

Tabla 29. Sílabas de S2 (c. s.)

G	Turno	Segmento	Duración
2	8	ce	41
1	11	nas	39
2	8	yo	37
2	8	treil	36
2	15	'dios	34
2	52	ke	32
2	2	'mi	32
2	8	dal	32
1	19	dre	30
2	15	'dios	29
2	23	te	5
1	13	a	5
2	21	de	4
2	37	po	4
1	7	a	4
2	29	mo	4
2	35	a	4
2	29	a	4
2	8	a	3
1	19	a	3

Tabla 30. Pausas de S2 (s.)					
G	Turno	Pausa	Duración	Inicio	Fin
1	31	##	1.26	132.67	133.93
1	13	##	1.02	74.32	75.34
1	13	##	1.01	71.50	72.51
2	29	##	0.92	110.32	111.24
2	8	##	0.81	42.54	43.35
1	13	##	0.81	77.04	77.85
2	23	##	0.70	90.68	91.37
2	60	##	0.69	199.72	200.41
2	2	##	0.67	18.52	19.19
1	31	##	0.64	136.02	136.66
1	31	#	0.14	129.20	129.34
1	31	#	0.13	129.88	130.01
2	2	#	0.13	8.18	8.31
2	23	#	0.13	91.56	91.69
2	15	#	0.12	78.32	78.44
2	56	#	0.12	192.52	192.64
1	19	#	0.12	111.28	111.39
2	8	#	0.11	43.55	43.66
2	33	#	0.11	135.39	135.50
2	29	#	0.10	107.06	107.15

Según las muestras, el *sujeto 2* presenta un comportamiento similar en ambas grabaciones, ya que los elementos largos y breves se reparten de manera equitativa en las tablas. Además, la distribución de los elementos de medidas extremas solo se acentúa en la parte central de la *grabación 2*. En general, los segmentos más largos o más breves aparecen esparcidos de manera uniforme a lo largo de los dos audios. A pesar de la coincidencia de comportamientos, establecer un perfil para el *sujeto 2* a partir de este corpus resultaría contraproducente, ya

que todas las realizaciones están condicionadas por el contexto.

4.1.3.3 Distribución temporal general por sujeto

Esta división del corpus pretende resaltar las cualidades propias de cada uno de los dos locutores como características identitarias; sin embargo, no es conveniente hacer ninguna afirmación al respecto, ya que el tiempo utilizado en cada segmento está condicionado por el contexto en que se realiza. Ahora bien, la porción del corpus ocupada por el *sujeto 2* es mayor que la del *sujeto 1*

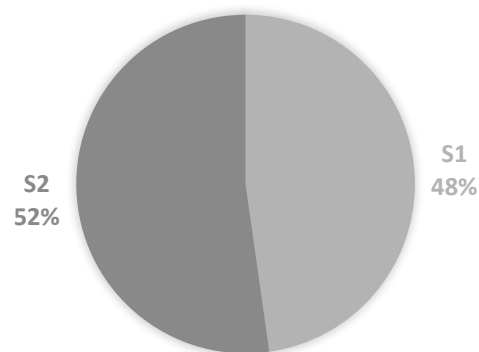


Figura 10. Proporción del corpus por sujeto

Lista	Segmentos	Sílabas	Pausas
S1	919	857	62
S2	1005	945	60

(ver figura 10) en cuanto a número de segmentos no en cuanto a cantidad de tiempo utilizada, pues el primer locutor

habla durante aproximadamente 5 minutos y el segundo por alrededor de 3 minutos, lo que confirma la tendencia del *sujeto 2* de hablar con más velocidad, y la tendencia del *sujeto 1* de hablar más lento. Además, como se aprecia en la tabla 31, quien realiza mayor cantidad de pausas es el primer sujeto, aunque la diferencia entre uno y otro es mínima, y también es el locutor que articula menos sílabas.

Como se dijo antes, los segmentos más largos del *sujeto 1* se encuentran en la grabación 1 y los más breves en la grabación 2, lo cual indica una marcada diferencia entre

Tabla 32. Variación de PS

Lista	Pausa (s.)	Sílaba (c. s.)	Segmentos
S1	0.07 a 2.54	4 a 49	919
S2	0.10 a 1.26	3 a 41	1005

las realizaciones de una conversación y otra, mientras que el *sujeto 2* presenta mayor variabilidad en sus enunciaciones, según los segmentos de medidas extremas. Además, si se confronta el rango de variación entre ambos sujetos (ver tabla 32), se aprecia que la variación de medidas es más amplia para el *sujeto 1* que para el *sujeto 2*, y que el primero, en general, utiliza más tiempo que el segundo para realizarlos. Sin embargo, la diferencia entre ambos rangos es tan pequeña que no se considera lo suficientemente contundente como para distinguir a un hablante de otro.

Para poder considerar la distribución temporal y la velocidad de habla como características identitarias, los locutores tendrían que comportarse de la misma manera en ambas grabaciones. En el caso del *sujeto 2* se cumple con ese requisito, pero no ocurre lo mismo con el *sujeto 1*, por lo que se descarta la posibilidad de que se trate de un fenómeno que identifique a un hablante. La velocidad de habla y la distribución del tiempo responden, más bien, a las particularidades del

contexto en el que se realizan y a las respuestas dadas por el interlocutor, ya que en ningún momento se mantiene la igualdad entre una realización y otra, sino que se trata de una sucesión de cambios de modulación.

4.1.4 Duración promedio total

Tomar el corpus entero como punto de partida implica que todas las realizaciones de los hablantes parten desde el mismo punto medio para calcular su duración, es decir, se presupone la existencia de un promedio universal para

Lista	Duración promedio de sílaba (c. s.)	Duración promedio de pausa (c. s.)
T	17	46

la duración de los segmentos en el que no influye ni la situación, ni el cambio de locutor, ni la sucesión de turnos de habla. Según este cálculo, la duración promedio para una sílaba es de 17 centésimas de segundo (c. s.) y para una pausa es de 46 c. s. (ver tabla 33). A partir de esas medidas, cualquier segmento, cuya duración se encuentre por encima o por debajo de la media, podrá ser clasificado como largo o breve.

La lista total del corpus incluye 1954 segmentos, de los cuales 122 son pausas, la más larga dura 2.54 segundos (s.) y fue realizada por el *sujeto 1* en el *turno 8* de la *grabación 1*, mientras que la más breve solo dura 7 c. s. y fue realizada también por el *sujeto 1*, pero durante el *turno 55* de la *grabación 2*. Las 10 realizaciones más largas y las 10 más breves de pausas se muestran en la tabla 34. Los 1832 segmentos restantes, salvo uno (ver 3.2.3), son sílabas, la más larga de todas ellas es la tónica /'ser/ con una duración de 49 c. s. que fue articulada por el *sujeto 1* en el *turno 18* de la *grabación 1*, la más breve es la átona /a/ que dura solamente 3 c. s. y que fue articulada por el *sujeto 2* en el *turno 19* de la *grabación 1*. La lista con las 10 sílabas más largas y los 10 más breves aparece en la tabla 35.

G	S	T	Pausa	Duración	Inicio	Fin
1	1	8	##	2.54	45.29	47.83
1	1	18	##	1.42	97.32	98.74
1	2	31	##	1.26	132.67	133.93
1	1	18	##	1.26	100.46	101.72
2	1	5	##	1.15	35.05	36.20
2	1	3	##	1.03	26.04	27.06
1	2	13	##	1.02	74.32	75.34
1	2	13	##	1.01	71.50	72.51
2	1	14	##	1.00	74.67	75.67
1	1	14	##	0.93	89.58	90.51
2	2	56	#	0.12	192.52	192.64
1	2	19	#	0.12	111.28	111.39
2	2	8	#	0.11	43.55	43.66
2	1	28	#	0.11	102.45	102.56
2	2	33	#	0.11	135.39	135.50
2	1	28	#	0.10	97.44	97.55
2	1	38	#	0.10	158.23	158.33
2	2	29	#	0.10	107.06	107.15
2	1	1	#	0.07	0.28	0.35
2	1	55	#	0.07	189.98	190.05

G	S	T	Segmento	Duración
1	1	18	'ser	49
2	1	3	'kres	48
2	1	34	no	47
1	1	20	'si	47
2	1	11	'si	47
1	1	33	se	45
2	1	9	'si	44
1	1	6	do	44
1	1	4	'toi	44
1	1	16	'ke	43
2	2	37	po	4
1	1	6	p	4
1	2	7	a	4
2	2	29	mo	4
2	1	16	ro	4
2	2	35	a	4
2	1	34	e	4
2	2	29	a	4
2	2	8	a	3
1	2	19	a	3

Si se parte desde el punto de vista de las pausas se observa que la *grabación 1* es más lenta que la *grabación 2*, y que esta última es más variada que la primera, pero no aporta datos relevantes sobre el comportamiento de los participantes de la conversación. Por otro lado, la lista de sílabas indica que hay una marcada tendencia hacia los alargamientos por parte del *sujeto 1*, quien presenta también mayor variación que el *sujeto 2* en cuanto a las duraciones de sus segmentos. Sin embargo, se carece de muestras representativas sobre el comportamiento general del *sujeto 2*.

La lista total del corpus es quizá la que menos datos reveladores aporta, pues no pone de relieve ningún aspecto específico. Si se considerara válida esta perspectiva, se obtendría que la variación de velocidad y la distribución temporal se trata de un fenómeno azaroso en el continuum del tiempo. Hacer esa valoración implicaría omitir la relevancia de los cambios ocurridos conforme el paso

del tiempo y el intercambio de información, por lo que no solo no se tendría una visión global del fenómeno, sino que sería una interpretación demasiado simplista de los actos comunicativos.

4.2 Velocidad de elocución en los turnos de habla

Una vez hecha la distribución del tiempo en segmentos largos y breves, es posible mostrar las variaciones de cantidad dentro de los turnos de habla enteros, además, con las distintas versiones de la distribución del corpus se obtienen cuatro transcripciones distintas entre sí, en la mayoría de los casos. Si bien es cierto que, si se compara con el total del corpus, no son muchos los segmentos que presentan diferencias en la distribución de cantidad entre listas, pero esas ligeras variaciones señalan los puntos del discurso en los que no hay una marcada modificación, pues se mantienen cerca del punto medio. Por consiguiente, esas partes de la conversación no representan intenciones ni de focalizar información ni de ocultarla, sino que podrían considerarse parte de la producción normal de los hablantes.

Los resultados de turnos de habla, a diferencia de los de segmentos, solo se presentan con división entre *grabación 1* y *grabación 2*, ya que las intervenciones de los dos locutores tienen una secuencia lógica y es importante mantenerla de esa manera para poder apreciar tanto las diferencias en la velocidad de habla como los fragmentos del discurso que tienen alguna implicación pragmática debida a alguna modificación de duración y/o velocidad. Además, se incluyen muestras de las características generales de los turnos de habla realizados por cada locutor, ya que el espacio que ocupan dentro de la conversación y la manera en que se presentan es distinta, lo cual permite la comparación de las estrategias utilizadas a lo largo de las grabaciones para lograr o no sus objetivos particulares.

4.2.1 Grabación 1

Como se mencionó en el apartado 3.1, la duración de la *grabación 1* es de 2 minutos con 24 segundos, tiempo en el que se desarrolló la primera de dos conversaciones grabadas de la negociación entre un extorsionador y el familiar de una víctima de secuestro, esta conversación se llevó a cabo mediante intervenciones dispares por parte de ambos locutores, lo que provocó que la comunicación terminara de manera abrupta. Aproximadamente 1 minuto con 9 segundos del tiempo total de la llamada telefónica están ocupados por el *sujeto 1*, es decir, la víctima de la extorsión, mientras que unos 54 segundos son utilizados por el *sujeto 2*, es decir, el extorsionador. Los 21 segundos restantes son pausas largas y ruido que no forman parte de las intervenciones de los hablantes, sino que constituyen, la mayoría de las veces, los límites de los 33 turnos de habla, en los que se realiza un total de 686 segmentos.

