



Universidad de Concepción
Dirección de Postgrado
Facultad de Humanidades y Arte – Programa Magíster en Lingüística
Aplicada

DISEÑO, IMPLEMENTACIÓN Y VALIDACIÓN DE UN *SCRIPT* DE PRAAT
PARA MEDIR EL RENDIMIENTO DIADOCOCINÉTICO



Tesis presentada a la Facultad de Humanidades y Arte de la Universidad de Concepción para optar al grado académico de Magíster en Lingüística Aplicada

DANIELA CATALINA OSSES VALENZUELA
CONCEPCIÓN-CHILE
2021

Profesor Guía: Dr. Hernán Pérez Muñoz
Dpto. de Español, Facultad de humanidades y Arte
Universidad de Concepción

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento.



AGRADECIMIENTOS

En primer lugar, me gustaría agradecer profundamente a mis profesores guía Dr. Hernán Pérez y Dr. Domingo Román, quienes siempre estuvieron dispuestos a guiarme durante este proceso. Sus conocimientos y comprensión son invaluableles.

También, es importante agradecer a quien facilitó una parte importante de esta tesis, al diseñador web Mauricio Constanzo, por su disposición y profesionalidad. De la misma manera, agradecer a quienes aportaron con su dedicación, tiempo y conocimientos a etapas de este trabajo, al Dr. Mauricio Figueroa, cuya asesoría estadística fue elemental para proseguir, y a Roberto Jiménez, quien realizó una laboriosa y esencial contribución.

A mi generación del Magíster, por vivir en conjunto las etapas de este camino, apoyarnos y avanzar en un grato ambiente a pesar de la distancia. De la misma manera, agradecer a las de siempre, por todo.

Por último, a quienes me apoyaron durante el transcurso de esta tesis con cariño, contención y confianza incondicional. Gracias familia.

TABLA DE CONTENIDOS

AGRADECIMIENTOS.....	II
TABLA DE CONTENIDOS.....	III
ÍNDICE DE TABLAS.....	VI
ÍNDICE DE ILUSTRACIONES.....	VII
RESUMEN.....	IX
ABSTRACT	X
1. INTRODUCCIÓN	1
2. MARCO TEÓRICO.....	4
2.1 Diadococinesias	6
2.1.1 Indicadores diadococinéticos.....	7
2.1.2. Estímulos para obtener muestras de habla diadococinéticas	10
2.1.3. Metodologías para obtener la tasa de diadococinesias:	11
2.1.4. Enfoque del análisis	13
2.1.4.1 Análisis perceptivo	13
2.1.4.2 Análisis instrumental objetivo	15
3. PLANTEAMIENTO DEL PROBLEMA	23
3.1 Preguntas de investigación.....	26
4. OBJETIVOS.....	27
4.1 Objetivo general	27
4.2 Objetivos específicos	27

5. METODOLOGÍA	29
5.1 Diseño e implementación del <i>script</i> (etapa 1)	29
5.1.1 Segmentación automática de la señal	32
5.1.2 Cálculo de los indicadores.....	33
5.2 Validación del <i>script</i> propuesto (etapa 2)	36
5.2.1 Validación y estimación de la confiabilidad del <i>script</i>	36
5.2.1.1 Análisis no automatizado de las muestras de habla	37
5.2.1.2 Análisis mediante el programa <i>Motor Speech Profile</i>	38
5.2.1.3 Participantes	39
5.2.1.4 Corpus	40
5.2.1.5 Instrumento de recolección del corpus	40
5.2.2 Validación del análisis no automatizado	42
5.2.3 Análisis de la fiabilidad del instrumento	43
6. RESULTADOS.....	44
6.1 Resultados de la validación del <i>script</i>	44
6.1.1. Resultados del análisis de la validación del comportamiento del <i>script</i> sin aplicar corrección humana de la segmentación silábica automatizada .45	
6.1.1.1 Resultados de la validación de los datos del promedio de sílabas por segundo (PSS), sin aplicar corrección	45
6.1.1.2 Resultados de la validación de los datos del coeficiente de variabilidad del PSS, sin aplicar corrección.....	47
6.1.1.3 Resultados de la validación de los datos de la duración promedio de la sílaba (DPS), sin aplicar corrección.....	49
6.1.1.4 Resultados de la validación de los datos del coeficiente de variabilidad de la DPS, sin aplicar corrección.....	51
6.1.2 Resultados del análisis de validación del comportamiento del <i>script</i> , aplicando además la funcionalidad de corrección humana de la segmentación silábica automatizada	53
6.1.2.1 Resultados de validación de los datos del promedio de sílabas por segundo (PSS), con corrección.....	53
6.1.2.2 Resultados de la validación de los datos del coeficiente de variabilidad del PSS	58

6.1.2.3 Resultados de la validación de los datos de la duración promedio de la sílaba (DPS)	64
6.1.2.4 Resultados de la validación de los datos del coeficiente de variabilidad de la DPS	68
6.1.3 Comparación entre los resultados del <i>script</i> sin aplicar correcciones en la segmentación con los resultados del <i>script</i> aplicando además la funcionalidad de corrección no automatizada de la segmentación	75
6.1.3.1 Indicador promedio de sílabas por segundo (PSS)	75
6.1.3.2 Indicador coeficiente de variabilidad del PSS	76
6.1.3.3 Indicador duración promedio de la sílaba (DPS)	77
6.1.3.4 Indicador coeficiente de variabilidad de la DPS	78
6.2 Resultados del análisis de la fiabilidad inter-evaluador del análisis no automatizado	79
6.3 Resultados del análisis de la fiabilidad del <i>script</i>	85
7. DISCUSIÓN	91
8. CONCLUSIONES	101
9. PROYECCIONES	103
10. REFERENCIAS BIBLIOGRÁFICAS	104

ÍNDICE DE TABLAS

Tabla 6.1 Resultados del análisis estadístico descriptivo de los resultados del script sin aplicar correcciones no automatizadas del PSS	46
Tabla 6.2 Resultados ANOVA con correcciones de Greenhouse-Geisser	47
Tabla 6.3 Resultados del análisis estadístico descriptivo del script sin correcciones humanas del coeficiente de variabilidad del PSS.....	48
Tabla 6.4 Resultados del análisis descriptivo de los datos del script sin correcciones humanas de la DPS	50
Tabla 6.5 Resultados del análisis estadístico descriptivo del script sin correcciones humanas del coeficiente de variabilidad de la DPS	51
Tabla 6.6 Resultados de la comparación de efectos principales del coeficiente de variabilidad de la DPS.....	53
Tabla 6.7 Resultados del análisis estadístico descriptivo del script con corrección humana del PSS	54
Tabla 6.8 ANOVA de medidas repetidas	58
Tabla 6.9 Resultados del coeficiente de variabilidad del PSS	60
Tabla 6.10 Resultados del promedio de la duración silábica.....	64
Tabla 6.11 ANOVA de medidas repetidas	68
Tabla 6.12 Resultados del coeficiente de variabilidad de la duración silábica...	69
Tabla 6.13 Resultados del coeficiente de variabilidad de la duración silábica...	72
Tabla 6.14 Resultado ANOVA de medidas repetidas.....	75
Tabla 6.15 Resultados del ICC de la fiabilidad inter-evaluador.	84
Tabla 6.16 Resultados del ICC de cada indicador.	89

ÍNDICE DE ILUSTRACIONES

Figura 5.1 Diagrama de flujo del <i>script</i> propuesto.....	31
Figura 5.2 Informe final.....	36
Figura 5.3 Página web.....	41
Figura 6.4. Gráfico comparativo de cada una de las mediciones individuales del PSS para cada uno de los tres métodos.....	55
Figura 6.5 Histogramas de distribución del indicador PSS. Se muestran los datos obtenidos desde las tres metodologías analizadas (manual-script-MSP).	56
Figura 6.6 Gráficos QQ del indicador PSS.....	56
Figura 6.7 Gráfico comparativo de cada una de las mediciones individuales del coeficiente de variabilidad del PSS para cada uno de los tres métodos.....	59
Figura 6.8 Gráfico comparativo de cada una de las mediciones individuales del coeficiente de variabilidad del PSS para el script y en método no automatizado.	61
Figura 6.9 Histograma de la distribución y curva de normalidad del coeficiente de variabilidad del promedio de sílabas por segundo.	62
Figura 6.10 Gráfico QQ de la distribución del coeficiente de variabilidad del promedio de sílabas por segundo.....	63
Figura 6.11 Gráfico comparativo de cada una de las mediciones individuales de la DPS para cada uno de los tres métodos.....	65
Figura 6.12 Histogramas con curva de normalidad de la distribución del promedio de la duración silábica.....	66
Figura 6.13 Gráficos QQ de la distribución del índice promedio de duración silábica.	67
Figura 6.14 Gráfico comparativo de cada una de las mediciones individuales del coeficiente de variabilidad de la DPS para los tres métodos.....	70

Figura 6.15 Histogramas con curva de normalidad del coeficiente de variabilidad de la duración silábica.	71
Figura 6.16 Gráfico QQ de distribución del coeficiente de variabilidad de la duración silábica.....	71
Figura 6.17 Histogramas con curvas de normalidad de los datos del coeficiente de variabilidad del promedio de duración silábico luego del proceso de normalización.	73
Figura 6.18 Distribución del coeficiente de variabilidad de la duración silábica luego del procedimiento de normalización.	73
Figura 6.19 Gráfico comparativo de las mediciones del PSS realizadas por los dos evaluadores	80
Figura 6.20 Gráfico comparativo de las mediciones de la DPS realizadas por los dos evaluadores	81
Figura 6.21. Gráfico comparativo de las mediciones del coeficiente de variabilidad del PSS realizadas por los dos evaluadores	81
Figura 6.22 Gráfico comparativo de las mediciones del coeficiente de variabilidad de la DPS realizadas por los dos evaluadores	82
Figura 6.23 Gráfico comparativo de las mediciones del PSS de dos emisiones distintas realizadas por el mismo sujeto.....	86
Figura 6.24 Gráfico comparativo de las mediciones del coeficiente de variabilidad del PSS de dos emisiones distintas realizadas por el mismo sujeto	86
Figura 6.25 Gráfico comparativo de las mediciones de la DPS de dos emisiones distintas realizadas por el mismo sujeto.....	88
Figura 6.26. Gráfico comparativo de las mediciones del coeficiente de variabilidad de la DPS de dos emisiones distintas realizadas por el mismo sujeto	88
Figura 7.27 Comparación entre la segmentación automática del <i>script</i> y la segmentación con correcciones: Intensidad	92
Figura 7.28. Comparación entre la segmentación automática del <i>script</i> y la segmentación con correcciones: Relación señal-ruido	93

RESUMEN

El objetivo de esta investigación es implementar y validar un código de programación que ejecuta, en una muestra de habla digitalizada, los procedimientos necesarios para calcular automáticamente cuatro indicadores cuantitativos del rendimiento diadococinético: el promedio de sílabas por segundo, el coeficiente de variabilidad del promedio de sílabas por segundo, la duración promedio de la sílaba y el coeficiente de variabilidad de la duración promedio de la sílaba.

Para ello, primero, se ha diseñado e implementado un *script* de Praat que segmenta las sílabas presentes en una emisión diadococinética y que, con ello, calcula y entrega de manera automatizada los cuatro indicadores.

Luego, una vez implementado el *script*, con el fin de validarlo, se comparan los resultados obtenidos de la aplicación a una muestra de 30 hablantes de este recurso con los datos que provienen de otras dos metodologías que cumplen la misma función y que se aplican en la misma muestra: un análisis y cálculo no automatizado realizado por una persona y un análisis y cálculo automatizado mediante un programa comercial disponible en el mercado.

Los resultados del proceso de validación revelan similitudes en el rendimiento de las tres metodologías, lo que permite, como conclusión, ratificar la validez de la herramienta propuesta en este trabajo.

ABSTRACT

The aim of this study is to implement and validate a programming code that runs the procedures that are required to automatically calculate four quantitative indicators of diadochokinetic performance on a digitized speech sample: the average number of syllables per second, the coefficient of variation of the average number of syllables per second, the average syllable length, and the coefficient of variation of the average syllable length.

For this purpose, we have initially designed and implemented a Praat script that segments the syllables contained in a diadochokinetic utterance and thus automatically calculates and returns the four indicators.

Following the script deployment and for validation purposes, the results obtained from applying the script to a sample of 30 speakers were compared with data collected from two other methodologies that fulfil the same objective and that were both conducted on the same sample: a non-automated human-performed analysis and calculation and an automated analysis and calculation using a market-available software program.

The findings of the validation process reveal substantial parallels as to the overall performance of the three methodologies, thus supporting the validity of the tool presented in this study.

1. INTRODUCCIÓN

Los protocolos y recursos de evaluación para determinar las características de cada uno de los componentes del habla normal y patológica son variados. En esta variedad, comúnmente se pueden encontrar aquellas herramientas que se basan en la observación, la medición y el análisis del comportamiento diadococinético.

Con el pasar del tiempo, el procedimiento para ponderar cuantitativamente el rendimiento diadococinético se ha realizado a través de distintas metodologías de acuerdo con los recursos tecnológicos disponibles en la época y en el lugar en que se realiza la evaluación.

En la actualidad, el avance digital ha posibilitado la creación de herramientas que apoyan y agilizan esta tarea, lo cual beneficia a los investigadores y a los terapeutas, ya que así se obtienen resultados con mayor rapidez y de forma más objetiva. Muchas de estas herramientas se basan en distintos programas y plataformas computacionales. Algunos recursos ofrecen

algoritmos que realizan una parte del procedimiento, otros recursos lamentablemente por distintas razones no son accesibles a todos los potenciales usuarios y otros están disponibles, pero como programas desarrollados de forma comercial, cuya adquisición y uso involucra un alto costo monetario asociado.

En este contexto, el propósito de esta investigación es diseñar, implementar y posteriormente validar una herramienta tecnológica que realice automáticamente y semi automáticamente el cálculo de cuatro indicadores cuantitativos del rendimiento diadococinético. Además, la motivación principal de este estudio es generar y entregar a la comunidad académica y terapéutica un recurso útil y eficiente, pero que sea de código libre, abierto y sin costo monetario para los usuarios.

Es por ello, que la herramienta que se propone se trata de un *script* que se implementa y ejecuta dentro del ambiente de *software* del programa Praat (Boersma y Weenink, 2020).

Además de implementar este recurso, esta investigación comprende también la realización de un proceso de validación de la herramienta. Para ello, se recurre a varias estrategias metodológicas para comparar los resultados de la aplicación del *script* a una muestra real con los resultados de otros métodos aplicados a la misma muestra.

Con respecto a la organización de este texto, en el capítulo 2 (Marco Teórico), se presentan los antecedentes bibliográficos que respaldan esta

investigación y, desde los cuales además se deriva el problema de investigación, que se explica con detalle en el capítulo 3. A continuación, en el capítulo 4, se presentan los objetivos que guían el desarrollo de la investigación y, consecuentemente, el contenido que se expondrá en los siguientes capítulos. Después, en el capítulo 5, se detalla la metodología de trabajo, es decir, cada uno de los pasos, estrategias y procedimientos que se realizaron durante la investigación con el propósito de lograr el cumplimiento de los objetivos. El capítulo siguiente (capítulo 6) presenta los resultados de los distintos análisis estadísticos realizados para evaluar los datos y así probar la fiabilidad de la herramienta. En el capítulo 7 (Discusión), se discuten algunos aspectos del proceso de validación, principalmente, y se intenta también exponer algunas reflexiones sobre el alcance, las proyecciones y las limitaciones de la investigación realizada. Finalmente, en el capítulo 8, se presentan las conclusiones de este trabajo.

2. MARCO TEÓRICO

El habla es un proceso complejo que requiere la interacción de diversos sistemas, como el fonatorio, el neuromuscular y el respiratorio (Alshahwan *et al.*, 2020). Estos sistemas poseen correlatos neurológicos que controlan e impulsan su funcionamiento. Una alteración en cualquiera de estos correlatos repercute negativamente en la realización y ejecución del habla. En muchos casos, esta alteración deriva en algún trastorno del habla. Las características de estos trastornos varían dentro de un amplio espectro, lo que conlleva múltiples dificultades para quienes trabajan con ellos tanto en el ámbito de la investigación como en el de la terapia.

El proceso diagnóstico-terapéutico y el seguimiento de los usuarios con dificultades en ámbitos motores orales es una tarea compleja por la variabilidad inherente de la sintomatología, el desarrollo y la etiología, entre otros aspectos. Es por ello que, con el fin de apoyar la labor de los especialistas, resulta de vital importancia desarrollar herramientas de medición y evaluación que permitan conocer y comprender con mayor detalle y precisión las características del comportamiento de todos y cada uno de estos trastornos.

Para poder obtener más información sobre las características que presentan los trastornos motores del habla, así como del sustrato afectado, los especialistas se valen de una variedad de instrumentos para evaluar el comportamiento motor oral. Dada la alta variabilidad de los trastornos motores del habla, su evaluación incluye mediciones de todos los sistemas que participan en la producción del habla. Los datos que se obtienen de cada instrumento se combinan para realizar un análisis lo más completo posible con el fin de tomar decisiones terapéuticas apropiadas. Las distintas herramientas que se utilizan en la evaluación de un correlato específico de cada sistema poseen un grado de sensibilidad diferente según la patología estudiada (Duffy, 2013), por lo que, al momento de evaluar, se debe contar con un amplio rango de mediciones. Con el fin de lograr este objetivo, los profesionales dedicados al tema se apoyan en diversos indicadores del comportamiento del habla como, por ejemplo, el tiempo máximo fonatorio; los movimientos orales, mandibulares y linguales; la velocidad de lectura oral; los índices de rendimiento diadococinético (Toledo *et al*, 2011), entre otros. Esta investigación se focaliza en este último indicador del comportamiento del habla.

2.1 Diadococinesias

Las diadococinesias (conocidas por su sigla en inglés como DDK), en general, se definen como la capacidad de cualquier persona para producir movimientos rápidos de cualquier parte de su cuerpo de manera opuesta, repetitiva (Wang *et al.*, 2009) y regular. En específico, cuando esto se refiere a movimientos de la musculatura involucrada en la producción del habla, se les denomina diadococinesias orales. Las diadococinesias orales se definen como la emisión repetitiva y sin errores de una unidad articulatoria, ya sea monosilábica o multisilábica (Ziegler, 2002).

Con el fin de recabar información cuantitativa del rendimiento individual durante la ejecución esta tarea, los profesionales se valen de uno o varios indicadores cuantitativos derivados de la producción diadococinética. Para obtener estos indicadores, los evaluadores deben enfrentarse a varias decisiones que permiten delimitar de manera más precisa el alcance de su análisis. Entre ellas se encuentran, por ejemplo, a) cuál o cuáles indicadores diadococinéticos utilizar, b) cuál es el formato que debe dársele a los estímulos requeridos para evaluar la producción motora del habla, c) cuál debe ser la metodología para extraer los datos y realizar los cálculos y d) cuál debe ser el enfoque de análisis. A continuación, se desarrollarán estos cuatro puntos mencionados.

2.1.1 Indicadores diadococinéticos

A partir de la producción oral de las diadococinesias se pueden realizar diferentes cálculos para obtener variados indicadores cuantitativos que aportan información para caracterizar la emisión del habla. Tradicionalmente se ha asociado la producción diadococinética con la cuantificación del promedio de producciones (Wang *et al.*, 2009) en una unidad de tiempo, indicador conocido como la tasa de diadococinesias. Los resultados de la medición de la tasa de diadococinesia del habla aportan información relevante sobre la integración y maduración neuromuscular (Lundeen, 1950; Duffy, 2013).

La caracterización de la producción diadococinética de un hablante, que considera tanto la información proveniente de la tasa diadococinética como de su variabilidad en el tiempo, se requiere al momento de realizar una diferenciación entre hablantes con trastornos motores del habla y aquellos sin afección (Karlsson *et al.*, 2020). De este modo, se obtiene información valiosa al momento de tomar decisiones para determinar la presencia o ausencia de alteraciones del control motor del habla, o bien, de su grado de deterioro, si es que lo manifiesta (Wang *et al.*, 2004).

Además de la tasa de diadococinesias, existen otros indicadores cuantitativos provenientes de la producción diadococinética del habla que aportan información adicional sobre la ejecución motora del hablante. Varios de estos indicadores han surgido como consecuencia del avance de la tecnología,

que ha otorgado más herramientas para poder analizar una muestra de habla de manera más precisa y rápida. Un ejemplo de ello se encuentra en los indicadores propuestos en el programa *Motor Speech Profile* (desde ahora MSP) desarrollado comercialmente por la empresa *PentaxMedical* (ex *KayPentax*, ex *Kay-Elometrics*), el cual, sin embargo, cabe mencionar que tiene un costo monetario asociado para poder ser adquirido, que además es bastante elevado.

