



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

Frecuencia fonemática del mapudungun



Tesis para optar al grado de Magíster en Lingüística Aplicada

ALDO PATRICIO BERRÍOS CASTILLO
CONCEPCIÓN-CHILE
2023

Profesor Guía: Dr. Gastón Salamanca Gutiérrez
Departamento de Español, Facultad de Humanidades y Arte
Universidad de Concepción



© 2023 Aldo Berríos Castillo

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

AGRADECIMIENTOS

Agradezco a la dirección y comité del Programa de Magíster en Lingüística Aplicada, así como a la dirección de Postgrado de la Universidad de Concepción, por concederme el honor de estudiar bajo el programa de becas de postgrado. Extiendo, de manera especial, mis agradecimientos al profesor Dr. Gastón Salamanca Gutiérrez por alentarme a proseguir estudios de postgrado y por su guía irrestricta durante la generación, desarrollo y conclusión de este proyecto. De igual modo, agradezco el conocimiento y apoyo recibido del profesor Dr. Mauricio Figueroa Candia, que me han permitido avanzar en mi formación. También agradezco al cuerpo académico —profesores y profesoras— del programa de Magíster en Lingüística Aplicada por la dedicación y excelencia que propician el máximo desarrollo de sus estudiantes. Agradezco también a mi compañera, María Jesús, por su apoyo y guía durante todas las etapas de este proceso.

Ka femngechi mañüm piafiñ tañi pu afkadinietew feytachi rüpi tañi witrañpüramngetuael ta mapudungun fillke püle Wallmapu mew —tañi pu kompañ Viktor engu Alina, konkülelu tati Kimeltuwe pingechi trokiñche, ka ti pu peñi ka pu lamngen wenuntuniefilu ka kimeltukefilu ta Mapudungun Konse püle.



Tabla de Contenido

Índice de tablas.....	x
Índice de ilustraciones.....	xii
Resumen.....	xiv
Abstract.....	xv
1. Introducción.....	1
2. Fundamentos teórico-conceptuales.....	4
2.1. Fonología y fonostadística.....	4
2.2. Frecuencia fonemática y lexicón mental.....	6
2.3. Lingüística computacional y frecuencia de fonemas.....	8
2.4. Tipología fonológica y morfológica.....	10
3. Estado del arte.....	14
3.1. Estudios de fonostadística en lenguas de alta vitalidad.....	14
3.2. Estudios y reportes de frecuencia fonémica en lenguas en peligro.....	20
3.3. Estudios y reportes de frecuencia fonémica en mapudungun.....	23
4. Pregunta y objetivos de investigación.....	30
4.1. Pregunta de investigación.....	30
4.2. Objetivos de investigación general y específicos.....	31
5. Metodología.....	32
5.1. Diseño investigativo.....	32
5.1.1. Alcance de la investigación.....	32
5.1.2. Unidad mínima de análisis y variables.....	33
6. Criterios para el filtrado de las unidades.....	35
6.1. Morfología sufijal.....	36
6.1.1. Sufijos derivativos y flexivos.....	37
6.1.1.1. -n inf.....	37
6.1.1.2. -tu vbr.....	40
6.1.1.3. -le est.....	41
6.1.1.4. -(ü)l caus.....	41
6.1.1.5. -(ü)w refl.....	42
6.1.1.6. -nge vbr.....	42
6.1.1.7. -ma apl.....	43

6.1.1.8. -künu pers.....	43
6.1.1.9. -(ü)m caus.....	44
6.1.1.10. -tüku col.....	44
6.1.1.11. -ka cont.....	45
6.1.1.12. -we loc.....	45
6.1.1.13. Otros sufijos de alta, mediana y baja productividad.....	46
6.1.2. Sufijos improductivos: -dü, -kü, -tu.....	48
6.1.2.1. Sufijo improductivo -dü.....	48
6.1.2.2. Sufijo improductivo -kü.....	50
6.1.2.2.1. Seudoafijo -kü unido a lexemas reduplicados.....	51
6.1.2.2.2. Seudoafijo -kü y lexemas que remiten a sonidos.....	51
6.1.2.2.3. Seudoafijo -kü y otros casos.....	52
6.1.2.3. Sufijo improductivo -tu.....	52
6.2. Composición nominal.....	54
6.3. Extranjerismos.....	56
6.3.1. Cognados con el quechua.....	56
6.3.2. Cognados con el castellano.....	57
6.4. Fluctuación y alternancia fonemática.....	58
6.4.1. Introducción.....	58
6.4.2. Alternancias fonemáticas.....	62
6.4.2.1. Vocales.....	62
6.4.2.1.1. Alternancia /ə/ > /i/.....	62
6.4.2.1.2. Alternancia /ə/ > /u/.....	63
6.4.2.1.3. Alternancia /ə/ > /e/.....	63
6.4.2.1.4. Alternancia /ə/ > /a/.....	63
6.4.2.1.5. Alternancia /ə/ > /o/.....	64
6.4.2.2. Exploración del efecto del criterio RIV y alternativas.....	64
6.4.2.3. Consonantes.....	65
6.4.2.3.1. Alternancia /k/ > /l/, /ɲ/, /ŋ/ (en contexto asimilatorio), Ø.....	65
6.4.2.3.2. Alternancia /l/ > /l/, /n/.....	66
6.4.2.3.3. Alternancia /ɲ/ > /n/, /ɲ/, /l/.....	66
6.4.2.3.4. Alternancia /ŋ/ > /f/, /m/, /n/, /ɲ/, /p/, /j/, Ø.....	66

6.4.2.3.5. Alternancia $\widehat{tʃ}/ > \widehat{tʃ}/, /t/, /m/, /l/, /f/, \emptyset$	67
6.4.2.3.6. Alternancia $/\theta/ > /s/, \widehat{tʃ}/, /n/, /j/, /l/, /f/, /ɲ/, /p/, \emptyset$	67
6.4.2.3.7. Otras alternancias.....	68
6.4.2.3.8. Alternancia en el posmargen: $/ɲ/ > /n/, /m/, /j/$ (en contexto asimilatorio), \emptyset	68
6.4.2.3.9. Alternancia en el posmargen: $/f/ > /m/, /l/, \emptyset$	69
6.4.2.3.10. Alternancia en el posmargen: $/w/ > /w/, /l/$	69
6.4.3. Fluctuación fonemática.....	69
6.4.3.1. Vocales.....	70
6.4.3.1.1. Fluctuación entre $/o/ \sim /u/$	70
6.4.3.1.2. Otras fluctuaciones vocálicas.....	71
6.4.3.2. Consonantes.....	71
6.4.3.2.1. Fluctuación $/l/ \sim /ʎ/ \sim /θ/$	71
6.4.3.2.2. Fluctuación $/ɲ/ \sim /ʃ/ \sim \emptyset$	71
6.4.3.2.3. Fluctuación $/θ/$	72
6.4.3.2.3.1. Fluctuación $/θ/ \sim /ɲ/$	72
6.4.3.2.3.2. Fluctuación $/θ/ \sim /ʃ/$	72
6.4.3.2.3.3. Fluctuación $/θ/ \sim /f/ \sim /ʎ/$	72
6.4.3.2.4. Fluctuación $\widehat{tʃ}/ \sim /t/$	73
6.4.3.2.5. Otras fluctuaciones entre fonemas consonánticos.....	73
6.4.3.2.6. Otras fluctuaciones con elisión o adición de uno o varios segmentos.....	73
6.4.3.2.7. Fluctuaciones léxicas de patrones fonológicos mayores.....	74
6.5. Metátesis.....	74
6.5.1. Metátesis de $/k/ \sim /θ/$	74
6.5.2. Metátesis de $/ɲ/ \sim /w/, /u/$	74
6.5.3. Metátesis de $/ə/ \sim /e/, /i/, /w/$	75
6.5.4. Metátesis de $/w/ \sim /l/$	75
6.5.5. Metátesis silábicas con fluctuación fonemática.....	75
6.6. Reiteración fonológica y reduplicación.....	75
6.7. Interjecciones.....	76
6.8. Falta de definición.....	77
7. Análisis y Resultados.....	78

7.1. Introducción.....	78
7.2. Carga de datos y preparación preliminar.....	79
7.3. Procesamiento y transformación.....	81
7.3.1. Fonemización automática.....	81
7.3.2. Creación de tabla de fonemas.....	83
7.3.3. Conteo de los fonemas.....	84
7.4. Visualización.....	86
7.5. Resultados.....	87
7.5.1. Cuadro fonológico.....	87
7.5.2. Preparación preliminar.....	88
7.6. Corpus lexicográfico.....	89
7.7. Corpus fonológico.....	91
7.8. Determinación de proporciones.....	93
7.8.1. Ratio consonante/vocal.....	93
7.8.2. Ratio segmentos resonantes/no-resonantes.....	97
7.8.3. Ratio segmentos esperados y observados.....	100
7.8.4. Ratio segmentos sonoros y áfonos.....	101
7.9. Recuento estadístico por clase natural.....	103
7.9.1. Vocales.....	103
7.9.1.1. Dirección de la lengua.....	104
7.9.1.2. Altura de la lengua.....	106
7.9.2. Consonantes.....	109
7.9.2.1. Consonantes por lugar de articulación.....	110
7.9.2.2. Consonantes por modo de articulación.....	113
7.10. Clasificación en categorías.....	117
8. Discusión.....	122
8.1. Resumen de los hallazgos.....	123
8.2. Comparación con otros estudios de frecuencia fonemática.....	128
8.2.1. Comparación con Márquez (2016).....	129
8.2.2. Comparación con Lagos (2001).....	133
8.2.3. Comparación con Álvarez-Santullano (1986).....	135
8.3. Frecuencia de /ə/.....	137

8.4. Criterios de filtrado.....	142
9. Conclusiones y proyecciones.....	145
10. Referencias.....	148
Anexos.....	161
1. Anexo sobre sufijos.....	161
1.1. Listado de sufijos identificados.....	161
1.2. Formas alternantes con la vocal /ə/ en el radical verbal.....	166
1.2.1. Formas seleccionadas sin la vocal /ə/.....	166
1.2.2. Formas seleccionadas que contienen la vocal /ə/.....	169
2. Anexos sobre préstamos.....	170
2.1. Préstamos del castellano.....	170
2.2. Cognados del quechua.....	174
3. Anexos sobre alternancia y fluctuación fonemática.....	176
3.1. Alternancia /ə/ > /i/.....	176
3.2. Alternancia /ə/ > /u/.....	178
3.3. Alternancia /ə/ > /e/.....	180
3.4. Alternancia /ə/ > /a/.....	180
3.5. Alternancia /ə/ > /o/.....	181
3.6. Alternancia /k/ > /l/, /ɲ/, /ŋ/ (en contexto asimilatorio), Ø.....	181
3.7. Alternancia /l/ > /l/, /n/.....	182
3.8. Alternancia /ɲ/ > /n/, /ɲ/, /l/.....	183
3.9. Alternancia /ŋ/ > /f/, /m/, /n/, /ɲ/, /p/, /j/, Ø.....	183
3.10. Alternancia /t͡s/ > /t͡ʃ/, /t/, /m/, /l/, /f/, Ø.....	184
3.11. Alternancia /θ/ > /s/, /t͡ʃ/, /n/, /j/, /l/, /f/, /ɲ/, /p/, Ø.....	186
3.12. Otras alternancias.....	187
3.13. Alternancia en el posmargen.....	188
3.13.1. Alternancia en el posmargen /ɲ/ > /n/, /m/, /ŋ/ (en contexto asimilatorio), Ø.....	188
3.13.2. Alternancia en el posmargen /f/ > /m/, /l/, Ø.....	189
3.13.3. Alternancia en el posmargen /w/ > /w/, /l/.....	189
3.14. Fluctuación /o/ ~ /u/.....	190
3.15. Otras fluctuaciones vocálicas.....	190

3.16. Fluctuación /l/ ~ /ʎ/ ~ /θ/.....	192
3.17. Fluctuación /x/ ~ /ʃ/.....	192
3.18. Fluctuación /θ/ ~ /x/.....	193
3.19. Fluctuación /θ/ ~ /ʃ/.....	193
3.20. Fluctuación /θ/ ~ /f/ (en posmargen), /ʎ/.....	195
3.21. Fluctuación /tʃ/ ~ /t/.....	195
3.22. Otras fluctuaciones entre fonemas consonánticos.....	195
3.23. Otras fluctuaciones con elisión y adición.....	197
3.24. Fluctuaciones léxicas de patrones fonológicos mayores.....	198
4. Reduplicaciones y reiteraciones de patrones fonológicos.....	202
4.1. Reduplicaciones.....	202
4.2. Entradas con un patrón fonológico que plasma reiteración de fonemas.....	207
5. Ejercicio que explica la ratio frecuencia observada / frecuencia esperada.....	209



Índice de tablas

Tabla 3.1: Comparación de las frecuencias absolutas (f. abs.) y relativas (f. rel.) de los trabajos que han reportado esta información respecto de los fonemas del mapudungun.	28
Tabla 6.1: Segmentación de los pronombres.	48
Tabla 6.2: Segmentos que son RIV, respecto de los que son RIT.	60
Tabla 6.3: Fluctuaciones fonemáticas.	61
Tabla 6.4: Número total de alternancias donde /ə/ se releva como RIV.	62
Tabla 6.5: Valores absolutos y relativos del efecto de aplicación del criterio RIV para las alternancias con /ə/.	64
Tabla 7.1: Extracto de 10 entradas seleccionadas al azar del corpus lexicográfico, creado por medio de la transformación de la planilla de datos proveniente de nuestra etapa de muestreo.	81
Tabla 7.2: Tabla de equivalencia de caracteres.	82
Tabla 7.3: Extracto de 15 entradas del corpus fonológico, a partir de cuatro entradas léxicas seleccionadas al azar.	84
Tabla 7.4: Vocales del mapudungun.	87
Tabla 7.5: Consonantes del mapudungun.	87
Tabla 7.6: Número de entradas según su estatus de aceptadas y no pertinentes.	88
Tabla 7.7: Estadígrafos de la longitud fonológica de los lexemas aceptados.	90
Tabla 7.8: Distribución de los lexemas de acuerdo a la clase de palabra a la que pertenecen.	90
Tabla 7.9: Jerarquía de los fonemas del mapudungun según sus frecuencias absolutas y relativas.	91
Tabla 7.10: Distribución de frecuencias absolutas, según modo de articulación, y posición en el lexema.	95
Tabla 7.11: Segmentos fonemáticos resonantes y no resonantes en mapudungun.	98
Tabla 7.12: Frecuencias absolutas y relativas de los segmentos fonemáticos resonantes y no resonantes en mapudungun.	98
Tabla 7.13: Ratio de frecuencias observadas y esperadas en el conteo general (“gen.”) de todos los segmentos, en las vocales (“voc.”), y en las consonantes (“cons.”).	101
Tabla 7.14: Jerarquía de los fonemas vocálicos.	103
Tabla 7.15: Frecuencias absolutas y relativas de vocales según rasgos articulatorios.	108

Tabla 7.16: Jerarquía según las frecuencias absolutas y relativas de las consonantes mapuche....	109
Tabla 7.17: Frecuencias absolutas y relativas de las consonantes según rasgos articulatorios.....	116
Tabla 7.18: Categorías de agrupamiento de los fonemas mapuche según su frecuencia (cuatro grupos).....	118
Tabla 7.19: Cinco categorías de agrupamiento de los fonemas mapuche según su frecuencia, comparado con la ratio de frecuencia observada/esperada.....	120
Tabla 7.20: Número de realizaciones y sus porcentajes por categoría de frecuencia de fonemas..	121
Tabla 8.1: Frecuencias fonemáticas de tres corpus de mapudungun, expresadas como la ratio de la frecuencia observada / frecuencia esperada (ratio Obs./Esp.).	128
Tabla 8.2: Comparación de las frecuencias de fonemas en el corpus del presente estudio y en el reportado en Márquez (2016), utilizando la ratio de frecuencia observada / esperada.	130
Tabla 8.3: Cinco categorías de agrupamiento de los fonemas mapuche según su frecuencia fonemática reportada en este estudio y en el de Márquez (2016).	132
Tabla 8.4: Comparación de las frecuencias de fonemas en el corpus del presente estudio y en el reportado en Lagos (2001).	134
Tabla 8.5: Comparación de las frecuencias de fonemas en el corpus de este estudio y en el reportado en Álvarez-Santullano (1986).	135
Tabla 8.6: Propuesta de representación del reemplazo del léxico por medio de préstamos y derivados gramaticales, y cómo podría afectar las distribuciones de las vocales en las unidades lexemáticas.....	140

Índice de ilustraciones

Gráfico 7.1: Frecuencia de las distintas longitudes fonológicas de los lexemas.	89
Gráfico 7.2: Frecuencias absolutas y relativas de los fonemas del mapudungun.	92
Gráfico 7.3: Frecuencia de los fonemas vocálicos y consonánticos en las posiciones inicial, media y final.	94
Gráfico 7.4: Distribución de las frecuencias fonemáticas en tres posiciones del lexema.	95
Gráfico 7.5: Distribución de los segmentos resonantes y no resonantes en la frecuencia fonemática.	99
Gráfico 7.6: Visualización de las frecuencias relativas de los seis fonemas del mapudungun.....	104
Gráfico 7.7: Frecuencias absolutas y relativas de las vocales del mapudungun según el parámetro articulatorio de dirección de la lengua.	105
Gráfico 7.8: Distribución de las frecuencias absolutas y relativas de las vocales, según altura....	107
Gráfico 7.9: Frecuencia relativa de las consonantes del mapudungun.	110
Gráfico 7.10: Frecuencias relativas de las consonantes del mapudungun según lugar de articulación.....	111
Gráfico 7.11: Distribución de las frecuencias relativas de las consonantes del mapudungun según lugar de articulación.	112
Gráfico 7.12: Frecuencias relativas de las consonantes del mapudungun según modo de articulación.	113
Gráfico 7.13: Distribución de las frecuencias relativas de las consonantes del mapudungun según modo de articulación.	114
Gráfico 7.14: Análisis de agrupamiento con cuatro centros.	118
Gráfico 7.15: Análisis de agrupamiento con cinco centros.	119
Gráfico 8.1: Representación de la configuración de los datos de frecuencias fonemática y rasgos de zona articulatoria junto con resonancia.	126
Gráfico 8.2: Comparación de las frecuencias, por medio de la ratio Obs./Esp., de nuestro estudio y las reportadas en Márquez (2016).....	129
Gráfico 8.3: Análisis de agrupamiento con cinco centros, según reporte de frecuencias en Márquez (2016).	131

Gráfico 8.4: Comparación de las frecuencias, por medio de la ratio Obs./Esp., de nuestro estudio y las reportadas en Lagos (2001).133

Gráfico 8.5: Comparación de las frecuencias, por medio de la ratio Obs./Esp., de nuestro estudio y las reportadas en Álvarez-Santullano (1986).136



Resumen

El propósito de este estudio es determinar la frecuencia fonemática del mapudungun. Para conseguir este objetivo, se creó un corpus electrónico que contiene unidades lexemáticas monomorfémicas, a partir de las 9703 entradas del *Diccionario mapudungún-español* de F. de Augusta (1916/2017), obra reconocida por su gran extensión, calidad y vigencia.

El muestreo de las unidades léxicas se realizó de manera exhaustiva en atención a criterios que aparecen en la literatura sobre fonestadística y al carácter aglutinante y polisintético de la morfología del mapudungun. Con este fin, se utilizó el lenguaje de programación R, para asegurar la replicabilidad y difusión de los resultados, y se consideraron variables articulatorias y cocientes que dan cuenta de la interacción de las unidades de análisis.

Los resultados dieron cuenta de una distribución típica para frecuencias fonemáticas, donde una gran proporción de las ocurrencias se adscribe a unos pocos fonemas. Los fonemas con mayores frecuencias fueron las vocales centrales y anteriores, y la consonante oclusiva velar áfona /k/. Entre las vocales, llamó la atención la altísima frecuencia de /ə/, seguida de /a/ y los otros fonemas vocálicos: /e/, /i/, /o/ y /u/. Entre las consonantes, destacó la alta frecuencia de /k/, seguida por la frecuencia moderada de /w/, /l/, /p/ y /m/. De baja frecuencia fonemática, fueron las consonantes /ŋ/, /ʎ/, /n/, /f/, /ʃ/, /j/, /ɰ/, /p/, /tʃ/ y /θ/. Por último, entre los fonemas de muy baja frecuencia se encuentran las interdental /l̪/, /n̪/ y /t̪/, junto con los fonemas /t/, /f/, /tʃ/ y /s/.

Además de caracterizar la carga funcional de los fonemas del mapudungun, y aportar con ello a los estudios sobre la fonología segmental de esta lengua, se espera contribuir a otras áreas del saber, como la tipología lingüística y los estudios de adquisición del mapudungun como primera lengua. De igual modo, la descripción de la metodología utilizada en este estudio espera ser una contribución para futuros trabajos que comparen lenguas en virtud de sus frecuencias fonemáticas.