En cuanto a las características generales de las intervenciones de los dos hablantes, se ha de mencionar que el *sujeto 1* habla durante 16 turnos, en los que realiza un total de 347 segmentos, 23 pausas vacías y 324 sílabas, mientras que el *sujeto 2* habla durante 17 turnos, en los que produce 339 segmentos, 16 pausas vacías y 323 sílabas. Lo anterior implica que, a pesar de que el *sujeto 2* toma la palabra en más ocasiones, el *sujeto 1* utiliza 15 segundos más para hablar,. Además, la duración promedio de los turnos de habla del *sujeto 1* es de 4.31 segundos, lo cual supera por más de un segundo a la duración promedio de los turnos su interlocutor, que es de 3.15 segundos. Cabe resaltar que el rango de duración de los turnos del primer locutor va desde las 14 centésimas de segundo (c. s.) hasta los 15.18 segundos, mientras que el rango de duración de los turnos de *sujeto 2* varía desde las 15 c. s. hasta los 16.40 segundos.

El contenido de los turnos de habla también es muy distinto entre un hablante y el otro, pues el *sujeto 1* puede realizar en un solo turno desde 1 hasta 78 segmentos, por su parte, el *sujeto 2* tiene

un rango de variación un poco más amplio, pues sus turnos pueden contener entre 2 y 95 segmentos. La cantidad de segmentos es de vital importancia para el cálculo de la velocidad de habla, pues de esta depende la distribución del tiempo en cada segmento. De tal manera que, la velocidad de elocución promedio para el *sujeto 1* es de 5.03 segmentos por segundo (s./s.), pero puede variar desde 2.89 s./s. hasta 6.99 s./s., mientras que la velocidad promedio del *sujeto 2* es de 7.16 s./s., y el rango de velocidades de sus turnos se coloca incluso más arriba, yendo desde los 5.28 s./s. hasta los 13.34 s./s. Toda esta información se puede ver condensada en las tablas 36 y 37.

	Promedio	Rango
Total de segmentos	21.69	1 a 78
Pausas	1.31	0 a 6
Duración (s.)	4.31	0.14 a 15.18
Velocidad (s./s.)	5.03	2.89 a 6.99

G1S2	Promedio	Rango
Total de segmentos	19.94	2 a 95
Pausas	1.00	0 a 6
Duración (s.)	3.15	0.15 a 16.40
Velocidad (s./s.)	7.16	5.28 a 13.34

4.2.1.1 Tempo

La interacción entre los dos locutores durante la primera grabación ocurre de manera distinta en cuanto a las características de sus turnos de habla (número de segmentos, pausas insertadas y duración). Es fácil detectar las diferencias entre uno y otro ya que, como se ve en las tablas 36 y 37, las medidas se mantienen dentro de dos rangos de variabilidad distantes. Dichas diferencias indican las estrategias utilizadas por cada hablante dentro de la conversación, como ya se había esbozado en el apartado 4.1, el *sujeto 1*, generalmente, tiene turnos de habla más largos que el *sujeto 2*, tanto en número de segmentos como en duración, aunque sea este último quien realiza la participación más larga. Los dos locutores pueden agregar entre 0 y 6 pausas, pero es el primer

sujeto quien utiliza mucho más tiempo de silencio y en más ocasiones, lo cual afecta directamente el resultado perceptivo de velocidad de elocución, es decir, el tempo.

Tabla 38. Turnos de habla del sujeto 1 en la grabación 1

Turno	Total de segmentos	Pausas	Duración (s.)	Duración de pausas (s.)	Velocidad (s./s.)	Velocidad de articulación (s./s.)
2	20	2	4.53	0.81	4.41	4.83
4	47	3	8.71	1.81	5.40	6.38
6	55	3	11.07	2.13	4.97	5.82
8	13	1	4.50	2.54	2.89	6.11
10	24	1	3.90	0.36	6.15	6.48
12	42	2	7.15	1.17	5.88	6.69
14	19	2	3.67	1.26	5.18	7.07
16	4	0	0.86	0	4.63	4.63
18	78	6	15.18	4.24	5.14	6.58
20	4	0	0.95	0	4.20	4.20
22	5	0	1.22	0	4.10	4.10
24	4	0	0.92	0	4.35	4.35
27	1	0	0.14	0	6.99	6.99
29	25	1	4.84	0.54	5.16	5.57
32	1	0	0.15	0	6.83	6.83
33	5	0	1.19	0	4.19	4.19

Tabla 39. Turnos de habla del sujeto 2 en la grabación 1

Turno	Total de segmentos	Pausas	Duración (s.)	Duración de pausas (s.)	Velocidad (s./s.)	Velocidad de articulación (s./s.)
1	8	0	1.28	0	6.23	6.23
3	2	0	0.33	0	6.11	6.11
5	31	2	5.27	0.96	5.89	6.74
7	20	0	2.61	0	7.65	7.65
9	17	0	2.33	0	7.30	7.30
11	16	0	3.02	0	5.29	5.29
13	95	6	16.40	3.97	5.79	7.16
15	12	1	1.65	0.18	7.28	7.51
17	7	0	0.86	0	8.17	8.17
19	27	1	3.35	0.12	8.07	8.05
21	13	0	1.88	0	6.90	6.90
23	5	0	0.65	0	7.65	7.65
25	2	0	0.15	0	13.34	13.34

Tabla 39. Turnos de habla del sujeto 2 en la grabación 1

Turno	Total de segmentos	Pausas	Duración (s.)	Duración de pausas (s.)	Velocidad (s./s.)	Velocidad de articulación (s./s.)
26	12	1	2.27	0.56	5.28	6.43
28	7	0	0.89	0	7.85	7.85
30	2	0	0.29	0	6.78	6.78
31	63	5	10.26	2.44	6.14	7.42

En las últimas dos columnas de las tablas 38 y 39 se presentan las velocidades de cada turno de habla, para el cálculo de la penúltima, nombrada simplemente «velocidad» se toman en cuenta las pausas, mientras que en la «velocidad de articulación» solo se consideran los sonidos articulados. Las diferencias entre una columna y otra, en su mayoría, solo son de 2 segmentos/segundo (s./s.), pero puede ocurrir un contraste muy radical como el del *turno 8* de la tabla 38, en el que, si una pausa tan larga no existiera, habría sido percibido como un mensaje rápido y no uno lento. Sin embargo, podría ocurrir que, si se quitara una pausa muy breve, el turno sea perceptivamente más lento, como en el *turno 19* de la tabla 39. Por lo tanto, para el análisis de estos resultados solo se toma en cuenta la columna «velocidad», ya que omitir el tiempo añadido por los silencios cambia considerablemente la sensación perceptiva que provocan los turnos de habla.

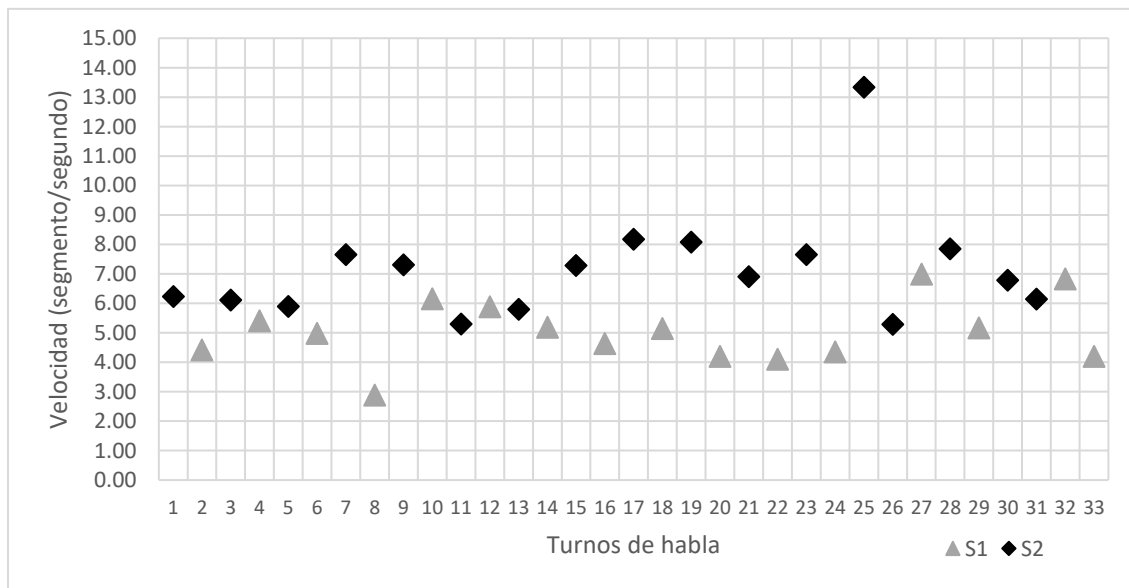


Figura 11. Velocidad de los turnos de habla para cada sujeto en la grabación 1

La naturaleza contrastiva de las intervenciones de ambos sujetos se puede observar en la figura 11, donde la línea formada, imaginariamente, por la unión de los puntos del *sujeto 2* se mantiene durante casi toda la conversación por encima de la línea del *sujeto 1*. Como ya se anunciaba desde el apartado 4.1, el *sujeto 2* habla mucho más rápido que el *sujeto 1*, además, cuando este último alcanza sus puntos más bajos, el otro llega los picos más altos, lo cual indica que en esas secciones de la grabación se está llevando a cabo un enfrentamiento. No obstante, también hay momentos en que ambos sujetos presentan casi la misma velocidad, aunque estos instantes de acuerdo mutuo son muy breves.

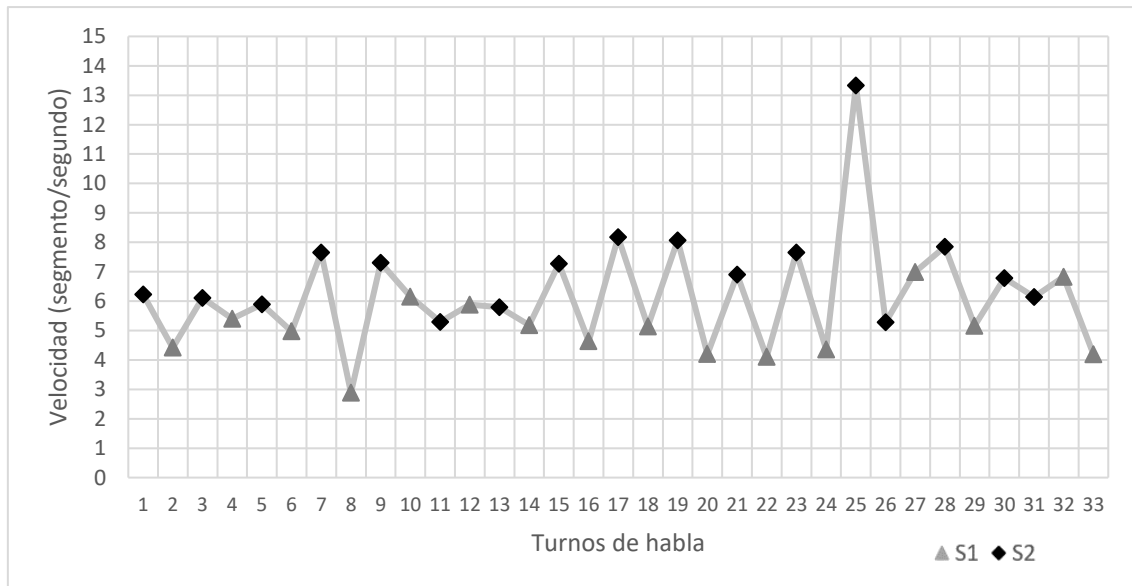


Figura 12. Velocidad en el continuum del tiempo de la grabación 1

El flujo de la conversación varía a lo largo de todo el audio, como se ve en la figura 12, durante casi toda la conversación, la sucesión de turnos de habla forma picos y valles prolongados, mientras que en ciertas secciones la línea en zigzag es menos amplia. Cuando existe una marcada diferencia entre turnos de habla contiguos se puede hablar de puntos de desacuerdo o de conflicto, en los que necesariamente se resalta la diferencia de velocidad entre el turno de habla pronunciado

y el turno de habla inmediatamente anterior. Por otro lado, cuando existe cierto grado de acuerdo común, el contraste entre las velocidades de turnos de habla contiguos disminuye o se anula.