El *MSP* contiene varios protocolos de análisis del comportamiento motor del habla, entre los que se halla el que analiza las diadococinesias. En específico, el protocolo que viene en este programa para realizar un análisis cuantitativo de la producción diadococinética del habla, entrega cinco índices:

- a) DDKavp /ms/ (índice de duración promedio de la sílaba). Este indicador mide y entrega el valor promedio de la duración de la sílaba, es decir, es el tiempo promedio, medido en milisegundos (ms), entre cada una de las unidades CV (esto es, cada vocalización /pa/). Según el manual del programa, el DDKavp es inversamente proporcional al DDKavr.
- b) DDKavr (promedio de sílabas por segundo o *tasa de diadococinesias*) /s/s/. Este indicador es el promedio de la producción de sílabas en una unidad de tiempo durante la emisión. Definido de manera más específica, es el número de combinaciones CV (sílabas) por segundo. Según en

manual del programa, el valor de este indicador está inversamente relacionado con el valor del DDKavp.

c) DDKcvp (coeficiente de variabilidad del período DDK o porcentaje de variabilidad de la duración de las sílabas) /%/. Este parámetro mide, en términos porcentuales, el grado de variación de la duración de las sílabas. Así, según el manual del programa, si la vocalización CV se repite con poca variación en su duración medida en milisegundos, entonces este valor es bajo. Si la duración de cada sílaba es muy variable, este valor aumenta.

d) DDKjit (perturbaciones del ciclo DDK o porcentaje de variabilidad del período ciclo a ciclo) /%/. Según lo que expresa el manual del MSP, este parámetro mide, en términos porcentuales, el grado de variación ciclo a ciclo del período, sin especificar de manera más clara a qué se refiere con esto. Según el manual, si la vocalización CV se repite con poca variación ciclo a ciclo, entonces este valor es bajo.

e) DDKcvi (coeficiente de variación de la intensidad máxima DDK) /%/. Este parámetro mide, en términos porcentuales, el grado de variación de la intensidad en el pico de amplitud de cada vocalización CV. Según el manual del programa, si las sílabas se repiten con poca variación en la intensidad de cada una, entonces este valor es bajo. Por otro lado, si un hablante varía la intensidad las sílabas de su emisión, este número

aumenta. Este parámetro está evaluando la capacidad del sujeto para mantener una amplitud constante (es decir, la intensidad) durante la tarea de combinación de ciclos CV.

Cabe destacar que estos indicadores descritos son susceptibles a ser utilizados por parte del evaluador según su conveniencia o necesidad, sin que sea obligatoria la selección de todos o de alguno en particular para realizar una caracterización del comportamiento motor del habla.

2.1.2. Estímulos para obtener muestras de habla diadococinéticas

La medición de la tasa de diadococinesias del habla frecuentemente se realiza utilizando dos tipos de estímulos: la producción de unidades consonante-vocal (por ejemplo, /pa/ en la producción /pa.pa.pa.pa.pa/) y/o la emisión de secuencias silábicas, usualmente trisílabos (por ejemplo, /pa.ta.ka/ en la producción continua y seguida de esta forma). La revisión de la literatura revela que varios autores se valen de los segmentos consonánticos /p/, /t/ y /k/ asociados al segmento vocálico /a/ para realizar sus investigaciones (Alshawan *et al.*, 2020; Diepeveen *et al.*, 2019; Fletcher, 1972). La posibilidad de ejecutar movimientos musculares opuestos de forma alternada que sucede al combinar una consonante oclusiva y una vocal avala esta selección. Además, cada uno de estos segmentos (/p/, /t/ y /k/) aporta información sobre un grupo muscular

específico, siendo estos los labios, la mandíbula y la lengua (Duffy, 2013). Al respecto, cabe destacar que, en Chile, Toledo *et al.* (2011) realizan una investigación para establecer los parámetros de habla normales en adultos entre 40 y 69 años, divididos en tres grupos etarios, el primero de 40 a 49 años, el segundo de 50 a 59 y el último, de 60 a 69 años. Entre los aspectos evaluados se encuentra la repetición de sílabas /pa/, /ta/ y /ka/. Los autores concluyen, entre otros hallazgos, que la secuencia consonante-vocal /pa/ es la única que establece una diferencia significativa entre los evaluados.



2.1.3. Metodologías para obtener la tasa de diadococinesias:

Una vez que se obtiene la muestra de habla a analizar, históricamente se ha procedido a extraer el dato cuantitativo mediante algún método de medición y cálculo, entre los que se destacan principalmente: a) el enfoque tradicional, b) el método “*time-by-count*” y c) el método de contabilizar las producciones realizadas durante un ciclo respiratorio.

El enfoque tradicional consiste en contar el número de sílabas dado un tiempo específico, generalmente cinco a diez segundos (Gadesmann y Miller, 2008). Por ejemplo, este procedimiento es aplicado por Albright (1948), quien estudia las habilidades motoras de dos grupos de hablantes, uno con articulación estándar versus otro grupo con una articulación alterada. Para lograrlo, se apoya

de distintas mediciones, entre las que se encuentra la contabilización de los estímulos /la/, /ta.ka/ y /mu/. El autor concluye que la prueba establece una diferencia significativa entre los grupos evaluados.

Por otro lado, Fletcher (1972) propone el “time-by-count”, una metodología para evaluar el rendimiento diadococinético mediante la cuantificación del tiempo requerido para producir un número específico de sílabas. Esta propuesta nace desde la modificación del método tradicional, el cual requiere mayor cantidad de etapas procedimentales para la obtención de datos. Así, el examinador enfoca su atención solo en detener el tiempo del cronómetro, en lugar de contabilizar las producciones y establecer la duración de la emisión (considerar que la propuesta se realiza en una época anterior al auge tecnológico digital). Fletcher, además, hace una evaluación de su propuesta comparando los resultados que se obtienen al utilizar un cronómetro para determinar el tiempo versus los resultados temporales obtenidos de un análisis realizado mediante una impresión oscilográfica del mismo corpus. Su investigación se basa en una muestra de 24 niños y 24 niñas entre 6 y 13 años, a quienes se les solicitó 20 repeticiones de cada una de las sílabas /pa/, /ta/, /ka/, /fa/ y /la/, 15 repeticiones de los bisílabos /pa.ta/, /pa.ka/ y /ta.ka/; y 10 de la secuencia /pa.ta.ka/. Luego de hacer la comparación de los resultados de la tasa de diadococinesias obtenidos desde ambos análisis, concluye que no existen diferencias significativas en sus resultados.

La tercera metodología para obtener el cálculo de la tasa de diadococinesias considera la contabilización de la totalidad de los estímulos producidos en un ciclo respiratorio. Para lograr obtener el indicador, el especialista cuantifica el número de producciones y el tiempo total de la muestra. Por ejemplo, se puede estipular que un hablante produce 32 sílabas en un tiempo de 8.68 segundos (Gadesman y Miller, 2008).

2.1.4. Enfoque del análisis

La evaluación de la ejecución de la producción diadococinética se puede realizar a través de un análisis que se vale de la percepción que el profesional califica de la ejecución del hablante o bien, mediante la utilización de instrumentos objetivos que apoyan la evaluación.

2.1.4.1 Análisis perceptivo

La cuantificación de la tasa de diadococinesias mediante una evaluación subjetiva del habla es ampliamente utilizada por los profesionales que se dedican a acompañar a los usuarios que presentan afecciones motoras. Su difundida utilización se condice con la accesibilidad y bajos costos asociados a este tipo de métodos perceptivos. Sin embargo, el análisis manual no automatizado es caracterizado como: “*subjective, time intensive, and error-prone*” (Wang et al., 2019). Una de las repercusiones de esto es que incide en la sensibilidad del índice a través del tiempo, por ejemplo, pierde resolución para realizar

comparaciones en el seguimiento del comportamiento de un mismo paciente durante distintas etapas de su tratamiento. Es por ello que Gadesman y Miller (2008) proponen evaluar la fiabilidad de los indicadores por medio de la comparación entre mediciones perceptuales realizadas por profesionales y un grupo control sin conocimiento de la prueba en comparación a los resultados obtenidos a través de una metodología objetiva (análisis espectrográfico). Los participantes analizaron en dos ocasiones las muestras de habla utilizando las tres metodologías (conteo de sílabas por un lapso definido, medición del tiempo que cada usuario se demora en producir un número específico de sílabas y la cantidad de sílabas por ciclo respiratorio). Además, los investigadores evaluaron la confiabilidad intra-evaluador y una serie de características perceptuales de los hablantes (ritmo, impresión general de la producción, entre otros) determinadas por los mismos observadores. Gadesman y Miller (2008) concluyen que no hay diferencias estadísticamente significativas entre los resultados obtenidos por los participantes al utilizar los tres métodos comparados (la tradicional, el *"time by count"* y la cantidad de sílabas por ciclo respiratorio). Respecto a la comparación con métodos objetivos, reportan diferencias estadísticamente significativas entre los resultados que se obtienen utilizando una metodología objetiva en comparación a los resultados de los participantes al utilizar la contabilización de sílabas por ciclo respiratorio. Además, informan sobre bajos valores correlación de los datos al analizar la confiabilidad de cada evaluador y entre los evaluadores. En este contexto, entonces, los autores concluyen que las herramientas objetivas

proveen mayor confiabilidad en los resultados, tanto al comparar los resultados que obtienen dos evaluadores al analizar la misma muestra de habla, así como aquel que realiza una misma persona en dos períodos distintos. Así, los métodos objetivos permiten mejorar la proyección y comparación de los indicadores, aportando precisión frente a las decisiones terapéuticas y de seguimiento.

2.1.4.2 Análisis instrumental objetivo

El análisis cuantitativo instrumental de la producción diadococinética del habla ha avanzado significativamente con el aumento y auge de la tecnología digital. Su contribución se expresa en la difusión de herramientas que ayudan a los terapeutas, como es el caso del programa *MSP*, que como ya se expresó anteriormente (subapartado 2.1.1), provee un análisis de la ejecución diadococinética del habla mediante cinco indicadores.

El protocolo para usar este programa consiste en solicitar una producción silábica de forma rápida y repetida por alrededor de ocho segundos, la cual es grabada y analizada automáticamente por el programa, que genera además una serie de gráficas que ayudan a interpretar los resultados, los cuales son comparados con valores de referencia propios del programa que fueron medidos, obtenidos y normalizados en una población norteamericana hablante nativa de inglés.

Distintos autores han utilizado el MSP para realizar sus investigaciones y con ello obtener información de personas con diversas características normales y patológicas. Por ejemplo, un estudio realizado por Wang *et al.* (2009) utiliza el MSP para analizar una población diagnosticada con disartria atáxica. En este estudio se toma una muestra de 21 participantes que presentan esta alteración y se comparan sus resultados con un análisis manual realizado por los mismos autores. El estudio concluye que existen discrepancias entre los resultados aportados por ambas metodologías, pero que éstas no presentan una diferencia significativa. Es importante mencionar que, dadas las características propias de los sujetos investigados, un tercio de la muestra no fue posible de analizar debido a su alta irregularidad.

El programa MSP es utilizado también en un estudio realizado por Çiyiltepe y Çifçi (2019), quienes comparan el rendimiento diadococinético de menores entre 7 y 12 años con y sin tartamudez utilizando este software. Participaron 36 menores entre la edad objetivo, 19 de ellos sin alteración del habla. Todos los participantes emitieron los estímulos /pa/, /ta/, /ka/, /pa.ta/ y /pa.ta.ka/ lo más rápido posible. Luego, estas muestras fueron procesadas en el programa. El estudio concluye que existen diferencias de rendimiento entre ambos grupos que son reflejo de las alteraciones motoras del habla, ya que el grupo con tartamudez presenta valores significativamente más bajos.

Dado que no existe tanta información acerca de valores normativos de la producción diadococinética del habla para el MSP que puedan servir de referencia para realizar más estudios comparativos, hay algunas investigaciones que han centrado sus esfuerzos en generar este tipo de datos. En esta línea, Pérez *et al.* (2015) realizan en nuestro país una investigación para determinar los valores normativos en hablantes del español de Chile con el fin de ser utilizados para realizar estudios con el programa MSP. Así, entregan una serie de valores referenciales de los cinco índices del software, propios para la población del país. Para ello, el estudio se vale del método de muestra probabilística espejo, considerando 39 personas entre 18 y 61 años de ambos sexos, excluyendo a quienes presenten alteración neurológica o motora del habla. A cada participante se le solicitó la emisión de sílaba /pa/, la cual fue grabada y analizada por el programa, para finalmente obtener los índices característicos de esta población. Estudios similares han sido realizados por John *et al.* (2014), quienes procedieron a normalizar datos obtenidos del software para una población hablante de canarés.

Debido a la amplia difusión del uso de los índices diadococinéticos en distintas poblaciones, es que uno de los propósitos actuales de los investigadores es establecer los datos normativos diadococinéticos para cada idioma. Así, Alshahwan *et al.* (2020), realizan un trabajo para medir la tasa promedio del árabe en dos variantes, saudí y bahreiní. Para ello, realizan una investigación que incluyó a 80 participantes adultos (40 hombres, 40 mujeres), representantes de

ambos dialectos. Se solicitó a los participantes, la producción de sílabas /ba/, /da/, /ga/ y de la secuencia /ba.da.ga/. El análisis acústico se realizó con la metodología *time-by-count*, contabilizando series de 20 sílabas o 20 secuencias de sílabas. Las mediciones se realizaron con el programa Praat, mediante la observación, demarcación y contabilización visual de las producciones. Posteriormente, se extrajo la información de tiempo mediante un *script*. El análisis estadístico de los datos concluyó que los hablantes del dialecto saudí presentan una tasa DDK más alta que los del bahrainis en monosílabos, pero que no hay diferencias entre trisílabos, ni entre sexos. Finalmente, entre los monosílabos, /ga/ es el que manifiesta el promedio DDK más bajo.

2.1.4.2.1 Análisis instrumental objetivo: algoritmos

Los índices diadococinéticos forman parte de la evaluación que realizan los profesionales al relacionarse con personas con sospecha de alteración motora. Es por ello que Diepeveen *et al.* (2019) proponen un protocolo que presenta subtests para generar una evaluación multidimensional de niños con alteraciones motoras orales. Uno de los subtest propuestos es aquel que determina el rendimiento promedio diadococinético. El protocolo postula la evaluación de sílabas /pa/, /ta/, /ka/, de bisílabos /pa.ta/ y /ta.ka/, y de la secuencia /pa.ta.ka/. En este trabajo, se realizó un análisis acústico utilizando un *script* de Praat desarrollado por uno de los autores, el cual marca el inicio y fin de cada sílaba o secuencia. El *script*, además, extrae información relativa al número

de emisiones y duración total de cada estímulo. Esta investigación concluye que los niños menores de 30 meses tienen muchas dificultades para producir las diadococinesias, incluso hasta los 3 años, por lo que plantean que este tipo de pruebas se debería aplicar sólo a partir de esta edad. Además, exponen que existe mayor agilidad al primer intento que a los siguientes, siendo la instrucción “lo más rápido posible” aquella que produce una emisión más rápida, pero con una diferencia mínima.

La automatización del trabajo para poder calcular los indicadores progresa con el avance del tiempo, generando instrumentos cada vez más especializados. En este contexto, una población típicamente objetivo de este tipo de estudios es aquella que comprende a personas que han sido diagnosticadas con enfermedad de Parkinson (EP), quienes cursan con sintomatología que incluye alteraciones del habla. Por ello, Montaña *et al.* (2018) proponen un estudio en el cual extraen datos asociados al *voice onset time* en muestras de diadococinesias para generar una mayor cantidad de información cuantitativa sobre las características de usuarios con EP. Participaron en este estudio, 54 voluntarios, 27 de ellos diagnosticados con esta enfermedad y los restantes sin alteración de origen neurológico. Se les solicitó a los informantes la emisión constante y repetida de cada una de las sílabas /pa/, /ta/ y /ka/, las cuales fueron grabadas y procesadas por un algoritmo creado por los autores para extraer la información relativa al VOT por medio de 12 índices. Además, se midió la tasa de diadococinesias, entendida en este trabajo como el número de sílabas producidas en un segundo.

Los autores concluyen que los indicadores más sensibles para pesquisar sintomatología en etapas iniciales son la tasa diadococinética, el porcentaje de variabilidad vocálica, la asimetría espectral (grado de inconsistencia entre los espectros de las producciones) y los valores promedio de frecuencia, los que discriminan características propias de esta condición versus la población sin patología.

Dentro de la creación de herramientas tecnológicas para facilitar la medición del comportamiento motor del habla, se encuentran trabajos como los de Rozenstoks *et al.* (2020), quienes analizan el rendimiento diadococinético de adultos diagnosticados con esclerosis múltiple (EM) y lo comparan con un grupo control. Entre las características que poseen las personas con esclerosis múltiple frecuentemente se encuentran alteraciones a nivel del control motor oral. Consecuentemente, la disartria atáxica, la espástica o bien la mixta presentan una alta prevalencia entre estas personas. Para realizar el análisis, los autores crean y utilizan un algoritmo de segmentación automática basado en un enfoque de Redes Neuronales Convolucionales. Luego de procesar las muestras de habla, se calculó la tasa de diadococinesia contabilizando las repeticiones en un segundo y la regularidad diadococinética como la desviación estándar de las pausas entre sílabas consecutivas. Así, los autores concluyen que el algoritmo presentado identifica los límites de las sílabas con alta precisión, con lo cual encuentran diferencias entre la población con EM y el grupo control, ya que los

primeros presentan bajas tasas DDK. Además, establecen que, para esta población, el estímulo sonoro es más sensible que su contraparte áfona.

Karlsson *et al.* (2020) proponen un estudio que vincula el rendimiento de producciones diadococinéticas como un predictor del nivel de deterioro que personas con disartria hipocinética presentan en la lectura en voz alta. Para lograrlo convocaron a 68 personas diagnosticadas con Enfermedad de Parkinson, quienes fueron divididas en dos grupos. Al primero se les solicitó sus muestras de habla mientras continuaban con su tratamiento farmacológico y al segundo, se les grabó en dos ocasiones; la primera mientras continuaban con sus medicamentos y la segunda, sin ellos. A todos los participantes se les solicitó la lectura en voz alta del mismo texto y la producción repetitiva de las sílabas /pa/, /ta/ y /ka/. Luego, se categorizaron perceptivamente las características de la lectura de cada voluntario (articulación, prosodia, respiración y voz) y estos resultados fueron comparados con un análisis acústico (46 características, la mayoría relativas al dominio de la duración) extraídas desde la producción diadococinética. Este análisis consistió en la segmentación manual de cada producción silábica y luego, la extracción automática de los datos evaluados. Entre los resultados de la investigación, destaca que las características acústicas provenientes de la secuencia /ka.ka.ka/ son las que presentan mayor similitud con el grado de deterioro articulatorio percibido por los evaluadores y, por lo tanto, son predictivas del comportamiento del componente articulatorio de los evaluados. Además, los autores aseveran, como otro hallazgo que, si bien el

indicador tradicionalmente utilizado para cuantificar la ejecución diadococinética, (la tasa de diadococinesias) es altamente indicativo de ausencia o presencia de alteración del habla, por sí solo no entrega información suficiente sobre otras dimensiones del habla, como, por ejemplo, la precisión articulatoria, la respiración o la prosodia. Así, es preciso complementar los resultados de este indicador con los que proporcionen otros índices para obtener más información sobre las características del evaluado.



3. PLANTEAMIENTO DEL PROBLEMA

Como ya se ha establecido, existe una gran variedad de procedimientos y protocolos para evaluar a las personas que presentan una posible alteración motora del habla. Entre ellos se encuentran aquellos que se valen de la cuantificación del rendimiento diadococinético, así como de los indicadores derivados de él. A través del tiempo, este cálculo se ha realizado en la práctica clínica basándose en la percepción que el evaluador tiene del rendimiento de la persona que ejecuta la tarea. Sin embargo, con el avance de la tecnología, se han creado herramientas que no reemplazan el trabajo de los profesionales, pero que sí aportan mayor precisión en la medición, que permiten también homologar los resultados y que, además, se constituyen en un claro apoyo frente a decisiones diagnósticas, terapéuticas, de seguimiento y de investigación.

De esta forma, se han destacado investigaciones que utilizan elementos propios de la tecnología y la computación para proponer, entre otros, algoritmos que segmentan automáticamente las señales para facilitar la extracción de los datos, algoritmos que automáticamente calculan índices asociados a una población específica (VOT), o bien, programas que realizan segmentación y

cálculo para un amplio espectro de usuarios, pero con un alto costo monetario asociado, lo que obviamente repercute en la accesibilidad para todos los especialistas.

Esto revela la necesidad de diseñar e implementar instrumentos que permitan realizar un análisis objetivo y cuantitativo del comportamiento motor del habla, pero que además sean de libre acceso, sin costo para los usuarios, y que incorporen también una metodología única.

Entre el espectro de soportes informáticos que facilitan en general el análisis y procesamiento cuantitativo e instrumental del correlato acústico de la voz y el habla en todas sus dimensiones se encuentra el software *Praat*, desarrollado por Boersma y Weenink (2020). Este recurso informático se trata de un programa de libre acceso (sin costo monetario asociado) que permite realizar manipulación de señales acústicas, edición de sus componentes, extracción de datos, síntesis de las señales, entre otras varias funciones. Este software es gratuito y tiene un uso bastante difundido en la investigación de las últimas décadas en áreas relacionadas al ofrecer una amplia gama de utilidades, las cuales, además, se alimentan con las sugerencias y contribuciones que realizan los propios usuarios.

Dentro de las múltiples funcionalidades de Praat, una de ellas, permite, mediante un ambiente de programación, crear complementos del software para poder realizar tipos específicos de análisis que el programa no considera en su

diseño original. Mediante esta funcionalidad, el usuario puede crear lo que en terminología computacional se denomina un *script*, que consiste en una serie de rutinas e instrucciones codificadas en un lenguaje de programación *ad hoc* que permiten automatizar procesos de análisis. Praat actúa entonces como un entorno de programación para la implementación y la ejecución de estas rutinas.