Abstract

This work is aimed at establishing the phoneme frequency of Mapudungun (ISO: ARN, Southern Cone of South America). An electronic corpus with monomorphemic lexemes was prepared based on the 9703 entries in Augusta's *Diccionario mapudungún-español*, first published in 1916, a publication that is acknowledged by its breadth, quality and validity.

The lexical units were thoroughly sampled following literature-based criteria on phonostatistics and in view of the polysynthetic and agglutinative properties of Mapudungun. R programming language was used in order to ensure replicability of our results. Articulatory variables as well as indicators based on different quotients were considered in order to explore and describe the interaction between the units under analysis.

The results showed a typical distribution curve for phoneme frequencies, in which a large proportion of occurrences corresponded to a few phonemes. The most frequent phonemes were central and back vowels, as well as the voiceless velar stop consonant /k/. Among vowels, the high frequency of /ə/ was of remarkable note, followed by /a/ and the other vocalic segments: /e/, /i/, /o/ and /u/. Among consonants, /k/ was of high frequency, followed by the moderate frequency of /w/, /l/, /p/ and /m/. The consonants /ŋ/, /ʎ/, /n/, /f/, /tʃ/, /j/, /ɬ/, /ɲ/, /tʃ/ and /θ/ were of low frequencies. Finally, the interdental phonemes /l̪/, /ɲ̪/, /t̪/, together with /t/, /ʃ/, /tʃ/ and /s/ were the less frequent phonemes in our study.

In addition to describe the functional load of the phonemes in Mapudungun, which contributes to the phonological study of this language, this work could also contribute to other research areas, such as language typology, and acquisition of Mapudungun as a first language. Similarly, the methodology that is described in this study could contribute to future works that compare languages in terms of their phoneme frequencies.

*“Lo primero se ha de notar que hay en esta lengua una vocal demás de las cinco que tienen las lenguas española, y latina, [...] pronuncian un sonido medio entre la e y la u y **usan muy frecuentemente de él**”*

—Valdivia (1606, p. 1)

“So we should get back to work now realizing the magnitude of our undertaking. If we have to waste time making lists, OK, we’ll waste that time. But we do not want to waste time making the wrong lists because information incorrectly correlated can lead only to major undertakings that waste our time to no avail.”

—Lightner (1978, p.8)¹



1 Extraído de Lightner, T. M. (1978). Let’s hear it one more time on cluster frequencies. *Linguistics*, 16(211).

1. Introducción

¿Por qué realizar un estudio sobre frecuencia fonemática? Entre las motivaciones para un estudio de esta índole está, en primer lugar, la necesidad de contar con una **descripción lingüística** de una lengua en particular. En palabras de Tambovtzev (2010), cada idioma tiene su propio *sound picture* o *cuadro sonoro*, vale decir “la distribución de los sonidos del habla en la cadena sonora” (p. 99). En este sentido, la frecuencia de ocurrencia de fonos y fonemas se cuenta dentro de esta caracterización fonético-fonológica de una lengua específica, que permite dar cuenta del rendimiento funcional de este tipo de unidades (Alarcos, 1965).

De la mano con la descripción de una lengua específica, otras aplicaciones del estudio de frecuencia fonemática son las de índole **tipológica**. En efecto, en la búsqueda de universales lingüísticos, y más específicamente universales fonológicos, lingüistas como Zipf (1935), Trubetzkoy (1939), además de Greenberg (1966/2005), observaron que las unidades fonológicas caracterizadas como ‘no marcadas’ (*unmarked*, en inglés) tenían una alta frecuencia de ocurrencia. Esto se cumple en el conteo de fonemas del húngaro, finlandés y otras lenguas (Greenberg, 1966/2005, p. 18), respecto de rasgos como la activación de las cuerdas vocales, glotalización y aspiración, en el caso de las oclusivas, y entre consonantes orales y nasales, entre otros. Más recientemente, por medio de la frecuencia fonémica en listados léxicos de 6901 lenguas, Everett (2018) evidenció que, a pesar de la diversidad de los inventarios fonológicos, existe una convergencia en la frecuencia de uso de algunas consonantes particulares, estimulando efectivamente la aplicación del estudio fonostadístico en la tipología lingüística.

Además de estas aportaciones, el estudio de la frecuencia fonemática ha contribuido históricamente a algunas áreas de la **tecnología** en campos diversos como la criptografía, la disposición de teclados para máquinas de escribir, la generación de escrituras taquigráficas y para desarrollos tecnológicos como el fonógrafo y las telecomunicaciones. Ahora, si bien estos aportes se desplegaron fuera de la fonología, puesto que el conteo se realizaba a partir de unidades de escritura, fueron, sin duda, un antecedente de la lingüística considerada empírica, al unir áreas del estudio científico como la estadística con el estudio del lenguaje. Más relacionado

con los avances tecnológicos actuales, el estudio de la frecuencia fonemática puede apoyar la investigación en sistemas de reconocimiento automático del habla y síntesis de voz (Guirao y García, 1990; Nahar et al., 2012; Silber-Varod, Latin y Moyal, 2017).

Otra área donde existe una contribución del estudio de la frecuencia fonemática es la **adquisición del lenguaje**. Una teoría que desarrolla este vínculo es la llamada *Fonología del Input*, de Zamuner (2003), la cual vincula la frecuencia de sonidos en el idioma materno con la probabilidad de ser articulados por los niños: “*high frequency sounds in the ambient language are more likely to be produced by children than are low frequency sounds*” (p. 9). Como señala Cohen (2015), en un estudio sobre la adquisición del fonema fricativo uvular sonoro /ʁ/ del hebreo, “se ha postulado la noción de adquisición como dependiente de la frecuencia, lo que sugiere que mientras más frecuente sea un segmento en una determina posición, su adquisición es más rápida en ese contexto” (p. 8). Ahora, si bien este estudio en particular no confirmó que la frecuencia del fonema en determinados contextos se correlacione con el orden de adquisición, el interés por la frecuencia fonémica en ciertas posiciones dentro de la estructura léxica, como factor que pudiese influir en el orden de adquisición fonemática, refuerza la idea de que los estudios fonoestadísticos tienen un área de aplicación en la adquisición de la primera lengua. Cabe señalar que, aunque esta teoría es de formulación relativamente reciente, ya en 1946 Irwin y Chen estudiaron la adquisición del inglés en niños pequeños por un periodo de dos años y medio. El desarrollo de la frecuencia fonemática medida en las producciones de los infantes permitía medir el dominio de la producción de vocales y consonantes, encontrando que en los primeros 30 meses de vida los pequeños producen más a menudo sonidos vocálicos que consonánticos, y solo después de dos años y medio la frecuencia de ocurrencia de consonantes se aproxima a la de las vocales (p. 124).

Además de los ámbitos descritos, los estudios de frecuencia fonémica se vincularon tempranamente con la **enseñanza y aprendizaje de una segunda lengua**. En efecto, a principios del siglo XX, Dewey (1923) realizó un estudio estadístico del léxico, sílaba y sonidos del inglés, con el objetivo de aportar de manera práctica a la enseñanza de esta lengua, así como a la pedagogía de su escritura y la creación de materiales de estudio. Asimismo, en el prefacio de la

edición revisada del texto, señala que estuvo al tanto de que su estudio se utilizó en Japón en contextos pedagógicos de enseñanza del inglés, además de otros usos más relacionados con el desarrollo tecnológico de la época.

También ha habido una contribución del estudio de la frecuencia fonemática en **trastornos clínicos** como la afasia. En este contexto, autores como Stenneken et al. (2005), Robson et al. (2003) y Hernández et al. (2012) estudiaron la relación entre la frecuencia fonémica y silábica de una lengua, y los tipos de producciones de pacientes con afasia. Así, por ejemplo, Stenneken et al. (2005) encontraron que sus resultados son compatibles con lo reportado en la literatura respecto de este tipo de relación. En la misma línea, Robson et al. (2003), tras estudiar un tipo particular de afasia, concluyeron que existe una gran influencia de la frecuencia fonémica en los patrones de producción de esos pacientes.

Por último, aunque no agotando el campo de aplicaciones y contribuciones posibles, se puede mencionar la relación que se ha establecido entre la frecuencia fonemática y el **lexicón**; esto es, el repertorio léxico mental al que accede un hablante cuando hace uso de una lengua. En efecto, se ha planteado que los patrones fonotácticos más comunes pueden afectar la posibilidad de que la comunidad de hablantes establezca y mantenga las unidades léxicas (Martin, 2007), y que la frecuencia de ciertos patrones puede afectar la estructura fónica del léxico (Bybee, 2004).

En definitiva, en atención a la diversidad de aplicaciones que puede tener el análisis y reporte de la frecuencia fonemática de una lengua particular, se responde la pregunta que inicia esta sección y que motiva el presente estudio. A continuación, se desplegarán los fundamentos teórico-conceptuales necesarios para desarrollar una investigación de esta índole.

2. Fundamentos teórico-conceptuales

En este apartado, se presentan los fundamentos teórico-conceptuales que orientan el marco de referencia de esta investigación. Concretamente, se despliegan cuatro temáticas. En primer lugar, revisamos conceptos y teorías de la ciencia lingüística en los campos de la fonología y fonostadística. En segundo lugar, profundizamos en la relación entre frecuencia fonemática y acceso al lexicón mental. En tercer lugar, se abordan los avances en la tecnología computacional y el resurgimiento de la llamada Lingüística del Corpus, y cómo esto influye en un estudio sobre frecuencia de fonemas. Finalmente, nos referimos a la tipología fonológica y morfológica de la lengua mapuche.

2.1. Fonología y fonostadística

A Saussure (1945) le debemos haber definido un objeto de estudio para la lingüística, cristalizado en el concepto de *lengua*: “sistema que no conoce más que su orden propio y peculiar” (p. 50). Opuesto a *lengua*, está el concepto de *habla*, o ‘actos lingüísticos’, que Coseriu (1986) define como la “realidad concreta del lenguaje” (p. 16). Saussure ejemplifica esta oposición entre *lengua*, como sistema, y *habla*, como manifestación concreta, comparándola con la confección de un tapiz. Para entender cómo aparece la obra de arte textil en el telar, importa más la oposición cromática de los hilos, que “los procedimientos con que se han obtenido los colores” (p. 60). En definitiva, la lingüística se enfoca en la lengua como objeto de estudio, por medio de una abstracción de los aspectos concretos e individuales de los actos lingüísticos, con el fin de acceder al sistema lingüístico común (Coseriu, 1986).

En cuanto al estudio de los sonidos de las lenguas, esta distinción entre elementos concretos y otros funcionales tiene su manifestación más prominente en la separación entre unidades *fonéticas* y *fonológicas*, conceptos desarrollados exhaustivamente por la Escuela de Praga. De este modo, por un lado, los fonos, como unidades de estudio de la fonética, son manifestaciones de la cadena del habla medibles, almacenables y reproducibles por medios tecnológicos. Por otro lado, los fonemas son los elementos funcionales de la lengua que pueden considerarse como

unidades abstractas o psicológicas. En palabras de Zipf (1935), el fonema es “*the smallest unit of distinctive significance*” (p. 54), lo que enfatiza esta capacidad de los fonemas de permitir la distinción entre una palabra y otra en una lengua determinada. Es la unidad fonológica, entonces, la que genera un contraste en aquellos contextos donde el intercambio de un sonido por otro provoca el cambio de significado de una unidad léxica.

La introducción de los conceptos de la fonología en los estudios lingüísticos abrió posibilidades de estudio antes inimaginadas, al permitir establecer leyes y explicaciones a fenómenos como los cambios fonológicos sincrónicos. Otra repercusión ha sido el establecimiento de sistemas de escritura práctica para las lenguas, como el mapudungun². En palabras de Trubetzkoy (1931/2001), “*the purpose of a practical writing system is to reproduce not all the sounds that are spoken but only those oppositions with phonological value*” (p. 4). Bajo este criterio rector, desde la década de los años ochenta se propuso una serie de sistemas de transcripción fonémica para el mapudungun (Croese et al., 1978; Catrileo, 1984; Hernández, 1984), que se consolidaron finalmente en el sistema de escritura práctica conocido como Alfabeto Unificado para esta lengua (Hernández, 1986).

En este contexto, entonces, junto con establecer cuál es el repertorio de fonemas de una lengua —y, con ello, proveer los insumos necesarios para eventuales propuestas de grafemarios prácticos—, determinar la frecuencia de aparición de los fonemas constituye uno de los objetivos centrales de esta nueva disciplina —la fonología—, fundada por la Escuela de Praga (Obediente, 2007). Es lo que se conoce como fonoestadística. En efecto, en *Grundzüge der Phonologie* (1939, p. 256), Trubetzkoy introduce el término *Lautstatistiks* o fonoestadística, para aludir a la investigación cuantitativa de los fonemas y sus combinaciones en una determinada lengua. Según el lingüista ruso, esta disciplina se ocupa, por un lado, de registrar la frecuencia de ocurrencia de un elemento fonológico específico y, por otro lado, de relevar el rendimiento funcional de ese elemento en diferentes combinaciones.

2 Se ha indicado que el mapudungun no desarrolló de manera interna un “sistema de registro gráfico” (Hernández, 1984, p. 189), aseveración que podría ser matizada a partir de las consideraciones de Loncon (2017) respecto de las inscripciones gráficas portadoras de conocimiento mapuche en alfarería, textiles y joyas de plata.

En cuanto a la primera ocupación de la fonostadística, se ha evidenciado que el sistema cognitivo humano hace uso de frecuencias al clasificar y procesar categorías lingüísticas abstractas, como las fonológicas, tanto en el plano de la percepción, como en el de la producción (Frisch, 2015). Ahora bien, es importante precisar que el concepto de frecuencia podría entenderse también como el número de repeticiones de un fenómeno concreto por medida de tiempo; por ejemplo, el conteo de instancias de un elemento fonético durante un discurso de 5 minutos, que fuese analizado desde el punto de vista acústico; no obstante, en el caso del análisis de corpus como diccionarios, se cuentan las ocurrencias de cada unidad funcional (fonema), información que, además, puede analizarse estadísticamente.

En cuanto a la segunda ocupación de la fonostadística (la carga o rendimiento funcional de las oposiciones fonológicas), es importante registrar la frecuencia de fonemas en unidades lingüísticas mayores como la sílaba, morfemas y unidades léxicas. Como señala Frisch (2011), existen muchos tipos de frecuencias y dependiendo del dominio en que se realiza el conteo de ocurrencias, diferentes niveles de generalizaciones fonológicas son relevantes para su estudio (p. 2138-2139). En definitiva, los patrones fonológicos pueden examinarse en unidades mayores como grupos consonánticos sub-silábicos, sílabas, morfemas y lexemas. En nuestro caso, como hemos señalado, determinaremos la frecuencia fonémica en unidades monomorfémicas lexemáticas.

2.2. Frecuencia fonemática y lexicón mental

El lexicón es el diccionario mental de los hablantes. La frecuencia fonemática está vinculada de manera directa con la naturaleza de este bagaje léxico, pues las unidades que lo componen están constituidas, precisamente, por fonemas. En otras palabras, para comprender aspectos del fenómeno del acceso léxico es importante utilizar la categoría fonemática y su correspondiente cómputo. En esta línea, Stevens (2002) propone un modelo de acceso léxico que considera como parte de su diseño los segmentos con rasgos distintivos que permiten diferenciar las palabras que

tienen representación en ese lexicón, independiente del hecho de que esas unidades tienen correlatos acústicos y articulatorios.

Ahora, con respecto a cómo los hablantes almacenan las unidades léxicas en este repositorio mental, Haspelmath y Sims (2002) señalan que las palabras monomorfémicas (por ejemplo, morfemas léxicos sin flexión o derivación) se almacenan y acceden directamente desde el lexicón. Para las formas complejas derivadas o flexionadas y los correspondientes morfemas gramaticales que los constituyen, existen tres posibilidades respecto de cómo pueden ser parte de este repertorio mental.

La primera posibilidad es la del denominado ‘lexicón morfémico’, la cual sostiene que los hablantes memorizan solo morfemas, y ninguna palabra compleja. La información que recaba este lexicón es mínima y predecible. Las desventajas teóricas de una configuración morfémica para el lexicón radican en que no es posible segmentar algunos elementos como la flexión fusional, acrónimos, o incluso aquellos que no tienen una forma fonológica, como los morfemas cero o nulos. Otra desventaja mencionada por Haspelmath y Sims es que el significado de algunos lexemas derivados no es predecible a partir de los elementos constituyentes. En este sentido, y como se verá en el apartado *2.4. Tipología morfológica y fonológica*, es posible considerar como una ventaja la tipología aglutinante del mapudungun, pues permite en la mayor cantidad de los casos acceder de manera transparente al significado discreto de las unidades que componen un complejo morfémico.

El segundo tipo de lexicón especifica que las entradas se componen de formas de palabras, sean simples o complejas. Esta postura resuelve cómo se almacenarían derivados léxicos a partir de morfemas improductivos (para el mapudungun, véanse los mencionados por Smeets, 2008, p. 310), pero también provoca un desafío para los idiomas con morfología aglutinante y polisintética. Para estas lenguas, no es posible saber si los hablantes memorizan formas complejas polimorfémicas de manera completa o no. Otro desafío a este tipo de lexicón es que no diferencia entre morfemas léxicos y morfemas gramaticales, distinción que se comprueba en

los patrones de distribución de algunas unidades fonológicas que, por ejemplo, no aparecen en ciertas categorías morfológicas.

La tercera perspectiva señalada por los autores refiere a una postura intermedia entre el lexicon morfémico y el de formas de palabras, donde tanto lexemas complejos como morfemas simples pueden constituir entradas léxicas. En este lexicon, denominado ‘moderado’, el acceso léxico puede tomar dos rutas: la ruta *directa* y la *descomposicional*, lo que está determinado por un conocimiento subconsciente de los hablantes respecto de la frecuencia de uso de una palabra y otros parámetros. El lexicon de formas de palabras moderado puede ser una forma de conciliar las dificultades de las otras dos propuestas. Ahora bien, las implicaciones concretas de asumir alguna de estas perspectivas en la determinación de la frecuencia de fonemas en el mapudungun se despliegan en el apartado metodológico; específicamente, 5.2.3. *Filtrado*.

2.3. Lingüística computacional y frecuencia de fonemas

En la actualidad, el uso de corpus electrónicos ha permitido construir herramientas de análisis que han propiciado una expansión en la comprobación de hipótesis sobre las estructuras lingüísticas (Scholz et al., 2022). En efecto, la inclusión de herramientas estadísticas y computacionales para el desarrollo de la ciencia del lenguaje ha permitido retomar y ahondar en preguntas que tienen una larga trayectoria. De igual modo, ha permitido que idiomas con una baja representación en los círculos académicos sean explorados con diversas aplicaciones y renovadas perspectivas. Ejemplo de lo anterior son los trabajos de González (2016) con la lengua cabécar, el estudio fonoestadístico de Márquez (2016) sobre el mapudungun, y la investigación translingüística de Macklin-Corde y Round (2020) que evaluó modelos estadísticos de frecuencia fonológica utilizando un corpus de 166 lenguas australianas.

Así, entonces, la disponibilidad de recursos informáticos en las últimas décadas ha permitido el desarrollo de técnicas que utilizan el corpus como base para el desarrollo de descripciones estructurales de las lenguas naturales. Por su naturaleza, estas indagaciones se gestaron al alero de corrientes más asociadas al estructuralismo norteamericano de los años 1920 al 1950, cuya

metodología consiste, precisamente, en la comprobación de hipótesis de las estructuras lingüísticas por medio de corpus (Scholz et al., 2022). Como indica Sampson (2001) en la introducción a su *Empirical Linguistics*, a fin de profundizar nuestra comprensión del lenguaje, se deben aplicar técnicas empíricas a los aspectos observables del fenómeno de estudio. En esta misma línea, este investigador justifica el uso de corpus para el análisis lingüístico, por cuanto la intuición del hablante puede no corresponder a la realidad lingüística.

Para el conteo de frecuencia de fonemas, se pueden utilizar dos tipos principales de corpus: de *tipos* y de *muestras*. Esta distinción fue establecida por el filósofo Charles Peirce en su publicación de 1906 *Prolegomena to an Apology for Pragmaticism*. Respecto de cómo entender la proposición bajo un prisma pragmático, propone dividir la naturaleza de un signo entre una forma abstracta, que denomina *tipo* (*type* en inglés), y una forma más concreta, que denomina *muestra* (*token* en inglés). De acuerdo con esto, las *muestras* son las instancias de la forma *tipo* que ocurren efectivamente en un texto determinado. A modo de ejemplo, en la oración (1) del diccionario *Ta iñ Mapun Dungun*, de Antinao (2014), el *tipo* (de palabra) *chi* ‘el, la’ ocurre en la oración con dos *muestras*.