4.2.1.2 Valores de relevancia adquiridos

Para complementar y afirmar las declaraciones hechas en el punto anterior, es necesario revisar el contenido de los turnos de habla de las secciones donde hay mucho o poco contraste en las velocidades de turnos de habla contiguos. Considero «conflicto» siempre que los dos hablantes realicen intervenciones con velocidades muy distantes entre sí, es decir, con diferencias mayores a los 2 segmentos/segundo (s./s.), en las que se trate de convencer o persuadir al receptor de dirigir la conversación hacia el propósito particular del emisor sin obtener respuesta positiva. De manera similar, considero «acuerdo» cuando las intervenciones de los hablantes presenten velocidades similares, con una diferencia no mayor a 2 s./s., y el emisor logre respuesta positiva, parcial o total, por parte del receptor.

4.2.1.2.1 Conflicto

El primer momento de la *grabación 1* en el que se presenta conflicto entre los hablantes ocurre entre los *turnos 7 y 9*, donde, como se puede observar en la tabla 40 y en su correspondiente transcripción ortográfica, el *sujeto 2* interrumpe al *sujeto 1*, después de que realice una pausa muy larga e intente poner una excusa. El *sujeto 2* esperaba una respuesta rápida y concisa, no una que durara el doble, que no le dijera nada, y que le hiciera perder el tiempo. Aunque el *sujeto 1* logró ganar tiempo, no consiguió que el *sujeto 2* disminuyera su velocidad, por lo que la prolongación de su participación le dio espacio a su contrincante para manipular la conversación.

- G1S2T07:** *Por eso, puto, o sea, ¿vas a poner en riesgo la vida de tu hijo?*
G1S1T08: *No, ps es lo que no no quiero. || Estoy hacien-*
G1S2T09: *Te dije que no quería que coliera nadie, ¿no, puto?*

Tabla 40. Sección 1 de conflicto de la grabación 1

Sujeto	Turno	Duración (s.)	Velocidad (s./s.)	Transcripción (PG)
S2	T7	2.61	7.65	po'reso'puTOsea'basapo'neREN'riesgola'bidade'tuixo
S1	T8	4.50	2.89	NOp'SESlo'keno'KIERO##es'toia'SIEN
S2	T9	2.33	7.30	te'diXEke'noke'rikete'ko'liera'nadieno'puto

En la transcripción de la tabla 40 se pueden detectar los momentos en que los mensajes se aceleran o se desaceleran, el *sujeto 2*, acentuando la velocidad, lanza una amenaza leve en el *turno 7*, rápida para que suene ligera, mientras que en el *turno 9*, diluye el último insulto con una veloz pregunta. Entre estos dos sucesos, el *sujeto 1*, no da una respuesta rápida a la pregunta formulada en el *turno 7*, sino que responde con un «no» largo, para después esconder una negación diciéndola rápido y alarga el verbo «querer» que afecta dicha negación. Finalmente, el *sujeto 1*, después de una pausa prolongada, trata de acelerar su turno de habla, pero es interrumpido antes de que pueda terminar su oración.

El segundo momento en el que se presenta conflicto entre los hablantes ocurre entre los *turnos 15 y 21*, como se observa en la tabla 41 y en su transcripción ortográfica, donde el *sujeto 2* está completamente cerrado al diálogo, tanto que solo emite el mismo insulto en cuatro ocasiones, más una amenaza, mientras que el *sujeto 1* trata comprender lo que dijo mal y hacer que su interlocutor se tranquilice. En comparación con el punto de conflicto anterior, el *sujeto 1* aumenta su velocidad en un intento fallido de controlar la situación, pero no es lo suficientemente veloz como para alcanzar o superar al *sujeto 2*, aunque quizá, si lograra hacerlo, la comunicación se cortaría de inmediato. De nuevo, la ganancia de tiempo del *sujeto 1* no es significativa y el *sujeto 2* logra presionarlo.

G1S2T15: *A chingar a su madre.*

G1S1T16: *Pero ¿por qué?*

G1S2T17: *A chingar a su madre.*

- G1S1T18: *¿Por qué? Dame chance, te llevo eso, com'e 'bias dicho. || Es to-, es to'lo que puedo hacer. || O, o dame chance, deja me... Ps mis familiares quedaron de traerm'y yo les dije que tenía'l plazo | hasta hoy, eh, | ellos quedaron que m'iban a traer.*
- G1S2T19: *A chingar a su madre. / La cabeza de tu hijo va a amanecer ahí frente a tu casa.*
- G1S1T20: *No seas así.*
- G1S2T21: *A chingar a su madre, puto.*

Tabla 41. Sección 2 de conflicto de la grabación 1

S	T	Duración (s.)	Velocidad (s./s.)	Transcripción (PG)
S2	T15	1.65	7.28	acin'garas#aCIN'garasu'maDRE
S1	T16	0.86	4.63	'peROpor'KE
S2	T17	0.86	8.17	acin'garaSU'madre
S1	T18	15.18	5.14	por'KEdami'CANsete'yebo'eSO#'komeBIAS'diCO##ES'toE S'toloke'puedoa'SER##oo'dame'CANse'dexaME#psMISfam i'liareske'DANdeTRA'ERmiyoles'dixekete'nial' <u>PLASO</u> #'ast a'OIE#'eyosmedi'xeronke'mibanatra'ER
S2	T19	3.35	8.07	aCIN'garasu'maDRE#laka'besade'tuixobamane'sera'iFRE Nteatu'kasa
S1	T20	0.95	4.20	no'SEsa'SI
S2	T21	1.88	6.90	aCIN'garasu'MADre'yale'dixe'PUto

En general, las breves participaciones del *sujeto 2* no presentan momentos en los que se acentúe algún cambio de velocidad, pero en el *turno 19* lanza una amenaza rápida como la del *turno 7*, aunque mucho más directa, la respuesta de su interlocutor no es la que se esperaría ante algo tan contundente porque ocurre muy rápido. Por otro lado, las intervenciones del *sujeto 2* presentan una velocidad regular entre 4 s./s. y 6 s./s. con sílabas muy largas, solo en el *turno 18* agrega tiempo excesivo con muchas pausas y dos secciones con significado muy largas, pero la respuesta que recibe no es la que él quería, ya que, además de un insulto, da una amenaza

El último momento de conflicto de la *grabación 1* es, justamente, el final de la conversación, desde el *turno 28* hasta el *turno 33*, como se puede corroborar en la tabla 42 y en la transcripción ortográfica de estos turnos de habla. Esta sección es una de las más conflictivas porque no solo se

roba el turno de palabra, sino que, también, se anula la posibilidad de tomar la palabra. Lo único que logra decir el *sujeto 1* es una súplica interrumpida durante el *turno 29*; en el *turno 33* logra hablar, pero ya no es escuchado porque el *sujeto 2* cortó la comunicación en cuanto terminó el *turno 31*.

- G1S2T28:** *A chingar a tu madre.*
G1S1T29: *Dame tiempo, dame un ratito más, mi familia me va a traer, | yo los 'toy...*
G1S2T30: *(¿Sale?)*
G1S2T31: *Por eso, | yo quiero con bien a mi familia, hijo de tu puta madre. || Mi familia, escúchalo, soy la familia, puto. || Con nosotros no se anda con mamadas, | valiste verga.*
G1S1T32: *(Les...)*
G1S1T33: *No, dame chance.*

Tabla 42. Sección 3 de conflicto de la grabación 1

S	T	Dur. (s.)	Vel. (s./s.)	Inicio (s.)	Fin (s.)	Transcripción (PG)
2	28	0.89	7.85	123.08	123.97	acin'garatu'maDRE
1	29	4.84	5.16	124.05	128.90	'dame'TIEMPO'dame'daMEUNra'tito'MASmMifa'miliame'baTRA'ER#'YOLOS'toi
2	30	0.29	6.78	124.24	124.53	'saLE
2	31	10.26	6.14	128.90	139.16	po'reso#YO'kiero#yo'kieroYO'kierokon'bienamifami'l ixodetu'puta'maDRE##mifa'milies'kucaLO'SOIlafa'milia'puTO##kono'soTROS'nose'andaKONma'maDAS#ba'liste'berGA
1	32	0.15	6.83	129.47	129.62	les
1	33	1.19	4.19	139.61	140.80	NO'dame'CANSE

El *turno 30* ocurre al principio del *turno 29* (figura 13), pero el *sujeto 2* no interrumpe al *sujeto 1* sino hasta su intervención en el *turno 31*. Igualmente, el *turno 32* ocurre al principio del *turno*

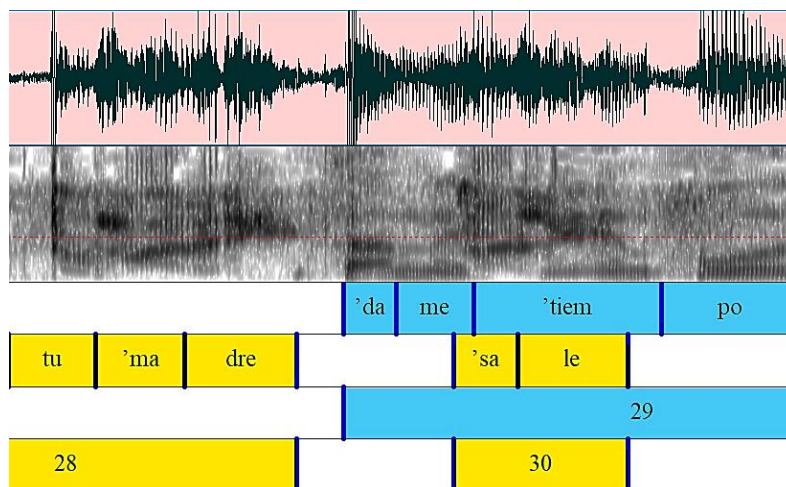


Figura 13. Solapamiento de los turnos 29 y 30 de la grabación 1

31 (figura 14), pero el *sujeto 1* no vuelve a tomar la palabra para ser escuchado.

Las velocidades del *sujeto 2* siempre se mantienen sobre 6 s./s y jamás tiene fragmentos con significado alargados, solo acorta el tiempo utilizado en sus mensajes para evitar la posibilidad de interrupción y réplica.