Así, es plausible plantearse la posibilidad de diseñar e implementar un *script* de Praat que segmente y etiquete automáticamente muestras de habla y que, además, haga las mediciones y realice los cálculos necesarios para generar un perfil clínico cuantitativo del rendimiento diadococinético.

Así también, una vez creado el *script*, resulta evidente la necesidad de evaluar su funcionalidad, eficiencia y eficacia mediante su aplicación a una muestra real y la consecuente ejecución de algún tipo de procedimiento de validación de sus resultados.

3.1 Preguntas de investigación

En relación con todos los antecedentes proporcionados y con el problema ya expuesto, esta investigación pretende responder a dos preguntas:

- a) ¿Es posible implementar, mediante los recursos de programación que provee el programa Praat, una aplicación de libre acceso que evalúe de forma automatizada o semi automatizada el rendimiento diadococinético?
- b) ¿El *script* de Praat que se propone es igual de preciso en la medición y el cálculo de los indicadores que el programa comercial y/o un análisis y medición visual y manual no automatizada?

4. OBJETIVOS

4.1 Objetivo general

Diseñar, implementar y validar un recurso informático que posibilite la obtención semi automatizada de cuatro índices usados para la medición del rendimiento diadococinético: a) promedio de sílabas por segundo, b) el coeficiente de variabilidad del promedio de sílabas por segundo, c) la duración promedio de la sílaba y d) el coeficiente de variabilidad de la duración promedio de la sílaba de manera automática, sin costo asociado y de libre disposición.

4.2 Objetivos específicos

- a) Diseñar un *script* de Praat que segmente de manera automatizada y semi automatizada las sílabas presentes en una producción diadococinética.
- b) Diseñar un *script* de Praat que calcule los indicadores diadococinéticos a) promedio de sílabas por segundo, b) el coeficiente de variabilidad del promedio de sílabas por segundo, c) la duración promedio de la sílaba y d) el coeficiente de variabilidad de la duración promedio de la sílaba.

- c) Calcular mediante el *script* el rendimiento diadococinético de personas sin patología neurológica de base.
- d) Comparar los resultados de los indicadores obtenidos por el *script* con otras metodologías.



5. METODOLOGÍA

Debido a las características de este estudio y al alcance de sus objetivos, la metodología involucra dos etapas de trabajo: la etapa 1 consiste en el proceso de diseño e implementación del *script* y la etapa 2 comprende todo el proceso de validación de este.

Asimismo, la validación del *script* (etapa 2) involucra tres procedimientos paralelos, pero que de diferentes formas apuntan a un mismo propósito: a) validación del *script* por medio de la comparación de sus resultados con los de otros métodos, b) la validación del análisis no automatizado y c) el análisis de la fiabilidad del instrumento.

5.1 Diseño e implementación del *script* (etapa 1)

La etapa de desarrollo del *script* consiste entonces en el diseño de una serie de rutinas que extrae de manera automatizada y semi automatizada de un archivo de audio los datos del dominio de la duración y que luego calcula de

manera automatizada los indicadores necesarios para caracterizar el rendimiento diadococinético.

En el diagrama de flujo presentado en la Figura 1, se ilustran el diseño original del *script* antes de ser implementado. En el diagrama, se muestra de manera ordenada el detalle de todas las tareas que debe realizar el *script* para poder cumplir con su función, además del orden en que estas tareas deben ser ejecutadas.

A continuación, se explicará con más detalle dos de las tareas que, de entre todas, son las esenciales: la segmentación automatizada de la señal y el cálculo de los indicadores.



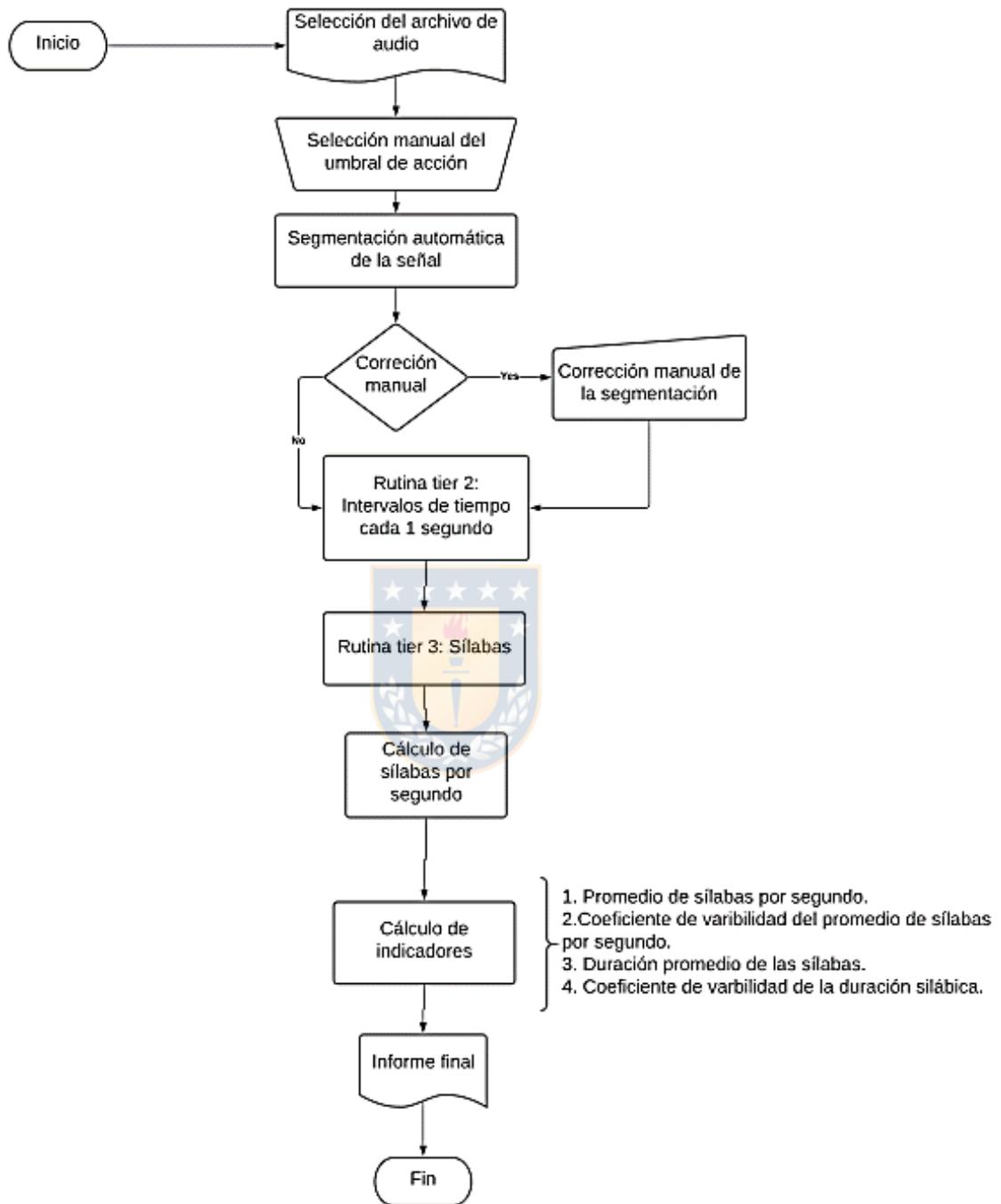


Figura 5.1 Diagrama de flujo del *script* propuesto.

5.1.1 Segmentación automática de la señal

Para definir los límites de cada sílaba, el *script* hace uso de la función *To TextGrid (silence)*, la cual detecta, segmenta y separa intervalos catalogados como “silencio” de aquellos catalogados como “sonoros”, tomando como referencia la medición y el análisis de los datos relativos a la intensidad (Boersma y Weenink, 2020) y a la duración mínima de un segmento que el programa debe estimar como válida. Estos datos varían de acuerdo con las características de cada uno de los componentes de la emisión, con las diferencias propias de la calidad de las señales según su origen, así como con la alta variabilidad de las emisiones características de las personas con trastornos motores del habla (Karlsson, 2020). Por esta razón, en el *script*, antes de comenzar a ejecutar el proceso de segmentación automatizada, se abre una ventana que le propone al usuario (analista) tres opciones. Estas tres opciones configuran el análisis con distinta sensibilidad (de acuerdo con la variación del parámetro *Silence threshold*) para la segmentación automática. Las opciones son: a) “alto”, que determina un umbral de 20 dB; b) “medio” que configura un umbral de 15 dB, y c) “bajo” que determina un umbral de 10 dB¹. La elección de una de estas tres opciones genera

¹ Para esta investigación, se consideraron sólo los resultados que se obtuvieron mediante el umbral “medio”, ya que se observó que en general era el que tenía mejor rendimiento de acuerdo con la calidad de la muestra obtenida para este trabajo.

como resultado una segmentación distinta de la muestra de habla, aunque también esto puede variar dependiendo de la calidad de la señal evaluada.

Es relevante mencionar que, posteriormente a la segmentación automatizada, el resultado de este paso es susceptible de ser evaluado por la persona que hace uso del *script*, y, si así lo considera necesario, el usuario puede modificar la segmentación y realizar eventuales correcciones de forma manual.

5.1.2 Cálculo de los indicadores



Para cuantificar el rendimiento diadococinético se utilizan en este proyecto 4 indicadores:

- a) El promedio de sílabas por segundo (PSS): que entrega información sobre la cantidad de sílabas promedio que la persona produce durante una unidad de tiempo establecida, en este caso, un segundo. Este indicador es similar al DDKavr del MSP, ya que de acuerdo con lo que indica el manual del MSP se puede concluir que ambos miden lo mismo. Este indicador es el que se conoce ampliamente como la “tasa de diadococinesia”, que es además el que da origen y nombre al protocolo de evaluación y que ha sido utilizado tradicionalmente durante muchos años en la práctica clínica.

- b) El coeficiente de variabilidad² del promedio de sílabas por segundo: que provee, en porcentaje, el grado de variabilidad de la producción de la cantidad de sílabas por segundo entre cada una de las unidades de tiempo. De acuerdo con lo que indica el manual del MSP, este indicador debería ser similar al DDKjit, y de esa forma se entendió al inicio de esta investigación. Sin embargo, como se verá más adelante, los resultados de este indicador comparados con los de DDKjit, son tan diferentes que es posible pensar que la forma de cálculo que realiza el MSP es completamente diferente. Lamentablemente, esto no resulta posible saberlo, debido a que el MSP no declara de forma explícita sus fórmulas de cálculo.
- c) El promedio de duración de la sílaba (DPS): que consiste en el cálculo del valor promedio, en milisegundos, de la duración de las sílabas presentes en la producción diadococinética. Este indicador es similar al DDKavp del MSP, ya que de acuerdo con lo que indica el manual del MSP se puede concluir que ambos miden lo mismo.
- d) El coeficiente de variabilidad del promedio de duración de la sílaba: que compara la duración de todas las sílabas de la muestra y pondera su variabilidad. Es decir, mide cuán constante es la duración de las sílabas

² Cabe recordar que el coeficiente de variabilidad de cualquier grupo de datos se calcula dividiendo la desviación estándar por el promedio, lo que arroja un valor entre 0 y 1, donde cero se interpreta como ausencia total de variabilidad y 1 como variabilidad absoluta. Además, para hacer más interpretable este dato, se suele multiplicar por 100 para transformarlo a su equivalente en porcentaje.

considerando todas las emisiones. Este indicador es similar al DDKcvp del MSP, ya que de acuerdo con lo que indica el manual del programa se puede concluir que ambos miden lo mismo.

Para lograr calcular los indicadores, primero es necesario que el *script* genere automáticamente los *tiers* o estratos de segmentación requeridos para contener los límites y los valores de la unidad de tiempo (1 segundo) y los límites y los valores de la unidad de habla (sílabas). Luego, el *script* mide y extrae los datos relativos a la duración de cada intervalo consonante-vocal, excepto de la primera producción³, para así también calcular su promedio, su desviación estándar y su coeficiente de variabilidad. De la misma forma, logra determinar la cantidad de sílabas presentes en cada unidad de tiempo y calcula su promedio, su desviación estándar y su coeficiente de variabilidad.

Una vez realizados los cálculos, el *script* despliega una ventana (Figura 5.2) donde es posible visualizar los resultados del análisis en formato de informe clínico. Así se entrega la información obtenida de la muestra para cada indicador y, a su lado, los valores normativos de referencia (Pérez et al., 2015) de 3 de los 4 indicadores.

³ Para extraer los datos, Ackermann *et al.* (1995) sugieren que la primera producción silábica sea descartada dado que usualmente las personas producen la primera combinación consonante-vocal con una duración mayor a las sucesivas. Así se evita una alteración de los resultados que puedan devenir de este efecto.

INFORME DE RENDIMIENTO DIADOCOCINÉTICO DEL HABLA (DDK)			
Promedio de sílabas por segundo (PSS)	:	6.46 s/s	(ref. 4 ± 1 s/s)*
Coef. de variabilidad del PSS	:	2.9%	
Duración promedio de la sílaba (DPS)	:	156 ms	(ref. 263 ± 66 ms.)*
Coef. de variabilidad de la DPS	:	6.7%	(ref. < 5,5%)*
Segundos analizados	:	8	(ideal 8 seg.)
Número de sílabas analizadas	:	56	(ideal 30 sil.)
(*) Valores de referencia para población adulta			

Figura 5.2 Informe final.

5.2 Validación del *script* propuesto (etapa 2)

La segunda etapa de la metodología involucra tres procedimientos diferentes, pero que apuntan hacia un mismo fin. Los tres procedimientos se realizan con el objetivo de lograr validar el *script* propuesto en esta investigación. A diferencia de la primera, que tiene un carácter distinto; en esta segunda etapa se debe declarar que se propone un diseño de estudio con un enfoque metodológico de tipo cuantitativo no experimental transversal descriptivo para los tres procedimientos.

5.2.1 Validación y estimación de la confiabilidad del *script*

Para validar su fiabilidad, se realizó una comparación entre los resultados del *script* y otras dos metodologías: a) una segmentación y cálculo no

automatizado de los mismos indicadores realizado por un analista humano y b) los resultados que se obtienen del análisis similar al utilizar un programa de carácter comercial, que en este caso es el MSP. Los tres métodos se aplicaron a una misma muestra de 30 informantes (para más detalle sobre la muestra y el corpus ver los subapartados 5.2.1.3 y 5.2.1.4).

5.2.1.1 Análisis no automatizado de las muestras de habla

El análisis no automatizado de todo el corpus lo realizó la autora de esta investigación. El procedimiento de análisis contempló los siguientes pasos y criterios. Primero, se crea un objeto *TextGrid* desde el objeto *Sound* con 2 *tiers* o estratos de segmentación. Uno para consignar la información silábica y otro en el que se definen los intervalos y valores para la unidad de tiempo, que en este caso debe corresponder a 1 segundo por intervalo. Luego, se realiza la segmentación silábica que considera la información proporcionada por el oscilograma y espectrograma y que se analiza de forma visual. El límite de cada intervalo se ubica luego de la vocal, justo antes de que comience la producción de la oclusiva, segmentando toda la señal. Desde ambos *tiers* se obtienen los datos necesarios para realizar los cálculos matemáticos que derivan en los cuatro indicadores. En una plantilla *Excel*, se digita la información de cada archivo, tanto

su nombre como sus resultados, y en esta misma planilla se realizan los cálculos mediante las funciones que para ello provee el mismo programa.

5.2.1.2 Análisis mediante el programa *Motor Speech Profile*

Para obtener los indicadores mediante esta aplicación, se siguen las instrucciones que provee el manual del mismo programa y se registran los valores. El MSP realiza el análisis de cada muestra de habla mediante la apertura del archivo de audio y la posterior ejecución de la función automatizada que da los resultados de cada indicador. El MSP cuenta con la opción de cambiar el umbral de sensibilidad en caso de que exista alguna dificultad con la segmentación automática por defecto. Así, se obtienen los resultados de los indicadores, los cuales son presentados en un informe para su posterior análisis o para el uso que el analista estime conveniente. Cabe señalar que, para realizar la comparación necesaria para el análisis de validación, solo se consideraron los indicadores DDKavr, DDKcvp, DDKavp del MSP, los cuales son análogos a los que para esta investigación se consideran como: la duración promedio de la sílaba, el coeficiente de variabilidad de la duración silábica y el promedio de sílabas por segundo, respectivamente.

5.2.1.3 Participantes

El corpus de habla se obtuvo mediante la grabación de las emisiones de 30 participantes voluntarios seleccionados por conveniencia. Para colectarlas, se creó una página web, en donde los informantes, luego de aceptar participar en el proyecto mediante un consentimiento informado, procedían a ejecutar la tarea, mientras la aplicación guardaba un registro de audio digitalizado de dicha emisión.

Los informantes corresponden a adultos entre 21-60 años, cuya edad promedio es 32,9 años, residentes de las ciudades Arauco (1), Santiago (3), Ancud (3), Curicó (10) y Concepción (13). De ellos, 20 pertenecen al sexo femenino y 10 al masculino. Los criterios de inclusión para participar de la investigación consistían en ser adultos entre 18 a 60 años que (a) no presentaran algún trastorno neurológico, cognitivo, psiquiátrico o de aprendizaje, (b) sin presencia de sintomatología actual o previa de trastorno de la voz, del habla o de la fluidez, (c) no haber recibido terapia de la voz o el habla y (d) ausencia de alteración auditiva o infección de oído.

5.2.1.4 Corpus

El corpus está compuesto de 2 grabaciones consecutivas emitidas por cada voluntario. En total, se obtuvieron 60 muestras de habla. Para realizar el análisis comparativo entre metodologías, se utilizaron solo 30 de ellas, una de cada voluntario. Cada una de ellas contiene la emisión repetitiva, regular y sin errores de la sílaba /pa/ por al menos 8 segundos. Además, los participantes utilizaron sus propios dispositivos electrónicos para grabar sus muestras, ya sea celular o computador personal. La segunda emisión se utilizó para el análisis de fiabilidad del *script* como instrumento (Capítulo 5.2.3).



5.2.1.5 Instrumento de recolección del corpus

Con el fin de recabar las muestras de habla necesarias para lograr esta investigación, se creó, en conjunto con un diseñador, una página web (disponible en www.grabacionhabla.cl) que recolectó las producciones de los voluntarios (Figura 4.3). La página recoge datos desde un formulario basado en lenguaje *php*. En esta plataforma, además, el usuario, antes de grabar, debe llenar un formulario de consentimiento informado, y leer las instrucciones para realizar la

tarea, así como realizar la grabación de las muestras de habla. Finalmente, la página web almacena automáticamente el audio.

Plataforma de grabación en línea

El objetivo de este consentimiento es informar sobre las características y propósito de la investigación titulada "Diseño, implementación y validación de un script de Praat para medir el rendimiento diadococinético" para optar al grado de Magíster en Lingüística Aplicada de la fonodisloga Daniela Osses Valenzuela, a cargo del profesor Dr. Hernán Pérez M.

Objetivo
Crear una herramienta informática que automatice la evaluación del rendimiento diadococinético como parte de los procedimientos para caracterizar los trastornos motores del habla. Para lograr este fin, es necesario analizar grabaciones de habla de distintas personas.

Requisitos para poder participar
(a) No presentar algún trastorno neurológico, cognitivo, psiquiátrico o de aprendizaje.
(b) Sin presencia de sintomatología actual o previa de trastorno de la voz, del habla o de la fluidez.
(c) No haber recibido terapia de la voz o el habla.
(d) Ausencia de alteración auditiva o infección de oído.

Procedimiento
Se solicitará que ingrese información personal (edad, sexo, ciudad de residencia y datos sobre su salud) y luego, procederá a grabar según las indicaciones. Se garantiza que su participación es confidencial y sólo se utilizarán los datos con fines investigativos. Las grabaciones se analizarán de forma anónima. Su participación es voluntaria, pero se requiere de autorización. Usted tiene la libertad para decidir si quiere participar del estudio y puede retirar su autorización en cualquier momento, sin necesidad de justificarse.

Figura 5.3 Página web.

Los archivos de audio del corpus poseen una duración mínima de 8 segundos y máxima de 12 segundos, con un promedio de 8,6 segundos. Además, presentan una cantidad mínima de sílabas de 32 y máxima de 63, con un promedio de 46 sílabas. Cada uno de estos archivos se almacena

automáticamente en formato WAV digitalizado con una frecuencia de muestreo de 44 kHz.⁴

5.2.2 Validación del análisis no automatizado

Con el propósito de evaluar la fiabilidad de la segmentación y el cálculo no automatizado realizado por la investigadora, un total de 9 muestras escogidas al azar (30% del total del corpus que se declara en capítulo 5.1.2.4) fueron seleccionadas para que un evaluador externo, quien cuenta con experiencia en el análisis acústico del habla, estudiante de posgrado del programa de Magíster en Lingüística Aplicada de la Universidad de Concepción y de la asignatura de Fonética Acústica, hiciera el mismo tipo de análisis. Esto, con el objetivo de ver si el evaluador externo llegaba a los mismos resultados que había llegado la autora. Así, se compararon sus resultados con el fin de determinar la fiabilidad inter-evaluador.

⁴ Si bien, por rigor metodológico aquí se mencionan algunas características relativas a la calidad y la resolución del archivo de audio como corresponde; en este caso, por basarse el procedimiento de medición en la obtención de valores de la dimensión temporal se asume que la calidad y el formato del archivo de audio no es relevante ni afecta el valor de los datos.