(1) *Münaüin chi kawellu ñi ütayam chi lelfün meu.*

‘Solté el caballo para que apaciente por el campo.’ (p. 180)

Siguiendo esta distinción, se ha propuesto desde épocas tempranas en la fonostadística que los estudios de frecuencia fonemática pueden realizarse a partir de dos tipos de corpus: corpus de textos continuos y entradas de diccionarios (Trubetzkoy, 1939). Las motivaciones para utilizar unos u otros insumos son que la estadística en fonología, por un lado, “debe mostrar la frecuencia de ocurrencia de un elemento fonológico específico de una lengua dada...” (p. 259) y, por otro lado, debe mostrar la “carga o rendimiento funcional de tal elemento o de una oposición fonológica específica” (p. 259). Así, para la primera tarea, se examinan textos continuos; y para la segunda, se hace uso de diccionarios. Es decir, se utilizan corpus de *muestras* y de *tipos*, respectivamente.

2.4. Tipología fonológica y morfológica

La *tipología fonológica* se puede definir como el estudio de la distribución y comportamiento de los sonidos en las lenguas humanas, de modo que se ocupa de la variación y frecuencia de patrones fonológicos entre ellas (Gordon, 2016). Una temática central en lo que al ámbito tipológico se refiere son los llamados *universales fonológicos*. Para Trubetzkoy (1931/2001), en efecto, “la búsqueda de leyes generales es una característica de todas las ramas de la fonología” (p. 3). En este contexto, una forma de abordar estos universales fonológicos es mediante el estudio de la complejidad fonológica de las lenguas del mundo, lo que se puede realizar tanto a través de la comparación de los tamaños de los inventarios fonémicos (Maddieson, 2013a, 2013b), así como mediante la determinación de si las eventuales diferencias en estos inventarios tiene repercusiones en otras características de una lengua.

Así, a partir de una muestra de 209 lenguas, entre las que se incluye el mapudungun, Crothers (1978) especifica los tipos de sistemas vocálicos más naturales en las lenguas del mundo, reportando que los sistemas pentavocálicos, hexavocálicos y trivocálicos son los más comunes. En esta misma investigación, Crothers desarrolla un sistema de notación que permite, por una parte, agrupar las lenguas en distintos perfiles (tipologías), y, por otra, definir un parámetro tipológico. Dicho sistema de notación consiste en dos cifras separadas por el signo ‘:’, donde la primera cifra corresponde al número total de vocales en el sistema, y la segunda al número de vocales que el autor define como *interiores* o *centralizadas*, en oposición a las vocales *periféricas*, es decir, /i, e, a, o, u/ (p. 100). En el caso del mapudungun, entonces, la notación tipológica que corresponderá es 6:1, puesto que tiene seis segmentos que componen el total de su sistema vocálico y uno de ellos es de índole interior (/ə/). Este ordenamiento es el segundo más común, con una frecuencia relativa del 13,9% de las lenguas estudiadas por Crothers, y es superado solamente por el sistema pentavocálico sin vocales interiores (notación: 5:0), el cual tiene una frecuencia relativa de un 26,3%.

Algunos de los universales propuestos por Crothers (1978) respecto de las vocales, y que interesan al presente estudio, son: que todas las lenguas tienen /i/, /a/ y /u/ (p. 115), que las

lenguas con seis o más vocales tienen /ɔ/ y además /i/ o /e/ (p. 116), que el contraste entre cinco cualidades vocálicas básicas es la norma para las lenguas humanas, y que, en general, los sistemas más comunes son aquellos que se acercan a este número de vocales básicas (p. 117).

A diferencia de las tendencias vocálicas universales halladas en la muestra estudiada por Crothers, las tendencias en el caso de las consonantes no se presentan con la misma claridad, debido a la mayor complejidad que tienen los inventarios de este tipo de segmentos respecto de los sistemas vocálicos, tal como señala Burquest (1993). Dicho esto, una generalización prominente que este autor, entre otros, plantea a partir de sus propias investigaciones es que en los inventarios consonánticos son esperables los fonemas oclusivos sordos, nasales y fricativos (como /s/ y /h/); y añade que pequeñas variaciones son posibles en algunas lenguas.

Gordon (2016) señala que los universales fonológicos pueden ser vulnerables a la refutación, lo que hace necesario indagar, en su lugar, en patrones estadísticos, y de ese modo explicar las frecuencias relativas (p. 9). En este sentido, dicho autor estudia la hipótesis de que la frecuencia fonemática en los inventarios lingüísticos de las lenguas del mundo está relacionada con la frecuencia de los fonemas que se producen efectivamente en esas lenguas. De acuerdo con su estudio, que involucró un corpus de una treintena de lenguas (corpus de diccionarios y textos continuos), existe una gran correlación entre la frecuencia fonemática interlingüística de las consonantes y su respectiva frecuencia intralingüística. En su estudio, la frecuencia intralingüística se computa por medio de la ratio del número de instancias de fonemas observados en relación con el número esperable, si todos esos segmentos ocurrieran con la misma frecuencia, lo que permite analizar su relación (Gordon, 2016).

Un concepto que se ha relevado en la teorización sobre principios lingüísticos universales es el de *markedness* o *marcadez*. Según Hume (2011), el concepto de *marcadez* en fonología puede tener dos usos comunes. Por un lado, está el uso de tipo descriptivo que refiere a “*an abstract relation holding over members of a set of observations displaying asymmetry, such that one subset is unmarked and the other is marked*” (p. 79), y un segundo uso de tipo teórico: “*a*

universal principle or laws that guide language acquisition, loss, inventory structure, process, rules, etc. toward the unmarked form.” (p. 80).

A este segundo uso de *marcadez* es al que adscribe Greenberg (1966/2005) cuando entrega las características de los rasgos *marcado* y *no marcado* en fonología. Dentro de las tendencias que establece este autor, se encuentra que el segmento no marcado prevalece en las posiciones donde ocurre neutralización de fonemas. Asimismo, indica que las categorías no marcadas tienen una mayor frecuencia que las marcadas, como es el caso de las vocales breves frente a las vocales largas. En la misma línea, el número de fonemas que tiene un rasgo particular marcado nunca es mayor que el número de fonemas con el rasgo no marcado; a saber, por ejemplo, consonantes nasales versus consonantes orales. Junto con ello, al establecer generalizaciones de implicatura condicional, se debe relevar que el rasgo marcado es el que implica los rasgos no marcados (la existencia de consonantes nasales implica la existencia de consonantes orales, por ejemplo). Finalmente, el símbolo fonémico escogido entre una serie de alófonos suele ser el que representa el rasgo no marcado.

La evidencia de que los hablantes son sensibles a las frecuencias de patrones fonológicos en distintos niveles de categorías lingüísticas (Munson, 2001) puede entregar nuevas luces respecto de la teorización en torno al concepto de *marcadez* como principio o tendencia universal, aun cuando Hume (2011) advierte que este tema sigue siendo controversial. En esta misma línea, Andrews (1990) refuerza la idea de que hay una tendencia general a que el elemento no marcado en una oposición a menudo ocurra de manera más frecuente que el elemento marcado —tanto en un sentido estadístico como distribucional—, pero que esta tendencia no corresponde a un axioma de la teoría que incluye este concepto (p. 138).

Con respecto a la morfosintaxis, por su parte, Pavey (2010) define la tipología lingüística como la organización de las lenguas en distintos tipos, en virtud de la presencia o ausencia de ciertos rasgos morfológicos o sintácticos (p. 312). Así, la tipología morfológica clasifica las lenguas en virtud de dos escalas o dimensiones: de *síntesis* y de *fusión*. La escala de síntesis, siguiendo la propuesta de la autora, corresponde, a grandes rasgos, al número de morfemas que pueden

sucederse en una palabra. En un extremo de esta escala, se encuentran las lenguas de morfología *aislante*, con palabras monomorfémicas sin afijación; y en el otro extremo, las lenguas de morfología *polisintética*, caracterizadas por palabras polimorfémicas.

Por otro lado, la escala de fusión corresponde al número de significados que puede contener un solo morfema. En un extremo, en las lenguas de morfología *aglutinativa*, cada morfema tiene un solo significado gramatical. En cambio, en una lengua de morfología *fusional*, un morfema puede plasmar muchos significados. Pavey advierte, no obstante, que las escalas de síntesis y fusión son guías que facilitan el análisis y descripción lingüísticos, y no indicadores categóricos: “*languages will usually not fall completely into a particular type, but we can use this typology to make general statements*” (p. 312).

En el caso de la lengua mapuche, los extremos de estas escalas de fusión y síntesis permiten describir y categorizar internamente la lengua; específicamente, por el tipo de morfología que tienen las distintas clases de palabras. En efecto, por un lado, los sustantivos y algunos tipos de palabras de clase cerrada, como pronombres, adjetivos y demostrativos no derivados, pueden describirse como de morfología más aislante, al no admitir afijos de flexión³ y muy pocos de derivación. En cambio, la clase abierta de los verbos es de morfología claramente polisintética, al admitir una mayor riqueza flexiva y derivativa, con decenas de afijos verbales. Asimismo, la morfología verbal del mapudungun es mayoritariamente aglutinante, puesto que es posible desagregar las unidades de significados de los afijos verbales en sus constituyentes discretos. A modo de ejemplo, el complejo morféxico *-ymi* ‘segunda persona singular de indicativo’ puede analizarse y separarse en los siguientes morfemas y significados: *-y* ‘indicativo’, *-m* ‘segunda persona’, *-i* ‘singular’.

Estos aspectos tipológicos son importantes de considerar, puesto que tienen consecuencias al momento de establecer cuáles y cuántas serán las unidades fonológicas segmentales que entrarán

3 Esta aseveración se hace con la salvedad de que es posible expresar caso oblicuo con la forma enclítica *mu* de la adposición *mew*, lo que podría interpretarse como una estructura morfológica flexiva que interactúa en nominales.

en el cómputo en las distintas clases de palabras que componen la lengua mapuche. Como se señaló en el apartado 2.2. *Frecuencia fonemática y lexicón mental*, la naturaleza aglutinante del mapudungun es una ventaja al momento de identificar las unidades para la constitución del corpus, por cuanto permite un acceso transparente a la relación entre significado y forma.

Así las cosas, y como corolario de esta breve descripción del marco de referencia de nuestra propuesta investigativa, nuestro estudio contribuirá a respaldar o cuestionar algunas de las hipótesis que se han planteado con respecto a las tendencias de las lenguas en el plano fonológico, así como también aportará con nuevos insumos para la discusión respecto del lugar que ocupa la unidad fonémica en la configuración del lexicón mental.

3. Estado del arte

3.1. Estudios de fonestadística en lenguas de alta vitalidad

El estudio de la frecuencia de fonemas en lenguas particulares tiene larga data, siendo posible vincular su inicio con trabajos realizados en lenguas globales, como el inglés y otras lenguas indoeuropeas, desde fines del siglo XIX. En efecto, décadas antes de la publicación del *Cours de Linguistique Générale* de Saussure, Dwight (1874) describió desde un punto de vista articulatorio los sonidos del idioma inglés. Asimismo, tabuló la frecuencia relativa de decenas de miles de sonidos que aparecían en textos literarios de diverso origen y periodos. Con el fin de no distorsionar sus resultados, tuvo la cautela de no seleccionar pasajes de textos que tuvieran palabras con una alta frecuencia. Dwight también entregó la proporción de vocales respecto de las consonantes y comparó sus resultados con otras lenguas indoeuropeas, aunque sin explicitar de dónde extrajo esta última información.

Posteriormente, y también adelantándose a los hitos históricos que marcarían el devenir de la lingüística como ciencia, Dewey (1923) publicó un reporte del conteo léxico, silábico y de sonidos del idioma inglés. Entre las motivaciones que el autor señala para realizar su trabajo, se cuentan la falta de materiales que permitiesen asegurar una descripción representativa de este

tema (p. 6), además de que muchos de sus precursores confundían el conteo de sonidos con el de las grafías. En su trabajo también advierte sobre el efecto de distorsión que puede tener en los resultados el conteo de frecuencia fonémica a partir de un léxico muy frecuente, es decir, en textos continuos. En este contexto, las cuatro etapas del proceso investigativo postuladas por Dewey son relevantes, en el sentido de que coinciden con las fases definidas en proyectos más modernos:

The stages of the investigation may be subdivided into: i) Counting, recording, arranging, and listing the 100,000 words, 2) Transcribing the different words recorded in a phonetic notation based on the Revised Scientific Alphabet, 3) Analysing this phonetic transcription with respect to syllables, reducing the data thus secured from the medium precision (48 sound) basis of the Revised Scientific Alphabet to the minimum precision (41 sound) basis of the phonetic print alphabet adopted; arranging, charting, and summarizing the results in various ways, 4) Similar analysis, arrangement, and charting with respect to separate single sounds. (p. 10)

Posterior al trabajo de Dewey, el conteo estadístico de frecuencias fonémicas tuvo un repunte y contribuyó a los estudios tipológicos y de universales lingüísticos. Así, por ejemplo, en la obra *Psycho-Biology of Language*, el precursor de la lingüística cuantitativa, G. Zipf, realizó estudios estadísticos de palabras, sílabas y sonidos. Bajo un enfoque empírico, propuso de manera tentativa algunas explicaciones para fenómenos de la expresión lingüística. Respecto del conteo de sonidos, reportó las frecuencias relativas de los fonemas en una veintena de lenguas, mayoritariamente indoeuropeas, pero también en algunas lenguas siníticas, en el idioma húngaro y en algunos estadios lingüísticos pretéritos, como el latín, inglés anglosajón y alto alemán antiguo. Otros estudios tempranos de la frecuencia fonemática en lenguas europeas son los del húngaro (Lotz, 1952), italiano (Busa et al., 1962) y la investigación comparativa entre el inglés, francés, alemán y español de Delattre (1965). Esta sobrerrepresentación de las lenguas indoeuropeas frente a otras familias y lenguas del mundo es común en los estudios sobre frecuencia fonológica, como se discutirá más adelante.

Como se desprende del apartado 2.1. *Fonología y fonoestadística*, Trubetzkoy (1939) propuso algunas coordenadas fundamentales para abordar el estudio de la frecuencia fonológica de las lenguas, en general, y de las lenguas de alta vitalidad, en particular. Una de ellas es la distinción fundamental en el tipo de insumo lingüístico y el tipo de tarea del estudio estadístico fonológico. Para este autor, en primer lugar, la estadística en fonología debe mostrar la frecuencia de un elemento fonológico de una lengua dada, ya sea en tanto fonema específico o como parte de una combinación de fonemas. Por otro lado, debe indicar “*the importance of the functional load of such an element or of a specific phonological opposition*” (p. 256). Para el primer caso, son útiles los textos continuos, y para el segundo, se hace uso de diccionarios.

Este prominente investigador señala, además, la dificultad de arribar a una explicación para ciertas tendencias que se producen en las frecuencias de los fonemas, especialmente porque este fenómeno es “*the result of a whole sequence of propelling forces*” (p. 264). Por lo mismo, indica que lo realmente importante al realizar conteos de fonemas es establecer la relación de las cifras con las cifras teóricamente expectables en sus respectivos contextos de distribución. Este giro en el enfoque lleva a la generación de diferentes indicadores estadísticos como razones matemáticas y otros cocientes que permiten comparar las cantidades resultantes. Asimismo, señala que también debe considerarse el reporte de la distribución de los fonemas en determinadas posiciones (p. 265), llegando a tabulaciones específicas que tengan un mayor grado de detalle y se concentren no solo en los fonemas, sino en las oposiciones, pues “*it should always be remembered that it is not the phonemes but the oppositions that represent the true object of phonological study*” (p. 266). Por último, menciona aspectos problemáticos para la estadística fonológica que no trabaja con textos continuos, como son la presencia de vocabulario muy técnico, específico o en desuso y terminología específica para una zona (vocabulario dialectal). Trubetzkoy se pregunta si debe incluirse ese vocabulario en un conteo de frecuencia de fonemas a partir del léxico (p. 268).

Como se aprecia en este breve recorrido, el interés por la frecuencia de fonemas tiene una larga data y se ha concentrado de manera prominente en las lenguas indoeuropeas, en general, y en

particular, en el inglés. A continuación, señalaremos el desarrollo de este tipo de investigación en otras lenguas globales y nacionales.

En cuanto a la investigación estadística en la lengua castellana, además de los conteos realizados en el estudio de Zipf (1935) y Zipf y Rogers (1939, citado en Alarcos, 1965), el lingüista español Tomás Navarro (1946), en su *Estudios de Fonología Española*, analiza “el volumen que cada fonema representa en el conjunto del material sonoro que la lengua utiliza” (p. 15). Se reportan en este trabajo la escala de frecuencia de los fonemas españoles y cuatro categorías fonemáticas según sus frecuencias. Por su parte, Alarcos (1965), en su libro *Fonología Española*, dedica un capítulo a la frecuencia de los fonemas y señala que se distancia de los trabajos ya mencionados de Zipf y Rogers (1939) y de Navarro (1946) en cuanto considera el fonema como unidad de análisis, y no sus alófonos. Además, Alarcos (1965) agrupa los fonemas de acuerdo con sus frecuencias relativas en cinco categorías: 1.- superior a 10%, 2.- entre 6% y 10%, 3.- entre 3% y 6%, 4.- entre 1% y 3%, y 5.- inferior a 1%.

En la década en que Alarcos publica su libro, se comienza a producir un cambio en la forma de llevar a cabo la ciencia lingüística. La aplicación de nuevas tecnologías constituirá una característica de los estudios de estadística fonológica en adelante. Así, el resurgimiento de los estudios basados en corpus que se da en la década del sesenta es posible gracias a la incorporación de los computadores en el ámbito lingüístico y al desarrollo de importantes corpus lingüísticos digitales, principalmente en inglés (Parodi, 2008). Este desarrollo permite la automatización de procesos de análisis y recolección de datos para estudios de distribución y frecuencia fonémica, lo que se evidencia, por ejemplo, en la ampliación del foco de estudio a unidades mayores como la sílaba.

Efectivamente, y como ejemplo de lo anterior, Lloyd y Schnitzer (1967) mencionan las ventajas de las mejoras relacionadas con el uso de herramientas de procesamiento de información a alta velocidad, porque permiten el análisis de grandes volúmenes de datos de manera eficiente, rápida y acuciosa. En este mismo trabajo, los autores abordan la estructura silábica, usando un corpus de léxico del castellano formado a partir del *Diccionario de la Lengua Española*, en su edición

XVII, de 1956, y que incluyó más de 70.000 entradas (p. 60). El estudio arrojó que había una predilección por las sílabas CV en todas las posiciones dentro de la palabra (inicial, media, final). También dentro de este mismo campo de interés, Guerra (1983) reporta los resultados de un estudio estadístico diacrónico de la sílaba del castellano, en cinco momentos de la historia de esta lengua.

En los últimos años, el foco de los estudios sobre la frecuencia fonémica en castellano se ha ampliado al mapeo de las frecuencias fonémicas en distintas variedades de esta lengua, con trabajos en español de Chile (Pérez, 2003), México (Pineda et al., 2004) y Colombia (González y Mejía, 2011).

Si bien es evidente la presencia de muchos y diversos estudios en el campo de la fonoestadística sobre lenguas globales, como el inglés y el castellano, también se han realizado trabajos en otras lenguas, especialmente desde el auge y uso masivo de los medios informáticos. Por ejemplo, Malécot (1974) da cuenta de los resultados del análisis, mediante herramientas computarizadas, de la distribución y frecuencias de fonemas en el francés moderno, donde el objetivo del estudio es la caracterización de su inventario fonológico.

Smirnova y Chistikov (2011), en tanto, realizan un estudio sobre la frecuencia y distribución de las unidades fónicas del ruso, con herramientas de transcripción fonológica automatizada. El corpus lo conforman una combinación de textos literarios y transcripciones de entrevistas y cátedras. De acuerdo con la estructura fonológica de esta lengua eslava, se da cuenta de la frecuencia de ocurrencia de combinaciones fonológicas (grupos consonánticos, por ejemplo) y se reporta también la proporción consonántica, que es de 1,35 (57,4%, consonantes; y 42,6%, vocales).

Respecto del árabe, en un estudio reciente, Nahar et al. (2012) dan cuenta de la frecuencia fonémica y la realización de parámetros acústicos concomitantes (duración, etc.), con miras al desarrollo de herramientas de reconocimiento de habla y síntesis de voz. Los autores señalan como motivación del estudio la falta, precisamente, de un análisis estadístico de los fonemas

árabes en textos continuos de habla. El estudio calcula, asimismo, la distribución probable de ocurrencia de los fonemas de esta lengua.

Munthuli et al. (2013), por su parte, presentan una metodología para conocer la distribución fonemática del tailandés, a partir de corpus escritos. En un segundo estudio (Munthuli et al., 2015), realizan comparaciones entre tres grandes corpus textuales, con pruebas estadísticas, para establecer si los resultados en cuanto a frecuencia fonémica y distribución en sílabas del tailandés son divergentes con significatividad estadística. Los resultados indican que los tipos de sílabas y la asignación de tonos tienen una alta correlación en los cuatro corpus; no obstante, existen discrepancias en cuanto a la frecuencia fonémica en algunas vocales y en las consonantes respecto de su posición en ataque y coda.

Otras lenguas que han sido objeto de estudio reciente, con el fin de desarrollar un estudio cuantitativo aplicado de su fonología, son el japonés (Tamaoka y Makioka, 2004); setswana, de la subfamilia bantú (Basadi Palai y O'Hanlon, 2004); telugu, de la familia drávida (Kalyani y Sunitha, 2009); y hebreo, de la familia semítica (Cohen, 2015; Silber-Varod et al., 2017), entre otras.