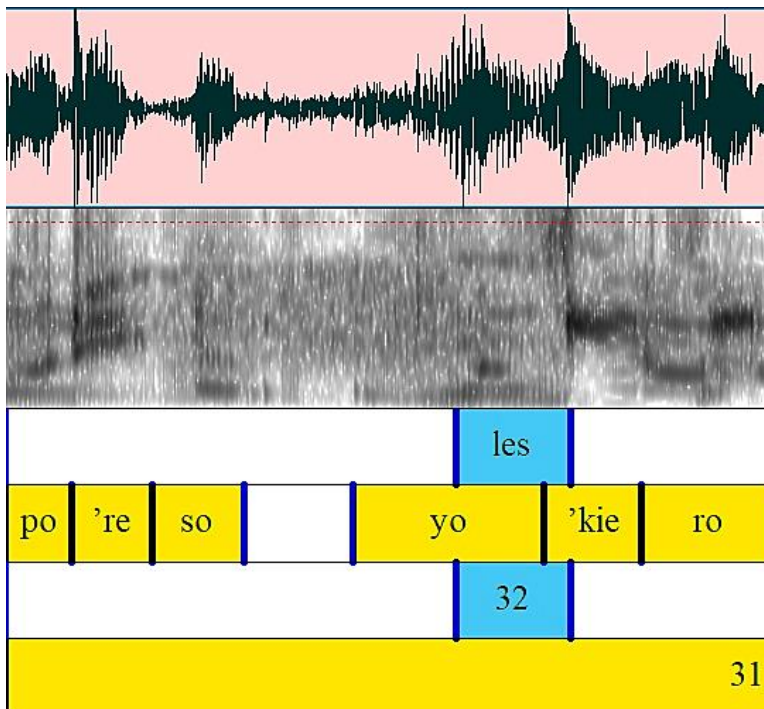


Figura 14. Solapamiento de los turnos 31 y 32 de la grabación 1

4.2.1.2.2 Acuerdo

El primer punto de acuerdo entre los dos hablantes ocurre al principio de la conversación entre los *turnos 3* y *6*, como se observa en la tabla 43 y en su transcripción ortográfica correspondiente. Tanto los turnos de habla del *sujeto 1* como los del *sujeto 2* se mantienen dentro de un rango de velocidad normal, alrededor de los 5 s./s., es decir, no hablan ni muy rápido ni muy lento, además, sus turnos son más o menos del mismo número de segmentos y las duraciones de los *turnos 4* y *5* son bastante cercanas en comparación con las otras. El *turno 3* contiene una pregunta breve consecutiva a la apertura de la llamada telefónica, por lo que es comprensible que presente una diferencia de duración muy grande con el resto, y el *turno 6* corresponde a la causa del primer conflicto de esta grabación, ya que es un intento, largo y lento, de excusarse.

G1S2T03: ¿Cuánto?

G1S1T04: Ciento diez y es un 'sfuerzo. || 's q' nadie me quier' prestar, fue to'lo que. |
Ya 'stoy | des'sperado, fue to'lo que 'cancé te a xun-, te alcancé a juntar.

G1S2T05: *De huevos, cabrón, ¿al chile nomás / ciento diez? / ¿Tú crees que yo vine por 110 mil pesos, cabrón?*

G1S1T06: *Tú échame la mano, ps sim / no puedo. // Nadie, nadie me quiere prestar, dicen, “es, es sábado” y ps / p' recurrí a mis familiares y dicen / “no tengo, quizás hasta el lunes”.*

Tabla 43. Sección 1 de acuerdo de la grabación 1

Sujeto	Turno	Duración (s.)	Velocidad (s./s.)	Transcripción (PG)
S2	T3	0.33	6.11	'KUANto
S1	T4	8.71	5.40	'SIENto'DIEsie'SUNs'FUERSO##sk'nadime'KIERP RES'TARfue'toloKE#YAS'TOI#dsspe'radoFUE'toloke kan'sete axunte alkan'sea XUN TAR
S2	T5	5.27	5.89	de'UEBOSka'BRONal'cileno'MAS#'SIENto'DIE#tu' k resKE yo'bine po 'sientoDIEmil'PEsska'BRON
S1	T6	11.07	4.97	'tuecamela'MANOpsSIM#no'PUEDO##'naDIE'NADie me'kiere PRES 'TARdiseneSE'SAbaDOIPS#preku'ria misfami'LI ARES I'DISEN#no'TENgoki'sasasta'luNE S

En los dos turnos del *sujeto 1* se presentan muchas pérdidas de sonidos, lo cual debería acelerar su dicción, pero también añade un gran número de pausas, por lo que logra compensar el tiempo que perdió al elidir sonidos. Por ello, las pausas no provocan el efecto perceptivo de lentitud gracias a la aceleración por pérdida de sonidos. De manera similar, el *sujeto 2* baja su velocidad porque hay muchos segmentos alargados en sus turnos de habla, incluso tiene una palabra entera alargada: «huevos». Ambos hablantes alcanzan velocidades similares porque están en un punto en el que pueden negociar, o al menos es lo que se esperaría apenas iniciada la llamada. Sin embargo, el *sujeto 1* repitió en el *turno 6* el mismo contenido que en el *turno 4* con menos velocidad, y, como ya se ha visto, las excusas lentas provocan conflicto.

El segundo y último momento en que hay acuerdo entre los dos hablantes ocurre entre los *turnos 10 y 14*, como se puede ver en la tabla 44 y en la transcripción ortográfica, donde el *sujeto 1* acelera la velocidad que había presentado en turnos anteriores y el *sujeto 2*, a su vez, desacelera, para tener una velocidad normal aproximada a los 5 s./s. En todos estos turnos de habla se presentan muchos

fragmentos con significado con una marcada aceleración, pero también aparecen algunos donde la velocidad disminuye. En el *turno 11* del *sujeto 2* la atención se centra al final de la participación «más, cabrón», es decir, el *sujeto 1* no está cumpliendo con las expectativas del *sujeto 2*. De manera similar, en el *turno 13*, el *sujeto 2*, obviando el insulto, focaliza dos puntos de interés: «gente» (compañeros de este hablante) y «tu hijo» (del extorsionado). Lo cual debería poner suficiente peso sobre el *sujeto 1* para que acepte las exigencias de su interlocutor.

- G1S1T10: *Estoy haciendo to'lo posible, nadie me está coliendo, | tú lo 'tás viendo, nadie.*
- G1S2T11: *No, mi'jo, échale ganas, yo sé que tienes más, cabrón.*
- G1S1T12: *Dame chance de moverme, | pus si | la familia me dijo | te vamos a apoyar, al rato te llevamos y los estoy esperando.*
- G1S2T13: *Mire, hijo de su puta madre, || al chile está hablando con la muñeca, culero. || A mí mamadas en la punta de la verga. || Yo quería 3 millones para sacar a mi gente, | tú te pasas de verga y 'horita | mismo vas a encontrar la cabeza de tu hijo | ahí frente a tu casa, hijo de tu puta madre.*
- G1S1T14: *Pero si no me estoy pasando de | de nada. || No... es... no... en...*

Tabla 44. Sección 2 de acuerdo de la grabación 1

Sujeto	Turno	Duración (s.)	Velocidad (s./s.)	Transcripción (PG)
S1	T10	3.90	6.15	<i>es'toiaSIEN'TOlopo'sible'nadieMES'tako'LIANDO#'tulosTAS'BIENdo'NADIE</i>
S2	T11	3.02	5.29	<i>'NOMixo'ecale'gaNAS'YOseke'tiene'MASKA'BRON</i>
S1	T12	7.15	5.88	<i>'dame'CANsedemo'BERME#pu'SI#LALAfa'miliame'dixo#te'bamosapo'YAral'ratoteye'bamoSISLOS'toispe'RANDo</i>
S2	T13	16.40	5.79	<i>'mirixodesu'PUta'MADRE##al'cilesta'blandokonlamu'ñekaku'lero##a'mima'madasenla'PUNtadela'BERga##'yoke'ria'TRESmi'yones'parasa'karami'XENTE#'tute'pasade'BERgaIO'riTA#'MISmo'basenkon'TRARlaka'besade'TUIXO#a'ifrentatu'KAsixodetu'PUta'mADRE</i>
S1	T14	3.67	5.18	<i>per'noMES'toipa'SANdode#de'NADA##no'esno'en</i>

El *sujeto 1* en sus tres turnos de habla solo busca excusas y en el último de ellos focaliza la palabra «nada», o sea, en ningún momento logra dar un mensaje claro, lo que eventualmente

provoca la ira del *sujeto 2*. En los dos segmentos de la conversación en los que se logra un instante de acuerdo, el *sujeto 1* tiene pérdida de sonidos, oraciones incompletas y pausas que compensan el tiempo que debían ocupar los sonidos no articulados, lo cual indica un discurso inconexo sin planeación ni seguridad. Al igual que en el ejemplo anterior, el diálogo mínimo de negociación se rompe cuando el *sujeto 1* no responde de la manera que el *sujeto 2* esperaba después del mensaje del *turno 13*, entonces, surge un conflicto.

4.2.2 Grabación 2

La *grabación 2*, presenta una duración de 3 minutos con 41 segundos, de los cuales aproximadamente 1 minuto con 37 segundos están ocupados el *sujeto 1*, mientras que alrededor de 1 minuto con 44 segundos son utilizados por el *sujeto 2*, y los 20 segundos restantes corresponden a ruido y pausas largas. Esta segunda grabación tiene un total de 1238 segmentos realizados, distribuidos en 70 turnos de habla en esta grabación. Por lo tanto, se podría decir que en el tiempo de la segunda llamada se logró entablar una comunión mucho más eficiente, en la que el intercambio de información fue mayor, con lo cual se alcanzó el objetivo del *sujeto 1* de conseguir tiempo.

En este audio el *sujeto 1* habla en 34 turnos, en los que realiza un total de 572 segmentos durante aproximadamente 1 minuto con 38 segundos, que es lo que dura la suma de todas sus intervenciones. Mientras que el *sujeto 2* habla por 1 minuto con 44 segundos, repartidos en 36 turnos, en los que produce 666 segmentos. En la *grabación 2* no hay una marcada diferencia de duración total entre un sujeto y otro, ya que el *sujeto 2* solo supera por 6 segundos al *sujeto 1*, lo notable aquí es que quien acapara un poco más de tiempo es el *sujeto 2*, lo cual habla de una competencia por la palabra más equitativa. La duración promedio por turno de habla para los dos

sujetos es de 2.88 segundos (s.), aunque los rangos de duración son más dispares que en el caso de *grabación 1*, ya que los turnos del *sujeto 1* duran entre 16 c. s. y 11.21 s., mientras que los del *sujeto 2* presentan duraciones desde las 7 c. s. hasta los 15.17 s.

Por otro lado, el *sujeto 1* presenta desde 1 hasta 68 segmentos por turno de habla, mientras que los turnos de *sujeto 2* contienen desde 1 hasta 94 segmentos. Además, la diferencia en la velocidad de elocución promedio de los dos locutores en esta segunda llamada es ligeramente más pequeña que en la *grabación 1*, pues el primer hablante presenta una rapidez general de 5.51 segmentos/segundo (s./s.), mientras que la velocidad del segundo locutor resulta en 6.88 s./s. Del mismo modo, los rangos de variación de las velocidades son cercanos, a diferencia del caso de la *grabación 1*, ya que el *tempo* del *sujeto 1* en cada turno de habla puede ir desde 2.54 s./s. hasta 8.75 s./s., mientras que el *sujeto 2* alcanza velocidades desde 3.8 s./s. hasta 14.16 s./s. Toda esta información se puede ver condensada en las tablas 45 y 46.

Tabla 46. Datos generales de los turnos de S1 en G2		
	Promedio	Rango
Total de segmentos	16.82	1 a 68
Pausas	1.15	0 a 6
Duración (s.)	2.88	0.16 a 11.21
Velocidad (s./s.)	5.51	2.54 a 8.75

Tabla 45. Datos generales de los turnos de S2 en G2		
	Promedio	Rango
Total de segmentos	17.54	1 a 94
Pausas	1.19	0 a 7
Duración (s.)	2.88	0.07 a 15.76
Velocidad (s./s.)	6.89	3.80 a 14.16

4.2.2.1 Tempo

En la segunda grabación, a diferencia de la primera, se observa que las medidas más bajas de los dos locutores son bastante cercanas, pero el *sujeto 2* sigue teniendo una mayor amplitud en sus rangos de variabilidad. En cuanto a las medidas promedio, se debe resaltar que en la *grabación 2* los turnos de ambos hablantes suelen tener prácticamente la misma cantidad de segmentos e

incluso duraciones similares. Lo cual indica que en esta conversación el intercambio se llevó a cabo de una manera más o menos equilibrada, a pesar de que el *sujeto 2*, en general, presente velocidades superiores a las del *sujeto 1* como ocurre en la *grabación 1*.