5.2.3 Análisis de la fiabilidad del instrumento

Para determinar la confiabilidad del instrumento, se hizo un análisis de correlación comparando el rendimiento de cada voluntario en su primera producción con una segunda ejecución de la producción diadococinética. Ambas muestras de habla se obtuvieron bajo las mismas condiciones (como se explica en el capítulo 5.1.2.4) y fueron analizadas por el *script* utilizando el umbral “medio”, realizando las correcciones cuando la segmentación no cumplía con los criterios expuestos.



6. RESULTADOS

La versión 1.0 del *script* se encuentra operativa y disponible para su revisión y uso en: <<https://github.com/heperez-udec/Praat-Script-analisis-diadocokinethic-rate>>.

6.1 Resultados de la validación del *script*

Los datos obtenidos desde las tres fuentes o metodologías de análisis se digitalaron, organizaron y posteriormente fueron analizados, comparados e interpretados mediante procedimientos de análisis estadístico de tipo descriptivo e inferencial. Primero, se aplicaron los procedimientos clásicos para calcular las medidas de tendencia central y de dispersión. Luego procedimientos para determinar las condiciones de normalidad y homogeneidad de la varianza de los datos para, posteriormente, aplicar la prueba de ANOVA de medidas repetidas con el fin de determinar si existen diferencias en los resultados según la metodología utilizada. Para realizar todos los cálculos estadísticos se importaron los datos a *R: A language and environment for statistical Computing* (Versión 4.0.3) y al IBM SPSS *Statistics* (Version 26).

La presentación de los resultados está dividida en dos partes. La primera contiene la información de los resultados que se obtuvieron de aplicar el *script* a

las muestras de habla sin realizar ninguna intervención o corrección humana de la segmentación que este realiza de manera automatizada. Posteriormente, en la segunda parte, se exponen los resultados obtenidos al aplicar el *script*, pero utilizando además la funcionalidad considerada en el diseño del *script* que permite al analista hacer una revisión visual y una corrección manual de la segmentación automática, previas al cálculo de los indicadores, si es que el usuario lo estima pertinente.

6.1.1. Resultados del análisis de la validación del comportamiento del *script* sin aplicar corrección humana de la segmentación silábica automatizada

6.1.1.1 Resultados de la validación de los datos del promedio de sílabas por segundo (PSS), sin aplicar corrección

Como se observa en la Tabla 6.1, en general, los resultados de la estadística descriptiva para este indicador, comparando las tres metodologías, son bastante similares. El análisis entrega un valor promedio de 5,7 sílabas por segundo (en adelante s/s) para dos de ellas y 5,66 s/s para el *script*. Asimismo, los tres métodos arrojan valores de desviación estándar bastante similares. Este mismo patrón se observa en los valores máximos y mínimos de las tres metodologías.

Tabla 6.1 Resultados del análisis estadístico descriptivo de los resultados del *script* sin aplicar correcciones no automatizadas del PSS

	Manual	Script	MSP
N	30	30	30
Promedio (%)	5.70	5.66	5.70
Mediana (%)	5.79	5.7	5.78
Desviación Estándar (%)	0.747	0.771	0.728
Mínimo (%)	4.08	4.01	4.07
Máximo (%)	7.18	7.19	7.17
Asimetría	-0.304	-0.314	-0.305
Curtosis	-0.322	-0.092	-0.238
Shapiro-Wilk W	0.980	0.979	0.978
Shapiro-Wilk p	0.821	0.810	0.777

Con el fin de determinar si los datos presentan una distribución normal, se aplicó la prueba de normalidad de Shapiro-Wilk (Tabla 6.1), la que arroja un resultado de $W(30): 0.98, p = 0.82$ para la metodología manual, para el *script* de $W(30): 0.98, p = 0.81$ y, finalmente para el MSP $W(30): 0.98, p = 0.78$. De acuerdo con esto, se concluye que los datos presentan una distribución normal.

Para determinar el estado de la homogeneidad de la varianza, se aplicó la prueba F-test cuyos resultados entre la medición manual y el *script* sin corrección es de $F(29, 29) = 0.939, p = 0.867$, manual y MSP $F(29, 29) = 1.054, p = 0.889$ y entre los valores del *script* comparados con los del MSP $F(29, 29) = 1.105, p = 0.789$. Estos resultados no revelan diferencias estadísticamente significativas entre las varianzas de los grupos estudiados.

Posteriormente, para determinar, mediante una prueba estadística, si existen diferencias en los resultados que dependan de la metodología utilizada,

se aplicó la prueba ANOVA de medidas repetidas. No obstante, previamente, como corresponde a este tipo de análisis, se revisó el valor de esfericidad que entrega la prueba de Mauchly, la cual dio un valor de $\chi^2(2) = 31.91$, $p < 0.001$, bajo el nivel de significancia, es decir, no cumple con el supuesto de esfericidad. En consecuencia, el resultado de la prueba ANOVA se debe interpretar con las correcciones de Greenhouse y Geisser, prueba que entregó un valor con un tamaño de efecto medio (como se observa en la tabla 6.2) de $F(1.19, 34.52): 2.27$, $p = 0.137$. Por lo tanto, la prueba no reporta diferencias estadísticamente significativas entre las metodologías utilizadas.

Tabla 6.2 Resultados ANOVA con correcciones de Greenhouse-Geisser

Origen		gl	F	Sig.	Eta parcial al cuadrado
Metodología	Greenhouse-Geisser	1,19	2,274	,137	0,074
Error	Greenhouse-Geisser	34,52			

6.1.1.2 Resultados de la validación de los datos del coeficiente de variabilidad del PSS, sin aplicar corrección

En la Tabla 6.3, se presentan los resultados del análisis de estadística descriptiva de los datos para este indicador. En general, lo que se puede comentar es que el valor promedio es de alrededor de un 3,3% de variabilidad

para el método manual y un valor de 4,5% para el *script*. También, se observa que el valor mínimo obtenido de ambas metodologías es similar (1,03% y 1,1%), en cambio, los valores de desviación estándar y máximo difieren de manera apreciable entre ambas metodologías.

Tabla 6.3 Resultados del análisis estadístico descriptivo del *script* sin correcciones humanas del coeficiente de variabilidad del PSS

	Manual	Script
N	30	30
Promedio (%)	3.33	4.5
Mediana (%)	3.13	3.7
Desviación Estándar (%)	1.47	3.9
Mínimo (%)	1.03	1.1
Máximo (%)	7.05	22.2
Asimetría	0.992	3.5
Curtosis	0.749	15.5
Shapiro-Wilk W	0.921	0.634
Shapiro-Wilk p	0.029	< 0.01

Para determinar formalmente si existe una distribución normal de los datos, se analiza el resultado de la prueba de normalidad de Shapiro-Wilk (Tabla 6.3) que arroja un resultado para el método no automatizado de $W(30)$: 0.92, $p < 0.05$ y para el *script* $W(30)$: 0.63, $p < 0.01$. Por lo descrito, se concluye que los datos no están normalmente distribuidos.

Debido a que los datos no cumplen con la condición de distribución normal, se optó por aplicar la prueba no paramétrica de los rangos con signos de Wilcoxon para determinar si existen diferencias significativas entre los

resultados obtenidos mediante ambas metodologías. La prueba reveló que no existen diferencias estadísticamente significativas del resultado del coeficiente de variabilidad del promedio de sílabas por segundo que se obtiene utilizando, ya sea el *script* (*Mdn*: 3,7, *n*: 30) o un análisis no automatizado (*Mdn*: 3,13, *n*: 30), z : -1.738, $p = 0.082$, con un tamaño de efecto (r) pequeño de 0,22.

6.1.1.3 Resultados de la validación de los datos de la duración promedio de la sílaba (DPS), sin aplicar corrección



Los datos de la Tabla 6.4, muestran los resultados de la estadística descriptiva para este indicador, comparando las tres metodologías. Se destaca un valor promedio de 179 milisegundos para dos de los tres métodos, mientras que el valor que se obtiene del *script* es de 180 milisegundos. De la misma manera, los tres métodos presentan valores de desviación estándar y valores mínimos y máximos bastante similares en los tres grupos.

Tabla 6.4 Resultados del análisis descriptivo de los datos del *script* sin correcciones humanas de la DPS

	Manual	Script	MSP
N	30	30	30
Promedio (%)	179	180	179
Mediana (%)	174	174	173
Desviación Estándar (%)	25.0	26,3	24.9
Mínimo (%)	139	139	139
Máximo (%)	244	247	246
Asimetría	0.795	1.054	0.883
Curtosis	0.281	0.93	0.468
Shapiro-Wilk W	0.950	0.916	0.936
Shapiro-Wilk p	0.173	0.021	0.070

Con el fin de determinar si los datos presentan una distribución normal, se analizó la prueba de normalidad de Shapiro-Wilk, la cual arroja valores sobre los de significancia en dos de las tres metodologías, específicamente, los datos provenientes de un análisis no automatizado ($W(30): 0.95, p = 0.17$) y aquellos obtenidos desde el MSP ($W(30): 0.94, p = 0.07$). En cambio, a partir de los resultados del *script* ($W(30): 0.92, p < 0.05$) se concluye que los datos no presentan una distribución normal.

Posteriormente, debido a que los datos no tienen una distribución normal, se aplicó la prueba de Friedman para determinar mediante una prueba de estadística inferencial si se observan diferencias estadísticamente significativas entre los resultados que puedan ser atribuibles al tipo de metodología utilizada. El resultado de esta prueba fue de $\chi^2(2, n: 90) = 3.53, p = 0.17$, lo que revela que

no existen diferencias estadísticamente significativas entre los resultados de las mediciones de las metodologías utilizadas.

6.1.1.4 Resultados de la validación de los datos del coeficiente de variabilidad de la DPS, sin aplicar corrección

La Tabla 6.5 muestra los datos del indicador según la metodología utilizada para el análisis. Se desprende de los resultados de la estadística descriptiva que los valores del *script* ($\bar{x} = 10.16$, DE: 7.9) presentan diferencias en comparación con las otras dos metodologías, con una mayor dispersión de datos. Se observa que este conjunto de datos presenta un valor máximo mucho mayor a los que se obtienen desde el análisis manual y del MSP.

Tabla 6.5 Resultados del análisis estadístico descriptivo del *script* sin correcciones humanas del coeficiente de variabilidad de la DPS

	Manual	Script	MSP
N	30	30	30
Promedio (%)	7.26	10.16	7.39
Mediana (%)	7.13	7.70	7.20
Desviación Estándar (%)	1.42	7.9	2.34
Mínimo (%)	4.63	5.2	3.74
Máximo (%)	10.5	47.1	17.7
Asimetría	0.474	3.764	2.90
Curtosis	0.123	16.43	12.9
Shapiro-Wilk W	0.97	0.55	0.72
Shapiro-Wilk p	0.576	< 0.01	< 0.01

Para determinar formalmente si existe una distribución normal de los datos, se analiza el resultado de la prueba de normalidad de Shapiro-Wilk (Tabla 6.5), lo que arroja un resultado para el método no automatizado de $W(30): 0.97$, $p= 0.58$, para el *script* $W(30): 0.55$, $p < 0.01$ y para el MSP $W(30): 0.72$, $p < 0.01$. Por lo descrito, se concluye que los datos provenientes desde el *script* y el MSP no están normalmente distribuidos.

Posteriormente, para determinar, mediante una prueba de estadística inferencial, si existen diferencias en los resultados que puedan ser atribuibles a la metodología utilizada, se aplicó la prueba Friedman debido a que los datos no presentan una distribución normal. El resultado de esta prueba fue de $\chi^2(2, n: 90) = 8.47$, $p < 0.05$, lo que revela que existen diferencias estadísticamente significativas entre las metodologías utilizadas. Para determinar con exactitud los grupos que presentan diferencias, se presentan los resultados de la comparación de los efectos principales entre las metodologías con la corrección de Bonferroni en la Tabla 6.6. A partir de estos datos, se concluye que los grupos que presentan diferencias estadísticamente significativas son los resultados del *script* en comparación al MSP y del *script* comparado con el método no automatizado.

Tabla 6.6 Resultados de la comparación de efectos principales del coeficiente de variabilidad de la DPS

Metodologías	Estadístico de prueba	Error estándar	Estadístico de prueba estándar	Sig.	Sig. Con ajuste de Bonferroni
Manual – MSP	0,033	0.258	0.129	0.879	1.0
MSP - <i>Script</i>	0,667	0.258	2.582	0.010	0.029
Manual - <i>Script</i>	-0.633	0.258	-2.453	0.014	0.043

6.1.2 Resultados del análisis de validación del comportamiento del *script*, aplicando además la funcionalidad de corrección humana de la segmentación silábica automatizada

6.1.2.1 Resultados de validación de los datos del promedio de sílabas por segundo (PSS), con corrección

Como se observa en la Tabla 6.7, en general, los resultados de la estadística descriptiva para este indicador, comparando las tres metodologías son bastante similares, casi idénticos. El análisis entrega un valor promedio de 5,7 s/s para todas y cada una de ellas. Asimismo, los tres métodos arrojan valores de desviación estándar bastante similares (0,75 para dos de ellos y 0,73 para uno). Además, como se puede observar los valores mínimos y máximos son casi iguales en los tres grupos.

Tabla 6.7 Resultados del análisis estadístico descriptivo del *script* con corrección humana del PSS

	Manual	Script	MSP
N	30	30	30
Promedio	5.70	5.70	5.70
Mediana	5.79	5.79	5.78
Derivación estándar	0.747	0.747	0.728
Mínimo	4.08	4.08	4.07
Máximo	7.18	7.18	7.17
Asimetría	-0.304	-0.296	-0.305
Curtosis	-0.322	-0.317	-0.238
Shapiro-Wilk W	0.980	0.981	0.978
Shapiro-Wilk p	0.821	0.850	0.777

En el gráfico de la Figura 6.4, se puede apreciar visualmente con mayor detalle la similitud de los resultados del cálculo del PSS. En el gráfico, se presentan los resultados de las 30 mediciones de manera individual y se comparan los tres métodos. Se incluyen además los resultados del análisis del *script* sin la aplicación de correcciones a la segmentación (que corresponde a los datos etiquetados como “script s/c”). Se puede observar que la similitud de las mediciones es tal que las líneas prácticamente se traslapan y una de ellas no alcanza a apreciarse claramente debido a esto. Los únicos datos que presentan una trayectoria distinta son los del *script* sin correcciones, pero tampoco resultan ser tan diferentes. La mayor diferencia entre las mediciones no alcanza a ser más de 6 décimas de sílaba en la penúltima medición.

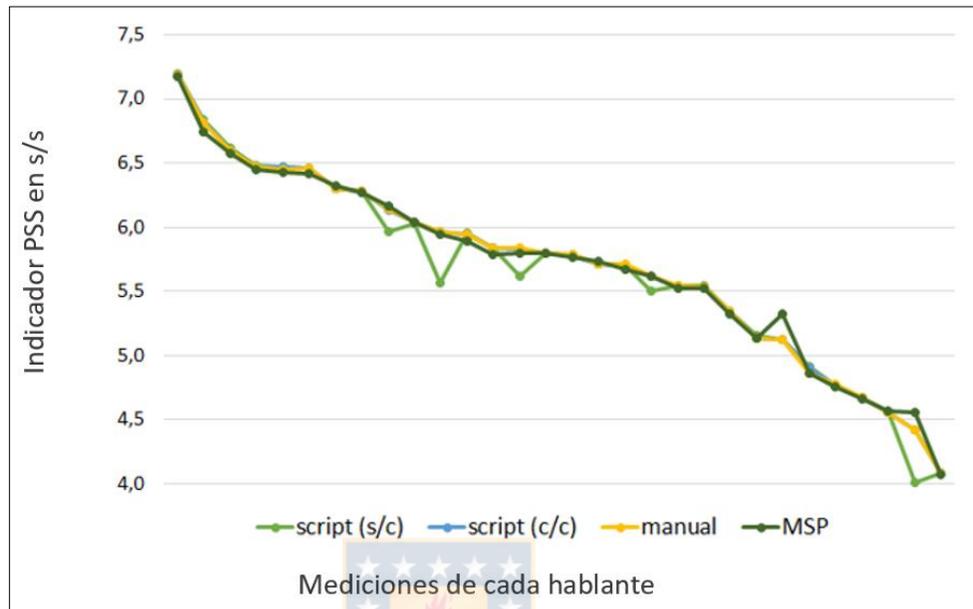


Figura 6.4. Gráfico comparativo de cada una de las mediciones individuales del PSS para cada uno de los tres métodos.

De acuerdo con los resultados que arroja el análisis descriptivo, se puede decir que en general los tres grupos se comportan de forma bastante similar, es decir, en principio no se ve que el factor método de análisis afecte los resultados.

Como se puede observar en el histograma de la Figura 6.5 y en el gráfico de densidad de la Figura 6.6, en los tres grupos los datos muestran una distribución normal, lo que además coincide con los resultados de la prueba de Shapiro-Wilk (Tabla 6.7). que arroja un resultado para el método no automatizado de $W(30): 0.98, p = 0.82$, para el script $W(30): 0.98, p = 0.85$ y para el MSP $W(30): 0.98, p = 0.85$.

(30): 0.98 , $p = 0.78$. Por lo descrito, se concluye que los datos provenientes desde las tres metodologías están normalmente distribuidos.

Tanto el histograma como el gráfico QQ muestran una distribución normal, en las tres metodologías, concentrados los resultados principalmente en el valor de 6 s/s, todos dentro de un rango de entre 4 s/s como mínimo y de 8 s/s como máximo.

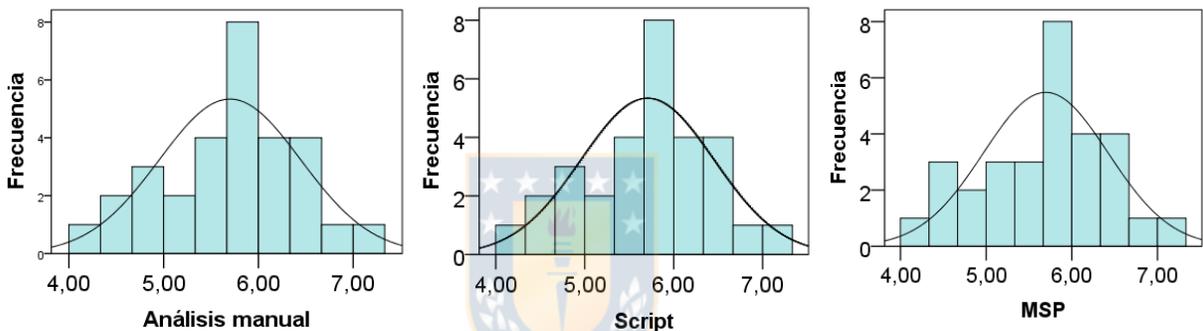


Figura 6.5 Histogramas de distribución del indicador PSS. Se muestran los datos obtenidos desde las tres metodologías analizadas (manual-script-MSP).

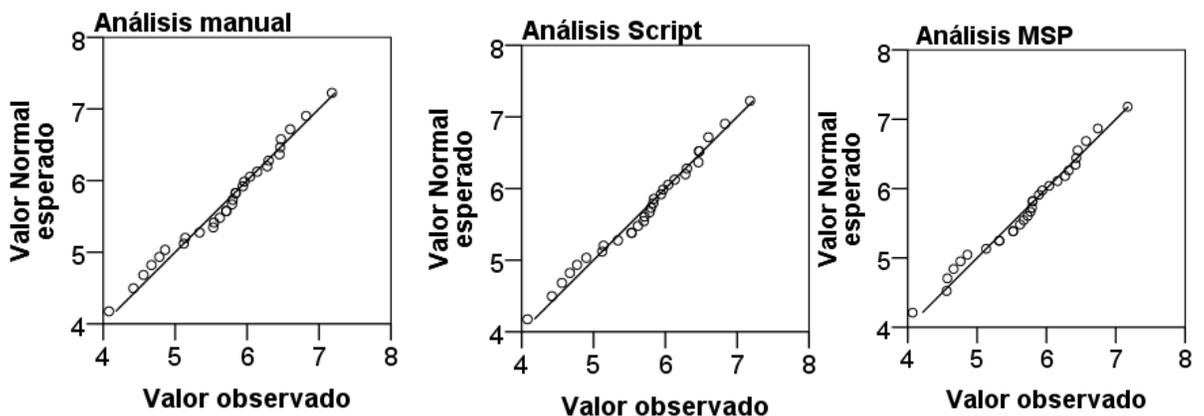


Figura 6.6 Gráficos QQ del indicador PSS.

Para determinar el estado de la homogeneidad de la varianza, se aplicó la prueba F-test cuyos resultados entre la medición manual y el *script* $F(29, 29) = 0.999$, $p = 0.999$, manual y MSP $F(29, 29) = 1.054$, $p = 0.889$ y entre los valores del *script* comparados con los del MSP $F(29, 29) = 1.006$, $p = 0.988$ no revelan diferencias estadísticamente significativas entre las varianzas de los grupos estudiados.

Luego, para determinar, mediante una prueba estadística, si existen diferencias en los resultados que dependan de la metodología utilizada, se aplicó la prueba ANOVA de medidas repetidas. No obstante, previamente, como corresponde a este tipo de análisis, se revisó el valor de esfericidad que entrega la prueba de Mauchly, la cual dio un valor de $\chi^2(2) = 65.54$, $p < 0.001$, bajo el nivel de significancia, es decir, no cumple con el supuesto de esfericidad. En consecuencia, el resultado de la prueba ANOVA se debe interpretar con las correcciones de Greenhouse y Geisser, prueba que entregó un valor con un tamaño de efecto bajo (como se observa en la tabla 6.8) de $F(1.05, 30.47): 0.198$, $p = 0.671$.