Como es posible desprender de los trabajos y estudios desplegados en este breve recorrido, ha existido un interés por estudiar el fenómeno de la frecuencia fonémica y su influencia en distintas temáticas. Este interés, especialmente en lenguas globales como el inglés y el español, se ha traducido en una amplia bibliografía de estudios realizados con diversas metodologías y aplicaciones en otros campos. Caso similar es posible reportar para otras lenguas con una vitalidad asegurada, donde no solo se han indicado las unidades fonológicas y su frecuencia, sino que también se ha informado respecto de la frecuencia fonológica en unidades combinatorias mayores como la sílaba. Este no ha sido el caso, en general, en las lenguas consideradas en peligro.

3.2. Estudios y reportes de frecuencia fonémica en lenguas en peligro

Unesco (2003) define una lengua en peligro como aquella que “se encuentra en vías de extinción” (p. 2). Por medio de diferentes criterios, como la transmisión intergeneracional, número absoluto de hablantes, cambios en su uso, entre otros, ha diseñado instrumentos que permiten evaluar y clasificar las lenguas del mundo según su grado de vitalidad. Con esta tipología a la vista, ponemos en foco estudios estadístico-fonológicos que dan cuenta del interés académico respecto de las lenguas consideradas en peligro, entre las que, por cierto, se incluye la que constituye nuestro foco investigativo: el mapudungun.

Un aspecto que es pertinente relevar una vez que se revisa la bibliografía sobre estas lenguas, es que dichos estudios no presentan el mismo alcance y profundidad si se comparan con los estudios sobre lenguas globales, como el inglés y el castellano, y de otras lenguas nacionales de comunicación mayoritaria. En efecto, muchos de los trabajos en los que se mencionan datos de índole estadística sobre lenguas en peligro corresponden a apartados breves dentro de perfiles de tipo fonológico, siguiendo el interés fijado desde Trubetzkoy (1939) de caracterizar las lenguas con este tipo de información; y en línea con ello, poco se ha ahondado en otras aplicaciones e implicancias de estas consideraciones. Con todo, es posible mencionar algunas contribuciones prominentes que se han ocupado de este tema.

Fuera de nuestro continente, Tambovtsev (2003), bajo la premisa de que la estructura de la frecuencia consonántica en la cadena hablada es un buen indicador de la cercanía tipológica entre lenguas, estudia la frecuencia fonemática del euskara, lengua seleccionada por ser considerada aislada genéticamente. Al no encontrar estudios anteriores de frecuencia fonemática de esta lengua vulnerable (Salminen, 2010), el autor realizó este cálculo a partir del análisis de textos literarios, reportando la frecuencia absoluta y relativa de vocales y consonantes, además de su cociente. Con esa información, comparó y midió la distancia fonoestadística del euskara con 40 lenguas indoeuropeas.

En 2020, Macklin-Cordes y Round, a partir de un corpus de 166 lenguas de pueblos originarios de Australia, re-evalúan las distribuciones matemáticas que modelan frecuencias fonémicas; por ejemplo, la distribución de Zipf y la de Yule-Simon. Los autores relevan como una ventaja de su estudio el contar con lenguas que no corresponden a las ya muy focalizadas lenguas euroasiáticas.

A diferencia de los estudios referidos, que mencionan las frecuencias fonológicas o de sílabas como parte de investigaciones mayores, González (2016) realiza un estudio focalizado en los aspectos fonoestadísticos del cabécar, lengua hablada en Costa Rica, perteneciente al grupo chibcha. El autor señala que su estudio se adscribe a la lingüística de corpus —el que cataloga como un campo emergente en Costa Rica— y que busca llenar un vacío de estudio, porque, en general, las investigaciones fonoestadísticas “son bastante escasas en lo que a lenguas aborígenes se refiere” (p. 1).

El estudio de González reporta cómo se conformaron los corpus electrónicos desde donde se extrajeron las unidades léxicas, además del procedimiento de análisis, que incluyó herramientas de análisis estadístico automatizado. Los *software* utilizados por el autor fueron Fieldworks Language Explorer (FLEx), principalmente para el segmentado y etiquetado morfológico de las unidades léxicas y extracción de listas léxicas, y MS Excel, para el conteo automatizado de los segmentos. Las fases de su investigación son coincidentes con las que se han reportado en investigaciones para otras lenguas (véase Nahar et al., 2012, p. 50 y Márquez, 2016, descrita en la siguiente sección). Específicamente, su investigación se llevó a cabo en tres etapas:

- a) Digitalización de los textos y conformación del corpus electrónico
- b) Diseño e implementación del modelo de transcripción fonológica automático
- c) Cálculo de las frecuencias fonémicas

Además de tabular los resultados indicando las frecuencias absolutas y relativas de las unidades fonológicas, se incluyeron gráficos de descripción estadística. La mayor parte de los resultados son comparados con los resultados de la tipología lingüística; por ejemplo, se señala que el sonido más frecuente es la vocal baja central no redondeada /a/, lo que concuerda con lo señalado por Ladefoged y Ferrari (2012) respecto de los sonidos vocálicos y sus tendencias universales. Finalmente, se reportan las frecuencias desde el punto de vista de los parámetros articulatorios, como modo y punto de articulación de las consonantes, tipo de vocal según función del velo del paladar (oral o nasal), y dirección y altura de la lengua.

Con respecto a Sudamérica, y como parte de un estudio mayor de la estructura lingüística del guaraní coloquial, Gregores y Suárez (1967) reportaron la frecuencia fonemática en una sección dedicada a la fonología. La proporción de fonemas vocálicos respecto de la totalidad de los segmentos, un indicador que se ha utilizado en estudios de tipología lingüística (Maddieson, 2013b), es de 0,25. Asimismo, se calculó la proporción de segmentos obstruyentes respecto de los resonantes, la cual, en este caso, corresponde a 1,16.

Algunos estudios de tipo lingüístico, en general, y otros específicamente fonológicos de lenguas habladas en países del cono sur de América (Chile y Argentina), en particular, han reportado las frecuencias absolutas y relativas de sus fonemas. Sin embargo, se resiente el hecho de que, en muchos casos, al no ser la frecuencia fonemática el tema central, se omite la metodología empleada para su determinación, de modo que no se puede tener plena certeza, por ejemplo, del tipo de unidad de análisis o tipos de textos utilizados, cuáles fueron los protocolos y herramientas empleados para dicho análisis, etc. Un trabajo relevante realizado en este referente geográfico, para el cual aplica el alcance anterior, es el realizado por Fernández Garay (1998) sobre el tehuelche o aonek'ó 'a'jen, el cual se realiza dentro del “marco de la teoría funcionalista cuyo principal representante es André Martinet” (p. 52).

Otro trabajo prominente de esta índole, que recopila información sobre la frecuencia fonológica de 5 lenguas originarias de Chile —aymara, rapanui, mapudungun, kawésqar y yagan—, es el realizado por el Centro de Estudios de Lenguas de Tradicional Oral (CELTO), de la Universidad

de Playa Ancha. En él, sus autores y autora presentan las primeras descripciones fonológicas en Chile “bajo el marco teórico del funcionalismo martinetiano” (Clairis et al., 2001, p. 3). La influencia de Martinet en estos trabajos es posible que fuera clave para consignar los conteos de frecuencias, puesto que para el lingüista francés la frecuencia fonológica a partir de unidades léxicas de una lengua particular guarda relación no solo con la caracterización de esa lengua, sino que con su rendimiento funcional de las unidades segmentales (Martinet, 1964).

A modo de ejemplo, en el trabajo del CELTO, Riffo (2001) reporta para el rapanui las unidades fonológicas correspondientes a vocales y consonantes, según el orden en la palabra, sílaba y también respecto de parámetros articulatorios. Asimismo, a partir de su frecuencia, realiza una clasificación de los fonemas en *muy frecuentes*, *bastante frecuentes*, *frecuentes* y *poco frecuentes*. Por su parte, Clairis (2001) da cuenta de la frecuencia de los fonemas del kawésqar según su posición en la palabra, utilizando dos corpus, uno de enunciados y otro de lexemas aislados (el conteo se realizó de manera manual). En dicho estudio, señala que “la disminución de la frecuencia desde el fonema más frecuente hasta el que lo es menos, no sigue un ritmo regular” (p. 95), es decir, las seis unidades más frecuentes representan más del 70-80% de las frecuencias de ocurrencias, mientras que las otras unidades representan el 20-30% restante. Las otras lenguas indicadas en este trabajo siguen reportes similares respecto de la frecuencia de unidades fonológicas y su posición en la palabra.

3.3. Estudios y reportes de frecuencia fonémica en mapudungun

Es un hecho que los primeros reportes sobre la fonética y fonología del mapudungun que siguen metodologías como el distribucionalismo norteamericano no dan cuenta de las frecuencias fonémicas absolutas o relativas en sus estudios.

Suárez (1959), quien trabajó a partir de los materiales recopilados por Rodolfo Lenz (1895-1897), no reporta un conteo de la ocurrencia de las unidades estudiadas, a pesar de que una de las motivaciones para realizar su estudio es poner en foco que el estatus de fonema de ciertos sonidos es dudoso, precisamente, debido a su baja frecuencia: “*sounds of low frequency, and*

morphemes of few occurrences seem not to have been sufficiently rechecked” (p. 177). En este contexto, realiza descripciones generales y cualitativas que dan cuenta de la baja frecuencia de ciertos fonemas o estructuras.

Echeverría (1964) y Echeverría y Contreras (1965), con la colaboración de hablantes de mapudungun, entre quienes se contaba una entonces estudiante de Pedagogía en Inglés, María Catrileo, publicaron dos artículos con el objetivo de describir la fonología segmental de esta lengua. No obstante, al igual que en el caso de Suárez, no se indicaron datos respecto de las frecuencias relativas o absolutas de los segmentos, o de los tipos de sílabas que se describen como más y menos frecuentes en el idioma.

El primer autor que entrega información cuantitativa sobre frecuencias de segmentos fonémicos, que ha sido posible rastrear para este trabajo, es Lagos (1984), quien indica la frecuencia relativa de las vocales (50,43%) y consonantes (49,56%) del mapudungun, bajo un breve apartado titulado *Frecuencia fonemática*. Lagos registra datos como el primer y último lugar en el rango de frecuencias, donde el primer lugar corresponde a la vocal /e/, con un 16,08%, y el último lo ocupa el fonema /t/, con un 0,86% (p. 44). Además, señala que Croese y él mismo ya habían realizado estudios de frecuencia de ocurrencia con anterioridad, pero con resultados diferentes.

Años más tarde, en su tesis para optar al grado de Magíster en Artes con Mención en Lingüística en la Universidad de Concepción, Álvarez-Santullano (1986) presenta una descripción de los fonemas segmentales del tse-sungun, variante del mapudungun, con la colaboración de hablantes nativos “residentes en un área ubicada a unos 30 km. al oeste de la ciudad de Osorno” (p. 6). En este trabajo, junto con el inventario de fonemas y alófonos, se aporta de manera relevante con la frecuencia y gradación de los fonemas, pero, lamentablemente, como señala Márquez (2016), no se explicitan “los procedimientos con que la frecuencia de fonemas fue medida, por lo que se entiende que se trató de un conteo general de los datos analizados previamente” (p. 12).

En 2001, en la obra publicada por CELTO, ya descrita en la sección anterior, Lagos presenta nuevamente un resumen esquemático o cuadro estadístico, bajo el apartado *2.4 Frecuencia de*

fonemas, donde señala frecuencias absolutas y relativas del conteo fonemático. Los resultados totales de fonemas vocálicos y consonánticos corresponden a los reportados en el trabajo de 1984.

Tendrían que pasar varias décadas para encontrar el primer trabajo que se centra en el conteo de frecuencia fonológica del mapudungun como objeto de estudio. Así, Márquez (2016) —como parte de su tesis para optar al grado de licenciado en educación— publica *Frecuencia de fonemas y consideraciones teórico-metodológicas en un corpus del mapudungun*, cuyo objetivo fue determinar la frecuencia fonológica de la lengua, a partir de distintas variables, y proponer un criterio metodológico para obtener los datos.

La metodología desplegada por Márquez consiste en utilizar dos diccionarios publicados a mediados de la década de los noventa por María Catrileo (1996) y Hernández y Ramos (1997), respectivamente, de los que se extrajo un listado léxico a partir de sus entradas. Posteriormente, se realizó un filtrado, a fin de eliminar repeticiones, castellanismos, además de algunos morfemas gramaticales. De este modo, se filtraron sufijos como *-we* ‘lugar, instrumento’, y *-fe* ‘agente’, así como el elemento léxico de alta frecuencia *che* ‘persona’, que puede estar presente en formas compuestas o como parte incorporada en un verbo. La misma acción se realizó para entradas que en su forma de cita tuvieran el sufijo *-n* ‘infinitivo’. Luego del filtrado, se generó un listado final de 1081 palabras, que incluye ambos corpus. Determinada así la lista léxica, se realizó la transcripción fonológica, para lo cual se utilizó el Alfabeto Fonético Internacional (AFI). Finalmente, para el recuento de la frecuencia, se utilizó MS Excel, por medio de una disposición especial de las celdas y el uso de fórmulas y funciones como la de “autosuma” que incorpora esta herramienta.

El resultado es una planilla de datos que contiene columnas para las palabras, las transcripciones fonológicas de las mismas y cada uno de los fonemas del mapudungun, con sus respectivas tabulaciones de frecuencias. Salvo pequeñas diferencias, esta tabla es muy similar, en cuanto a su disposición, a la diseñada para la cuantificación de los fonemas del cabécar (González, 2016, p. 21).

El siguiente paso del estudio fue dar cuenta de los resultados, por medio de estadística descriptiva, reportando la frecuencia absoluta y porcentual de los fonemas, además de la inclusión de gráficos de barra ordenados por el rango de frecuencia de los segmentos. En primer lugar, se separaron los resultados por corpus, para luego entregarlos de manera acumulada. En específico, se reportaron los resultados para los fonemas en general y luego para los consonánticos y los vocálicos. Los datos arrojados fueron comparados con los resultados de los estudios de Lagos (2001) y Álvarez-Santullano (1986), de lo que es posible desprender ciertas discrepancias, especialmente en el cociente consonante/vocal. En cuanto al parámetro *dirección de la lengua*, “las vocales con mayor frecuencia son aquellas de articulación central /a/ y /ə/, en la posición intermedia de frecuencia se encuentran las vocales anteriores /e/ e /i/, y en el extremo de menor frecuencia se ubican las vocales de articulación posterior /u/ y /o/” (p. 42), lo cual resulta interesante si se considera la alternancia de /o/ y /u/ en sílaba átona, reportada por Echeverría (1964, p. 46) y Salas (1992a, p. 89).

Como es posible evidenciar a partir de la bibliografía en torno a la fonoestadística en la lengua mapuche, exceptuando el trabajo de Márquez (2016), no se han realizado otros estudios que hayan abordado de manera exclusiva la frecuencia de fonemas a partir de unidades léxicas —o de cualquier unidad para ese efecto—, ni la frecuencia de ocurrencia en unidades combinatorias mayores como la sílaba.

En la **Tabla 3.1**, se recopilan los resultados de las frecuencias de fonemas de los trabajos mencionados en esta sección. Una de las dificultades de comparar los distintos resultados es que los inventarios fonológicos no se corresponden completamente, por presentar diferencias dialectales y de interpretación de ciertos segmentos. En el caso del estudio de Álvarez-Santullano (1986), en particular, al estudiar una variante de la lengua mapuche que ha sido caracterizada como más separada o aislada que las otras (Salas, 2006), es razonable esperar inventarios diferentes. En efecto, la investigadora da estatus fonológico a la lateral retrofleja sonora /ɭ/, lo que, como es previsible, tiene un efecto en la frecuencia de ocurrencia. Asimismo, esta variante meridional no tiene como unidades funcionales las consonantes oclusiva interdental átona /ɬ/,

nasal interdental sonora /ɲ/, fricativa alveopalatal áfona /ʃ/, ni la fricativa velar sonora /ɣ/. Por su parte, Lagos (2001) tampoco establece como unidades fonológicas las consonantes fricativa alveopalatal áfona /ʃ/, ni la fricativa velar sonora /ɣ/.

En la misma **Tabla 3.1**, a partir de la descripción articulatoria contenida en los trabajos de Álvarez-Santullano (1986) y Lagos (2001), se reemplaza el símbolo original de los autores, por el símbolo AFI.



Tabla 3.1: Comparación de las frecuencias absolutas (f. abs.) y relativas (f. rel.) de los trabajos que han reportado esta información respecto de los fonemas del mapudungun.

ÁS			L			M-C			M-HR			M-CHR		
AFI	f. abs.	f. rel. (%)	AFI	f. abs.	f. rel. (%)	AFI	f. abs.	f. rel. (%)	AFI	f. abs.	f. rel. (%)	AFI	f. abs.	f. rel. (%)
a	–	8,96	e	6509	16,08	a	358	11,67	a	407	11,09	a	765	11,35
w	–	8,27	a	4549	11,24	ə	267	8,70	k	301	8,20	ə	554	8,22
e	–	7,93	i	4546	11,23	e	244	7,95	e	292	7,96	e	536	7,95
u	–	7,75	n	2751	6,80	k	235	7,66	ə	287	7,82	k	536	7,95
n	–	7,75	t	2343	5,79	i	185	6,03	i	242	6,59	i	427	6,34
k	–	7,58	m	2285	5,65	u	172	5,60	u	216	5,89	u	388	5,76
i	–	7,06	k	2097	5,18	w	155	5,05	w	189	5,15	w	344	5,10
m	–	6,03	u	1704	4,21	o	133	4,33	o	171	4,66	o	304	4,51
w	–	4,65	w	1578	3,90	m	131	4,27	p	149	4,06	m	273	4,05
p	–	4,48	o	1524	3,77	p	122	3,98	m	142	3,87	p	271	4,02
o	–	3,96	l	1484	3,67	n	109	3,55	l	137	3,73	l	242	3,59
l	–	3,27	ɲ	1222	3,02	l	105	3,42	n	121	3,30	n	230	3,41
ʈ	–	3,10	f	1182	2,92	f	88	2,87	ʎ	115	3,13	ʎ	200	2,97
f	–	3,10	ʈ	1003	2,48	ɲ	87	2,83	ɻ	111	3,02	ɻ	195	2,89
ɣ	–	3,10	w	998	2,47	ʎ	85	2,77	ʈ	106	2,89	ɲ	183	2,72
j	–	2,24	p	947	2,34	ɻ	84	2,74	ɲ	96	2,62	ʈ	183	2,72
ʎ	–	2,06	ʈ	911	2,25	ɳ	78	2,54	j	91	2,48	f	178	2,64
t	–	1,72	j	810	2,00	ʈ	77	2,51	ɳ	91	2,48	ɳ	169	2,51
ʈ	–	1,55	ɻ	646	1,60	j	76	2,48	f	90	2,45	j	167	2,48
ɳ	–	1,55	θ	460	1,14	ʈ	74	2,41	ʈ	87	2,37	ʈ	161	2,39
ɲ	–	1,20	ʎ	305	0,75	ð	52	1,69	ð	80	2,18	ð	132	1,96
θ	–	1,03	ɳ	216	0,53	t	47	1,53	t	52	1,42	t	99	1,47
ʎ	–	0,86	s	201	0,50	ɳ	36	1,17	ɳ	37	1,01	ɳ	73	1,08
s	–	0,34	ʎ	126	0,31	ʎ	32	1,04	ʎ	22	0,60	ʎ	54	0,80
ʎ	–	0,34	ɳ	43	0,11	s	19	0,62	s	19	0,52	s	38	0,56
ɻ	–	–	ɻ	35	0,09	ɣ	9	0,29	ɣ	8	0,22	ɣ	16	0,24
ɳ	–	–	ɣ	–	–	ɻ	9	0,29	ɣ	7	0,19	ɻ	13	0,19
ɣ	–	–	ɳ	–	–	ɣ	0	0,00	ɻ	4	0,11	ɣ	8	0,12
ɻ	–	–	ʎ	–	–	ʎ	–	–	ʎ	–	–	ʎ	–	–
	–	99,88%		40475	100%		3069	100%		3670	100%		6739	100%

Leyendas: AS: conteo reportado en Álvarez-Santullano (1986); L: conteo reportado en Lagos (2001); M-C: conteo reportado en Márquez (2016), con base en el *Diccionario lingüístico-etnográfico de la lengua mapuche* (Catrileo, 1996), M-HR: conteo reportado en Márquez (2016), con base en *Mapuche. Lengua y Cultura. Diccionario Mapudungun-español-inglés* (Hernández y Ramos, 1997), M-CHR: conteo acumulado reportado en Márquez (2016) que se sustenta en los dos diccionarios antes mencionados.