Sorpresivamente, el locutor que realiza la mayor cantidad de pausas en esta ocasión es el *sujeto 2*, con 20 pausas, mientras que el *sujeto 1* solo realiza 16. No obstante, el tiempo en que este último se mantiene en silencio es ligeramente mayor que el del segundo locutor, ya que su silencio total dura 14.75 s., y el del *sujeto 2* dura 12.7 s. Estas mismas pausas son las que ayudan a determinar la medida de velocidad de habla de ambos hablantes. En las últimas dos columnas de las tablas 47 y 48 se pueden observar dichas medidas. Si no se tomara en cuenta el tiempo de las pausas, la velocidad promedio del *sujeto 1* pasaría de 5.51 segmentos/segundo a 5.92 sílabas/segundo, y la del *sujeto 2* de 6.89 segmentos/segundo a 7.24 sílabas/segundo, es decir, la velocidad general de las participaciones sería perceptivamente más rápida sin los silencios.

Sin embargo, hay casos en los que la ausencia de pausas altera el resultado del cálculo de velocidad de manera contraria. En el *turno 38* del *sujeto 1* y los *turnos 33* y *56* del *sujeto 2* pasa que, al no incluir las pausas, la velocidad de articulación es más lenta. Por lo tanto, la presencia de pausas no necesariamente implica que la sensación perceptiva sea más lenta, puede ocurrir que un silencio rápido aumente el dinamismo de una enunciación, aunque la tendencia general sugiera que las pausas vuelven lenta la forma de hablar.

Tabla 47. Turnos de habla del sujeto 1 en la grabación 2

Turno	Total de segmentos	Pausas	Duración (s.)	Duración de pausas (s.)	Velocidad (s./s.)	Velocidad de articulación (s./s.)
24	1	0	0.39	0	2.54	2.54
26	1	0	0.35	0	2.82	2.82
69	1	0	0.28	0	3.51	3.51
47	2	0	0.53	0	3.80	3.80
45	1	0	0.25	0	3.99	3.99

Tabla 47. Turnos de habla del sujeto 1 en la grabación 2

Turno	Total de segmentos	Pausas	Duración (s.)	Duración de pausas (s.)	Velocidad (s./s.)	Velocidad de articulación (s./s.)
61	1	0	0.25	0	4.03	4.03
20	2	0	0.49	0	4.07	4.07
49	2	0	0.46	0	4.32	4.32
59	1	0	0.21	0	4.67	4.67
14	15	2	4.16	1.54	3.61	4.96
22	2	0	0.39	0	5.09	5.09
63	6	0	1.11	0	5.41	5.41
53	3	0	0.54	0	5.53	5.53
57	3	0	0.52	0	5.75	5.75
3	37	3	7.64	2.07	4.85	6.10
43	11	1	2.12	0.54	5.18	6.32
9	51	2	8.88	1.19	5.74	6.37
40	1	0	0.16	0	6.43	6.43
38	12	2	1.82	0.27	6.60	6.44
11	45	3	8.60	2.10	5.23	6.46
55	12	2	2.20	0.67	5.44	6.52
32	31	3	5.56	1.42	5.57	6.76
41	8	0	1.17	0	6.83	6.83
28	50	6	8.24	1.92	6.07	6.96
65	22	1	3.58	0.60	6.14	7.04
30	68	6	11.21	2.45	6.06	7.08
36	32	1	4.83	0.50	6.62	7.16
16	12	0	1.66	0	7.21	7.21
51	19	1	3.19	0.72	5.96	7.30
1	37	2	4.98	0.30	7.43	7.49
34	35	1	4.80	0.26	7.30	7.49
5	37	3	5.90	1.49	6.27	7.71
70	2	0	0.24	0	8.41	8.41
18	9	0	1.03	0	8.75	8.75

Tabla 48. Turnos de habla del sujeto 2 en la grabación 2

Turno	Total de segmentos	Pausas	Duración (s.)	Duración de pausas (s.)	Velocidad (s./s.)	Velocidad de articulación (s./s.)
2	94	7	15.76	3.22	5.96	6.94
4	5	0	0.66	0	7.59	7.59
6	2	0	0.32	0	6.19	6.19
7	3	0	0.46	0	6.59	6.59

Tabla 48. Turnos de habla del sujeto 2 en la grabación 2

Turno	Total de segmentos	Pausas	Duración (s.)	Duración de pausas (s.)	Velocidad (s./s.)	Velocidad de articulación (s./s.)
8	43	3	8.23	1.44	5.23	5.89
10	9	0	1.37	0	6.58	6.58
12	1	0	0.07	0	14.16	14.16
13	52	3	9.07	1.22	5.73	6.24
15	26	3	4.40	0.70	5.90	6.20
17	7	0	1.06	0	6.61	6.61
19	19	1	3.02	0.22	6.30	6.44
21	3	0	0.42	0	7.18	7.18
23	15	3	2.45	1.06	6.12	8.60
25	9	2	1.37	0.40	6.57	7.25
27	1	0	0.15	0	6.73	6.73
29	42	5	6.76	1.78	6.21	7.42
31	36	1	4.99	0.50	7.21	7.79
33	21	1	3.00	0.11	7.00	6.92
35	55	2	8.04	0.61	6.84	7.13
37	8	1	1.42	0.48	5.65	7.51
39	5	0	0.74	0	6.78	6.78
42	33	2	4.74	0.53	6.96	7.37
44	5	0	0.60	0	8.31	8.31
46	30	2	3.87	0.48	7.75	8.25
48	18	1	2.61	0.24	6.90	7.16
50	2	0	0.28	0	7.04	7.04
52	35	1	4.94	0.17	7.09	7.14
54	2	0	0.31	0	6.45	6.45
56	13	1	2.03	0.12	6.42	6.30
58	25	1	3.30	0.27	7.58	7.94
60	18	1	3.09	0.69	5.83	7.08
62	17	2	2.47	0.51	6.88	7.65
64	4	0	0.64	0	6.27	6.27
66	1	0	0.26	0	3.80	3.80
67	1	0	0.10	0	9.70	9.70
68	6	0	0.79	0	7.55	7.55

El contraste de velocidades de turnos de habla contiguos se puede apreciar en la figura 15, donde también es posible detectar los fragmentos de la conversación donde se presentan distintos

patrones de velocidad. Siguiendo las líneas imaginarias formadas por la unión de los puntos de cada locutor, se observa que el *sujeto 2* se mantiene ligeramente por encima del *sujeto 1*, y que su variación es menor que la de su interlocutor durante casi toda la grabación. Por su parte, el *sujeto 1* presenta cambios radicales de una sección de la conversación a otra, es decir, en algunas partes alcanza medidas muy bajas de velocidad de elocución, en otras logra medidas similares a las del *sujeto 2* y en algunos momentos supera ligeramente la rapidez de su interlocutor. Por lo tanto, la línea del *sujeto 1* es mucho más irregular, lo cual indica que tuvo que recurrir a diferentes cambios para mantener abierto el canal de comunicación.

Desde otro ángulo, en la figura 16 se ponen de relieve las cualidades de la interacción entre los dos hablantes. En general, las diferencias en las medidas de velocidad de elocución entre turnos de habla contiguos no son muy grandes, aunque sí hay segmentos de la grabación en los que la distancia entre picos y valles es muy pronunciada. Cabe destacar que la parte central de la conversación, entre los *turnos 27 y 42*, es la que presenta el más prolongado punto de acuerdo de todo el corpus. Por las características de la dinámica establecida en la *grabación 2*, se puede afirmar que la conversación tuvo el nivel de acuerdo común suficiente para que la conversación se mantenga estable a pesar de sus momentos de conflicto. Además, la duración del audio confirma que el *sujeto 1* alcanzó la meta de ganar tiempo, aunque no se puede saber si fue suficiente, y que el *sujeto 2* logró que se aceptaran las condiciones del intercambio, porque los puntos de acuerdo solo se pueden establecer bajo los términos de este locutor.

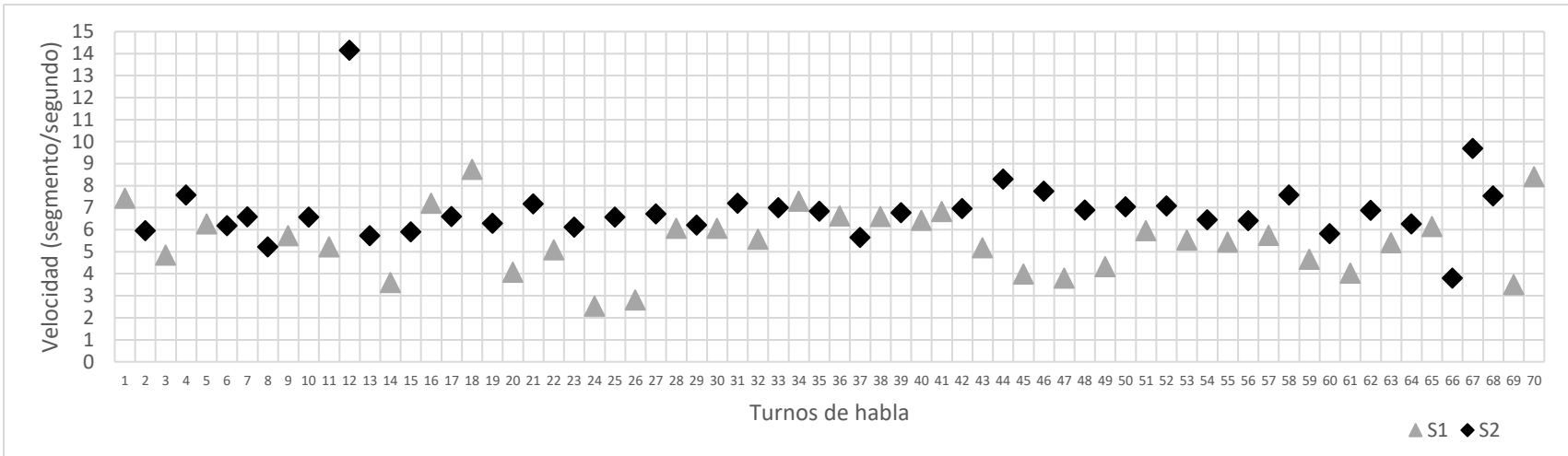


Figura 15. Velocidad de los turnos de habla para cada sujeto en la grabación 2

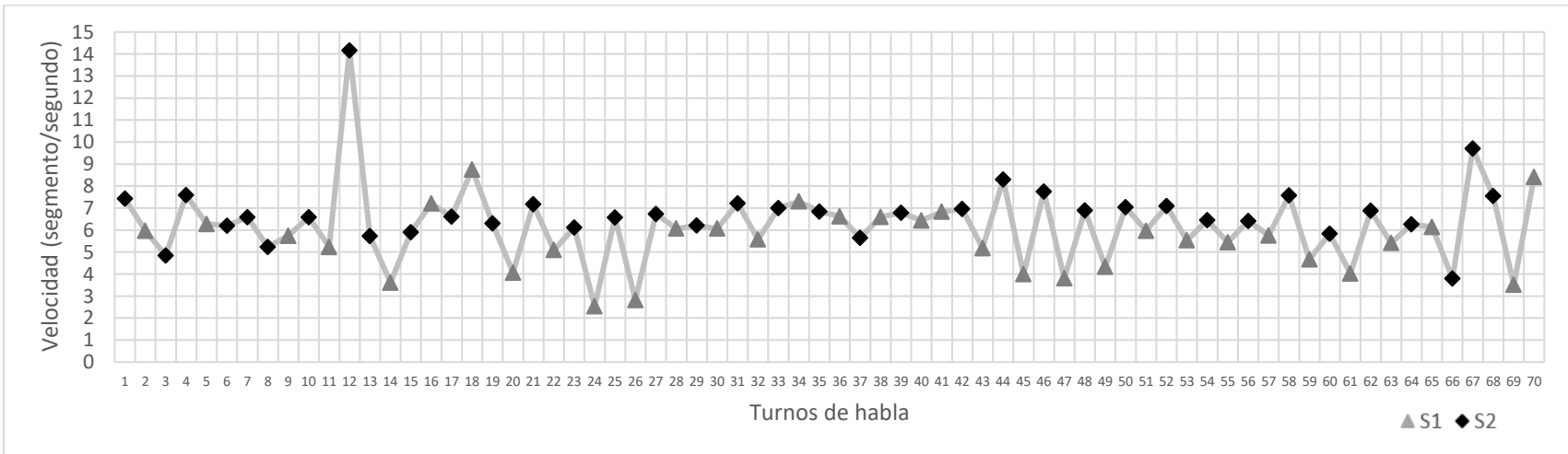


Figura 16. Velocidad en el continuum del tiempo de la grabación 2

4.2.2.2 Valores de relevancia adquiridos

Cuando las velocidades de dos turnos de habla contiguos superan los 2 s./s. considero que existe un punto de conflicto, en caso de ser menor a 2 s./s. considero que existe un punto de acuerdo (ver 4.2.1.2).