Tabla 6.8 ANOVA de medidas repetidas

Origen		gl	F	Sig.	Eta parcial al cuadrado
Metodología	Greenhouse-Geisser	1.051	0.198	0.671	0.007
Error	Greenhouse-Geisser	30.466			

En consecuencia, como se puede apreciar en la Tabla 6.8, aplicada la prueba de ANOVA de medidas repetidas, no se observan diferencias estadísticamente significativas entre los tres grupos de datos, es decir, el método segmentación, de medición y de cálculo del promedio de sílabas por segundo no afecta significativamente a los resultados.

6.1.2.2 Resultados de la validación de los datos del coeficiente de variabilidad del PSS

De acuerdo con los objetivos de este estudio, al inicio de la investigación se pretendía comparar todos los indicadores que arroja el análisis mediante el *script* con sus similares en el MSP. Así, a partir de lo que se puede entender de lo que dice el manual del MSP, el parámetro DDKjit mediría lo mismo que el indicador coeficiente de variabilidad del PSS. Sin embargo, como se aprecia en

el gráfico de la Figura 6.7, los resultados del MSP son tan evidentemente diferentes, que no es posible ni siquiera insinuar que ambos indicadores se estén calculando de la misma forma.

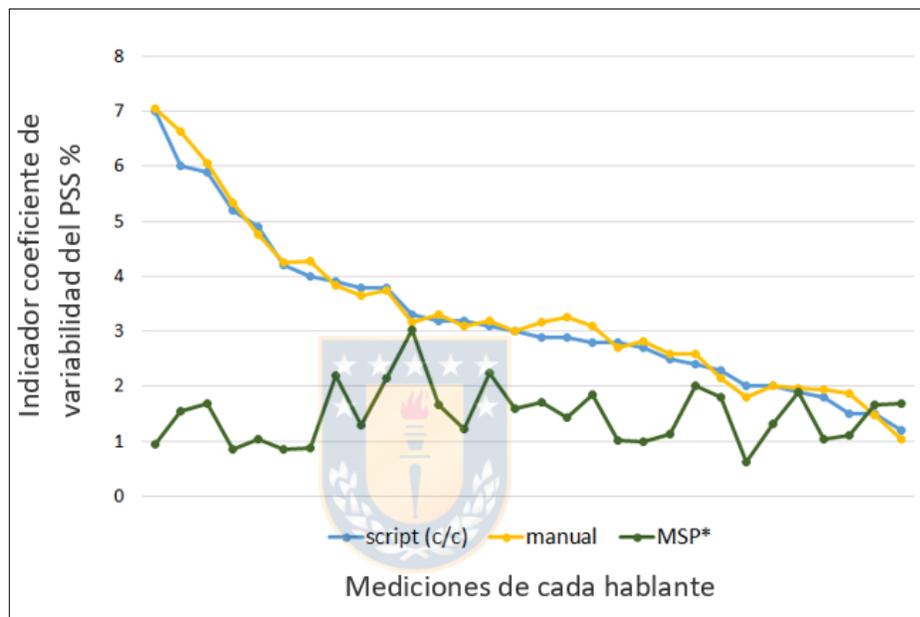


Figura 6.7 Gráfico comparativo de cada una de las mediciones individuales del coeficiente de variabilidad del PSS para cada uno de los tres métodos.

Como se observa en la Figura 6.7, varias de las mediciones del MSP presentan más de 3 puntos porcentuales de diferencia con respecto a los valores de los otros métodos (más de puntos en un par de casos), además de presentar un comportamiento muy heterogéneo.

Por lo tanto, en el análisis de validación del *script* para este indicador, solo se comparan los resultados del análisis no automatizado con el automatizado mediante el *script*.

En la Tabla 6.9, se presentan los resultados del análisis de estadística descriptiva de los datos para este indicador. En general, lo que se puede comentar es que el valor promedio es de alrededor de un 3,3% de variabilidad en ambos métodos. También, se observa que el valor mínimo obtenido es de un 1,03% de variabilidad y que el máximo es de un 7,05% para el método manual, mientras que, con el análisis del *script*, estos parámetros se dan en un rango bastante similar de 1,2% y 7,0% respectivamente.

Tabla 6.9 Resultados del coeficiente de variabilidad del PSS

	Manual	Script
N	30	30
Promedio %	3.33	3.26
Mediana %	3.13	2.95
Derivación estándar %	1.47	1.42
Mínimo %	1.03	1.20
Máximo %	7.05	7.00
Asimetría	0.992	0.956
Curtosis	0.749	0.619
Shapiro-Wilk W	0.921	0.930
Shapiro-Wilk p	0.029	0.048

En el gráfico de la Figura 6.8, se puede apreciar visualmente con mayor detalle la similitud de los resultados del cálculo de la variabilidad del PSS. En el gráfico, se presentan los resultados de las 30 mediciones de manera individual y se comparan los dos métodos. Se puede observar que la similitud de las mediciones es bastante. La mayor diferencia entre las mediciones no alcanza a ser de más de 6 décimas porcentuales, el resto nunca pasa más allá de las 2 o 3 décimas.

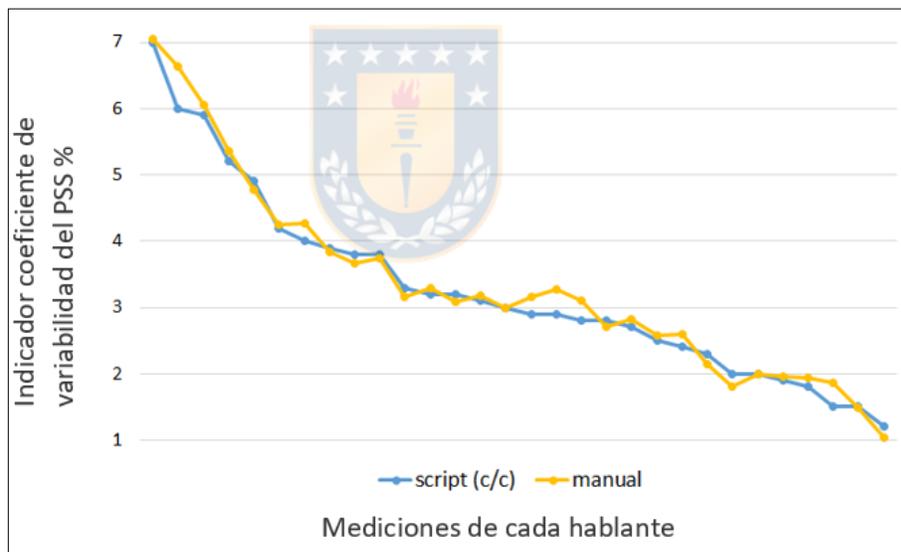


Figura 6.8 Gráfico comparativo de cada una de las mediciones individuales del coeficiente de variabilidad del PSS para el script y en método no automatizado.

De acuerdo con los resultados que arroja el análisis descriptivo, se puede decir que en general ambos grupos se comportan de forma bastante similar, es decir, en principio no se ve que el factor método de análisis afecte los resultados.

Se desprende del análisis visual del histograma de la Figura 6.9 y en el gráfico QQ de la Figura 6.10 que los dos grupos de datos muestran una distribución que no coincide categóricamente con una distribución normal. Los valores de asimetría y curtosis (Tabla 6.9) de ambas metodologías se encuentran sobre 0,5. Lo anterior se corrobora con el análisis formal de la distribución mediante la prueba de normalidad de Shapiro-Wilk (Tabla 6.9) un resultado para el método no automatizado de $W(30): 0.92, p < 0.05$, para el *script* $W(30): 0.93, p < 0.05$. Por lo descrito, se concluye que los datos no están normalmente distribuidos.

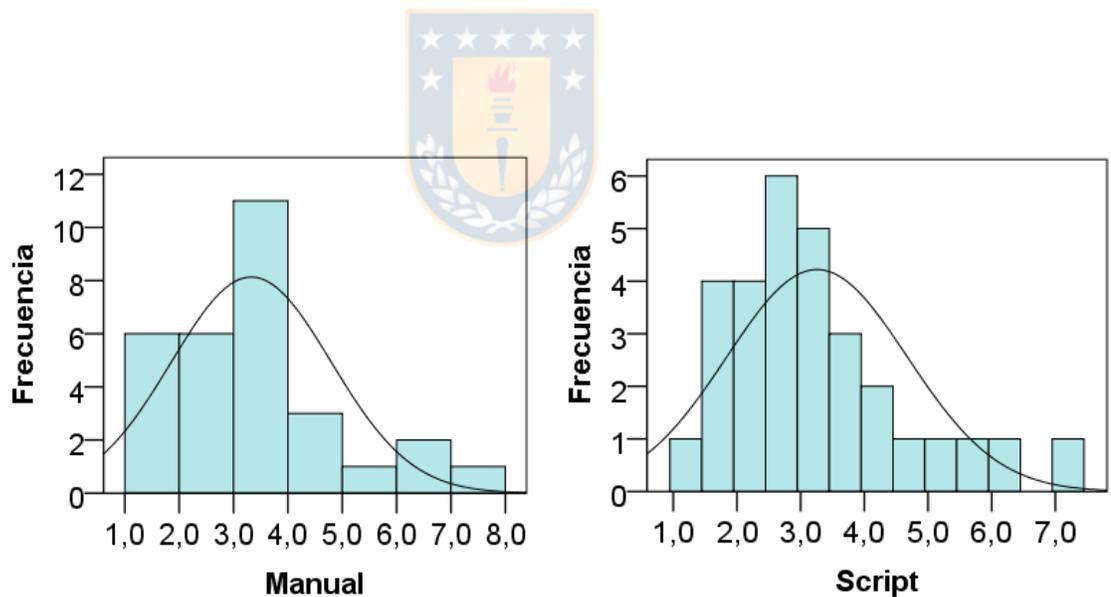


Figura 6.9 Histograma de la distribución y curva de normalidad del coeficiente de variabilidad del promedio de sílabas por segundo.

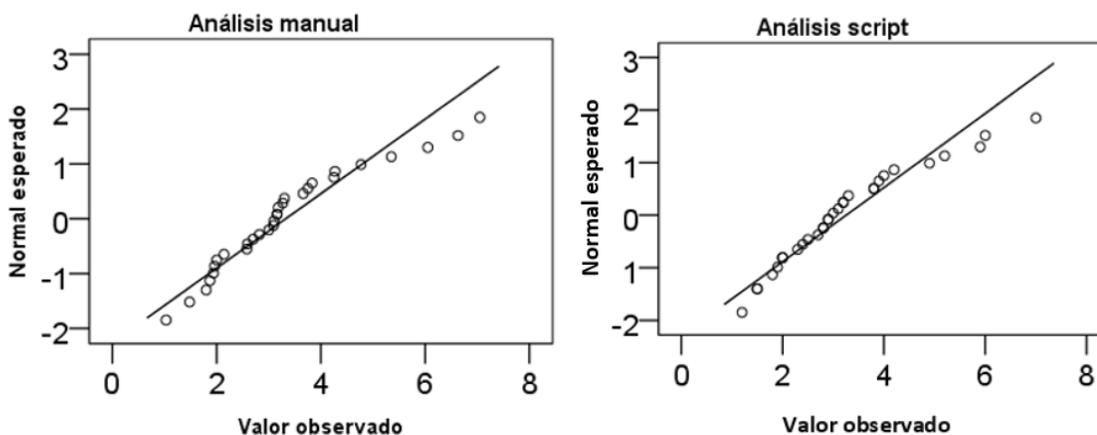


Figura 6.10 Gráfico QQ de la distribución del coeficiente de variabilidad del promedio de sílabas por segundo.

Debido a que los datos no cumplen con la condición de distribución normal, se optó por aplicar la prueba no paramétrica de los rangos con signos de Wilcoxon para determinar si existen diferencias significativas entre los resultados obtenidos mediante ambas metodologías. La prueba reveló que no existen diferencias estadísticamente significativas del resultado del coeficiente de variabilidad del promedio de sílabas por segundo que se obtiene utilizando, ya sea el *script* (*Mdn*: 2.95, *n*:30) o un análisis no automatizado (*Mdn*: 3.13, *n*: 30), z : -1.58, $p= 0.12$, con un tamaño de efecto (r) pequeño de 0.2.

6.1.2.3 Resultados de la validación de los datos de la duración promedio de la sílaba (DPS)

Los datos de la Tabla 6.10, muestran los resultados de la estadística descriptiva para este indicador, comparando las tres metodologías. Se destaca un valor promedio de 179 milisegundos para todas y cada una de ellas. De la misma manera, los tres métodos arrojan el mismo valor de desviación estándar (25 ms) y valores mínimos y máximos casi idénticos en los tres grupos.

Tabla 6.10 Resultados del promedio de la duración silábica

	Manual	Script	MSP
N	30	30	30
Promedio (ms)	179	179	179
Mediana (ms)	174	173	173
Desviación estándar (ms)	25.0	25.0	24.9
Mínimo (ms)	139	139	139
Máximo (ms)	244	245	246
Asimetría	0.795	0.901	0.883
Curtosis	0.281	0.465	0.468
Shapiro-Wilk W	0.950	0.937	0.936
Shapiro-Wilk p	0.173	0.074	0.070

En el gráfico de la Figura 6.11, se puede apreciar visualmente con mayor detalle la similitud de los resultados del cálculo de la DPS. En el gráfico, se presentan los resultados de las 30 mediciones de manera individual y se comparan los tres métodos. Se incluyen además los resultados del análisis del *script* sin la aplicación de correcciones a la segmentación (que corresponde a los datos etiquetados como “script s/c”). Se puede observar que la similitud de las mediciones es tal que las líneas prácticamente se traslapan y una de ellas no alcanza a apreciarse claramente debido a esto. Sólo una de las mediciones del método no automatizado y una del *script* (sin correcciones) manifiestan un valor notoriamente diferente al resto de los métodos, pero ninguna de las dos diferencias sobrepasa los 20 ms.

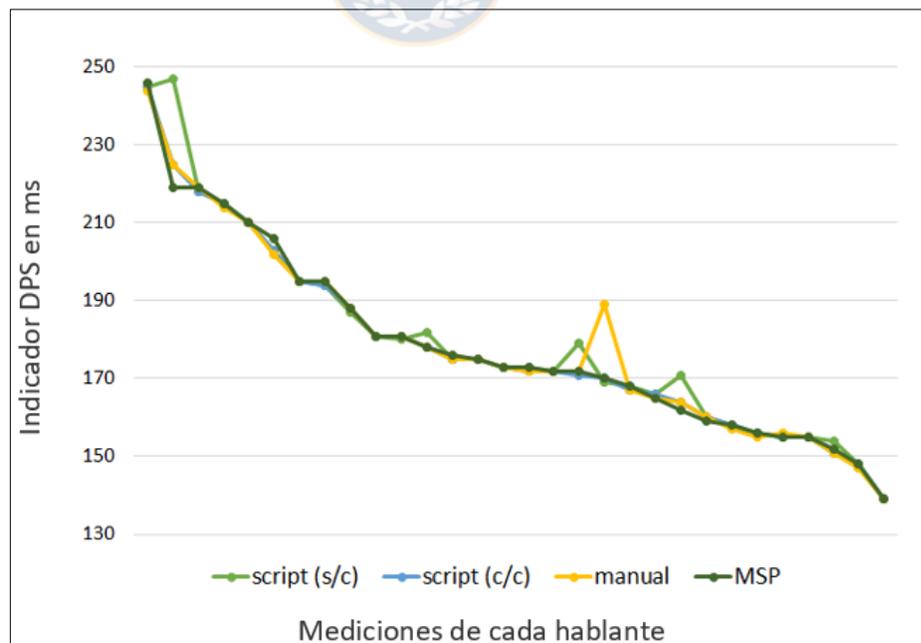


Figura 6.11 Gráfico comparativo de cada una de las mediciones individuales de la DPS para cada uno de los tres métodos.

De los resultados del análisis descriptivo se desprende que en general los tres grupos se comportan de forma bastante similar (casi idéntica), es decir, en principio no se ve que el factor método de análisis afecte los resultados.

Como se puede observar en el histograma de la Figura 6.12 y en el gráfico QQ de la Figura 6.13, en los tres grupos los datos muestran una distribución normal, lo que además coincide con los resultados de la prueba de normalidad de Shapiro-Wilk (Tabla 6.10) para el método no automatizado de $W(30)$: 0.95 , $p = 0.17$, para el *script* $W(30)$: 0.94 , $p = 0.07$ y para el MSP $W(30)$: 0.94 , $p = 0.07$. Ambos gráficos muestran una distribución normal de los datos de las tres muestras, los cuales se distribuyen dentro de un rango de entre 140 ms y 250 ms concentrados los valores principalmente en los 170 ms.

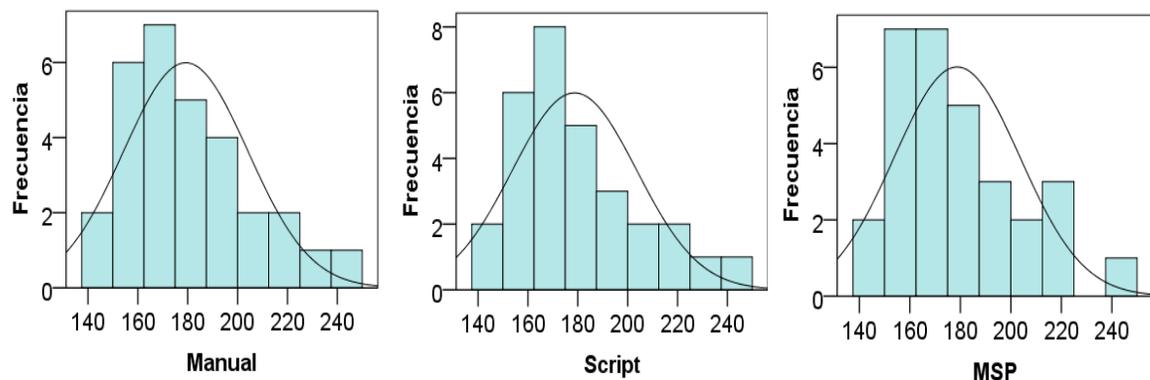


Figura 6.12 Histogramas con curva de normalidad de la distribución del promedio de la duración silábica

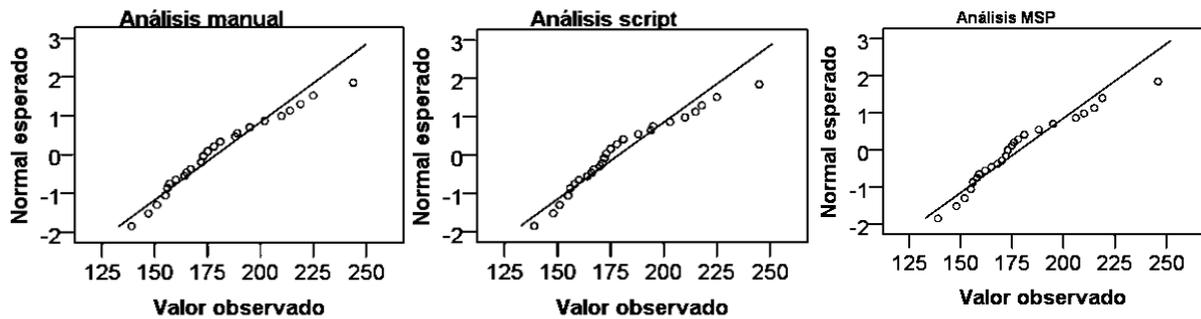


Figura 6.13 Gráficos QQ de la distribución del índice promedio de duración silábica.

Para determinar la homogeneidad de la varianza se aplicó la prueba F-test, cuyos valores entre el análisis manual y el *script* $F(29, 29) = 1, p = 1$, el manual y el MSP $F(29, 29) = 1.01, p = 0.99$ y entre los resultados del *script* comparados con los del MSP $F(29, 29) = 1.01, p = 0.99$ indican que la prueba no encuentra diferencias estadísticamente significativas entre las varianzas de los grupos estudiados.

Luego, para determinar, mediante una prueba de estadística inferencial, si se observan diferencias en los resultados que puedan ser atribuibles al tipo de metodología utilizada, se aplicó la prueba ANOVA de medidas repetidas. No obstante, previamente, como corresponde a este tipo de análisis, se revisó el valor de esfericidad que entrega la prueba de esfericidad de Mauchly, la cual dio un valor de $\chi^2(2) = 28.65, p < 0.001$ menor al de significancia, es decir, no cumple con el supuesto de esfericidad. En consecuencia, el resultado de la prueba ANOVA se debe interpretar con las correcciones de Greenhouse y Geisser,

prueba que, prueba que entregó un valor con un tamaño de efecto bajo (como se observa en la tabla 6.11) de $F(1.22, 35.35) = 0.55$, $p = 0.5$. Por lo cual, la prueba no reveló diferencias estadísticamente significativas atribuibles a la metodología de análisis de las muestras de habla para calcular este indicador.

Tabla 6.11 ANOVA de medidas repetidas

Origen		gl	F	Sig.	Eta parcial al cuadrado
Metodología	Greenhouse-Geisser	1.219	0.551	0.496	0.019
Error	Greenhouse-Geisser	35.353			

6.1.2.4 Resultados de la validación de los datos del coeficiente de variabilidad de la DPS

La Tabla 6.12 muestra los datos no normalizados del indicador según la metodología utilizada para analizar. Se desprende de los resultados de la estadística descriptiva para este indicador que los valores promedio son cercanos: 7,3% para el análisis manual, 7,2% para el *script* y 7,4% para el programa MSP.