En la **Tabla 3.1**, además, se registran las cifras absolutas de casi todos los trabajos —Álvarez-Santullano (1986) no las reporta—, las que se usaron para calcular los porcentajes en la frecuencia relativa, aproximando los resultados a dos decimales, con el objetivo de que la suma de las frecuencias relativas fuese 100%. En el caso de algunos porcentajes, puede haber discrepancias con los señalados en las obras originales, debido a esta forma de aproximar los decimales, pero en ningún caso esa diferencia debiera ser mayor a una décima. Por último, para el despliegue de esta tabla, debió corregirse un porcentaje en el caso del trabajo de Lagos (2001)⁴.

En resumen, y como ha sido posible apreciar, respecto de las lenguas en peligro, en general, y del mapudungun, en particular, es evidente que han sido poco representadas en la literatura sobre fonostadística. Salvo contadas excepciones, en los casos donde se mencionan reportes de frecuencia de fonemas, estos sistemáticamente forman parte de un estudio con un objetivo mayor; por ejemplo, caracterizar los fonemas, alófonos y fonotaxis de la lengua respectiva. Esta situación ha implicado, en la mayoría de los casos, que no sea posible identificar metodologías comunes que permitan comparaciones exhaustivas o dar cuenta de eventuales mejores criterios operativos. Asimismo, es notable la falta de estudios donde se analice la frecuencia fonémica en unidades combinatorias mayores y donde se aborde su relación con otros aspectos como la adquisición de la lengua como L1 o L2, el desarrollo de recursos tecnológicos, la enseñanza y aprendizaje de lenguas en contextos de revitalización, o las patologías clínicas como la afasia, aspectos que sí se han desarrollado respecto de lenguas globales.

En este contexto, nuestro estudio contribuye a aumentar el escaso caudal de conocimientos sobre la frecuencia de fonemas del mapudungun, a la vez que, a partir de sus resultados, se espera motivar la realización de trabajos que relacionen, precisamente, la frecuencia de fonemas con temáticas como las recién mencionadas.

4 Para la alveolar nasal sonora /n/, Lagos establece un 6,97% (2001, p. 256), pero al comprobar el cálculo a partir del conteo absoluto reportado, esta cifra debería ser 6,79%, aproximado a 6,8%. Otra anomalía es que la suma total de los valores absolutos para los segmentos consonánticos —es decir, 20.065— no corresponde al valor reportado de 20.055 (p. 253), pero esta diferencia de 10 unidades se puede considerar como nimia.

4. Pregunta y objetivos de investigación

Dado que la presente investigación tiene como propósito describir la frecuencia de las unidades fonológicas segmentales del mapudungun, no se formularán hipótesis, sino que se plasmará la pregunta que este estudio procura responder y los objetivos de investigación que se derivan de la misma.

4.1. Pregunta de investigación

A partir de un corpus electrónico preparado sobre la base del *Diccionario mapudungún-español* de Augusta (1916/2017), la presente investigación se propone responder la siguiente pregunta:

- ¿Cuál es la frecuencia de las unidades fonológicas segmentales que ocurren en ese corpus?

Según se menciona y detalla en el apartado 3.3. *Estudios y reportes de frecuencia fonémica en Mapudungun*, la pregunta sobre la frecuencia de fonemas de esta lengua sólo ha sido planteada como objetivo central de investigación en el trabajo de Márquez (2016). Sin embargo, y a diferencia de aquel estudio, en nuestro caso utilizaremos las entradas del *Diccionario mapudungún-español* de Augusta (1916/2017), obra que ha sido considerada como un texto de primer orden, todavía no superado en cuanto a su calidad lexicográfica, alcance informativo y gran número de entradas, manteniendo plena vigencia desde su publicación (Salas, 1985, 1992b). Por otro lado, estimamos que en nuestro estudio controlamos con mayor exhaustividad los criterios de inclusión de las unidades que proveen los insumos para la determinación de dicha frecuencia (Consúltese el apartado 5.2. *Constitución del corpus electrónico*).

4.2. Objetivos de investigación general y específicos

A partir de la pregunta de investigación, es posible formular el siguiente objetivo general:

- Determinar la frecuencia de los fonemas del mapudungun que ocurren en las unidades lexemáticas monomorfémicas, utilizando un corpus preparado a partir de las entradas del *Diccionario mapudungún-español* de F. Félix José de Augusta (1916/2017).

Los siguientes objetivos específicos tributan al objetivo general planteado:

1. Establecer la jerarquía general de los fonemas del mapudungun, a partir de su frecuencia de ocurrencia.
2. Determinar distintos indicadores de relación entre tipos de fonemas, como la ratio consonante / vocal, la proporción de segmentos resonantes respecto de los obstruyentes, la razón entre segmentos sonoros y áfonos, y la ratio entre frecuencias observadas y esperadas.
3. Realizar un recuento estadístico de las frecuencias de los fonemas del mapudungun en cuanto a clases naturales, es decir, según parámetros articulatorios para consonantes y vocales.
4. Clasificar en categorías de frecuencia *ad hoc* los fonemas del mapudungun.

Una vez logrados los objetivos planteados por este estudio, además de avanzar en la caracterización y descripción fonológica del idioma mapuche —lo que en sí mismo constituye una contribución al conocimiento respecto de esta lengua en peligro—, se espera aportar a diversos campos, como la tipología y la búsqueda de universales lingüísticos, al estudio sobre la adquisición del mapudungun como primera y segunda lengua, así como a la aplicación de desarrollos tecnológicos en áreas como el procesamiento lingüístico automático.

5. Metodología

En esta sección, se describirá con mayor detalle el diseño investigativo que se desplegó para responder la pregunta de investigación, lo que incluye el alcance del mismo. De igual modo, profundizaremos en la descripción de la unidad de análisis y la descripción de las variables de estudio.

5.1. Diseño investigativo

La presente investigación sigue un enfoque cuantitativo de alcance descriptivo. El diseño de la investigación es de tipo no experimental transeccional. La investigación no experimental corresponde a los “estudios que se realizan sin la manipulación deliberada de variables y en los que sólo se observan los fenómenos en su ambiente natural para después analizarlos” (Hernández-Sampieri et al., 2010, p. 149).

La elección de este tipo de estudio está motivada por la pregunta de investigación, que indaga respecto de la frecuencia fonemática del mapudungun, dimensión que es susceptible de ser medida, descrita y comparada. El aspecto transeccional de esta investigación no experimental refiere a que se recolectan datos en un momento determinado en el devenir de la lengua mapuche; es decir, se estudia el estado de la lengua de manera sincrónica, y no en su dimensión diacrónica o histórica (Saussure, 1945).

5.1.1. Alcance de la investigación

El alcance de esta investigación se define como de tipo descriptiva. De acuerdo con Hernández-Sampieri et al. (2010), “los estudios descriptivos buscan especificar las propiedades, las características y los perfiles” (p. 80) de cualquier fenómeno que se encuentra bajo análisis. El objetivo de este tipo de estudios no es necesariamente explicar la relación entre variables o predecir cómo se comporta esa relación, como sí sería en un estudio con un alcance predictivo o que incluya el sometimiento a prueba de hipótesis (Hernández-Sampieri et al., 2010). En

palabras de estos autores, un estudio de corte descriptivo “es útil para mostrar con precisión los ángulos o dimensiones de un fenómeno, suceso, comunidad, contexto o situación.” (p. 85). El alcance descriptivo, entonces, es un paso necesario para construir modelos relacionales más completos y explicativos de los fenómenos que se están analizando, y, por lo tanto, es un paso adecuado tras haberse explorado un tema en la literatura.

5.1.2. Unidad mínima de análisis y variables

En atención a la pregunta y objetivos de investigación planteados, se procede a la definición de la unidad mínima de análisis, que corresponde a los casos que serán materia del estudio.

La noción de *estructura fonológica de la palabra* o *phonological word structure* propuesta por Altmann (2005) es adecuada para establecer un diseño investigativo que permita responder a los objetivos y pregunta de investigación planteados en este estudio. De acuerdo con este autor, “*the entity word —being one that can be defined in different ways— displays different measurable properties*” (p. 191). También señala que algunos de los constituyentes de esta estructura fonológica de la palabra son, precisamente, unidades menores como los fonemas y las sílabas, y como tales, a partir de estos componentes es posible formular hipótesis que permitan describir y explicar los mecanismos subyacentes a estas estructuras fónicas.

En atención a lo señalado sobre la estructura fonológica de la palabra, se decidió trabajar con un diccionario específico, cuya extensión permite responder la pregunta y objetivos planteados. Como se mencionó en 4.1, se utilizaron las entradas léxicas del *Diccionario mapudungún-español* de Félix José de Augusta (1916/2017), cuya versión electrónica se encuentra alojada en el *Corpus Lexicográfico del Mapudungun* (Chandía, 2014), bajo una licencia *Creative Commons*, de acceso y utilización tal que permite “copiar y redistribuir el material en cualquier medio o formato y remezclar, transformar y construir a partir del material, con la condición de dar crédito al origen del material, no hacer uso comercial del mismo y finalmente compartir de la misma forma el material así derivado” (Creative Commons, 2022). Esta licencia y posibilidad de usar el material alojado en el *Corpus Lexicográfico del Mapudungun* fue verificada con el responsable

del sitio web, el académico Andrés Chandía, por medio de correspondencia electrónica. Estas entradas del diccionario fueron cotejadas en todo momento con la versión original de la obra (Augusta, 1916), y una vez filtradas permitieron decantar una serie de unidades —entiéndase, unidades lexemáticas monomorfémicas—, que nos permiten el análisis de acuerdo con los focos de interés de este estudio.

Este proceso de filtrado será descrito en extenso en el sección siguiente *6. Criterios para el filtrado de las unidades* y sigue algunas de las recomendaciones que se han propuesto en la literatura sobre fonestadística (por ejemplo, Altmann, 2005 y Márquez, 2016). Una vez realizado el proceso de filtrado, el listado así educido servirá para constituir un corpus lingüístico electrónico que dé cuenta del tipo de variables que son de interés en este estudio.

Con respecto a las variables, en un estudio descriptivo de la estructura fonológica, estas permiten realizar diferentes operaciones estadísticas, como mediciones de frecuencia de ocurrencia. En este contexto, las variables que serán consideradas en este estudio se vinculan con los parámetros articulatorios que se han propuesto para la clasificación de las consonantes (por ejemplo, modos y puntos de articulación) y las vocales (por ejemplo, altura y dirección de la lengua). Todos estos indicadores serán desarrollados en la sección *7. Análisis y Resultados*.

6. Criterios para el filtrado de las unidades

Esta sección tiene por objetivo dar cuenta de los pasos que se siguieron para alcanzar un listado de unidades monomorfémicas lexemáticas, de acuerdo con nuestro objetivo de investigación general. En este contexto, se propusieron diversos criterios sustentados en la estructura lingüística del mapudungun, especialmente respecto de fenómenos fonológicos y gramaticales. De igual modo, se consideraron aspectos sociolingüísticos, como el filtrado de los préstamos léxicos —es decir, su extracción de manera tal que no sean parte del corpus—, debido a su posible efecto en las distribuciones fonemáticas del mapudungun, y también respecto de variaciones fonológicas ocasionadas por el contacto lingüístico con una lengua de comunicación mayoritaria como el español.

En vista del carácter polisintético y aglutinante del mapudungun, muchas entradas léxicas se trataban de complejos morfémicos, por lo que se decidió controlar esta variable por medio de la identificación y segmentación de todas las unidades gramaticales presentes en el *Diccionario mapudungún-español* de Augusta (1916/2017). Así, además de entregarnos las unidades léxicas, este análisis morfológico nos permite definir un listado de morfemas tipo. La tarea no es menor, dado que del total de 9703 entradas léxicas, aproximadamente 6000 manifiestan algún grado de complejidad morfológica y cada una de ellas debe ser segmentada en sus constituyentes. Asimismo, la presencia de residuos en la forma de sufijos improductivos hace esta tarea particularmente desafiante.

La existencia de fenómenos como la composición nominal y verbal —además de la incorporación de sustantivos— requería también de segmentar las entradas considerando estas unidades. En algunos casos, la segmentación de algunos compuestos nos permitió seleccionar formas para nuestro conteo final que de otro modo no habría sido posible obtener, al no estar definidas como formas independientes en el resto del diccionario.

Respecto de los criterios fonológicos, reservamos en esta sección un apartado a la variación fonemática, debido a su alta frecuencia entre las entradas del diccionario. En cuanto a ello, cabe

señalar que en mapudungun es posible distinguir dos fenómenos: por un lado, la *fluctuación fonemática*, que entraña la posibilidad del hablante de realizar dos formas léxicas que se diferencian solamente en un segmento (por ejemplo, **wentru** /wenʧsu/ vs. **wentro** /wenʧso/ ‘hombre’) y este cambio fonémico puede atribuirse a dinámicas intrasistémicas (Salas, 1976, 2006; Zúñiga, 2000). Por otro lado, se manifiesta la llamada *alternancia fonemática*, que también genera un par de lexemas (por ejemplo, **pewëñ** /peweɲ/ vs **pewen** /pewen/ ‘araucaria’), pero, a diferencia de la fluctuación de fonemas, el cambio puede atribuirse al contacto lingüístico asimétrico con la lengua mayoritaria: el castellano (Henríquez y Salamanca, 2015).

Finalmente, abordamos los préstamos lingüísticos, puesto que pueden presentar estructuras anómalas a la fonología de la lengua que los recibe. Por este motivo, se decidió controlar y filtrar la presencia de este tipo de lexemas, cuando su forma evidenciaba que provenían del español o el autor indicaba esta situación de manera expresa. Cabe señalar que este filtrado de elementos foráneos es algo que la literatura ya había considerado (Márquez, 2016), y, en nuestro caso, identificamos además los préstamos provenientes de la lengua quechua.

Finalmente, en esta sección, también se tratan otros fenómenos y dificultades que se tuvieron al momento de escoger algunas entradas para nuestro conteo final; esto con el objetivo de evaluar la robustez de nuestros criterios y asegurar la replicabilidad de nuestra metodología.

6.1. Morfología sufijal

Se identificaron, segmentaron y analizaron más de 70 sufijos durante el proceso de muestreo del corpus lingüístico. Esta tarea persigue dos objetivos: 1) controlar el efecto que tendría la inclusión o exclusión de estas unidades en las entradas léxicas, especialmente debido a la alta productividad de muchas de ellas; y 2) facilitar la confección de un listado de *morfemas tipo* que servirá para un estudio futuro que considere el conteo de la fonología segmental en el dominio de la morfología flexiva y derivativa. Para este segundo objetivo, solo se deben incluir aquellas formas que no tengan una contraparte léxica en el cuerpo del diccionario. Por ejemplo, es posible

rastrear el origen del sufijo de persistencia **-künu** en el lexema **künu-n** ‘v. poner, dejar, hacer’, por lo que en ese caso, no se considera la inclusión de ese sufijo.

A continuación, se presentan los 12 sufijos con mayor ocurrencia, seguido de los otros morfemas ordenados según niveles de productividad. Cabe señalar que en la sección *Anexos* se puede consultar una tabla que contiene todos los sufijos, su número de ocurrencias y los afijos que se podrían considerar para un conteo fonemático de la morfología flexiva y derivativa del mapudungun. Posteriormente, se presentan tres morfemas identificados como sufijos improductivos o *seudosufijos*, relevados en su propia sección, debido al alto número de realizaciones en las que se manifiestan.

6.1.1. Sufijos derivativos y flexivos

6.1.1.1. -n INF⁵

El sufijo más productivo es el infinitivo **-n**, con un alomorfo **-ün** (/ən/) tras consonante. Se contabilizaron 6671 formas con este sufijo. Su alto número se debe a que es la forma verbal escogida para encabezar las entradas del diccionario.

- (1) ko-n
agua-INF
‘v. traer agua’

5 En esta sección se utilizan las siguientes abreviaturas y glosas gramaticales: ABJ ‘movimiento hacia abajo’, ADJ ‘adjetivizador’, ADN ‘movimiento hacia adentro’, ADV ‘adverbializador’, AFR ‘afirmativo’, AFU ‘movimiento hacia afuera’, AG ‘agente’, AND ‘andativo’, APL ‘aplicativo’, ARR ‘movimiento hacia arriba’, BEN ‘benefactivo’, CAUS ‘causativo’, CIRC ‘movimiento circular’, CIS ‘cislocativo’, COL ‘colocación’, CONT ‘continuativo’, DES ‘desiderativo’, DIM ‘diminutivo’, EST ‘estativo’, EXTR ‘extracción’, FUT ‘futuro’, GRP ‘grupalizador’, INF ‘infinitivo’, INM ‘inmediatez’, INTS ‘intensificador’, INTR ‘interruptivo’, LOC ‘locativo’, MOV ‘movimiento’, NEG ‘negación’, NOM ‘nominalizador’, PERS ‘persistencia’, PL ‘pluralizador’, PLPF ‘pluscuamperfecto’, PROG ‘progresivo’, PROX ‘proximidad temporal o espacial’, RECR ‘recreativo’, REFL ‘reflexivo’, REP ‘reportativo’, RES ‘residuo’, RST ‘restaurativo’, RPR ‘residuo pronominal’, SIML ‘evento simulado’, SOC ‘relación social’, TEMP ‘temporal’, TRAV ‘movimiento a través’, VAN ‘acción en vano’, VBR ‘verbalizador’, VEN ‘venitivo’.

- (2) lef-ün
 correr-INF
 ‘v. correr’

Dada la alomorfía del sufijo **-(ü)n** y su estatus de *forma de cita*, es necesario indicar los criterios que permiten establecer si el segmento vocálico de la forma <ün> (/ən/) es un fenómeno epentético o parte de la estructura fonológica del radical verbal.

En primer lugar, nos basamos en la propia segmentación que realiza Augusta (1916), al separar todo radical de sus sufijos derivativos. Para ilustrar este criterio de segmentación, nos serviremos del par de lexemas **af-ün** ‘concluirse’ y **afü-n** ‘cocer’, que se diferencian, precisamente, en la presencia o ausencia de este segmento /ə/ en el tema verbal.

Augusta, en su obra original (1916), despliega estas entradas del siguiente modo (escritura original):

- (3) **af**|n concluirse. | -í ñi duam = afeluwn. || -naqn, n., consumirse enteramente. (p. 3)
 (4) **afü**||kelleñfe*, la chicharra. || -ln, afümn. || -mn, cocer. || -n cocer, madurar. (p. 3)

Como se observa en (3), para expresar el verbo **af-ün** /af-ən/ ‘concluirse’, no escribe la vocal epentética. En el caso de **afü-n** /afə-n/ ‘cocer’, en cambio, la forma de cita incluye la vocal <ü> (/ə/) como parte de la estructura fonológica del radical. Además, tras esta forma se incluyen las derivaciones a partir del uso de sufijos y otros recursos morfológicos.

Por otro lado, en el caso de que no exista esta segmentación por medio de la barra vertical en las entradas del *Diccionario mapudungún-español* de Augusta (1916), nos valimos de la presencia o ausencia de la sexta vocal <ü> ~ <ə>, al expresar el radical derivado. Las siguientes entradas ejemplifican esta situación, escritas según la versión original del diccionario.

- (5) ewün·n+, n., tener arcadas (p. 42)
- (6) alülün, n., tener acedía; estar muy adolorido. (p. 6)

En estos dos casos, seleccionamos los lexemas **ewün'** (/ewəṅ/) 'tener arcadas' y **alülü** (/alələ/) 'tener acedía', en vista de la forma en que se transcriben en la obra original; es decir, sin <ü> (después de la consonante nasal interdental), para el primer caso, y con este grafema, en el segundo.

En algunos lexemas, se produce alternancia entre dos formas: una terminada en <ü> y otra que no contiene esa vocal. En esos casos, se opta por seleccionar la forma que el autor ha relevado, por ejemplo, al desplegar una definición bajo la entrada pertinente. Cuando no fue posible seguir ese criterio, se optó por la forma que indicaba vitalidad, es decir, la forma con la vocal <ü> ~ <ə>, como parte del radical. A continuación se ejemplifica cada criterio con entradas del diccionario (1916):

- (7) foltrain, foltrayün, tr., matar con los dientes (p. 50)
- (8) a. üf|n*, tr. apretar con boquis las pajas (al techar) (p. 271)
- b. üf|n*, tr. (=füyəmn) apretar con boqui u otro medio para atar.

En (7), seleccionamos la primera forma relevada según la disposición que sigue el autor, previa resolución del grupo vocálico, es decir, **foltray** /folʈsaj/. En (8), puesto que ambas formas son relevadas en la disposición del diccionario, seleccionamos **üf|n** /əfə/, por ser la forma que indica vitalidad (véase sección 6.4. *Fluctuación y alternancia fonemática*).

En otras palabras, este alcance supone que el autor, en algunos casos, relevó una forma que terminó encabezando la entrada. En un contexto más general, seleccionamos esta forma para nuestro corpus. Cuando esta situación no ocurre, es decir, se despliega cada forma en su respectivo artículo lexicográfico, operamos seleccionando la forma que indica vitalidad. Un listado completo de estas formas se encuentra en el *Anexo I.2*.