4.2.2.2.1 Conflicto

El primer punto de conflicto aparece entre los *turnos 15 y 18*, como se puede observar en la tabla 49 y en la transcripción ortográfica correspondiente. En esta sección, a pesar de que no se presenta ningún tipo de información que pudiera considerarse agresiva, ninguno de los dos está escuchando a su interlocutor, por lo que sus mensajes solo son interrupciones inconexas. Además, el *sujeto 1* le roba la palabra al *sujeto 2* en el *turno 18*, por lo que el conflicto se alarga hasta la sección entre los *turnos 23 y 26*, como se ve en la tabla 50 y en su transcripción ortográfica. En esta prolongación también ocurre una falta de entendimiento de ambas partes, mientras el *sujeto 2* pronuncia amenazas, el *sujeto 1* irrumpe con negaciones largas, sin que ninguno pueda establecer un acuerdo.

- G2S2T15:** *Yo te juro por Dios, /oilo bien, /te juro por Dios que tu hijo 'stá bien, cabrón, // ¿sale?*
G2S1T16: *Sí, no, yo quiero arreglarme contigo.*
G2S2T17: *No vamos a tocar a tu hi-*
G2S1T18: *Quiero arreglarme contigo.*

Tabla 49. Sección 1 de conflicto de la grabación 2

Sujeto	Turno	Duración (s.)	Velocidad (s./s.)	Transcripción (PG)
S2	T15	4.40	5.90	'YOte'XUropor'DIOS#'oilo'BIEN#te'xuroPOR'DIOSK E'TUIXOS'ta'BIENka'BRON##'sale
S1	T16	1.66	7.21	'SInoyo'kieroare'GLARmeKON'TIgo
S2	T17	1.06	6.61	NO'baMOSTo'kara'tui
S1	T18	1.03	8.75	kieroare'glarMEkon'tigo

- G2S2T23:** *Pero, //pero /te /pasas de verga, mi'jo.*
G2S1T24: *No.*

G2S2T25: *Y al chi- / va a valer / esto.*

G2S1T26: *No.*

Tabla 50. Sección 2 de conflicto de la grabación 2

Sujeto	Turno	Duración (s.)	Velocidad (s./s.)	Transcripción (PG)
S2	T23	2.45	6.12	'pero##'pero#te#'pasade'BERga'miXO
S1	T24	0.39	2.54	NO
S2	T25	1.37	6.57	ial'ci#'baba'LER#'esto
S1	T26	0.35	2.82	NO

En las dos primeras secciones de conflicto de la *grabación 2*, el *sujeto 2* realiza un gran número de pausas, en su mayoría breves, ello muestra poca seguridad y nulo poder de convencimiento. Por su parte, el *sujeto 1* alcanza velocidades de hasta 8 s./s. en la sección 1 (ver tabla 49), superando ampliamente a su interlocutor, y llega a puntos muy bajos en la segunda sección. En ningún caso le funcionó modificar su velocidad para persuadir al *sujeto 2*, pero logró mantener activo el canal de comunicación. Por otro lado, se puede apreciar que el *sujeto 2* trata de conciliar con su receptor acentuando el fragmento «por Dios» en el *turno 15* (ver tabla 49).

El siguiente punto de conflicto se presenta entre los *turnos 44* y *48*, como se puede comprobar en la tabla 51 y en su correspondiente transcripción. Se podría considerar que, según el contenido de los turnos de habla, en este fragmento del audio no existe conflicto, no obstante, en el *turno 47* se observa que el *sujeto 1* trata de prolongar su turno, pero el *sujeto 2* lo interrumpe. Por lo tanto, se presume que en el *turno 45* del *sujeto 1* tampoco se aceptan las condiciones del intercambio y que solo trata de que el *sujeto 2* crea que el pacto está hecho. Sin embargo, el *sujeto 2* no está poniendo atención a lo que quiere decir el *sujeto 1* porque en turnos anteriores ya ha hablado demasiado sin llegar a ninguna parte, por lo que impide su participación.

Se aprecia que el *sujeto 2* acelera durante gran parte de los *turnos 46* y *48*, mientras que el *sujeto 1* solo pronuncia dos afirmaciones muy largas, una de ellas seguida por un intento fallido de tomar la palabra: «em...». La información que el segundo locutor lanza al aire incluye información del

hijo: «porque tu chavo está»; una advertencia: «jugar al verga, mi'jo, eh»; y una oportunidad: «yo te vo'a dar».

- G2S2T44:** *Agarra'l pedo.*
G2S1T45: *Sí.*
G2S2T46: *Yo te voy a dar chance, / yo te vo'a dar chance, / pero no le quieras jugar al verga, mi'jo, eh.*
G2S1T47: *Sí, em...*
G2S2T48: *Porque tu chavo está bien paniquiado, / 'tá bien paniquia'o.*

Tabla 51. Sección 3 de conflicto de la grabación 2

Sujeto	Turno	Duración (s.)	Velocidad (s./s.)	Transcripción (PG)
S2	T44	0.60	8.31	a'gaRAL'pedo
S1	T45	0.25	3.99	SI
S2	T46	3.87	7.75	'yoTE'boiaDAR'CANse#yote'boadar'canse#'pero'N Ole'KIEraxu'garal'berga'mixo'e
S1	T47	0.53	3.80	'SIEM
S2	T48	2.61	6.90	'porketu'caboes'ta'BINpani'KIAdo#ta'BIENpani'KIA O

4.2.2.2.2 Acuerdo

El primer punto de acuerdo en la *grabación 2* ocurre entre los *turnos 5 y 10*, como se puede comprobar en la *tabla 52* y en su correspondiente transcripción ortográfica. A pesar de la hostilidad presente en estos turnos de habla, el *sujeto 1* logra mantener el turno de palabra ignorando los intentos de réplica del *sujeto 2* (*turnos 6 y 7*), mientras que este último trata de establecer los términos del intercambio. Aunque no haya respuestas positivas de ninguna de las partes, el hecho de que las intervenciones se realicen desde un nivel de igualdad, como un regateo, implica que ambos quieren llegar a un acuerdo. Sin embargo, el *sujeto 1* no propone un método de pago real, solo niega las posibilidades de realizar las sugerencias del *sujeto 2*.

- G2S1T05:** *¿Tú crees que, / tú crees que voy a poner en riesgo la vida de mi'jo? // Como q-, yo quiero arreglarme contigo.*
G2S2T06: *(Yo te...)*
G2S2T07: *(Por eso...)*

- G2S2T08:** *Por eso, cabrón, yo al chile sé que tú tienes billete, hijo, // eh. // Dile, | dile a tu jefa que venda'l X-trail, venda'l pinche Jetta.*
- G2S1T09:** *Pero ps, ¿quién me los compra así de momento, ps yo ofr- ofrecí mis cosas a mitad de precio. // No las quieren, o sea, la gente no tiene igual dinero. // Deja ver a...*
- G2S2T10:** *¿Por qué no hipotecan sus casas?*

Tabla 52. Sección 1 de acuerdo de la grabación 2

S	T	Dur. (s.)	Vel. (s./s.)	Inicio (s.)	Fin (s.)	Transcripción (PG)
1	05	5.90	6.27	32.00	37.90	tu'KRESKE#tu'KRESke'boia#'boiapo'neren'RIESgola'bidade'miXO##'komokyo'kieroare'GLARmeKON'TIGO
2	06	0.32	6.19	32.62	32.94	'yoTE
2	07	0.46	6.59	36.62	37.08	po'RESO
2	08	8.23	5.23	38.70	46.93	po'REsoka'BRONYO'yoyal'cileyo'seketu'tieNESbi'yeTIXO##e##'dile#'dileatu'xefake'BENDalecsTREIL'BENDAL'PINCE'yeTA
1	09	8.88	5.74	47.04	55.92	'perops'kienmelos'kompra'sideMO'MENtoPSyo'YOFROFRE'SIMIS'kosasami'tade'PREcio##'noLAS'KIEReno'seaLA'XENteno'tieni'GUALdi'NERO##'deXA'bera
2	10	1.37	6.58	55.59	56.95	por'kenoipo'teKANsu'kaSAS

Sin contar los solapamientos del *sujeto 2* (figura 17), la mayor parte de cada turno de habla está constituida por fragmentos con pequeñas duraciones. Las partes que el *sujeto 1* resalta son «contigo» y «yo ofrecí mis», es decir, está poniendo de relieve que pretende establecer un diálogo con su interlocutor y que ya intentó cumplir con los términos monetarios de la transacción. Por otro lado, el *sujeto 2* resalta fragmentos como «eso», «tu hijo» y «venda el pinche», o sea, primero trata de tomar la palabra con una intervención larga, sin lograrlo, y después utiliza a la víctima para presionar el *sujeto 1*, para, finalmente, proponerle opciones. El uso de las pausas es más o menos equitativo en esta sección, por lo que no existe ninguna cualidad relevante en cuanto a la sensación perceptiva que pudieran modificar en el oyente.

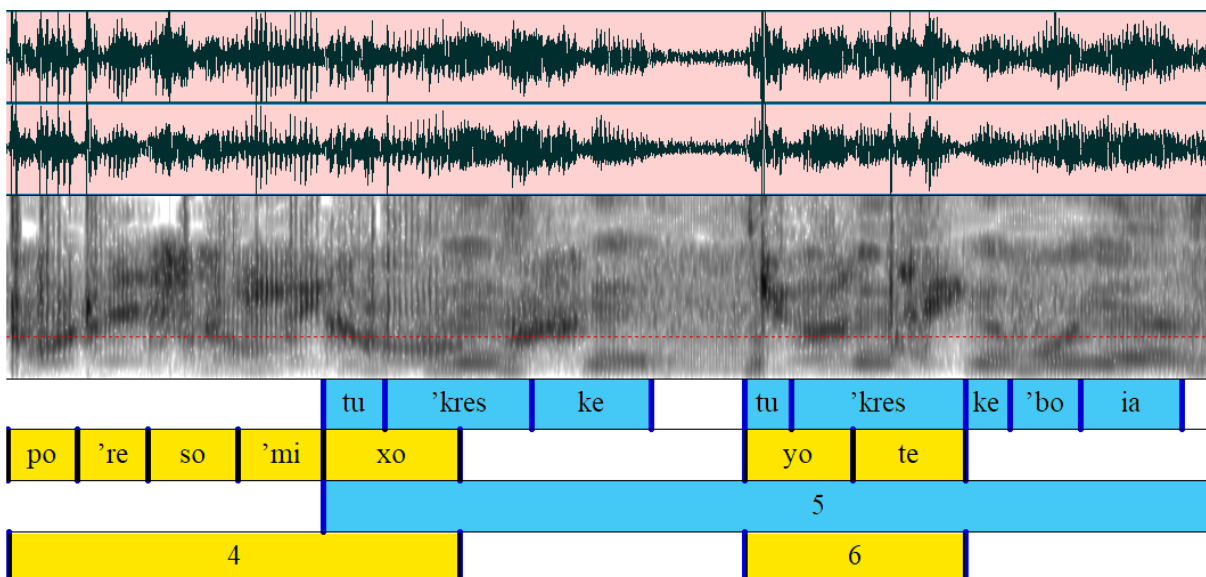


Figura 17. Solapamiento de los turnos 4, 5 y 6 de la grabación 2

El segundo momento en el que hay acuerdo entre los dos hablantes ocurre entre los *turnos* 33 y 36, como se observa en la tabla 53 y en su transcripción ortográfica. El diálogo que se desarrolla en esta parte del audio se da de manera casual, mientras que el *sujeto 2* advierte, el *sujeto 1* afirma que no irá en contra de lo que dice su interlocutor, aceptando así las condiciones del intercambio. Al igual que en el ejemplo anterior, las pausas no son un elemento relevante. Los elementos alargados del *sujeto 2* son «madre» (*sujeto 2*), acentuando la advertencia, y «por Dios» (*sujeto 1*), recurriendo a divinidades como comprobante de que no se traicionará el pacto.