Tabla 6.12 Resultados del coeficiente de variabilidad de la duración silábica

	Manual	Script	MSP
N	30	30	30
Promedio (%)	7.26	7.15	7.39
Mediana (%)	7.13	7.20	7.20
Desviación Estándar (%)	1.42	1.49	2.34
Mínimo (%)	4.63	4.50	3.74
Máximo (%)	10.5	10.7	17.7
Asimetría	0.474	0.451	2.90
Curtosis	0.123	0.221	12.9
Shapiro-Wilk W	0.97	0.97	0.72
Shapiro-Wilk p	0.57	0.74	<0.001

Sin embargo, se observa que los valores máximos y el valor de la desviación estándar difieren de manera algo más notoria, particularmente en el método MSP.

En el gráfico de la Figura 6.14, se puede apreciar visualmente con mayor detalle el comportamiento de los resultados del cálculo de la variabilidad de la DPS. En el gráfico, se presentan los resultados de las 30 mediciones de manera individual y se comparan los tres métodos. Se puede observar que para el método no automatizado y el *script* la similitud de las mediciones es bastante. No obstante, la medición con el MSP presenta más diferencias, una de ellas bastante pronunciada. El resto de las mediciones presenta algunos valores iguales y otros

diferentes, pero en la gran mayoría la diferencia no sobrepasa los 3 puntos porcentuales.

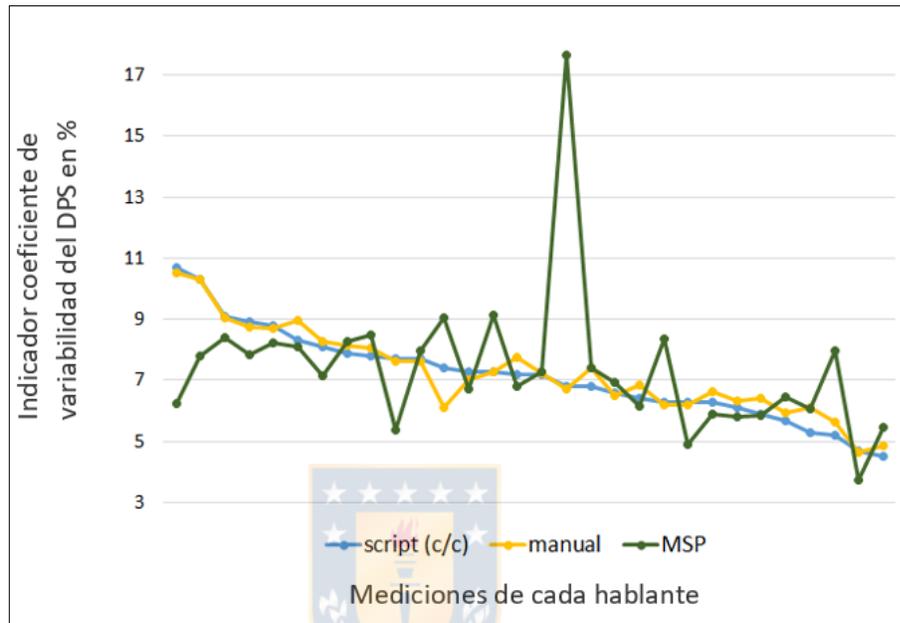


Figura 6.14 Gráfico comparativo de cada una de las mediciones individuales del coeficiente de variabilidad de la DPS para los tres métodos.

El análisis visual, tanto del histograma (Figura 6.15) como del gráfico QQ (Figura 6.16) muestra una distribución de los valores entre 6% y 8%. No obstante, el análisis realizado por el MSP indica la existencia de un valor atípico, el cual no ocurre en los resultados de las otras metodologías. Esta situación, sumada al análisis descriptivo de los datos, tanto el valor de la curtosis como de la asimetría, permite decidir eliminar el valor atípico presente en los resultados que arroja el programa comercial.

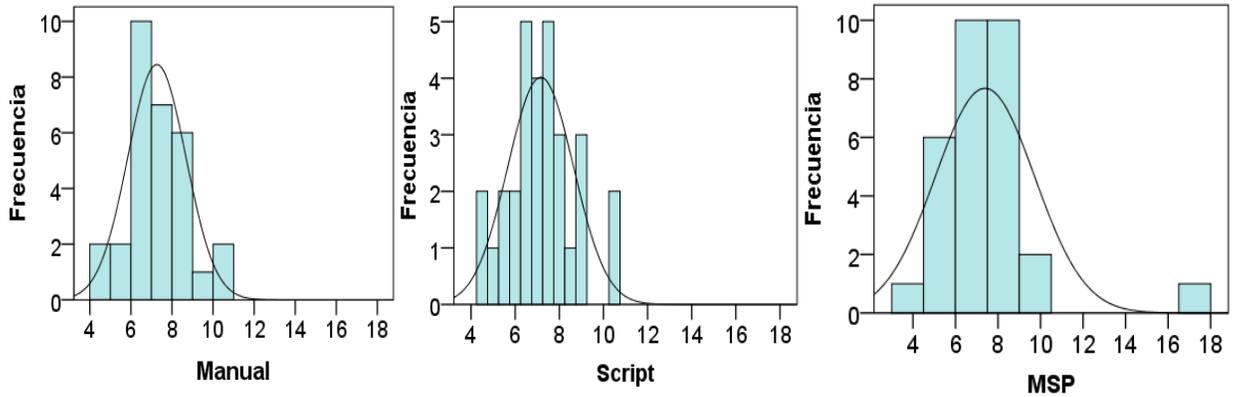


Figura 6.15 Histogramas con curva de normalidad del coeficiente de variabilidad de la duración silábica.

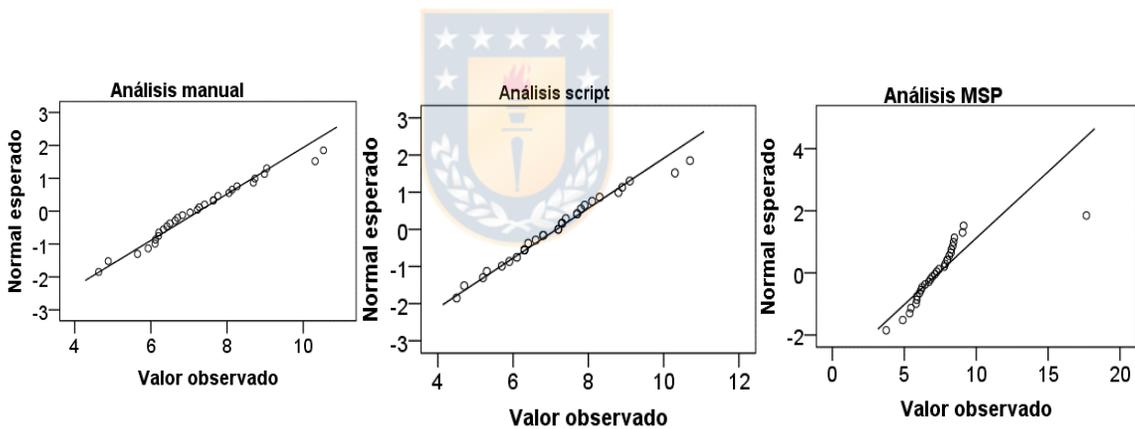


Figura 6.16 Gráfico QQ de distribución del coeficiente de variabilidad de la duración silábica

Los datos que se obtienen luego del proceso para eliminar el valor atípico se someten a las mismas pruebas de estadística descriptiva (Tabla 6.13). Basado en un análisis general, se desprende que los resultados de la estadística descriptiva para este indicador arrojan un valor promedio de 7,3% de variabilidad para el análisis no automatizado; de 7,2% para el *script* y un valor promedio en

el MSP de 7% de variabilidad. Los valores de la desviación estándar son de un 1,3% para el MSP, de un 1,2% para el *script* y de un 1,4% para el análisis no automatizado.

Tabla 6.13 Resultados del coeficiente de variabilidad de la duración silábica

	Manual	Script	MSP
N	29	29	29
Promedio (%)	7.28	7.17	7.03
Mediana (%)	7.22	7.20	7.13
Desviación estándar (%)	1.44	1.51	1.33
Mínimo (%)	4.63	4.50	3.74
Máximo (%)	10.5	10.7	9.12
Asimetría	0.430	0.420	-0.458
Curtosis	0.0244	0.108	-0.279
Shapiro-Wilk W	0.973	0.977	0.964
Shapiro-Wilk p	0.650	0.762	0.404

Para determinar si los datos se encuentran normalmente distribuidos, se presenta el histograma de la Figura 6.17 y el gráfico QQ de la Figura 6.18. Se concluye que en los tres grupos los datos muestran una distribución normal, lo que además coincide con los resultados de la prueba de normalidad de Shapiro-Wilk (Tabla 6.13), para el método no automatizado de $W(29): 0.97, p=0.65$, para el *script* $W(29): 0.98, p=0.77$ y para el MSP $W(29): 0.96, p=0.4$.

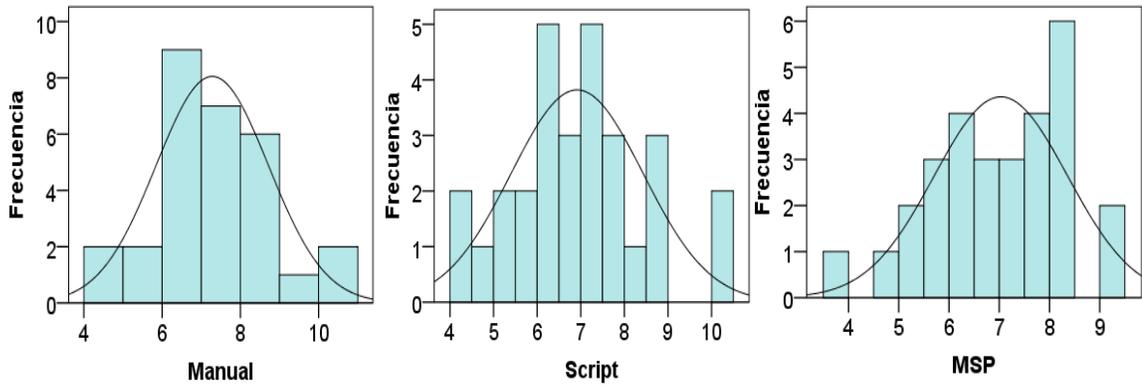


Figura 6.17 Histogramas con curvas de normalidad de los datos del coeficiente de variabilidad del promedio de duración silábico luego del proceso de normalización.

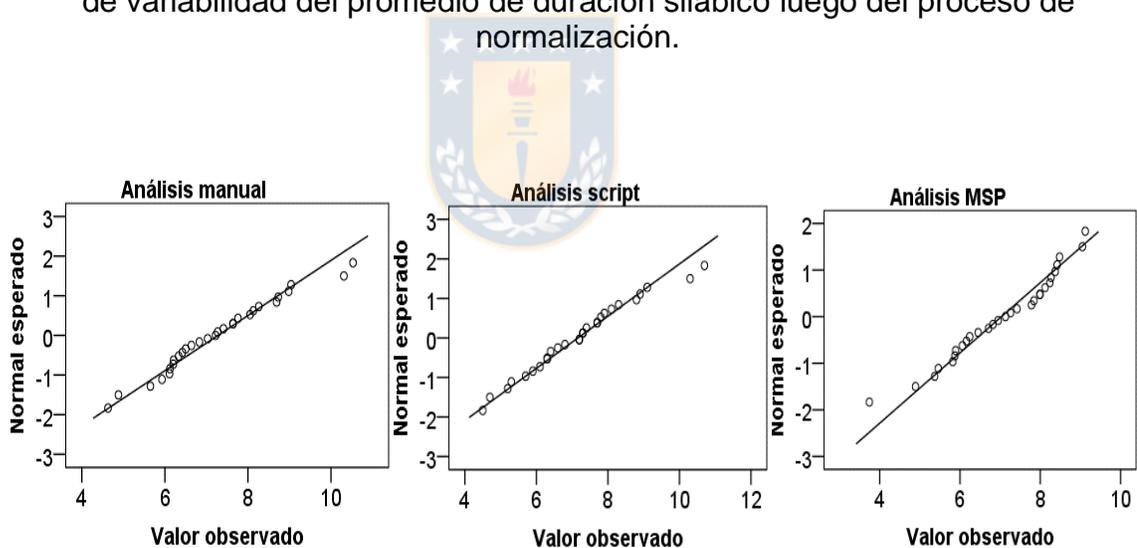


Figura 6.18 Distribución del coeficiente de variabilidad de la duración silábica luego del procedimiento de normalización.

Se aplicó la prueba F-test del paquete “car” en R para obtener información sobre la homogeneidad de las varianzas entre el análisis manual y el *script* $F(28, 28) = 0.901$, $p = 0.784$, el análisis manual y el MSP $F(28, 28) = 1.173$, $p = 0.676$ y entre los resultados del *script* comparados con los del MSP $F(28, 28) = 1.303$, $p = 0.489$). Todos ellos arrojan valores sobre 0,05, por lo que se puede concluir que la prueba no encuentra diferencias estadísticamente significativas entre las varianzas de las metodologías comparadas.

Posteriormente, para determinar, mediante una prueba estadística inferencial, si existen diferencias en los resultados que puedan ser atribuible a la metodología utilizada, se aplicó la prueba ANOVA de medidas repetidas. No obstante, previamente, como corresponde a este tipo de análisis, se revisó el valor de esfericidad que entrega la prueba de Mauchly, la cual dio un valor de $\chi^2(2) = 43.08$, $p < 0.001$ menor al de significancia, es decir, no cumple con el supuesto de esfericidad. En consecuencia, el resultado de la prueba ANOVA se debe interpretar con las correcciones de Greenhouse y Geisser, prueba que, como se observa en la Tabla 6.14, entregó un valor con un tamaño de efecto bajo de $F(1.11, 31.16) = 0.615$, $p = 0.456$. Al analizar el valor de significancia, se concluye que la prueba no reveló diferencias estadísticamente significativas atribuibles a la metodología de análisis de las muestras de habla para calcular este indicador.

Tabla 6.14 Resultado ANOVA de medidas repetidas.

Origen		gl	F	Sig.	Eta parcial al cuadrado
Metodología	Greenhouse-Geisser	1.113	0.615	0.456	0.021
Error	Greenhouse-Geisser	31.16			

6.1.3 Comparación entre los resultados del *script* sin aplicar correcciones en la segmentación con los resultados del *script* aplicando además la funcionalidad de corrección no automatizada de la segmentación

Con el fin de determinar las similitudes y diferencias que existen entre utilizar el *script* de manera totalmente automatizada, o bien, hacerlo aplicando la funcionalidad de este que permite alguna intervención en la segmentación, se presenta un análisis comparativo entre los resultados de ambas opciones.

6.1.3.1 Indicador promedio de sílabas por segundo (PSS)

Los datos de la estadística descriptiva y los valores de la prueba de Shapiro-Wilk del *script* sin correcciones ($\bar{x} = 5.66$, DE: 0.771) se pueden apreciar en la Tabla 6.1 y, los del *script* con correcciones ($\bar{x} = 5.7$, DE: 0.747), en la Tabla

6.7. En general, presentan valores relativamente similares. En ambos casos, la prueba para verificar si los datos presentan una distribución normal confirma este supuesto. De la misma manera, la prueba para determinar la homogeneidad de varianza entrega los siguientes valores $F(29, 29) = 0.94$, $p = 0.867$, por lo cual se concluye que no existe una diferencia estadísticamente significativa entre la varianza de ambos grupos.

Para evaluar la relación entre ambos grupos se aplicó la prueba T para dos muestras relacionadas. El resultado que se obtuvo indica que no existen diferencias estadísticamente significativas entre las medianas de ambos grupos: $T(29): 1.89$, $p = 0.069$, con un tamaño de efecto (d) medio (0.344).

6.1.3.2 Indicador coeficiente de variabilidad del PSS

Los datos de la estadística descriptiva y los valores de la prueba de Shapiro-Wilk del *script* sin correcciones ($\bar{x} = 4.5$, DE: 3.9) se pueden apreciar en la Tabla 6.3 y, los del *script* con correcciones ($\bar{x} = 3.26$, DE: 1.42), en la Tabla 6.9. En general, presentan valores con diferencias entre los grupos, con una amplia dispersión de datos. En ambos casos, la prueba para verificar si los datos presentan una distribución normal refuta este supuesto. De la misma manera, la prueba para determinar la homogeneidad de varianza entrega los siguientes

valores $F(29, 29) = 0.134$, $p < 0.001$, por lo cual se concluye que existe una diferencia estadísticamente significativa entre la varianza de ambos grupos.

Para evaluar la relación entre ambos grupos se aplicó la prueba de los rangos de signos de Wilcoxon para dos muestras relacionadas, debido a que los datos no están normalmente distribuidos. El resultado revela que existen diferencias estadísticamente significativas entre los datos del coeficiente de variabilidad del promedio de sílabas por segundo que se obtiene utilizando, ya sea el *script* sin corrección ($Mdn: 3.7$, $n: 30$) o efectuando una corrección manual ($Mdn: 2.95$, $n: 30$), $Z: -2.204$, $p < 0.05$, con un tamaño de efecto (r) moderado de 0.4.

6.1.3.3 Indicador duración promedio de la sílaba (DPS)

Los datos de la estadística descriptiva y los valores de la prueba de Shapiro-Wilk del *script* sin correcciones ($\bar{x} = 180$, $DE: 26,3$) se pueden apreciar en la Tabla 6.4 y, los del *script* con correcciones ($\bar{x} = 179$, $DE: 25$), en la Tabla 6.10. En general, presentan valores relativamente similares. La prueba para verificar si los datos presentan una distribución normal del *script* con correcciones confirma este supuesto, pero cuando se analizan los resultados del *script* sin correcciones se obtiene un valor bajo el de significancia, por lo que este grupo no cumple con el supuesto de normalidad de distribución de los datos. Además, los resultados de la prueba para determinar la homogeneidad de varianza

entregan un valor de $F(29, 29) = 0.898$, $p = 0.773$), por lo cual se puede concluir que no existe una diferencia estadísticamente significativa entre la varianza de ambos grupos.

Para evaluar la relación entre ambos grupos se aplicó la prueba de los rangos de signos de Wilcoxon para dos muestras relacionadas, debido a que los datos no están normalmente distribuidos. El resultado revela que no existen diferencias estadísticamente significativas entre los datos del indicador DPS que se obtienen utilizando, ya sea el *script* sin corrección ($Mdn: 174$, $n:30$) o efectuando una corrección manual ($Mdn: 173$, $n: 30$), $Z: -1.442$, $p = 0.129$, con un tamaño de efecto pequeño de $r: 0.263$.

6.1.3.4 Indicador coeficiente de variabilidad de la DPS

Los datos de la estadística descriptiva y los valores de la prueba de Shapiro-Wilk del *script* sin correcciones ($\bar{x} = 10.16$, $DE: 7.9$) se pueden apreciar en la Tabla 6.5 y, los del *script* con correcciones ($\bar{x} = 7.15$, $DE: 1.49$), en la Tabla 6.12. En general, presentan valores con diferencias entre los grupos, con una amplia dispersión de datos. La prueba para verificar si los datos presentan una distribución normal del *script* con correcciones (Tabla 6.12) confirma este supuesto, pero cuando se analizaron los resultados del *script* sin correcciones (Tabla 6.5) se obtuvo un valor bajo el de significancia, por lo que este grupo no cumple con el supuesto de normalidad de distribución de los datos. De la misma manera, la prueba para determinar la homogeneidad de varianza entrega los

siguientes valores $F(29, 29) = 28.666$, $p < 0.001$), por lo cual se concluye que existe una diferencia estadísticamente significativa entre la varianza de ambos grupos.

Para evaluar la relación entre ambos grupos se aplicó la prueba de los rangos de signos de Wilcoxon para dos muestras relacionadas, debido a que los datos no están normalmente distribuidos. El resultado revela que existen diferencias estadísticamente significativas entre los datos del indicador coeficiente de variabilidad de la DPS que se obtiene utilizando el *script* sin corrección ($Mdn: 7.70$, $n: 30$) en comparación a los resultados que se obtienen al efectuar una corrección manual ($Mdn: 7.2$, $n: 30$), $Z: -3.191$, $p < 0.001$, con un tamaño de efecto de alta magnitud de $r: 0.583$.

6.2 Resultados del análisis de la fiabilidad inter-evaluador del análisis no automatizado

De acuerdo con los objetivos de esta investigación, una parte del proceso debe estar orientado a determinar la fiabilidad del análisis no automatizado, debido a que, por tratarse de un análisis realizado visualmente por un humano, siempre existe la posibilidad del error o del sesgo natural en este tipo de procedimientos.

Es por ello, que una parte de la muestra fue analizada además por un evaluador externo quien siguió los mismos pasos que la evaluadora principal, pero sin ningún tipo de supervisión de parte de ella.

En el gráfico de la Figura 6.19, se presenta el detalle comparado de las mediciones del indicador PSS que ambos evaluadores realizaron en las mismas 9 muestras. Se observa que ambos grupos de datos presentan un comportamiento casi idéntico. Lo mismo ocurre con los datos que se observan en el gráfico de la Figura 6.20, pero que en este caso los datos corresponden a las mediciones de la DPS.