Creemos que estos criterios son suficientes para establecer si un lexema radical finaliza o no en esta vocal, para así poder distinguirlo de la epéntesis producto de la afijación del morfema de infinitivo **-(ü)n**.

6.1.1.2. -tu VBR

El segundo sufijo más productivo es el verbalizador **-tu**, que aparece 863 veces. Su función no se limita a la derivación verbal, sino que también entrega significados gramaticales como restaurativo, aplicativo y adverbializador. Estas diversas funciones permiten explicar su alta ocurrencia en las entradas del diccionario.

(9) müchay-tu-n
michay-VBR-INF
'v. teñir con michay (cierto arbusto)'

(10) amu-tu-n
ir-RST-INF
'v. irse de vuelta'



(11) lef-kon-tu-n
correr-ADN-APL-INF
'v. embestir, atacar'

(12) ponwi-tu
interior-ADV
'adv. en el interior'

6.1.1.3. -le EST

El tercer sufijo más productivo es el estativo **-le**, contabilizado 766 veces. Cuando es precedido por una raíz u otro morfema terminado en consonante, se realiza con el alomorfo **-küle**, aunque esto puede variar según el territorio.

(13) fūcha longko-tu-le-n
grande cabeza-VBR-EST-INF
'v. tener los cabellos largos'

(14) lef-küle-n
correr-EST-INF
'v. estar corriendo'

6.1.1.4. -(ü)l CAUS

El causativo **-l** es el cuarto sufijo más productivo, con 636 instancias. Tras consonante, se realiza como **-ül**. Su alta productividad se debe, posiblemente, a que es un recurso que permite aumentar la valencia de los verbos, sin la necesidad de recurrir a innovaciones léxicas. Por ejemplo, en (15) se observa que, a partir de un verbo intransitivo, el sufijo **-l** permite crear una contraparte transitiva, algo que en castellano requiere de dos lexemas independientes: *venir* y *traer*.

(15) kupa-n > kupa-l-ün
venir-INF venir-CAUS-INF
'v. venir' 'v. hacer venir, traer'

(16) llako-l-ün
tranquilizarse-CAUS-INF
'v. tranquilizar (a alguien)'

(17) man kug-ül-ün
derecha mano-CAUS-INF
'v. pasar la mano derecha'

6.1.1.5. -(ü)w REFL

El reflexivo **-(ü)w** es el quinto sufijo en cuanto al número de ocurrencias. Sus 619 instancias se realizan con distintos alomorfos: **-w**, **-uw**, **-üw** y **-wü**, que responden a la posibilidad de selección de dos vocales epentéticas: /ə/ o /u/, con el fin de evitar configuraciones silábicas anómalas. Su alta productividad se debe a que permite la creación de verbos intransitivos, con valor de *reflexivo* o *reciprocidad*, a partir de cualquier tema verbal (simple o complejo).

(18) lif-w-ün
limpio-REFL-INF
'v. limpiarse'

(19) che-w-ün
persona-REFL-INF
'v. hacerse persona acaudalada'

6.1.1.6. -nge VBR

El sufijo **-nge** es el sexto sufijo más productivo, con 303 realizaciones. Equivale en castellano al verbo 'ser', que al unirse con sustantivos, adjetivos, etc., permite expresar una situación permanente.

(20) kure-nge-n
esposa-VBR-INF
'v. tener esposa, ser casado'

(21) chofü-nge-n
flojo-VBR-INF
'v. ser flojo'

6.1.1.7. -ma APL

El aplicativo **-ma** es el séptimo en cuanto a productividad, con 266 instancias. Tiene los siguientes alomorfos: **-ma**, **-ñma**, **-üñma**. Este sufijo permite agregar un objeto en la estructura sintáctica del verbo; por esto, es común su uso con verbos impersonales que significan fenómenos atmosféricos.

(22) ngilla-tu-ñma-n
comprar-APL-APL-INF
'v. hacerle rogativas y sacrificios (a la divinidad)'



(23) mawün'-ma-n
lluvia-APL-INF
'v. lloverle a uno'

6.1.1.8. -künu PERS

El octavo sufijo con mayor productividad es el de persistencia **-künu** 'PERS', con 241 realizaciones. En algunas entradas, tiene un valor más bien adverbial, puesto que suele ocurrir con verbos que indican posturas.

(24) nūla-kūnu-n
abrir-PERS-INF
‘v. dejar abierto, abrir’

(25) payl’a-kūnu
de. espalda-ADV
‘adv. de espaldas’

6.1.1.9. -(ü)m CAUS

El causativo **-(ü)m** ‘CAUS’ es el noveno sufijo más productivo, con 237 instancias. Tiene un alomorfo **-üim** cuando sigue a raíces u otros morfemas terminados en consonante.

(26) narfü-m-ün
humedecerse-CAUS-INF
‘v. mojar, humedecer’



(27) nor-pūra-m-ün
derecho-ARR-CAUS-INF
‘v. enderezar, levantar derecho’

6.1.1.10. -tüku COL

El décimo sufijo más productivo es el que corresponde a la gramaticalización del verbo ‘colocar’ **tüku-n**, con las siguientes formas: **-tüku**, **-ntüku**, **-ñtüku** y **-üntüku**. Se contabilizaron 216 lemas con este sufijo. Compárese su significado de *colocación*, presente al lado derecho, en las siguientes duplas:

- | | | | |
|------|--|---|---|
| (28) | malal-ün
corral-INF
‘v. hacer un corral’ | > | malal-tüku-n
corral-COL-INF
‘v. cercar (= colocar el cerco)’ |
| (29) | tralof-ün
golpear-INF
‘v. golpear’ | > | tralof-tüku-n
golpear-COL-INF
‘v. dar un golpe fuerte (= encajar un golpe)’ |

6.1.1.11. -ka CONT

El undécimo sufijo más productivo es el continuativo **-ka**, que unido a temas verbales permite crear nuevos verbos, muchas veces con un significado de acción continuada o reiterativa. Suele formar ensambles morfémicos con los sufijos **-w** REFL y **-l** CAUS, también con una alta productividad. Se registran más de 214 instancias con este sufijo.

- (30) kiñe-ka-n
uno-CONT-INF
‘v. ser igual, lo mismo’



- (31) nie-ka-n
tener-CONT-INF
‘v. retener’

6.1.1.12. -we LOC

El duodécimo sufijo más productivo es **-we**, también con 214 realizaciones. Este sufijo se une a temas verbales para crear una nominalización con significado de locación, instrumento, y, en algunos casos, temporal, lo que posiblemente explica su alta productividad para la generación de neologismos.

- (32) trana-chadi-we
machacar-sal-LOC
'v. mortero para moler sal'
- (33) anü-m-ka-we
asentarse-CAUS-CONT-LOC
'v. plantío, huerto'
- (34) epu-we-mom
dos-LOC-NOM
'v. hace dos días'

6.1.1.13. Otros sufijos de alta, mediana y baja productividad

Otros sufijos con una alta productividad —de 100 a 162 instancias— son los de movimiento **-nag** 'ABJ' y **-kon** 'ADN', el sufijo de movimiento errático **-yaw** 'CIRC', los aplicativos **-el** 'APL' y **-ye** 'APL', el locativo **-pa** 'CIS', la gramaticalización del verbo **-entu** 'EXTR', el sufijo de agente **-fe** 'AG', y el sufijo de persistencia **-nie** 'PERS'.

Los sufijos que tuvieron un rango de productividad entre 55 y 98 realizaciones son el verbalizador y grupalizador **-ntu** 'GRP', el sufijo andativo **-me** 'AND', el pluralizador **-ke** 'PL', los sufijos de movimiento **-püra** 'ARR' y **-tripa** 'AFU', el locativo **-pu** 'LOC', el adjetivizador **-fal** 'ADJ', y el sufijo de movimiento interrumpido **-rpa** 'INTR'.

En un rango entre 10 y 48 instancias, se encuentran los siguientes sufijos: el adjetivizador **-chi** 'ADJ'; los nominalizadores **-peyüm** 'NOM', **-lu** 'NOM' y **-el** 'NOM'; un verbalizador improductivo **-kü** 'VBR'; los sufijos de movimiento interrumpido **-(ü)rpu** 'INTR', **-rume** 'INTR' y **-ru** 'INTR'; el sufijo de acción recreativa **-kantu** 'RECR'; el sufijo social **-wen** 'SOC'; el adverbializador **-kechi** 'ADV'; un nominalizador improductivo **-ñ** 'NOM'; los sufijos de negación **-no** 'NEG', **-ki** 'NEG' y **-la** 'NEG'; el benefactivo **-ñpe** 'BEN'; un sufijo diminutivo improductivo **-kiñ** 'DIM'; el sufijo de

futuro **-a** ‘FUT’; el sufijo de movimiento **-amu** ‘AND’; un nominalizador/adjetivizador improductivo **-nte** ‘NOM’; el intensificador **-kütuye** ‘INTS’; el sufijo de proximidad temporal **-pe** ‘PROX’; la gramaticalización del verbo **-wül** ‘INTS’ con significado de ‘golpear’; y el sufijo de aspecto progresivo **-meke** ‘PROG’.

En el último lugar, se encuentra una serie de sufijos con un rango entre una a nueve instancias: el intensificador **-(n)kepi** ‘INTS’, el simulativo **-faluw** ‘SIML’, el sufijo de acción en vano **-püda** ‘VAN’, los sufijos de movimiento interrumpido de manera continua o reiterada **-yeküme** ‘MOV’ y **-yeküpa** ‘MOV’, los nominalizadores **-mum** ‘NOM’, el sufijo de movimiento hacia el lugar de habla **-küpa** ‘VEN’, los sufijos afirmativos **-lla** y **-lle** ‘AFF’, un sufijo de acción en beneficio propio **-ñmu** ‘BEN’, un sufijo nominalizador temporal para acción de rasgo constante **-yüm** ‘TEMP’, un sufijo nominalizador **-kün** ‘NOM’, el sufijo de reporte o evidencialidad **-(ü)rke** ‘REP’, el sufijo de movimiento circular **-tiaw** ‘CIRC’, el sufijo de movimiento a través de un curso de agua o camino **-n’o** ‘TRAV’, un sufijo impersonal improductivo **-am** ‘NOM’, un sufijo pluralizador **-küf** ‘PL’, un verbalizador improductivo **-ku** ‘VBR’, un sufijo de inmediatez **-fem** ‘INM’, dos sufijos nominalizadores improductivos **-en** ‘NOM’ y **-tra** ‘NOM’, un nominalizador de acción pasada **-wma** ‘NOM’, el sufijo de pluscuamperfecto **-wye** ‘PLPF’ y un sufijo adverbializador temporal improductivo **-ülke** ‘ADV’.

Igualmente, en este último rango existe una serie de residuos de los que no es posible establecer un significado, pero en general remiten a nominalizaciones y otros procesos morfológicos. Estos sufijos son: **-y**, **-ü**, **-u**, **-mu**, **-pi**, **-ken’**, **-ngey**, **-ngell(u)**, **-üll**, **-ngelüm**, **-ngen**, **-all**, **-i**, **-lül**, **-muñ**, **-ñchü**, **-ngken**, **-cha**, **-che**, **-chil**, **-del**, **-dü** y **-di**.

Cabe mencionar, también, que son de baja productividad los residuos que resultan de la segmentación de las formas de los pronombres personales y posesivos: **-i**, **-n**, **-iñ**, **-u**, **-e**, **-m**, **-ng**, **-ün**, **-che**. Dicha interpretación se basa en la propuesta de Smeets (2008), contenida en la **Tabla 6.1**.

Tabla 6.1: Segmentación de los pronombres.

realización	segmentación	Significado	Realización	segmentación	significado
<i>iñche</i>	i-n-che	P. personal 1SG	<i>emu</i>	e-m-u	P. conmitativo 2DL
<i>iñchiñ</i>	i-n-che-iñ	P. personal 1DL	<i>emün</i>	e-m-ün	P. conmitativo 2PL
<i>iñchiw</i>	i-n-che-u	P. personal 1PL	<i>yu</i>	i-u	P. posesivo 1DL
<i>eymi</i>	e-i-m-i	P. personal 2SG	<i>Iñ</i>	i-iñ	P. posesivo 1PL
<i>eymu</i>	e-i-m-u	P. personal 2DL	<i>mi</i>	m-i	P. posesivo 2SG
<i>eymün</i>	e-i-m-ün	P. personal 2PL	<i>mu</i>	m-u	P. posesivo 2DL
<i>engu</i>	e-ng-u	P. personal DL	<i>mün</i>	m-ün	P. posesivo 2PL
<i>engün</i>	e-ng-ün	P. personal PL			

6.1.2. Sufijos improductivos: -dũ, -kü, -tu

Durante nuestro proceso de filtraje, se identificaron dos formas, **-dũ** y **-kü**, que posiblemente corresponden a *seudoafijos*, y un sufijo improductivo **-tu** VBR, que se adosa a radicales verbales con más de dos sílabas. En muchas instancias, su estatus era evidente al existir alternancia con otros lemas que se realizaban sin esa sílaba, aunque también se evidenciaron realizaciones donde no existía esa alternancia. Con el objetivo de evitar una sobredimensión de los fonemas que los componen, se tomó la decisión de controlar sus realizaciones durante el proceso de filtrado en todas las situaciones.

6.1.2.1. Sufijo improductivo -dũ

Se identificó un posible residuo morfológico **-dũ** en ciertas alternancias léxicas y al final de algunos verbos trisilábicos. Su estatus de *seudoafijo* se refuerza al considerar que existe un campo semántico que comparten un gran número de entradas, y que se puede definir como ‘sostener un daño superficial’.

Respecto de la selección de las formas, cuando fue posible identificar un par alternante sin el *seudoafijo* **-dũ**, se escogía este para ser parte del corpus, bajo el supuesto de que se trata

efectivamente de un vestigio morfológico y así evitar que su frecuencia alterara nuestro análisis. En el siguiente listado de formas, los lexemas seleccionados se colocan en negrita:

1. **chomü-dü-n** Cfr. **chomü-n** ‘secarse y caer (las flores u hojas)’
2. **kafü-dü-n** ‘v. raspar’ Cfr. **kaf-ün** ‘v. desbastar’
3. **nükü-dü-n** ‘v. humedecerse’ Cfr. **nükü-m-tu-n** ‘v. ponerse bajo techo para no mojarse’
4. **ngüchi-dü-n** ‘v. apachurrar, comprimir’ Cfr. **ngüri-n** ~ **ngüchi-n** ‘v. apretar, comprimir’
5. **ngüñkü-dü-n** Cfr. **ngüñkü-n** ‘v. refregar, frotar’
6. **wilngo-dü-n** ‘v. descoyuntarse (huesos)’ Cfr. **üllngüd** ~ **wilngüd** ‘s. coyuntura’
7. **wil’pü-dü-n** Cfr. **wil’pun-ün** ‘v. recocerse (papas, carne, etc.)’
8. **wiñu-dü-n** ‘v. carmenar, desenredar (lana)’ cfr. **wiñü** ‘adj. hebra (de hilo)’

En otros casos, no fue posible identificar una forma alternante sin la sílaba **-dü**. De todas maneras, con el objetivo de evitar una sobrecuantificación de un elemento que puede tratarse de un *seudoafijo*, se decidió controlar su realización, seleccionando solo la forma base, sin la aglutinación de este elemento. En el siguiente listado, las formas seleccionadas se escriben en negrita.

9. **chafo-dü-n** ‘v. destrozarse; destrozar’
10. **chalmu-dü-n** ‘v. tener pereza o debilidad’
11. **chepi-dü-n** ‘v. tragar mal’
12. **chewpü-dü-n** ‘v. exterminar; exterminarse’
13. **fochü-dü-n** ‘v. extraer chupando’
14. **iyudü-n** ‘v. pulir’
15. **küchü-dü-n** ‘v. refregar’
16. **llikü-dü-n** ‘v. macollarse’
17. **lollo-dü-n** ‘v. apostemarse (una lesión)’
18. **ñikü-dü-n** ‘v. dar comezón a uno, escocerle, producirle malestar’
19. **ngacho-dü-n** ‘v. doblarse, encogerse’
20. **ngafa-dü-n** ‘v. hacer una impresión (en la piel)’

- | | | |
|-----|-------------------|--|
| 21. | pakü-dü-n | ‘v. estar preñado el trigo’ |
| 22. | pewpü-dü-n | ‘v. dar un traspie’ |
| 23. | püllü-dü-n | ‘v. deshacerse algo (al apretar); ulcerarse’ |
| 24. | trike-dü-n | ‘v. hacer volteretas’ |
| 25. | willu-dü-n | ‘v. quitar pasando la mano’ |
| 26. | wiyu-dü-n | ‘v. andar en hilera’ |
| 27. | yellu-dü-n | ‘v. desatarse, disolverse’ |

6.1.2.2. Sufijo improductivo -kü

Al igual que en el caso del residuo **-dü**, en varios verbos trisilábicos se identificó una sílaba con la forma **-kü**, que podría corresponder a un sufijo improductivo.

En primer lugar, su forma remite a la raíz etimológica ***küw** ‘v. ir’, que se realiza en las entradas **küw-me-n** ‘v. ir, seguir adelante, proseguir’ (Augusta, 1916/2017, p. 91), además de las fluctuaciones **kuw-me-n** y **kow-me-n**, con el mismo significado. En fuentes históricas, también se registra como **cun** /kun/ ‘ir’ (Valdivia, 1606), **cüun** /küwn/ ‘proseguir’ (Febrés, 1764).

De igual modo, en la actualidad, es posible asociarlo a los sufijos compuestos **-yekowme** ~ **-yeküme** ‘ir en aumento, continuar (en la idea) ejecutándose la acción del verbo’ y **-yeküpa** ‘continuar la acción al venir’; además del verbo **küpan** ‘venir’ y **künapan** ‘venir acá’.

En algunos casos, la decisión respecto de cómo segmentar el radical se facilita debido a la existencia de un lema alternativo sin la sílaba **-kü**, con el mismo significado o un significado relacionado. En esos casos, al igual que con el sufijo improductivo **-dü**, se optó por seleccionar la forma sin **-kü**. En otros casos, independiente de no existir esta alternancia, se prefirió controlar la frecuencia de este pseudoafijo, seleccionando solamente la parte radical previa a su aglutinación. En adelante, el uso de negritas indica cómo se seleccionaron las formas.

Si bien no es posible establecer de manera clara su significado gramatical, es posible agrupar algunos de los lexemas que contienen **-kü** siguiendo principios semánticos y estructurales, como se hará a continuación.

6.1.2.2.1. Seudoafijo -kü unido a lexemas reduplicados

Existe un caso donde el seudoafijo **-kü** se une a bases reduplicadas y su significado remite a la expresión de onomatopeyas e interjecciones. En los ejemplos a continuación, la forma en negrita remite a la forma seleccionada durante este estadio para continuar en el proceso de selección. Vale señalar que la forma final del lexema que se establece para nuestro conteo puede simplificarse aún más una vez ejecutado el criterio de controlar la reduplicación.

- | | | | |
|-----|-----------------------|--|---|
| 1. | meme-kü-n | | < meme-n , meme-ka-n ‘v. balar’ |
| 2. | rürü-kü-n | ‘v. hacer bullicio, zurriar.’ | < rürü ‘s. zancudo’ |
| 3. | trorror-kü-n | ‘v. hacer sonar las coyunturas’ | < tror-kü-m ‘s. hernia, ruptura |
| 4. | yofyof-kü-n | | < yofyof-ün ‘v. hervir, borbollar’ |
| 5. | mama -kü-n | ‘v. Pangui. mugir la vaca.’ | |
| 6. | momo -kü-n | ‘v. mugir; pitar (vapor)’ | |
| 7. | pongpong -kü-n | ‘v. hincharse (la masa para hacer pan, cadáveres)’ | |
| 8. | rara -kü-n | ‘v. hacer ruido el mar, el viento, un arroyo o mucha gente.’ | |
| 9. | tratra -kü-n | ‘v. cacarear’ | |
| 10. | wawa -kü-n | ‘meter mucha bulla (gentío)’ Cfr. wakeñ ‘s. ruido’ | |

6.1.2.2.2. Seudoafijo -kü y lexemas que remiten a sonidos

Una asociación con sonidos también se presenta en estos lemas, pero sin haber una reduplicación en el lexema radical:

- | | | |
|----|---------------|---|
| 1. | mutrir-kü-n | < mutrir-ün ‘v. estimular el caballo haciendo sonar los labios.’ |
| 2. | trongong-kü-n | < trongong-ün ‘v. llamar golpeando’ |
| 3. | wangar-kü-n | < wangar-ün ‘v. mascar con ruido’ |

6.1.2.2.3. Seudoafijo -kü y otros casos

Junto con los lemas que entrañan cierto simbolismo fónico, existe un considerable número de lexemas con otros significados, con los cuales ocurre el seudoafijo **-kü**:

1. düching-kü-n ‘s. rayo del sol’ < **düching**-ün antü ‘v. radiar el sol’
2. eteng-kü-n < **eteng**-ün ‘v. exclamar enojo’ ‘v. radiar el sol’
3. nütrüng-kü-n < **nütrüng**-künu-n ‘v. extender a lo largo’
4. ngüde-kü-n-tüku-n ‘v. arrojarse’ < **ngüdef**-künu-n ‘v. encajar, meter’
5. trelüng-kü-n < **trelüng**-küle-n ‘v. estar aclarado’
6. trolong-kü-n < **trolong**-ün ‘v. golpear’
7. üyer-kü-n < **üyer**-ün ‘v. mostrar los dientes’
8. welüng-kü-n < **welüng**-ün ‘v. sentir aflicciones en el pecho’
9. weñang-kü-n < **weñang**-ün ‘v. entristecerse’
10. winüng-kü-n < **winüng**-ün ‘v. palpitar’
11. üchang-kü-n ‘v. despedazar con la boca’ < **üchañ**-ün ‘v. mascar y chupar el jugo’
12. **künung**-kü-n ‘v. trotar’
13. **liwang**-kü-n ‘v. echar las orejas hacia adelante (animales)’
14. **mellüng**-kü-n ‘v. mecer al chiquillo (con la cuna)’
15. **nüni**-kü-n ‘v. hormiguar, hervir, abundar’
16. **pidong**-kü-n ‘v. picotear’
17. **piltang**-kü-n ‘v. descuartizar.’
18. **rüken**-kü-n ‘v. hincharse, prenderse (el estómago)’

6.1.2.3. Sufijo improductivo -tu

Un gran número de entradas léxicas de tres sílabas se realizan con **-tu** ‘VBR’, donde no es posible identificar una alternancia con otros lexemas que no tengan esa sílaba.