G2S2T33: *Te juro / por Dios que, si vas con la tira, wey, va a valer madre e'to, eh.*

G2S1T34: *No, ¿cómo crees? Te juro por Dios que no he ido, / o sea, ¿tú crees que yo voy a poner en riesgo la vida de mi'jo?*

G2S2T35: *Bueno, tú sabrás, porque esos putos tú los conoces y yo tengo gente en todos lados, / y esos weyes lo que te van a decir / “tú sigue adelante, sigue adelante con la nego-”*

G2S1T36: *“A-, siga adelante”, ¿cómo voy a...? || Yo no vo-, yo no puedo poner en riesgo la vida de mi'jo, tú lo sabes.*

Tabla 53. Sección 2 de acuerdo de la grabación 2

Sujeto	Turno	Duración (s.)	Velocidad (s./s.)	Transcripción (PG)
S2	T33	3.00	7.00	te'XUro#por'DIOSkesi'BASKonla'Tira'UEIbabaLER' MADREto'e
S1	T34	4.80	7.30	NO'komo'KREStelo'XUroPOR'DIOSke'noe'iDO#'se atu'KRESkeyo'boiapo'neren'RIESgola'bidade'mixo
S2	T35	8.04	6.84	'bueno'TUsa'BRASPORkeSOS'PUTOS'tulosko'nose SIyo'TENgo'XENten'todo'lados#ieSO'ueyelokete'ban ade'SIR#tu'sigeade'LANte'sigeade'LANtekonlanego
S1	T36	4.83	6.62	A'siGA'LANTE'komo'boIA##'yonobo'yono'puedopo' neREN'riesgola'bidae'MIXo'tulo'SABES

La última parte donde se presenta acuerdo entre los dos sujetos aparece entre los *turnos* 54 y 56, como se puede constatar en la tabla 54 y en su correspondiente transcripción ortográfica. En este fragmento de la grabación se observa la rotunda aceptación de los términos establecidos por el *sujeto 2* para la realización de la transacción. En el *turno 54* se lanza una pregunta que espera una confirmación por parte del *sujeto 1* que sí ocurre, por lo que la respuesta dada en el *turno 56* es muy tranquila y contiene un mensaje para tranquilizar al *sujeto 1*.

- G2S2T54:** ¿Sale?
G2S1T55: Sí. // Dame, | dame chance nada más.
G2S2T56: Sí, no hay pedo, | tu hijo 'stá bien, cabrón.

Tabla 54. Sección 3 de acuerdo de la grabación 2

Sujeto	Turno	Duración (s.)	Velocidad (s./s.)	Transcripción (PG)
S2	T54	0.31	6.45	'saLE
S1	T55	2.20	5.44	SI##'dame#'dame'CANse'nada'MAS
S2	T56	2.03	6.42	'SI'noai'PEDO#i'TUIXOS'ta'BIENka'BRON

El *turno 55* del *sujeto 1* contiene mucho tiempo de silencio, pero no es algo intolerable para el *sujeto 2*, ya que ha recibido una respuesta positiva. Además, el *sujeto 1* compensa el tiempo de las pausas hablando rápido y el *sujeto 2* acentúa las palabras «tu hijo» como recurso conciliador entre las dos partes. De esta manera, esta última sección de acuerdo constituye el ideal del intercambio

de información de una transacción, el «propietario» de lo que se va a transferir, en este caso la víctima, solo espera que quien vaya a pagar acepte sus condiciones, y, cuando se logra, no es necesario recurrir ni a advertencias ni a amenazas.

4.2.2.2.3 Casos excepcionales

El fragmento de la *grabación 2* que se presenta en la tabla 66 y en su correspondiente transcripción ortográfica debería indicar conflicto debido al contraste tan grande entre las velocidades de los dos turnos de habla. Sin embargo, el contenido semántico muestra que en este diálogo se establece un acuerdo en el que el *sujeto 1* acepta la orden de darse prisa en conseguir el dinero del rescate. A pesar de que este ejemplo contradiga lo anterior respecto a las condiciones de velocidad de habla necesarias para establecer un conflicto o un acuerdo, confirma la idea de que el *tempo* en sí mismo no añade una carga significativa específica, sino que resaltan las formas que los demás elementos del discurso ya han dado a los enunciados.

G2S2T68: *En putiza, mi'jo.*

G2S1T69: *Sí.*

De manera similar, los dos elementos que presentan los picos más altos de velocidad en ambas grabaciones no adquieren ninguna carga de significado nueva por el hecho de presentar una aceleración muy alta, sino que, simplemente, tienen esa característica de rapidez porque son elementos inconclusos. En el *turno 25* de la *grabación 1* no se logra realizar una sentencia completa y en el *turno 12* de la *grabación 2* ni siquiera se articula una palabra entera.

G1S2T25: *(Quiero...)*

G2S2T12: *(Po-)*

5 Conclusiones

5.1 *Time is money*

Este estudio reveló que el extorsionado (*sujeto 1*) usa un tempo más lento que el de su interlocutor, lo cual muy probablemente obedezca al miedo, pero, al mismo tiempo, actúa como una defensa que le permite lograr resultados a su favor, ya que busca extender la duración de cada llamada telefónica para que la policía tenga suficientes elementos y/o ganancia de tiempo para localizar a los responsables del secuestro. Por otro lado, el extorsionador (*sujeto 2*) también está presionado por el miedo, por lo que utiliza la menor cantidad de tiempo posible para convencer a su interlocutor y evitar poner en riesgo su propia libertad.

El *sujeto 2* sabe que hablar rápido desde la posición de poder en la que se encuentra condiciona respuestas de menor velocidad sin importar la longitud del enunciado, por lo que el *sujeto 1* responde según lo que acaba de escuchar. Si el extorsionado actuara de manera consciente durante todas sus intervenciones, evitando alejarse demasiado de la diferencia de los 2 segmentos/segundo y sin rebasar la medida de velocidad de elocución de su interlocutor, podría alcanzar su propósito.

El contraste existente entre la velocidad de habla de turnos contiguos responde al punto de conformidad o acuerdo que existe entre las dos partes de la transacción. Un contraste marcado, es decir, un punto de «conflicto», anuncia la «victoria» del extorsionador. Dicha «victoria» no implica que se alcance el objetivo propio (monetario), sino el impedimento parcial o total de que su interlocutor alcance el suyo, es decir, provoca el retroceso del extorsionado, negándole la posibilidad de réplica. De manera similar, las velocidades contiguas cuya diferencia no es mayor a 2 segmentos/segundo (s./s.) implican la «victoria» del extorsionado, pues la estabilidad («acuerdo») quiere decir que la comunicación puede extenderse por más tiempo.

Como se vio en este caso, la abundancia de momentos conflictivos se genera debido a la tensión y el miedo, y si no se alcanzan muchos puntos de acuerdo, si las velocidades de ambos locutores no se acercan lo suficiente, no puede existir una negociación y ninguno podrá alcanzar sus objetivos. Los hablantes responden según lo que escucharon antes de intervenir, si reaccionan como un acto de reflejo lo más seguro es que se inicie un punto de conflicto, pero, si se consideran los aspectos particulares del acto comunicativo, se puede elegir una opción adecuada de respuesta.

5.1.1 Valores de relevancia adquiridos

La velocidad por sí misma no puede determinar absolutamente nada más que la rapidez concreta del habla, los valores de relevancia solo se adquieren cuando se puede contrastar una medida de velocidad con otra, siempre que el contenido del mensaje lo permita. No existe la posibilidad de establecer parámetros en los que la velocidad implique necesariamente alguna carga de significado, es decir, las velocidades lentas no siempre indican inseguridad al hablar, sino que también pueden ayudar en la exposición de algún tipo de información que requiera total atención y comprensión por parte del receptor; o puede indicar estados de ánimo diversos. Lo único que coincide en todas las velocidades de habla lentas es que serán oídas por completo, dependerá del receptor la respuesta que decida dar.

Del mismo modo, las velocidades altas jamás implican ninguna carga de significado por sí mismas, tienen que contrastarse, tanto pueden utilizarse para evitar que el receptor tome la palabra, como para que no escuche lo que se está diciendo, o, simplemente, se trata de un recurso involuntario que se adhiere a las realizaciones de los hablantes. Lo que tienen en común los usos de altas velocidades de habla es que se presenta una enorme cantidad de sonidos que no se pueden entender con claridad y siempre habrá pérdida de información, por lo que la respuesta del

interlocutor puede ser conflictiva.

5.1.2 Uso contrastivo del *tempo*

A partir de la interacción entre los dos locutores del corpus, se pueden formular algunas claves que para mantener abierto un canal comunicativo en un ambiente hostil. En primer lugar, responder muy lento después de una enunciación muy rápida puede ser la causa de un conflicto, es decir, hay que tratar de que la velocidad de cada turno de habla sea similar a la del turno inmediatamente anterior, pero jamás superarla. Hacer pausas está bien, siempre y cuando no sean muy largas, también es necesario evitar las palabras de negación o disyuntivas de una o dos sílabas con una duración exagerada, ya que los silencios y los sonidos prolongados pueden resultar incómodos y provocar ansiedad en el interlocutor.

Robar la palabra y/o impedir que el otro hable aumenta la tensión y puede frustrar al interlocutor, lo cual podría hacer que el diálogo se torne agresivo. Hablar demasiado rápido puede ocasionar lo mismo. Es importante nunca detenerse, siempre hablar, aunque el mensaje pueda llegar a resultar inconexo, porque hay que evitar el silencio y responder con la velocidad adecuada. Finalmente, lo primordial sería procurar nunca llevar la contraria, y, si pasa, habría que intentar que la velocidad sea similar a la del turno inmediatamente anterior.

Sin embargo, o es pertinente usar esto como una guía en situaciones de peligro, como en una extorsión, porque 1) los riesgos de cometer un error fatal son muy elevados, 2) las muestras obtenidas del corpus no son suficientes para determinar si los contrastes de rapidez son comunes a todo tipo de conversación, y 3) los contextos de comunicación siempre son distintos.

5.2 Pertinencia de la velocidad de habla

Después de la comparación y revisión de las medidas del tempo de los dos hablantes del corpus,

se concluye que la velocidad de habla no constituye una característica que pueda distinguir de manera contundente a un hablante de otro, ya que la rapidez de ambos es muy cercana. Por lo mismo, tampoco se puede establecer ningún patrón de comportamiento de aspectos temporales, es decir, las duraciones de los elementos y la rapidez de los turnos de habla no se repiten de manera sistemática, sino que la variación responde a la interacción que se da entre ambos. Además, la situación victimario-víctima condiciona las características más marcadas, o sea, que el extorsionador hable aceleradamente y que el extorsionado hable muy lento es otra cosa una manifestación acústica de la relación de poder entre los dos locutores.