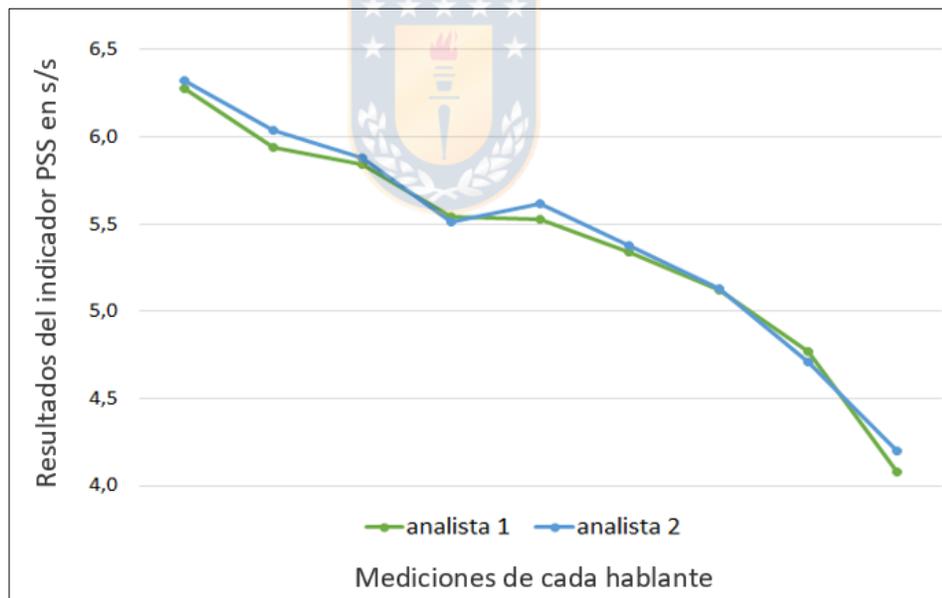


Figura 6.19 Gráfico comparativo de las mediciones del PSS realizadas por los dos evaluadores

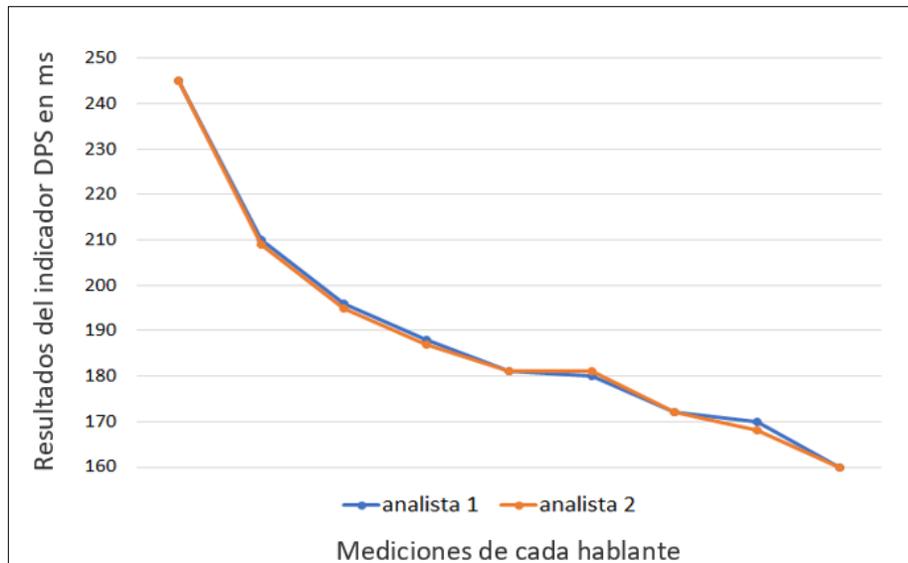


Figura 6.20 Gráfico comparativo de las mediciones de la DPS realizadas por los dos evaluadores

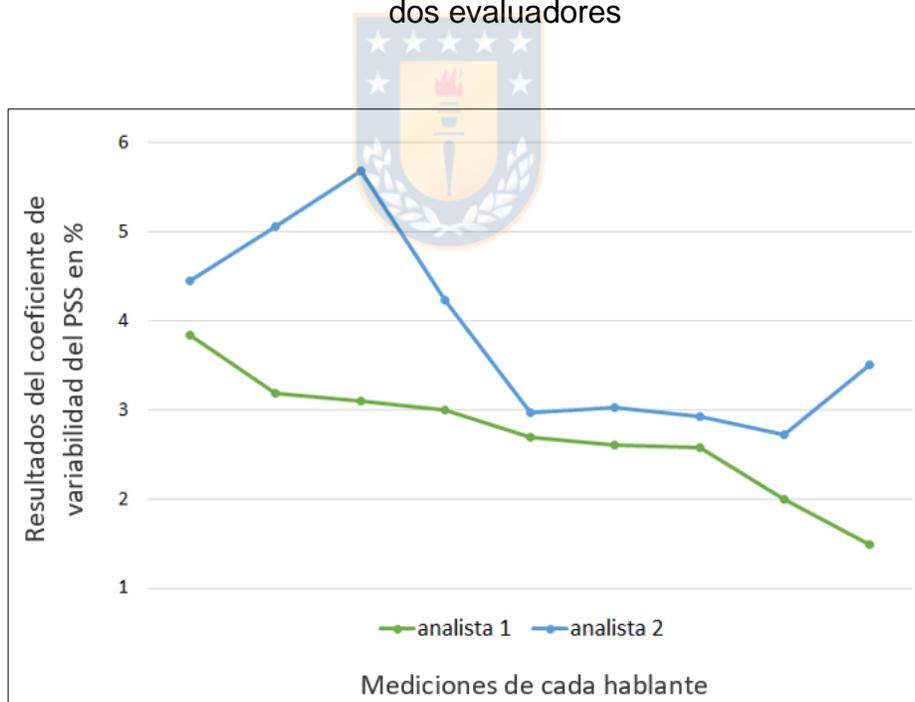


Figura 6.21. Gráfico comparativo de las mediciones del coeficiente de variabilidad del PSS realizadas por los dos evaluadores

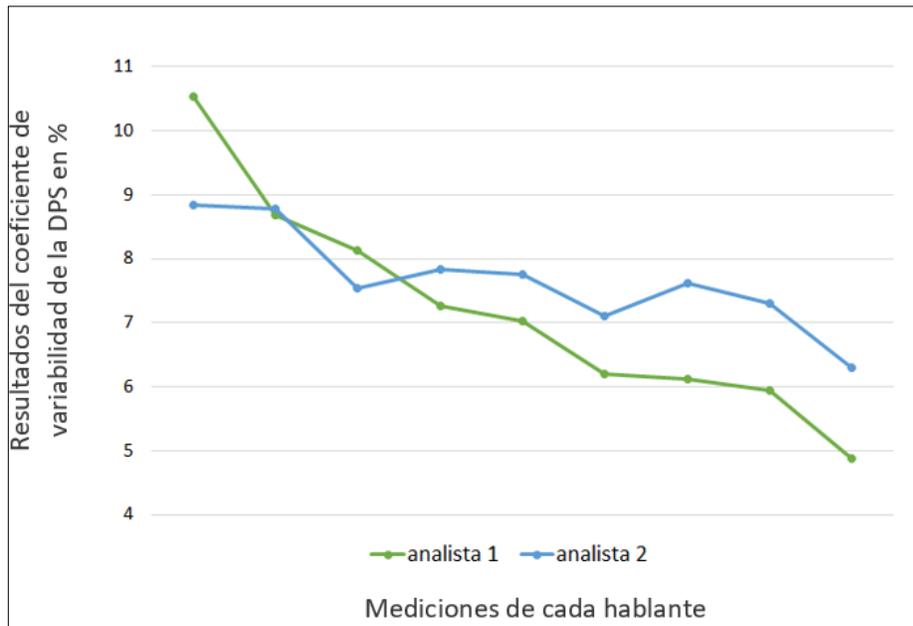


Figura 6.22 Gráfico comparativo de las mediciones del coeficiente de variabilidad de la DPS realizadas por los dos evaluadores

En el gráfico de la Figura 6.21, se presenta el detalle comparado de las mediciones del indicador coeficiente de variabilidad del PSS, que ambos evaluadores realizaron en las mismas 9 muestras. Curiosamente, en este caso, a diferencia de lo que ocurre con el PSS, se observa un comportamiento algo diferente de las mediciones de ambos evaluadores. No obstante, se debe destacar que la mayor diferencia no sobrepasa los 3 puntos porcentuales y el 50% de las diferencias entre las mediciones no alcanza a ser más de 1 punto porcentual.

En el gráfico de la Figura 6.22, se presenta el detalle comparado de las mediciones del indicador coeficiente de variabilidad de la DPS, que ambos

evaluadores realizaron en las mismas 9 muestras. En este caso se observa un comportamiento que tiende a ser similar de las mediciones de ambos evaluadores. Existen diferencias, pero ninguna de ellas sobrepasa los 2 puntos porcentuales.

Con el propósito de ratificar si el factor evaluador incide en los resultados del análisis, los datos entregados por ambos analistas fueron comparados mediante la prueba de correlación interclase 2 (ICC, por sus siglas en inglés).

Como se puede observar en la Tabla 6.15, los valores de la prueba indican una correlación excelente (sobre 0,9) para los dos indicadores promedio y entre “moderada” (entre 0,5 y 0,75) a “buena” (entre 0.75 a 0.9) para los coeficientes de variabilidad. Estos últimos indicadores se catalogan de esta manera dado los valores del intervalo de confianza, que entregan valores entre los expuestos.

Tabla 6.15 Resultados del ICC de la fiabilidad inter-evaluador.

Indicador	Coeficiente de correlación interclase	95% intervalo de confianza	
		Límite inferior	Límite superior
Promedio de sílabas por segundo	0,99	0,98	0,99
Coeficiente de sílabas por segundo	0,79	0,54	0,89
Promedio de duración silábica	0,98	0,97	0,99
Coeficiente de variabilidad del promedio de duración silábica	0,73	0,50	0,85

En consecuencia, los resultados del análisis estadístico, permiten decir que el evaluador humano no incide en los resultados del análisis no automatizado de los indicadores PSS y DPS, pero sí incide levemente en los coeficientes de variabilidad.

6.3 Resultados del análisis de la fiabilidad del *script*

De acuerdo con los objetivos de esta investigación, una parte del proceso debe estar orientado a determinar la fiabilidad del análisis automatizado, debido a que, por tratarse de un análisis realizado a un comportamiento humano, siempre existe la posibilidad de que los resultados puedan variar en distintas circunstancias.

Es por ello, que cuando se elicó el corpus, se solicitó explícitamente a los informantes que realizaran dos veces la emisión requerida. Esto con la idea de comparar los resultados de la medición de dos emisiones distintas, pero del mismo sujeto, realizadas en dos momentos distintos.

En el gráfico de la Figura 6.23, se presentan los resultados de la medición del PSS para el mismo sujeto, pero en dos momentos distintos. Se observa que, si bien hay diferencias, la mayoría de estas no superan un valor superior a media sílaba y la diferencia máxima no supera las dos sílabas, además de tratarse de solo dos casos de entre treinta.

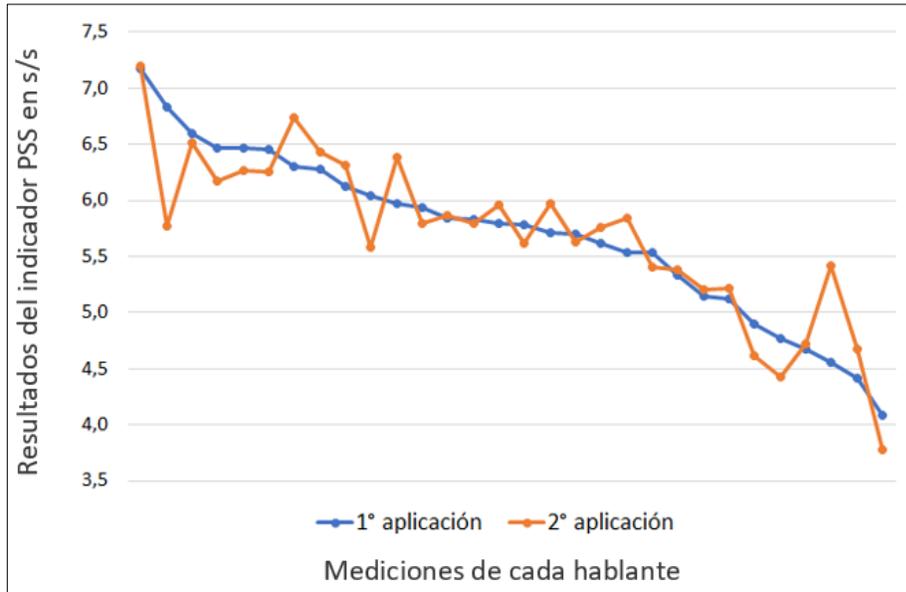


Figura 6.23 Gráfico comparativo de las mediciones del PSS de dos emisiones distintas realizadas por el mismo sujeto

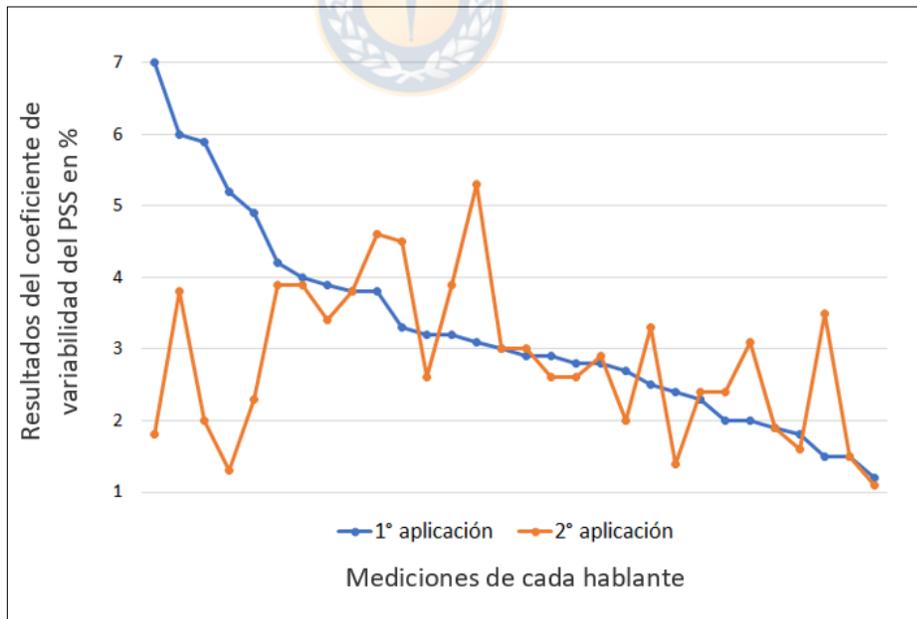


Figura 6.24 Gráfico comparativo de las mediciones del coeficiente de variabilidad del PSS de dos emisiones distintas realizadas por el mismo sujeto

En el gráfico de la Figura 6.24, se presenta el detalle comparado de las mediciones del indicador coeficiente de variabilidad del PSS para el mismo sujeto en dos emisiones distintas. Curiosamente, en este caso, a diferencia de lo que ocurre con el PSS, se observa un comportamiento algo diferente de los datos en ambas mediciones. No obstante, se debe destacar que el 80% de las diferencias entre las mediciones no alcanza a ser más de 2 puntos porcentuales.

En el gráfico de la figura 6.25, se presentan los resultados de la medición de la DPS para el mismo sujeto en dos momentos distintos. Se observa que, si bien hay diferencias, la mayoría de estas no superan un valor superior 20 ms. Sólo un caso supera este límite, pero no pasa más allá de los 25 ms.



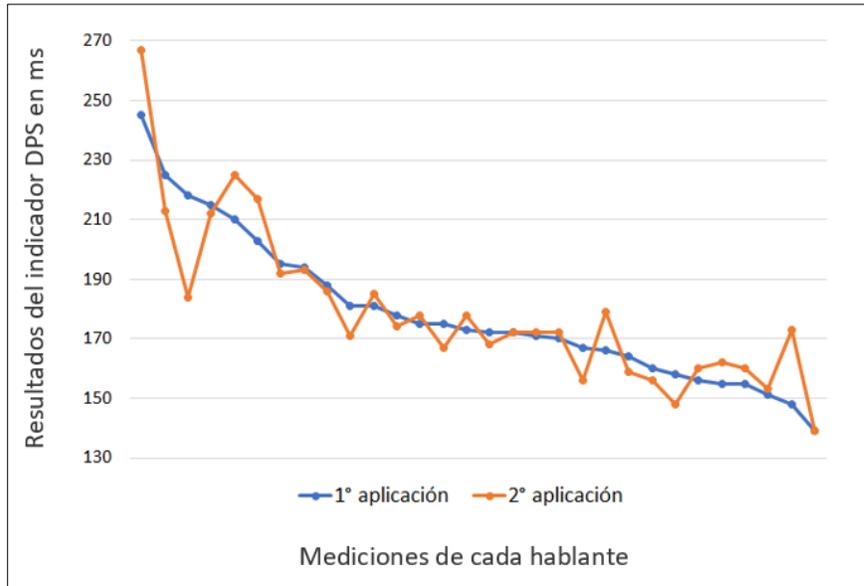


Figura 6.25 Gráfico comparativo de las mediciones de la DPS de dos emisiones distintas realizadas por el mismo sujeto

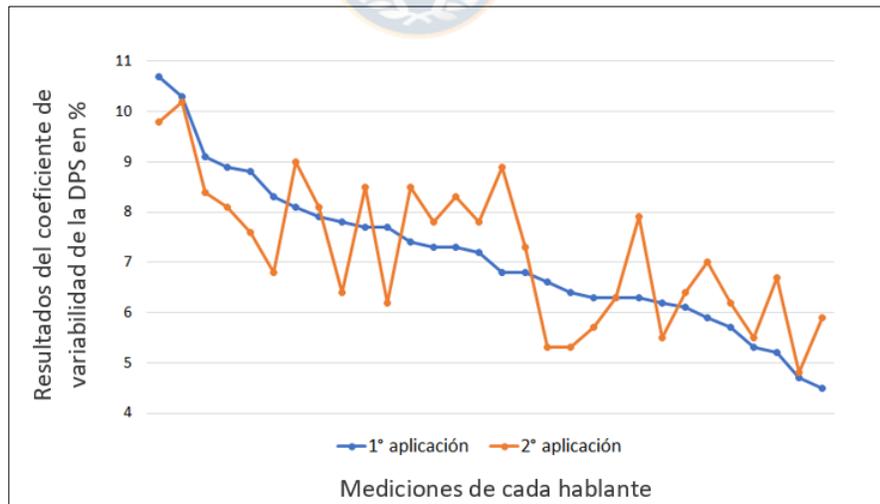


Figura 6.26. Gráfico comparativo de las mediciones del coeficiente de variabilidad de la DPS de dos emisiones distintas realizadas por el mismo sujeto

En el gráfico de la Figura 6.26, se presenta el detalle comparado de las mediciones del indicador coeficiente de variabilidad de la DPS para el mismo sujeto, pero de dos emisiones distintas. Se observa un comportamiento diferente de los datos en ambas mediciones, no obstante, ninguna de las diferencias entre las mediciones supera los 2 puntos porcentuales.

Para saber si la variable temporal incide en los resultados del análisis, los datos entregados por ambas emisiones fueron comparados mediante la prueba del Coeficiente de Correlación Interclase 2.

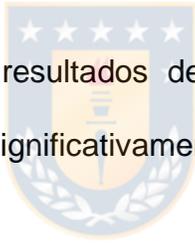
Tabla 6.16 Resultados del ICC de cada indicador.



Indicador	Coeficiente de correlación interclase	95% intervalo de confianza	
		Límite inferior	Límite superior
Promedio de sílabas por segundo	0,90	0,82	0,94
Coeficiente de sílabas por segundo	0,87	0,78	0,93
Promedio de duración silábica	0,91	0,84	0,95
Coeficiente de variabilidad del promedio de duración silábica	0,76	0,59	0,86

En la Tabla 6.16 se muestran los resultados de la aplicación de la prueba ICC para evaluar la fiabilidad test-retest. Para interpretar los datos, se analiza la información del indicador, así como del intervalo de confianza (Koo y Lee, 2016). Los resultados del indicador coeficiente de variabilidad del promedio de la duración silábica, se interpretan como de “moderado” (entre 0,5 a 0,75) a “bueno” (entre 0,75 a 0,9). Los valores que se observan del indicador coeficiente de sílabas por segundo, el promedio de sílabas por segundo y el promedio de duración silábica se categorizan como “buenos” (entre 0,75 a 0,9) a “excelentes” (mayor que 0,9).

Esto indica que los resultados del método de análisis automatizado mediante el *script* no varían significativamente dependiendo del momento en que se toma la muestra.



7. DISCUSIÓN

Los resultados de la aplicación del *script* sin corrección en comparación con las otras dos metodologías analizadas solo presentaron diferencias estadísticamente significativas en uno de los cuatro indicadores, el coeficiente de variabilidad de la DPS. Los otros tres indicadores no presentan diferencias estadísticamente significativas entre los métodos de análisis utilizados. Resulta llamativo, eso sí, que solo un indicador de los cuatro presente estos resultados, esto porque es bastante plausible suponer a priori y esperar que un método de análisis completamente automatizado de una muestra de habla debería presentar una alta probabilidad de error en una tarea de segmentación silábica automatizada, debido a la gran variabilidad inherente en este tipo de fenómenos.