Es probable que la forma radical cayera en desuso y por eso no se registra en Augusta (1916/2017). Independiente de la inexistencia de estas formas alternantes, se prefirió controlar la alta frecuencia de esta sílaba, seleccionado solo la parte del radical previa a su realización. Esto se hizo solo en el caso de lexemas trisilábicos o polisilábicos. A continuación se registran esas entradas:

1. **chüllkem**-tu-n ‘v. guarecerse de la lluvia’
2. **domdomü**-tu-n ‘v. rezongar’
3. **eyü**-tu-n ‘v. gemir’
4. **kawli**-tu-n ‘v. rasguñar’
5. **koa**-tu-n ‘v. bramar’
6. **kolle**-tu-n ‘v. llegar sin estar invitado’
7. **kupe**-tu-n ‘v. tr. parchar ropa o loza’
8. **küllpay**-tu-n ‘v. poner zancadilla a alguno’
9. **küña**-tu-n ‘v. Tomar de la mano’
10. **küyüm**-tu-n ‘v. consolar con comida, regalos’
11. **l’uka**-tu-n ‘v. insultar, retar con insultos’
12. **llalli**-tu-n ‘v. visitar’
13. **llapa**-tu-n ‘v. parchar’
14. **momol**-tu-n ‘v. amontonar palos y prenderles fuego’
15. **müñe**-tu-n ‘v. bañarse’
16. **müñu**-tu-n ‘v. realizar cierta práctica cultural’
17. **mütra**-tu-n ‘v. luchar’
18. **mütrig**-tu-n ‘v. machacar para desgranar’
19. **nikef**-tu-n ‘v. torcer la cara por un lado hacia arriba (probl)’
20. **ñido**-tu-n ‘v. masturbarse (probabl. mutuamente)’
21. **ngafül**-tu-n ‘v. revisar’
22. **poshkeñ**-tu-n ‘v. tomar por favorito’
23. **rale**-tu-n ‘v. parchar’
24. **rüfe**-tu-n ‘v. hocicar (el chancho).’

- | | | |
|-----|---------------------|--|
| 25. | rütre -tu-n | ‘v. pellizcar’ |
| 26. | shuche -tu-n | ‘v. descabecear (el trigo, la cebada)’ |
| 27. | üllwem -tu-n | ‘v. abstenerse (comida, sexo)’ |
| 28. | üyñif -tu-n | ‘v. recoger lo que dejaron otros’ |
| 29. | wifü -tu-n | ‘v. montar un caballo de un salto’ |
| 30. | wüfchi -tu-n | ‘v. acechar’ |
| 31. | wütrin -tu-n | ‘v. mirar’ |

6.2. Composición nominal

La composición nominal y verbal, además de la incorporación nominal en la morfología verbal, son fenómenos muy productivos en mapudungun (Zúñiga, 2014). Una de las dificultades que se presenta al momento de identificar estos compuestos es la forma en que se escribe la palabra, puesto que no existe una convención ortográfica estandarizada al respecto (Zúñiga, 2014). En efecto, una forma compuesta puede aparecer escrita junta, como en **topel’** /topel̥/ ‘s. parte dorsal del cuello’ (de **to** /to/ ‘musculatura a ambos lados del espinazo cervical’ y **pel’** /pe̥l̥/ ‘cuello, garganta’), o separada, como en **ülcha domo** /ə̃ʎtʃaθomo/ ‘s. niña joven’.

Durante el proceso de análisis, se consideraron todas las instancias de compuestos nominales, para poder luego reportar cuántas son y cómo se manifestaron. En específico, nos referiremos a los compuestos que ocurren en las palabras seleccionadas, con el fin de focalizarnos en cómo arribamos a algunos de los lexemas que forman parte de nuestro estudio.

En muchos de estos casos, fue necesario mantener uno de los elementos del compuesto al no aparecer en otro lugar del diccionario, pero controlando la realización del otro elemento. Por ejemplo, en **yupiange** /jupiaŋe/ ‘s. pomulos de la cara’, se mantuvo el lexema **yupi** /jupi/, a pesar de solo manifestarse en este caso. Lo anterior es sin perjuicio de que aquello que podía ser un sufijo se segmentara siguiendo los criterios para la parte de morfología derivativa y flexiva.

Otros ejemplos son:

1. **chomüll**-ko /tʃoməɫ-ko/ ‘s. caracol de mar’
2. **fid**-pilun /fiθ-pilun/ ‘s. cerumen’
3. **kachan**-l'awen' /kaʃan-l̥awen̩/ ‘s. cierta hierba’

En la mayoría de los casos, se seleccionó el primer elemento, pero también es posible que en algunos casos fuera el segundo elemento:

4. lollom-**udu** /loɫom-uθu/ ‘s. ciertos caracoles’
5. küla-**fodi** /kəla-foθi/ ‘s. cierta hierba’

Respecto de los compuestos verbales, estos se dividen en verbos seriales y en incorporación nominal. Se procedió de igual manera que en el caso de los sustantivos. Estos son algunos ejemplos:

1. **afkü**-duam-ün /afkə-θuamən/ ‘v. suspirar, tener pena’
2. **kutraf**·**kutraf**-tu-kug-ün /kuʃafkuʃaftu-kuɯ-ən/ ‘v. dar palmadas repetidas’
3. witra-**kütrafü**-n /wiʃa-kəʃafə-n/ ‘v. tener puntadas por correr mucho’
4. tren-**trepül**-ün /ʃen-ʃepəl-ən/ ‘v. ir dando traspiés’

De aproximadamente 2900 lexemas seleccionados, se identificaron alrededor de 164 compuestos, correspondiente a casi el 5,63% del total. En cuanto a las clases de palabras representadas, a grandes rasgos, 126 lexemas se escogieron de entradas correspondientes a sustantivos compuestos, y 32 a verbos compuestos (incluyendo incorporación nominal).

6.3. Extranjerismos

6.3.1. Cognados con el quechua

La influencia cultural y lingüística del imperio incásico —quechua-hablante— en la nación mapuche ha sido evidenciada en diversos trabajos (Véase Moulian et al., 2015). El quechua puede ser caracterizado como una familia lingüística (Adelaar, 2004, p. 168), divisible en dos grupos principales según la zona geográfica (Mineduc, 2021).

Tempranamente, en sus *Estudios Araucanos*, Lenz (1895-1897) señala que algunas palabras del mapudungun podrían provenir del quechua. Durante las últimas décadas, se ha abordado este tema con mayor profundidad, lo que incluye la creación de algunos repertorios léxicos, como los contenidos en Smeets (2008) y Moulian et al. (2015). A fin de tener nuestra propia perspectiva e inventario de los posibles préstamos provenientes del quechua, se cotejaron estos repertorios con las entradas del diccionario de Laime (2007).

Así, se registraron 24 vocablos que muy probablemente provienen de esta lengua. A modo de ilustración, se presenta una lista con los vocablos en mapudungun y sus glosas en español (primera y segunda columna, respectivamente), además de los ítems del quechua con los que se vinculan y sus glosas respectivas (tercera y cuarta columna, respectivamente). La definición en esta última lengua y su número de página remiten al diccionario antes mencionado.

1. /atʃawəʎ/	s. gallina	< /waʎpa/	s. gallina (p. 128)
2. /ampi/	v. aplicar remedio	< /hampij/	v. sanar aplicando remedios. (p. 34)
3. /awka/	adj. alzado, rebelde	< /awqa/	adj. enemigo; contrario (p. 16)
4. /tʃaʎwa/	s. pescado	< /tʃaʎwa/	s. pescado (p. 19)
5. /tʃaʎki/	s. cierta cecina	< /tʃʰarki/	s. cecina (p. 27)
6. /tʃilka/	s. libro, carta	< /qiʎqa/	s. escritura (p. 89)
7. /tʃina/	s. esclava, sirvienta	< /tʃina/	adj. hembra. fig. sirvienta (p. 21)
8. /tʃutʃoka/	s. maíz tostado	< /tʃutʃoka/	s. maíz medio hervido y secado
9. /kaka/	s. cierto familiar	< /kaka/	s. cierto familiar (p. 42)

10. /kamaɾ/	s. guardián, pastor	< /kamaj/	s. ocupación, trabajo (p.43)
11. /kaɾka/	v. asar	< /kankaj/	v. asar (p.43)
12. /ka.ɾa/	s. ciudad	< /pukara/	s. fortaleza, castillo (p. 79)
13. /kaɾfu/	s. pasto, hierba	< /q'atfu/	s. hierba verde, pasto (p. 96)
14. /kawətu/	s. catre de cama	< /kawitu/	s. hamaca, mecedora (p.43)
15. /minka/	v. cierta práctica social	< /mink'a/	s. cierta práctica social (p. 68)
16. /miɱki/	adj. dulce; miel	< /misk'i/	adj. dulce, sabor de miel (p. 68)
17. /pataka/	num. Ciento	< /patʃak/	adj. cien. (p. 75)
18. /poθkol/	s. levadura	< /p'usqu/	s. levadura, fermento. (p. 86)
19. /putʃu/	v. sobrar	< /putʃu/	s. residuo, sobra (p. 79)
20. /pətʃə/	adj. pequeño, poco	< /pisi/	adj. poco, escaso, limitado. (p. 79)
21. /ɾaki/	v. contar, calcular	< /rak'i/	s. partición, división. (p. 99)
22. /wampo/	s. canoa	< /wamp'u/	s. barco, embarcación (p. 129)
23. /waɾka/	num. mil	< /waranqa/	adj. mil (p. 131)
24. /jana/	s. criado	< /jana/	s. esclavo (p. 137)

6.3.2. Cognados con el castellano

Debido al contacto lingüístico entre el castellano y el mapudungun, se ha incorporado a esta última lengua una gran cantidad de términos, con la readecuación fonológica y gramatical correspondiente (Catrileo, 2010).

Se identificaron 124 posibles préstamos del castellano. Al igual que en el caso del quechua, no se seleccionó ninguno de estos términos para ser parte de nuestro corpus. A continuación se despliegan algunos ejemplos (se incluye una posible derivación del cognado a partir de una transcripción fonológica para el castellano):

1.	/kawtʃu/	< /'gautʃo/	adj. soltero, soltera
2.	/waɾia/	< /'guardia/	s. ciudad
3.	/kojotʃo/	< /ko'jotʃo/	s. coyochó

4.	/lofo/	< /'lobo/	adj. huidor
5.	/malkaθa/	< /'maskara/	s. careta
6.	/peŋa/	< /'peka/	adv. indiscretamente
7.	/pefejtʉ/	< /be'siʝo/	v. mimar
8.	/palao/	< /'palido/	adj. bayo
9.	/maθakal/	< /'maskara/	s. un adorno
10.	/lonʝkaŋ/	< /longa'nisa/	s. salchicha
11.	/tołto/	< /'toldo/	s. toldo
12.	/etʃal/	< /e'tʃar/	v. echar
13.	/empeŋa/	< /empe'ñar/	v. dar en prenda, empeñar
14.	/estipu/	< /es'tribo/	s. estribos
15.	/fawəl/	< /ba'ul/	s. baúl
16.	/fisaʝo/	< /bi'saro/	adj. perito, baquiano
17.	/fela/	< /'bela/	s. vela para dar luz

En un caso, si bien no se contó con una hipótesis, en virtud de la estructura fonológica atípica del lexema (esto es, grupo vocálico al final de palabra), se decidió excluirla del conteo final:

12. /templaria/ < ? (P.) s. vaso de plata para tomar chicha

El listado completo de posibles préstamos del castellano puede ser consultado en la sección de *Anexos*.

6.4. Fluctuación y alternancia fonemática

6.4.1. Introducción

A continuación se da cuenta de una serie de criterios de selección, motivados por la variación fonológica que se manifiesta en algunas de las entradas del *Diccionario mapudungún-español* de Augusta (1916/2017). El objetivo de esta sección es consignar de manera expresa el lema

escogido cuando había dos formas que se distinguían por un fonema y se asociaban a un mismo significado.

Como hemos señalado en algunos párrafos precedentes, dentro de la variación fonológica existe una distinción que realizan Henríquez y Salamanca (2015), al hablar de alternancias vs. fluctuaciones. La *alternancia fonemática* se entiende como “motivada por diversos factores sociolingüísticos” (p. 212) y como parte de procesos de cambio lingüístico en curso. Por otro lado, la *fluctuación fonemática* se entiende como un “fenómeno característico y generalizado del mapudungun” (p. 212), donde el hablante puede realizar un ítem léxico con uno u otro fonema (Salas, 2006).

Para el caso de las alternancias fonemáticas, el criterio aplicado selecciona, entre dos lemas, aquél que contiene un segmento relevado como *Realización Indicadora de Vitalidad* (RIV). Este concepto debe entenderse dentro del marco del contacto lingüístico asimétrico que se produce entre la sociedad mapuche y la sociedad mayoritaria hispano-hablante, donde se producen cambios en las estructuras lingüísticas de la lengua minoritaria por influencia de la lengua mayoritaria (Henríquez y Salamanca, 2015).

En el plano fonológico, esos cambios pueden determinarse por medio de la variación que sufren los segmentos vernaculares, denominados RIV, al actualizarse con rasgos fonéticos atribuibles a la lengua de comunicación mayoritaria, denominados *Realizaciones Indicadoras de Transferencia* (RIT) (Henríquez y Salamanca, 2015; Henríquez, 2016). A modo de ilustración de la *alternancia*, téngase presente el par de lemas **llüwa** (/ʎəwa/) ~ **lliwa** (/ʎiwa/), con el significado de ‘advertir’. La elección de un lexema entre ambas formas superficiales se realiza considerando la forma que contiene un segmento atribuible a una RIV; vale decir, en este caso, **llüwa** (/ʎəwa/), al realizarse con la vocal /ə/.

La **Tabla 6.2**, a continuación, enumera los segmentos que pueden ser caracterizados como RIV respecto de las RIT, además de incluir un conteo general de los casos recogidos en nuestro muestreo.

Tabla 6.2: Segmentos que son RIV, respecto de los que son RIT.

RIV	RIT	N° de casos
<i>en todo contexto</i>		
/ə/	/i, e, a, o, u/	148
/ʌ/	/l/, /ɲ/, /ɲ/ (en contexto asimilatorio), ∅	36
/l̥/	/l/, /n/	7
/ɲ̥/	/n/, /ɲ/, /l/	5
/ɲ/	/f/, /m/, /n/, /ɲ/, /p/, /j/, ∅	17
/t̥s/	/t̥j/, /t/, /m/, /l/, /f/, ∅	32
/θ/	/s/, /t̥j/, /n/, /j/, /l/, /f/, /ɲ/, /p/, ∅	21
/ɰ/, /f/, /t̥/	/l/, /j/, /t̥j/, ∅	9
<i>en el postmargen</i>		
/ɲ/	/n/, /m/, /ɲ/ (en contexto asimilatorio), ∅	20
/f/	/m/, /l/, ∅	6
/w/	/w/, /l/	4
Total:		305

Además de alternancias, durante el muestreo también se registraron 162 casos de lemas que evidencian *fluctuación fonemática*. Diversos pares de segmentos participaban de este tipo de variación, a saber: /o/ ~ /u/; /a/ ~ /e/; /l̥/ ~ /ʌ/; /ɰ/ ~ /f/, /θ/ ~ /ɰ/; /t̥j/ ~ /t/, entre muchos otros.

La **Tabla 6.3** registra los conteos de 58 tipos de fluctuaciones, ordenadas de acuerdo a la frecuencia de casos. El fonema a la izquierda (o símbolo de elisión, según corresponda) señala la selección realizada y, a la derecha, su alternativa. En *Anexos*, se despliegan los casos específicos, ordenados de acuerdo con los fonemas que participan en la variación.

Tabla 6.3: Fluctuaciones fonemáticas. Se indica el segmento seleccionado, su alternancia y el número de casos.

Fluct.	#	Fluct.	#	Fluct.	#	Fluct.	#
/θ/ ~ /ʃ/	23	/e/ ~ /a/	2	/l̥/ ~ /θ/	1	/p/ ~ /f/	1
/ɹ/ ~ /ʃ/	14	/t/ ~ /tʃ̃/	2	/f/ ~ /θ/	1	/p/ ~ /n/	1
/ʃ/ ~ /θ/	13	/ɲ/ ~ /n/	2	/ɬ/ ~ /θ/	1	/j/ ~ /m/	1
/o/ ~ /u/	10	∅ ~ /w/	2	/m/ ~ /l/	1	/l/ ~ /n/	1
/ɹ/ ~ /θ/	10	/j/ ~ ∅	2	/m/ ~ /n/	1	/ɲ/ ~ /ɬ/	1
/u/ ~ /o/	8	/m/ ~ ∅	2	/m/ ~ /ɲ/	1	/ɲ/ ~ /l/	1
/l̥/ ~ /ɬ/	6	/n/ ~ ∅	2	/f/ ~ /n/	1	/tʃ̃/ ~ /ɹ/	1
/a/ ~ /e/	5	/ɲ/ ~ ∅	2	/f/ ~ /p/	1	/w/ ~ ∅	1
/tʃ̃/ ~ /t/	4	/i/ ~ /e/	2	/f/ ~ /k/	1	/l/ ~ ∅	1
∅ ~ /j/	4	∅ ~ /a/	1	/k/ ~ /p/	1	/k/ ~ ∅	1
/w/ ~ ∅	3	/a/ ~ /o/	1	/k/ ~ /tʃ̃/	1	∅ ~ /k/	1
/θ/ ~ /ɹ/	3	/e/ ~ /i/	1	/w/ ~ /o/	1	∅ ~ /l/	1
∅ ~ /n/	3	/e/ ~ ∅	1	/w/ ~ /u/	1	∅ ~ /w/	1
/ʃ/ ~ /ɹ/	2	∅ ~ /i/	1	/n/ ~ /m/	1		
/a/ ~ ∅	2	∅ ~ /o/	1	/n/ ~ /ɲ/	1		

Además, se registraron 98 pares de lemas que manifestaban variación léxica a partir de patrones fonológicos mayores, los cuales iban desde alternancias o fluctuaciones de pares de fonemas en distintas posiciones en la estructura fonológica del lexema, hasta alternancias de sílabas.

En estos casos, la selección de un lema en particular está motivada por distintos criterios, como la disposición textual, la mayor relevancia lexicográfica, o el criterio de RIV. En *Anexo 3.25. Fluctuaciones léxicas de patrones fonológicos mayores*, es posible consultar este listado de lexemas y sus distintas variaciones.

A continuación, se describirán y ejemplificarán los tipos de variación fonológica encontrados. En primer lugar, se detallan las alternancias fonemáticas, posteriormente, se tratan las fluctuaciones.

6.4.2. Alternancias fonemáticas

Esta sección da cuenta de las alternancias fonemáticas identificadas durante el muestreo, donde se escogió la forma que contenía una RIV. En primer lugar, se indicarán las vocales y posteriormente las consonantes. Para cada alternancia, solo se presentan algunos ejemplos. En la sección de *Anexos*, es posible consultar los listados completos.

6.4.2.1. Vocales

Se identificaron 148 alternancias vocálicas⁶, donde se seleccionó el ítem léxico que contenía /ə/, por considerarse una *Realización Indicadora de Vitalidad* (RIV). La **Tabla 6.4** da cuenta del tipo y frecuencia de estas alternancias.

Tabla 6.4: Número total de alternancias donde /ə/ se releva como RIV.