5.3 Distribución temporal

En el proceso para determinar la distribución temporal interna de los turnos de habla, se detectó que tres de las cuatro versiones del corpus tienen problemas graves para el cálculo. El corpus por grabación y sujeto y el corpus total ignoran todo lo relacionado con el contexto de la situación comunicativa, es decir, estas dos posturas consideran que los fenómenos temporales son fortuitos y que no es relevante la situación comunicativa particular, ni si existe o no un interlocutor. Mientras que la postura del corpus por sujeto implica que la duración y el *tempo* son características identitarias de los hablantes y que siempre ocurren de manera similar en todos los contextos. Por ello, la postura del corpus por grabación resulta el camino más viable, ya que separa cada acto comunicativo y considera relevante la interacción entre ambos sujetos, porque dicha interacción condiciona las características de los mensajes que se intercambian.

Además, después de descartar aquellos elementos cuya duración no es marcadamente larga o breve, se encontró que los alargamientos de ciertas partes de los enunciados no siempre coinciden con elementos de mayor carga significativa, como insultos y súplicas, sino que pueden presentarse

como característica del habla del centro de México al final de un grupo fónico. Las secciones que aportan información relevante siempre aparecen al interior de los turnos de habla y son fragmentos con significado completo, o sea, palabras y frases.

Los segmentos por sí mismos no presentan ninguna carga significativa, es necesario verlos en el continuum del tiempo y contrastarlos con sus adyacentes. Si bastara con que un elemento presentara una duración determinada, todos los fenómenos temporales se considerarían fortuitos. Esta clase de análisis realizado en un corpus de mayor extensión podría confirmar los resultados obtenidos o señalar algunas otras consideraciones pertinentes.

6 Referencias

- Alarcos Llorach, E. (1974). *Fonología española*. Madrid: Gredos.
- Alto al secuestro. (2018). *Alto al secuestro*. Recuperado de <http://altoalsecuestro.com.mx/estadisticas/>
- Asuaje, R. A., Blondet, M. A., Mora, E., y Rojas, E. (2005). Codificación prosódica de la información incidental en el discurso espontáneo. Un estudio de caso. *Fermentum. Revista Venezolana de Sociología y Antropología*, 44, 449-460.
- Battaner, E., Gil J., Marrero, V., Llisterri J., Carbó, C., Machuca, M. J., Madrigal, N., Riera, M. y Ríos, A. (2003). VILE: Estudio acústico de la variación inter e intralocutor en español. *SEAF 2003. Actas del II Congreso de la Sociedad Española de Acústica Forense*, 59-70.
- Blondet, M. A., Méndez, J., y Mora, E. (2003). Breve descripción de la duración silábica como señal de frontera sintáctica en tres velocidades de elocución. *Lengua y Habla*, 8, 9-16.
- Cederborg, A. Á., y Chávez Aguilar, R. (1992). Modificaciones lingüísticas y conversacionales en el habla del profesor de español como L2. *Estudios de Lingüística Aplicada*, 15.
- Contreras, S. (2013). La velocidad de habla como elemento diferenciador entre el registro formal e informal del habla. *Lengua y Habla*, 17, 237-243.
- D'Introno, F., Teso, E. d., y Weston, R. (1995). *Fonética y fonología actual del español*. Madrid: Cátedra.
- Escandell Vidal, M. V. (2014). *La comunicación. Lengua, cognición y sociedad*. Madrid: Akal.
- Foucault, M. (2010). *Obras esenciales*. Madrid: Paidós.
- Gil Fernández, J. (2007). *Fonética para profesores de español: de la teoría a la práctica*. Madrid: Arco Libros.
- Gili Gaya, S. (1975). *Elementos de fonética general*. Madrid: Gredos.

- Grice, P. (2012). *La búsqueda del significado*. Madrid: Tecnos.
- Hidalgo Downing, R. (2003). *La tematización del español hablado*. Madrid: Gredos.
- Hidalgo Navarro, A., y Quilis Merín, M. (2004). *Fonética y fonología española*. Valencia: Tirant lo Blanch.
- Llisterri, J. (2016). Los elementos suprasegmentales. Recuperado de http://liceu.uab.es/~joaquim/phonetics/fon_prosod/suprasegmentales.html
- Machuca, M. J. (2009). Locución y prosodia en los medios de comunicación oral. En S. Alcoba (ed.), *Lengua, comunicación y libros de Estilo* (pp. 107-121). Barcelona: Premisas.
- Machuca, M., y De la Mota, C. (2006). Estrategias pragmlingüísticas orales: El énfasis en la publicidad. *Actas del XXXV Simposio Internacional de la Sociedad Española de Lingüística*, 1126-1142.
- Manuel Jerez, J. d. (2003). Medios audiovisuales e informáticos en la formación de intérpretes. En M. Á. Vega Cernuda (ed.), *Una mirada al taller de San Jerónimo: bibliografías, técnicas y reflexiones en torno a la traducción* (pp. 53-62). Madrid: Instituto Universitario de Lenguas Modernas y Traductores.
- Martínez Celdrán, E. (1983). *Fonética (con especial referencia a la lengua castellana)*. España: Teide.
- Martínez Sánchez, F., Meilán, J. J., Carro, J., Íñiguez, C. G., Millian Morell, L., Valverde, I. P., y López, D. E. (2016). Estudio controlado del ritmo del habla en la enfermedad de Parkinson. *Neurología*, 31, 466-472.
- Martínez Sánchez, F., Meilán, J. J., García Sevilla, J., Carro, J., y Arana, J. M. (2013). Análisis de la fluencia lectora en pacientes con la enfermedad de Alzheimer y controles asintomáticos. *Neurología*, 28, 325-331.

- Musselman Shank, R. (2006). El mantenimiento del turno como estrategia de dominio de la palabra. En B. P. (ed.), *Líderes lingüísticos: estudios de variación y cambio* (pp. 159-184). Ciudad de México: Colegio de México.
- Navarro Tomás, T. (2004). *Manual de pronunciación española*. Madrid: Ministerio de Educación y Ciencia.
- Quilis, A. (1981). *Fonética acústica de la lengua española*. Madrid: Gredos.
- Quilis, A. (1999). *Tratado de fonología y fonética españolas*. Madrid: Gredos.
- Schwab, S. (2015). Las variables temporales en el español de Costa Rica y de España: un estudio comparativo. *Revista de Filología y Lingüística de la Universidad de Costa Rica*, 41, 127-139.

Índice de tablas

Tabla 1. Alfabetos fonéticos para el español de México	26
Tabla 2. Segmentos	30
Tabla 3. Dos segmentos del <i>sujeto 2</i> en la <i>grabación 1</i> que presentan irregularidad en el etiquetado.....	31
Tabla 4. Dos segmentos del <i>sujeto 2</i> en la <i>grabación 1</i> etiquetados a partir del mismo promedio	32
Tabla 5. Pausas con etiquetas irregulares del sujeto 1 en la grabación 1.....	32
Tabla 6. Datos del <i>turno 29</i> de la <i>grabación 1</i>	34
Tabla 7. Transcripciones del <i>turno 29</i> de la <i>grabación 1</i>	34
Tabla 8. Duración promedio por grabación y sujeto.....	38
Tabla 9. Pausas de G1S1 (s.)	39
Tabla 10. Sílabas de G1S1 (c. s.).....	40
Tabla 11. Pausas de G1S2 (s.)	41
Tabla 12. Sílabas de G1S2 (c. s.).....	42
Tabla 13. Pausas de G2S1 (s.)	43
Tabla 14. Sílabas de G2S1 (c. s.).....	43
Tabla 15. Pausas de G2S2 (s.)	44
Tabla 16. Sílabas de G2S2 (c. s.).....	45
Tabla 17. Segmentos del corpus por grabación y sujeto.....	45
Tabla 18. Variación de PGS.....	46
Tabla 19. Duración promedio por grabación	47
Tabla 21. Sílabas de G1 (c. s.)	48
Tabla 20. Pausas de G1 (s.).....	48
Tabla 22. Pausas de G2 (s.).....	50
Tabla 23. Sílabas de G2 (c. s.)	50
Tabla 24. Segmentos del corpus por grabación	51
Tabla 25. Variación de PG.....	52
Tabla 26. Duración promedio por sujeto	53
Tabla 27. Pausas de S1 (s.)	54
Tabla 28. Sílabas de S1 (c. s.).....	54
Tabla 29. Sílabas de S2 (c. s.).....	55
Tabla 30. Pausas de S2 (s.)	56
Tabla 31. Segmentos del corpus por sujeto	57
Tabla 32. Variación de PS	57
Tabla 33. Duración promedio total	58
Tabla 34. Pausas de T (s.).....	59
Tabla 35. Sílabas de T (c. s.).....	59
Tabla 36. Datos generales de los turnos de S1 en G1	62
Tabla 37. Datos generales de los turnos de S2 en G1	62
Tabla 38. Turnos de habla del <i>sujeto 1</i> en la <i>grabación 1</i>	63
Tabla 39. Turnos de habla del <i>sujeto 2</i> en la <i>grabación 1</i>	63

Tabla 40. Sección 1 de conflicto de la <i>grabación 1</i>	67
Tabla 41. Sección 2 de conflicto de la <i>grabación 1</i>	68
Tabla 42. Sección 3 de conflicto de la <i>grabación 1</i>	69
Tabla 43. Sección 1 de acuerdo de la <i>grabación 1</i>	71
Tabla 44. Sección 2 de acuerdo de la <i>grabación 1</i>	72
Tabla 46. Datos generales de los turnos de S2 en G2.....	74
Tabla 45. Datos generales de los turnos de S1 en G2.....	74
Tabla 47. Turnos de habla del <i>sujeto 1</i> en la <i>grabación 2</i>	75
Tabla 48. Turnos de habla del <i>sujeto 2</i> en la <i>grabación 2</i>	76
Tabla 49. Sección 1 de conflicto de la <i>grabación 2</i>	80
Tabla 50. Sección 2 de conflicto de la <i>grabación 2</i>	81
Tabla 51. Sección 3 de conflicto de la <i>grabación 2</i>	82
Tabla 52. Sección 1 de acuerdo de la <i>grabación 2</i>	83
Tabla 53. Sección 2 de acuerdo de la <i>grabación 2</i>	85
Tabla 54. Sección 3 de acuerdo de la <i>grabación 2</i>	85

Índice de figuras

Figura 1. Niveles de anotación en Praat.....	27
Figura 2. Solapamiento.....	28
Figura 3. Solapamiento.....	28
Figura 4. Script.....	28
Figura 5. Datos para ejecución del script.....	29
Figura 6. Archivo de resultados.....	29
Figura 7. Transcripción automática según PGS.....	33
Figura 8. Distribución del corpus por grabación y sujeto.....	46
Figura 9. Proporciones del corpus por grabación.....	52
Figura 10. Proporción del corpus por sujeto.....	56
Figura 11. Velocidad de los turnos de habla para cada sujeto en la <i>grabación 1</i>	64
Figura 12. Velocidad en el continuum del tiempo de la <i>grabación 1</i>	65
Figura 13. Solapamiento de los turnos 29 y 30 de la <i>grabación 1</i>	69
Figura 14. Solapamiento de los turnos 31 y 32 de la <i>grabación 1</i>	70
Figura 15. Velocidad de los turnos de habla para cada sujeto en la <i>grabación 2</i>	79
Figura 16. Velocidad en el continuum del tiempo de la <i>grabación 2</i>	79
Figura 17. Solapamiento de los turnos 4, 5 y 6 de la <i>grabación 2</i>	84