Sin embargo, cabe destacar que, en rigor, la corrección humana de la segmentación automatizada se tuvo que realizar sólo en 5 de los 30 archivos de la muestra, lo que corresponde a un 16,6% del total. Para detallar la intervención, se presenta la figura 7.27 en donde se muestra la comparación entre ambos resultados. El *tier 2* muestra la corrección manual efectuada y, el *tier 3*, el resultado de la segmentación automática del *script* utilizando el umbral medio. Como se puede apreciar, la corrección se realizó para segmentar las sílabas que el *script* no detectó, aumentando de 31 (*tier 2*) a 34 unidades (*tier 1*), con el respectivo valor de duración. En este ejemplo, la intensidad de la producción

(dadas las características de la metodología de recolección) varía drásticamente entre las vocalizaciones, por lo tanto, el *script* configurado en el umbral medio no es capaz de identificar estas variaciones de intensidad. No obstante, cabe mencionar, que se hicieron otras pruebas de manera informal y se observó que al cambiar del umbral medio al umbral alto (tier 1), el *script* logra segmentar correctamente de forma completamente automatizada todas las sílabas, sin necesidad de una corrección humana.

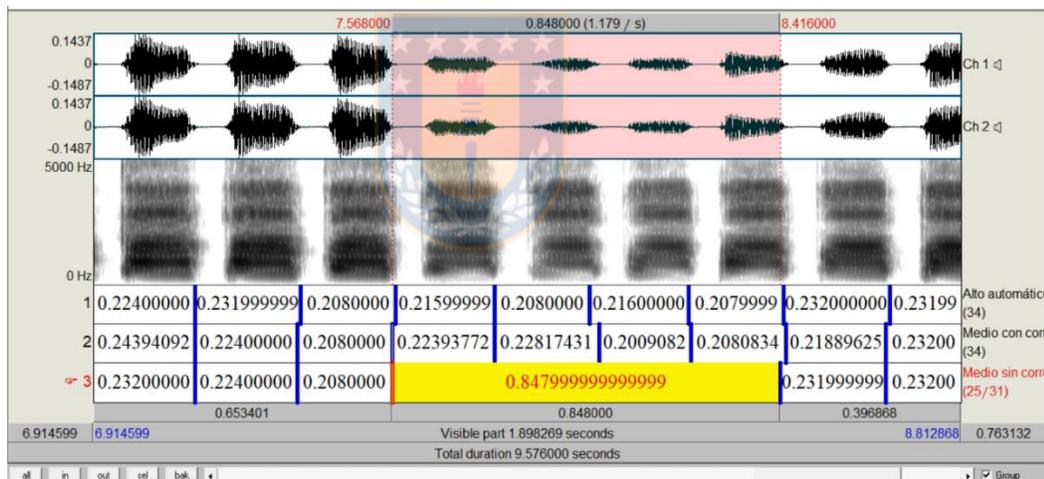


Figura 7.27 Comparación entre la segmentación automática del *script* y la segmentación con correcciones: Intensidad

El efecto que la variación de la intensidad tiene en el proceso de segmentación automatizada que realiza el *script* al utilizar el umbral medio se repite en tres de los cinco archivos. La segunda muestra de habla con correcciones presenta una contabilización final del *script* sin corrección de 55

sílabas en comparación a las 57 que se lograron con la corrección no automatizada. El último de ellos aumenta de 49 a 54 las producciones.

En la figura 7.28, se muestra el desempeño de la segmentación automática del *script* utilizando el umbral medio cuando la muestra de habla presenta una relación señal-ruido bajo los 10 dB en algunos puntos. Como consecuencia de esto, el *script* no logra detectar las variaciones que le permiten individualizar las producciones silábicas. En la figura, se muestra el caso de un archivo de audio en el que el proceso automático no logra segmentar la primera sílaba, contabilizando 56 (tier 3) unidades, en comparación con las 57 (tier 2) que se obtienen con una corrección manual. El segundo archivo que presenta características similares aumenta de 42 sílabas contabilizadas por el *script* sin corrección a 43 producciones cuando se realiza una corrección no automatizada de la segmentación.

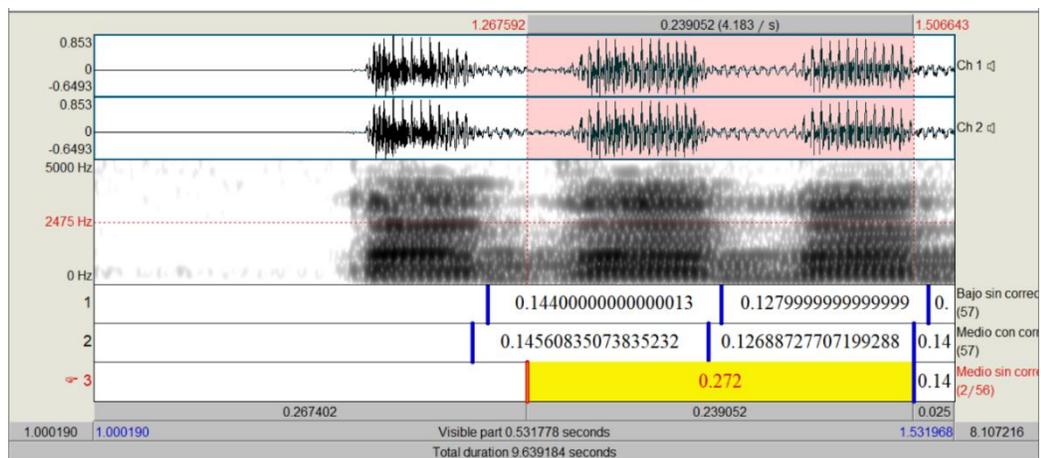


Figura 7.28. Comparación entre la segmentación automática del *script* y la segmentación con correcciones: Relación señal-ruido

Cabe destacar que, cuando la segmentación automática con un umbral medio no logra su cometido frente a un archivo con una relación señal-ruido bajo los 10 dB, es posible, gracias al diseño de este *script*, seleccionar un umbral bajo, así el *script* logra segmentar automáticamente sin errores todas las sílabas de la muestra de habla (tier 1).

Por estas razones, desde su origen, esta herramienta fue pensada, diseñada y está implementada para utilizarse considerando siempre la posibilidad de que el analista pueda hacer alguna intervención o corrección manual posterior a la segmentación automatizada.

Entonces, la segmentación automática no siempre logra delimitar con precisión todas las sílabas presentes en una muestra de habla. Esto debido principalmente a dos factores propios de las condiciones acústicas de toma de las muestras y de almacenamiento y registro de los archivos digitales de audio, estos son, la intensidad de las producciones y la relación señal ruido. No obstante, en ambos casos, mediante pruebas informales que se hicieron durante el transcurso de este estudio, se descubrió que es posible utilizar la opción que tiene el *script* de cambiar el umbral de sensibilidad para mejorar la precisión de la segmentación automatizada, por ejemplo, utilizar el umbral bajo cuando el motivo sea la relación señal ruido o utilizar el umbral alto, si las dificultades provienen de una intensidad baja presente en la muestra de audio. Aún así, si esto no resulta existe la opción de realizar las correcciones no automatizadas.

Por otra parte, el MSP tampoco logró segmentar bien todas las sílabas de una de las muestras de habla, lo que influyó en sus resultados. Esto es coincidente con lo que se indica en el estudio realizado por Wang *et al.* (2009), quienes mencionan que en su investigación el 11% de las muestras de habla de personas sin alteración neurológica no pudieron ser ejecutadas por el programa, valor que asciende a un tercio del total de los participantes con disartria atáxica en otro estudio similar (Wang *et al.*, 2008). Por lo mismo, es importante destacar que una gran ventaja que tiene el *script* propuesto en este trabajo es que permite corregir manualmente (de manera opcional) la segmentación en el caso de muestras que resulten difíciles de analizar para el procedimiento automatizado.

Es necesario aclarar que para realizar las comparaciones estadísticas del segundo apartado de la validación del *script* entre las metodologías se decidió utilizar los datos que se obtienen del *script* al utilizar un umbral de sensibilidad medio y al hacer uso de la opción de permitir correcciones, ya que, después de todo este proceso de análisis del funcionamiento del *script*, se estima que es probable que esta sea la opción preferida por los eventuales usuarios de él.

En general, los resultados, tanto del *script* con correcciones implementado en este trabajo, como del análisis no automatizado y del que entrega el MSP presentan valores bastante similares y en algunos casos casi idénticos. Lo cual además es ratificado por los resultados del análisis estadístico inferencial. Los

tres métodos no muestran diferencias estadísticamente significativas entre sus resultados.

En particular, el comportamiento que se observa del *script* con correcciones, en comparación con los resultados que se obtuvieron del análisis no automatizado, son altamente similares para cada uno de los cuatro indicadores estudiados. Los datos que provienen de la estadística descriptiva relativos al promedio, los valores mínimos y máximos, así como los de la desviación estándar presentan mínimas diferencias entre ambas metodologías.

Es llamativo que los indicadores que entregan información promedio presentan valores con mayor semejanza entre las metodologías que los resultados de los indicadores de variabilidad. Esto se replica tanto para los resultados del *script* con y sin correcciones.

De la misma manera, si se comparan los resultados que se obtienen utilizando el *script* con correcciones propuesto con los que resultan del análisis que realiza el MSP, se puede ver que ambas metodologías presentan un comportamiento relativamente similar. La mayor diferencia se observa en el indicador coeficiente de variabilidad de la duración silábica, pues el MSP no logró segmentar adecuadamente una de las muestras, lo que alteró en general los resultados del indicador. No obstante, cuando se realizó un procedimiento para remover el valor atípico y se volvieron a comparar los datos, se observa una mayor similitud entre de los resultados de ambas metodologías.

Por otra parte, se debe comentar que los datos obtenidos del análisis inter evaluador presentan una alta correlación para los indicadores promedio de duración silábica y promedio de sílabas por segundo, pero una correlación catalogada como moderada para el coeficiente del promedio de sílabas por segundo y el coeficiente de la duración silábica. Es llamativo que los coeficientes de variabilidad presenten un comportamiento sistemáticamente más variable que los indicadores que entregan valores promedio. No obstante, esto resulta bastante comprensible, cuando al analizar detalladamente la segmentación que hizo cada analista, se pueden ver pequeñas diferencias al decidir el punto de ubicación de los límites. Estas diferencias de apreciación no afectan tanto a las medidas de tendencia central, pero las medidas de dispersión, por su naturaleza, sí son más sensibles a estas leves diferencias. Es precisamente por ello que otra de las razones por las que se justifica la necesidad de diseñar e implementar un procedimiento automatizado como el propuesto en esta investigación es justamente neutralizar la natural variabilidad que puede presentar la medición realizada por humanos. Los análisis realizados manualmente por dos personas distintas están sujetas a la experiencia, al conocimiento y al criterio de los evaluadores, razón por la cual, resulta coherente que existan algunas diferencias entre lo realizado por dos o más evaluadores. Sin embargo, al utilizar un método objetivo, como un análisis automatizado, disminuye la variabilidad atribuible al factor personal individual.

La automatización del cálculo de los indicadores diadococinéticos se ha realizado con distintas plataformas y/o programas. Varios autores han propuesto modelos y herramientas que son capaces de realizar este procedimiento de forma automática. Por ejemplo, la bibliografía revisada menciona un *script* de Praat (Diepeveen *et al.*, 2019) que logra segmentar las sílabas de producciones diadococinéticas. Lamentablemente, no fue posible encontrar el código de este *script* ni tampoco los valores de referencia de su rendimiento. Además, según los antecedentes obtenidos, este *script* solo segmenta, no realiza cálculos de indicadores. Los autores solo documentan los resultados de la tasa de DDK para su análisis, sin mencionar el resultado del *script*. Hubiese sido interesante poder comparar el rendimiento este instrumento con el propuesto en esta investigación y poder evaluar así la eficacia de ambos para el proceso de segmentación.

Sin embargo, se puede encontrar otros casos de instrumentos de cálculo automatizado como, por ejemplo, una investigación realizada por Rozenstoks *et al.* (2020) basada en un modelo neural de detección de sílabas que, según sus autores, aporta una alta sensibilidad para la segmentación de estas unidades. En su estudio compararon el rendimiento de algoritmo con un análisis manual, cuyos resultados dan cuenta de una correlación del 98,9% en sujetos sin trastorno motor. De la misma forma, la correlación de la tasa diadococinética entre estas dos metodologías (manual y con el algoritmo) estudiadas en esa investigación presentan valores de correlación de un 0,94 (con un valor de significancia de $p > 0,005$). El *script* propuesto en este estudio presenta un valor similar de

significancia entre las metodologías al de Rozenstoks. Sus diferencias radican en la tecnología base para implementar el algoritmo y la cobertura o difusión que se puede desprender de ambos.

El *script* presentado provee de cuatro indicadores que entregan información que se relaciona con distintos factores. El más conocido y utilizado es aquel que se enfoca en la cantidad de producciones por unidad de tiempo (la tasa de diadococinesia), el cual, se puede medir y obtener incluso de forma manual, analógica y no automatizada, como se ha hecho tradicionalmente por muchos años. Sin embargo, la tasa provee de información sobre el estado general del control motor oral, sin indicar otras características como podría ser la ritmicidad o la regularidad de la producción, la precisión articuladora, la localización de las producciones en la unidad de tiempo, así como a su variabilidad. El *script* que aquí se propone, al incorporar más indicadores que solo la tasa, agrega información valiosa para el terapeuta o investigador sobre las características de su paciente y puede contribuir en la dimensión diagnóstica como la de seguimiento (Wang *et al.*, 2020).

El *script* está escrito en una plataforma de código abierto ampliamente difundida y utilizada por diversos profesionales e investigadores. Dadas las características del programa es posible que este *script* sea utilizado por distintas personas de forma libre, sin que por ello los usuarios tengan que realizar algún tipo de pago. Esto posibilita una cobertura que abarca a un amplio sector de

profesionales e investigadores, sin distinción. Además, por tratarse de una aplicación de código abierto queda a disposición de la comunidad para futuras y eventuales modificaciones según los requerimientos y la experiencia de aquellos que se interesen en mejorar este código.



8. CONCLUSIONES

Considerando los objetivos propuestos en esta investigación, se ha logrado diseñar e implementar con éxito un *script* de Praat capaz de segmentar de manera automatizada y semiautomatizada muestras de habla que contienen producciones diadococinéticas. Además, el *script*, a partir de dicha segmentación, calcula y entrega automáticamente cuatro indicadores cuantitativos del rendimiento diadococinético: a) la tradicional *tasa de diadococinesia* o promedio de sílabas por segundo, b) su correspondiente coeficiente de variabilidad, c) la duración promedio de la sílaba y también d) su correspondiente coeficiente de variabilidad.

Con el propósito de validar los resultados que el *script* entrega, este se aplicó a una muestra de habla de una población de adultos chilenos sin trastorno motor. Los resultados fueron comparados con otras metodologías que tienen el mismo fin y que se aplicaron a la misma muestra.

A partir de esta comparación se concluye que el análisis automatizado que realiza el *script* sin correcciones entrega resultados estadísticamente diferentes en uno de los cuatro indicadores en comparación a los que entrega un análisis similar realizado de forma no automatizada por un humano y a los resultados de un análisis similar que realiza automáticamente un programa comercial de alto costo monetario. Sin embargo, al hacer uso de la corrección no automatizada que

el mismo *script* presenta, se pueden obtener resultados sin diferencias estadísticamente significativas con otras metodologías, por lo cual, se recomienda efectuar un análisis visual de la segmentación, así como una corrección no automatizada si es necesario para lograr los resultados más precisos, o bien, cambiar el umbral de sensibilidad dependiendo de la calidad de la muestra de habla.



9. PROYECCIONES

Proyecciones que se desprenden de esta investigación se relacionan con la población estudiada y, si ya está claro que, si el *script* cumple con su objetivo, ahora resulta evidente que se debe determinar su comportamiento con personas que presenten alteraciones motoras del habla y que presenten, además, diversas características sociodemográficas y se ubiquen también en diferentes rangos etarios.

Otra de las proyecciones puede consistir en probar ampliar el espectro de estímulos analizables por el *script*. Entre ellos, se pueden considerar otras unidades consonánticas áfonas como, por ejemplo, /t/, /k/, o bien, sonoras como /b/, /d/, /g/. De la misma forma, es plausible diseñar un *script* que detecte unidades multi silábicas como /pa.ta.ka/ o /ba.da.ga/ o que considere también el comportamiento y variabilidad de las transiciones del segundo formante en la articulación de las vocales cerradas.

10. REFERENCIAS BIBLIOGRÁFICAS

- Ackermann, H., Hertrich, I. y Hehr, T. (1995). Oral diadochokinesis in neurological dysarthrias. *Folia phoniatica et logopaedica*, 47(1), 15-23.
- Albright, R. W. (1948). The motor abilities of speakers with good and poor articulation*. *Speech Monographs*, 15(2), 164-172.
<https://doi.org/10.1080/03637754809374954>
- Alshahwan, M. I., Cowell, P. E. y Whiteside, S. P. (2020). Diadochokinetic rate in Saudi and Bahraini Arabic speakers: Dialect and the influence of syllable type. *Saudi J Biol Sci*, 27(1), 303-308. doi:10.1016/j.sjbs.2019.09.021
- Boersma, P. y Weenink, D. (2020): Praat: doing phonetics by computer (v 6.1.16) [Computer program]
- Çiyiltepe, M. M. y Çifçi, B. (2019). Comparing diadochokinetic performances of stuttering and non-stuttering children between 7-12 ages. *Hearing Balance and Communication* <https://doi.org/10.1080/21695717.2019.1644859>
- Diepeveen, S., van Haften, L., Terband, H., de Swart, B. y Maassen, B. (2019). A Standardized Protocol for Maximum Repetition Rate Assessment in Children. *Folia Phoniatr Logop*, 71(5-6), 238-250. <https://doi.org/10.1159/000500305>
- Duffy, J. R. (2013). Motor speech disorders-E-Book: Substrates, differential diagnosis, and management. Elsevier Health Sciences.
- Fletcher, S. (1972). Time-by-count measurement of diadochokinetic syllable rate. *Journal of Speech and Hearing Research*, 15, 763-770.
- Gadesmann, M. y Miller, N. (2008). Reliability of speech diadochokinetic test measurement. *Int J Lang Commun Disord*, 43(1), 41-54.
<https://doi.org/10.1080/13682820701234444>
- John, J., Ganapathy, K., John, S. y Rajashekhar, B. (2014). Normative for Motor Speech Profile in Kannada-Speaking Adults. *Journal of Voice*, 28, 7-13.

- Karlsson, F., Schalling, E., Laakso, K., Johansson, K. y Hartelius, L. (2020). Assessment of speech impairment in patients with Parkinson's disease from acoustic quantifications of oral diadochokinetic sequences. *The Journal of the Acoustical Society of America*, 147(2), 839-851.
- Koo, T. K. y Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine*, 15(2), 155-163.
- Lundeen, D. (1950). The relationship of diadochokinesis to various speech sounds. *Journal of Speech and Hearing Disorders*, 15, 54-59.
<https://doi.org/10.1044/jshd.1501.54>
- Montana, D., Campos-Roca, Y. y Perez, C. J. (2018). A Diadochokinesis-based expert system considering articulatory features of plosive consonants for early detection of Parkinson's disease. *Comput Methods Programs Biomed*, 154, 89-97. <https://doi.org/10.1016/j.cmpb.2017.11.010>
- Pérez, H., Fernández, R. y Oliva, C. (2015). Valores normativos para el programa de análisis acústico del habla Motor Speech Profile en hablantes de español de Chile. *Revista Chilena de Fonoaudiología*, 14, 4-14.
- Rozenstoks, K., Novotny, M., Horakova, D. y Rusz, J. (2020). Automated Assessment of Oral Diadochokinesis in Multiple Sclerosis Using a Neural Network Approach: Effect of Different Syllable Repetition Paradigms. *IEEE Trans Neural Syst Rehabil Eng*, 28(1), 32-41. <https://doi.org/10.1109/TNSRE.2019.2943064>
- Toledo, L., Bahamonde, C., González, J., Martínez, M., Muñoz, M. y Muñoz, D. (2011). Parámetros del habla en adultos normales chilenos. *Revista Chilena de Fonoaudiología*, 10, 33-43.
- Wang, Y. T., Gao, K., Kloepper, A. M., Zhao, Y., Kuruvilla-Dugdale, M., Lever, T. E. y Bunyak, F. (2019). *DeepDDK: A Deep Learning based Oral-Diadochokinesis Analysis Software* 2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI),
- Wang, Y. T., Kent, R. D., Duffy, J. R. y Thomas, J. E. (2009). Analysis of diadochokinesis in ataxic dysarthria using the motor speech profile program. *Folia Phoniatr Logop*, 61(1), 1-11. <https://doi.org/10.1159/000184539>
- Wang, Y. T., Kent, R. D., Duffy, J. R., Thomas, J. E. y Weismer, G. (2004). Alternating motion rate as an index of speech motor disorder in traumatic brain injury. *Clin Linguist Phon*, 18(1), 57-84. <https://doi.org/10.1080/02699200310001596160>

Ziegler, W. (2002). Task-related factors in oral motor control: Speech and oral diadochokinesis in dysarthria and apraxia of speech. *Brain and language*, 80(3), 556-575