Alternancias	N.º de casos
/ə/ > /i/	86
/ə/ > /u/	31
/ə/ > /e/	12
/ə/ > /a/	11
/ə/ > /o/	8
Total	148

6.4.2.1.1. Alternancia /ə/ > /i/

Se identificaron 86 casos donde /ə/ es la *Realización Indicadora de Vitalidad* respecto de /i/. Por ejemplo:

6 Además de estas alternancias con otras vocales, se identificó una variación de /ə/ con una elisión a principio de palabra: /əwa/ ~ /wa/ 's. maíz'. En este caso en particular, también existía fluctuación fonemática con una vocal de RIT: /əwa/ ~ /uwa/, por lo que solo se incluyó este último ejemplo en esta sección. En *Anexos*, *Tabla 3.2. Alternancia /ə/ > /u/*, es posible consultar ambos casos.

- | | | |
|-------------------|-----------------------|--------------------------|
| 1. arüñ | /aɪəp/ > /aɪp/ | ‘v. ahumarse la comida.’ |
| 2. ñümü | /ɲəmə/ > /ɲəmi/ | ‘tr. coger, alzar.’ |
| 3. trinrü | /tʃinʃə/ > /tʃinʃi/ | ‘adj. crespo’ |
| 4. wakülpe | /wakəlpe/ > /wakilpe/ | ‘s. un pez, huaiquil’ |

6.4.2.1.2. Alternancia /ə/ > /u/

Se identificaron 31 casos de alternancia entre /ə/ y /u/. Por ejemplo:

- | | | |
|------------------|-------------------|----------------------------|
| 1. küllum | /kəɫum/ > /kuɫum/ | ‘v. lavarse la cara’ |
| 2. mümul | /məməɫ/ > /muməɫ/ | ‘adj. instruido, ejercido’ |
| 3. rüku | /ɹəku/ > /ɹuku/ | ‘s. pecho’ |
| 4. wafüm | /wafəm/ > /wafum/ | ‘s. colmillo’ |



6.4.2.1.3. Alternancia /ə/ > /e/

Se identificaron 12 casos donde /ə/ es segmento indicador de vitalidad, respecto de /e/. Por ejemplo:

- | | | |
|-----------------|-------------------|-------------------------|
| 1. arkün | /aɹkən/ > /aɹken/ | ‘s. baja (mar)’ |
| 2. aymüñ | /ajməɲ/ > /ajmeɲ/ | ‘adv. algo, un poco’ |
| 3. l’üm | /ləm/ > /lem/ | ‘v. relumbrar’ |
| 4. müfür | /məfəɹ/ > /mefəɹ/ | ‘v. agitar con la mano’ |

6.4.2.1.4. Alternancia /ə/ > /a/

Se identificaron 11 casos donde /ə/ es indicador de vitalidad respecto de /a/. Por ejemplo:

- | | | |
|---------------------|-------------------------|----------------------|
| 1. dellküntu | /θeɫkəntu/ > /θeɫkantu/ | ‘s. grasa del caldo’ |
| 2. ewün’ | /ewən/ > /ewan/ | ‘v. tener arcadas’ |
| 3. llüpüd | /ɫəpəθ/ > /ɫipaθ/ | ‘v. echarse (aves)’ |

6.4.2.1.5. Alternancia /ə/ > /o/

Se identificaron 8 casos donde el segmento /ə/ releva una mayor vitalidad respecto de /o/. Por ejemplo:

1. küngoy	/kəŋoj/ > /koŋoj/	‘s. monito del monte’
2. llangkü	/ɬaŋkə/ > /ɬaŋko/	‘v. caerse’
3. mütrü	/mətʃə/ > /mətʃo/	‘v. tr. sacudir’
4. rentrü	/ɾentʃə/ > /ɾentʃo/	‘adj. contrapesado’

6.4.2.2. Exploración del efecto del criterio RIV y alternativas

Con el fin de explorar el efecto que pudiese tener la decisión de utilizar el criterio RIV, se plasman en la **Tabla 6.5** los conteos absolutos y relativos de los fonemas vocálicos a partir del *Diccionario mapudungún-español* de Augusta (1916/2017), en tres casos: con RIV, con RIV y algunas RIT y solo con RIT.

Tabla 6.5: Valores absolutos y relativos del efecto de aplicación del criterio RIV para las alternancias con /ə/.

Caso 1 <i>criterio RIV</i>			Caso 2 <i>criterio RIV + algunas RIT</i>			Caso 3 <i>solo formas con RIT</i>		
Fon.	#	%	Fon.	#	%	Fon.	#	%
/ə/	1527	27%	/ə/	1527 - 51 = 1476	26%	/ə/	1527 - 148 = 1379	24%
/a/	1319	23%	/a/	1319 + 4 = 1323	23%	/a/	1319 + 11 = 1330	23%
/e/	870	15%	/e/	870 + 6 = 876	15%	/e/	870 + 12 = 882	15%
/i/	758	13%	/i/	758 + 30 = 788	14%	/i/	758 + 86 = 844	15%
/o/	631	11%	/o/	631 + 4 = 635	11%	/o/	631 + 8 = 639	11%
/u/	618	11%	/u/	618 + 7 = 625	11%	/u/	618 + 31 = 649	11%
Total	5723		total	5723		total	5723	

Como se observa, se consideran tres situaciones o casos. En primer lugar, nuestra forma de operar se identifica con el Caso 1, que considera el criterio de aceptar siempre la forma que indica vitalidad. El Caso 2 corresponde al mismo criterio RIV, pero permitiendo algunas formas

con RIT, debido a que tienen una relevancia lexicográfica al ser la forma que contiene la definición (Véase 4.3. *Fluctuación fonemática* para un ejemplo de este criterio). En el Caso 3, se distribuyen los conteos y porcentajes si consideráramos el criterio poco probable de escoger solo las formas que indican transferencia.

Es posible observar que las jerarquías no sufren modificaciones en ninguno de los tres casos. Asimismo, los porcentajes se mantienen muy similares, incluso tras haber escogido el criterio de RIT. La mayor diferencia se produce en el Caso 3, para el fonema /ə/, el cual ve reducido su valor relativo, acercándose al porcentaje de /a/. Esto se debe a que, en dicho caso, /ə/ no se escoge en los dobles bajo ninguna circunstancia. Ahora bien, independiente de sus valores absolutos o relativos, /ə/ mantiene su posición en los tres casos, de modo que, de acuerdo con esta tabulación, el criterio basado en las RIV no distorsiona, al menos de manera importante, los resultados de los conteos de frecuencia vocálica.

6.4.2.3. Consonantes

A continuación, se registran 157 casos de alternancias consonánticas. Al igual que para las vocales, se optó por la forma que contuviera una de las RIV. En general, respecto de las consonantes, estas realizaciones indicadoras de vitalidad se refieren, por un lado, a segmentos vernaculares del mapudungun no presentes en el inventario fonológico del español; y, por otro, a realizaciones con distribuciones fonotácticas que no existen o son atípicas en esa lengua. Es posible remitirse a la **Tabla 6.2**, ya mencionada, para un resumen de estos casos. A continuación, se presentan algunos ejemplos para cada alternancia, pero el listado completo se puede consultar en *Anexos*.

6.4.2.3.1. Alternancia /ʎ/ > /l/, /ɲ/, /ŋ/ (en contexto asimilatorio), ∅

La lateral palatal sonora /ʎ/ se considera una RIV cuando alterna en ítems léxicos con la lateral alveolar sonora /l/, la nasal palatal sonora /ɲ/, la nasal velar sonora /ŋ/ en contexto asimilatorio —es decir, antes de otra velar— y con la elisión. Se registraron 36 casos, donde 33 corresponden a alternancias con la lateral alveolar. Algunos ejemplos son:

1. illngaw	/iɬŋaw/ > /ilŋaw/	‘adj. algo dulce’
2. kollong	/koɬoŋ/ > /koloŋ/	‘s. disfrazado’
3. llüfke	/ɬəfke/ > /ləfke/	‘s. relámpago’

6.4.2.3.2. Alternancia /ɬ/ > /l/, /n/

En siete casos, /ɬ/ se relevó como RIV respecto de seis alternancias con el fonema /l/ y solo una con /n/. Algunos ejemplos son:

1. l’oj	/ɬoɬ/ > /loj/	‘adj. tonto.’
2. l’ol’on	/ɬoɬon/ > /lolon/	‘v. provocar vómito’
3. l’el’	/ɬel/ > /nel/	‘v. dejar ir’

6.4.2.3.3. Alternancia /ɲ/ > /n/, /ɲ/, /l/

Similar al caso anterior, /ɲ/ es una realización indicadora de vitalidad respecto de /n/, /ɲ/ y /l/. Se registraron en total cinco casos de este tipo de alternancia, dos con la nasal alveolar, dos con la nasal palatal y uno con la lateral alveolar. Presentamos tres ejemplos:

1. mül’pun’	/mɛɲpun/ > /mɛɲpun/	‘s. hollín’
2. ün’el’	/əɲel/ > /ənel/	‘s. sabor, aliño’
3. wen’i	/weɲi/ > /weɲi/	‘s. muchacho’

6.4.2.3.4. Alternancia /ŋ/ > /f/, /m/, /n/, /ɲ/, /p/, /j/, ∅

La alternancia que involucra la consonante nasal velar /ŋ/ se produjo en 17 casos. En la mayoría de ellos (12), /ŋ/ es una RIV al alternar con una RIT. Por ejemplo:

1. trongtrong	/tʂoŋtʂoŋ/ > /tʂontʂon/	‘s. bolsa’
2. pingliñ	/piŋliŋ/ > /piliŋ/	‘s. helada’
3. lüngfü	/lɛŋfə/ > /ləmfə/	‘s. raíz del cochayuyo’

En el contexto asimilatorio (antes de consonante velar), el fonema nasal velar /ŋ/ es considerado una RIT, frente a las realizaciones RIV, como /m/ y /n/, en este contexto fonológico. Dos de los cinco casos identificados son los siguientes:

- | | | |
|-----------------|-------------------|--------------------|
| 1. ñomke | /ŋomke/ > /ŋoŋke/ | ‘s. sobrecubierta’ |
| 2. penka | /penka/ > /peŋka/ | ‘v. el zapallo’ |

6.4.2.3.5. Alternancia /t͡s/ > /t͡ʃ/, /t/, /m/, /l/, /f/, ∅

Se identificaron 32 casos, donde la africada retrofleja sorda /t͡s/ alterna con las realizaciones RIT /t͡ʃ/, /t/, /m/, /l/, /f/ y elisión. A continuación se registran algunos ejemplos:

- | | | |
|--------------------|-------------------------|------------------|
| 1. kütrüŋ | /kət͡səŋ/ > /kət͡ʃəŋ/ | ‘s. manojo’ |
| 2. kültro | /kəl͡t͡so/ > /kəlto/ | ‘v. aflojarse’ |
| 3. trey | /t͡sej/ > /tej/ | ‘v. derrumbarse’ |
| 4. llafatra | /ʎafat͡sa/ > /ʎafat͡ʃa/ | ‘s. sapo’ |

6.4.2.3.6. Alternancia /θ/ > /s/, /t͡ʃ/, /n/, /j/, /l/, /f/, /p/, /p/, ∅

Se registraron 21 alternancias de la fricativa interdental sorda /θ/ con segmentos RIT. Uno de esos casos ocurre con la fricativa alveolar sorda. Dado que /s/ es considerada un fonema recientemente introducido al inventario fonológico del mapudungun, tras el contacto con el quechua y el castellano (Molineaux, 2021), podemos señalar que la ocurrencia de /θ/ y /s/ en un ítem léxico, indica vitalidad y transferencia, respectivamente. A continuación se indican algunos ejemplos:

- | | | |
|-------------------|-----------------------|--------------------------|
| 1. dechu | /θet͡ʃu/ > /set͡ʃu/ | ‘v. apuntalar, apoyar’ |
| 2. dicho | /θit͡ʃo/ > /sit͡ʃu/ | ‘v. indicar con el dedo’ |
| 3. donü | /θonə/ > /sonə/ | ‘s. arruga’ |
| 4. düching | /θət͡ʃiŋ/ > /sət͡ʃiŋ/ | ‘s. rayo del sol’ |

También, ocurrieron alternancias con otros fonemas, como /tʃ/, /n/, /j/, /l/, /f/, /ɲ/, /p/ y elisión, que se pueden consultar en la tabla respectiva en *Anexos (Tabla 3.11)*.

6.4.2.3.7. Otras alternancias

Por último, se registraron siete alternancias que involucran a la aproximante retrofleja sonora /ɰ/. Estas ocurrieron con la lateral alveolar sonora /l/ (cuatro casos), con la elisión (dos casos), y con la aproximante palatal sonora /j/ (un caso). De igual modo, se identificaron dos casos donde la fricativa alveopalatal sorda /ʃ/ y la oclusiva interdental sorda /t̪/ alternaban con una africada alveopalatal sorda /tʃ/. Algunos ejemplos de estas alternancias son los siguientes:

- | | | |
|-------------------|-----------------------|------------------|
| 1. rawilma | /ɰawilma/ > /jawilma/ | ‘s. cierta ave’ |
| 2. üfür | /əfəɰ/ > /əfəl/ | ‘v. sorber’ |
| 3. fot’üm | /foɰəm/ > /foɰəm/ | ‘s. un familiar’ |

6.4.2.3.8. Alternancia en el posmargen: /ɲ/ > /n/, /m/, /ŋ/ (en contexto asimilatorio), ∅

La nasal palatal sonora, si bien es un fonema en español, tiene un contexto fonotáctico en mapudungun que se considera indicador de vitalidad: el posmargen silábico. Se identificaron 20 realizaciones de este tipo, donde existen formas alternantes con realizaciones indicadoras de transferencia. Algunos ejemplos son los siguientes:

- | | | |
|-----------------|-------------------|------------------------|
| 1. dükoñ | /θəkoɲ/ > /θəkon/ | ‘s. coronta’ |
| 2. meñki | /meɲki/ > /menki/ | ‘s. penacho (de aves)’ |
| 3. pewañ | /peweɲ/ > /pewen/ | ‘s. araucaria’ |
| 4. umañ | /umaɲ/ > /uma/ | ‘v. alojarse’ |

6.4.2.3.9. Alternancia en el posmargen: /f/ > /m/, /l/, ∅

El fonema fricativo labiodental sordo /f/ en mapudugun, cuando ocurre en el posmargen, se considera una realización indicadora de vitalidad, puesto que dicho segmento en esta posición es anómalo en castellano. Identificamos seis casos donde /f/ en este contexto se considera una RIV respecto de otros segmentos o la elisión. Dos ejemplos son:

- | | | |
|-----------------|------------------|-----------------|
| 1. lelif | /lelif/ > /leli/ | ‘v. mirar’ |
| 2. illaf | /iɫaf/ > /iɫam/ | ‘v. despreciar’ |

6.4.2.3.10. Alternancia en el posmargen: /ɥ/ > /w/, /l/

El fonema aproximante velar sonoro /ɥ/ en mapudugun suele aparecer en coda silábica. Puesto que este segmento en español —alófono del fonema oclusivo velar sonoro— es atípico en este contexto silábico, es posible establecer su estatus de RIV en mapudugun en esta posición. Se registraron 4 casos, tres de los cuales son los siguientes:

- | | | |
|------------------|-------------------|--------------------------|
| 1. dagllu | /θaɥɫu/ > /θawɫu/ | ‘camarón de los esteros’ |
| 2. legpe | /leɥpe/ > /lewpe/ | ‘s. callana’ |
| 3. külleg | /kəɫeɥ/ > /kəɫel/ | ‘declive, cuesta’ |

6.4.3. Fluctuación fonemática

Se presentaron 162 casos de fluctuaciones fonemáticas durante el filtrado de las unidades léxicas. Este tipo de variación fonológica no puede resolverse por medio de la identificación de las realizaciones indicadoras de vitalidad, pues los segmentos que alternan entre los pares léxicos no responderían a la influencia de la estructura fonológica del castellano, sino que corresponderían a una dinámica intrasistémica. Es el caso, por ejemplo, cuando ambos segmentos pertenecen solo al inventario del mapudugun o a ambos sistemas, como en **payl'a** /pajɫa/ y **paylla** /pajɫa/ ‘adv. de espaldas’, con laterales que no ocurren en español chileno, y **weku** /weku/ ~ **weko** /weko/ ‘s. tío materno’, con vocales que ocurren en ambos sistemas.

En estos casos, se utilizaron otros criterios que permitieron arribar a una forma escogida. De particular ayuda fue evidenciar en la entrada del diccionario cuál forma desplegaba una definición y cuál era solamente una llamada a otra entrada. Por ejemplo, para el caso de **shongi** /ʃoŋi/ ‘s. arveja’, se seleccionó la forma con /o/, porque en el diccionario era la forma que contenía la definición; en cambio, la entrada **shungi** /ʃuŋi/ solamente remitía al lema **shongi** /ʃoŋi/. En otros casos, la forma seleccionada era aquella que no tenía una marca territorial, en contraposición a otra forma que sí la tenía. Por ejemplo, la entrada para **cheküm-** /tʃekəm/ ‘raíz de v. guarecerse de la lluvia’ no tiene marca dialectal; por lo tanto, es la seleccionada respecto de **chelküm-** /tʃelkəm/ que sí la tiene.

Cabe señalar que se anotaron cada una de estas fluctuaciones para efectos de replicabilidad de nuestros procedimientos y evaluación de los mismos. A continuación, se muestran solo algunos ejemplos, los cuales, sin embargo, dan cuenta de la diversidad de este fenómeno. En la sección *Anexos* se pueden consultar los listados de manera íntegra.



6.4.3.1. Vocales

6.4.3.1.1. Fluctuación entre /o/ ~ /u/

Se identificaron 18 lexemas donde hay fluctuación fonemática entre /o/ y /u/. El segmento /o/ se seleccionó 10 veces. Algunos casos son:

- | | | |
|------------------|-------------------|--------------------------|
| 1. weku | /weku/ > /weko/ | ‘s. tío materno.’ |
| 2. shongi | /ʃoŋi/ > /ʃuŋi/ | ‘s. arveja’ |
| 3. kotrür | /koʃəɹ/ > /kuʃəɹ/ | ‘adj. tostado’ |
| 4. ñadu | /ɲaθu/ > /ɲaθo/ | ‘s. cuñada de una mujer’ |

6.4.3.1.2. Otras fluctuaciones vocálicas

Se identificaron 17 casos donde hay fluctuación entre diversas vocales y también presencia de elisión. La variación más frecuente ocurrió entre /a ~ e/ (5 casos), seguida de /a/ ~ Ø, /e ~ a/ e /i ~ e/, con dos casos cada una. Diversas variaciones se manifestaron solo una vez: Ø ~ /a/, /a ~ o/, /e ~ i/, etc.

- | | | |
|--------------------|---------------------|-----------------|
| 1. alla | /aʎa/ > /eʎa/ | ‘adj. bonito’ |
| 2. kollella | /koʎeʎa/ > /koʎaʎa/ | ‘s. hormiga’ |
| 3. traylef | /ʧajʎef/ > /ʧajʎif/ | ‘v. despejarse’ |

6.4.3.2. Consonantes

En el caso de las consonantes, se siguieron criterios similares que para las vocales, registrándose en total 127 casos.

6.4.3.2.1. Fluctuación /ɭ/ ~ /ʎ/ ~ /θ/

Se identifican 7 casos donde se seleccionó /ɭ/, por sobre las consonantes también vernaculares /ʎ/ y /θ/.

- | | | |
|------------------|-------------------|-------------------|
| 1. kül'wi | /kəɭwi/ > /kəʎwi/ | ‘porotos’ |
| 2. payl'a | /pajɭa/ > /pajʎa/ | ‘adv. de espalda’ |

6.4.3.2.2. Fluctuación /ɻ/ ~ /ʃ/ ~ Ø

La aproximante retrofleja sonora /ɻ/ se intercambia con otros fonemas y elisión en 16 casos, con distintas selecciones dependiendo de los criterios expuestos al inicio de este apartado. Algunos ejemplos son:

- | | | |
|---------------------|-----------------------|-----------|
| 1. pishpishü | /piʃpiʃə/ > /piɻpiɻə/ | ‘v. piar’ |
| 2. ruka | /ɻuka/ > /ʃuka/ | ‘s. casa’ |

6.4.3.2.3. Fluctuación /θ/

Existen diversos casos donde /θ/ está en fluctuación con otros segmentos. Por su prominencia, se considera pertinente enumerarlos de manera separada. Así, en primer lugar, se relevan aquellos casos donde hay fluctuación con /ɬ/, luego con /ʃ/, y, finalmente, algunos otros casos, los cuales se exponen bajo un mismo subapartado. En total, se registraron 51 fluctuaciones fonemáticas con participación de /θ/, y las selecciones son diversas, según se detallan a continuación.

6.4.3.2.3.1. Fluctuación /θ/ ~ /ɬ/

Se registran 13 casos para esta alternancia. Los lemas con /θ/ se seleccionaron 3 veces; en cambio, se seleccionaron 10 casos con /ɬ/. A continuación algunos ejemplos:

- | | | |
|-----------------|-------------------|-------------------|
| 1. radal | /ɬaθal/ > /ɬaɬal/ | ‘s. radal, nogal’ |
| 2. küru | /kuɬə/ > /kuθə/ | ‘adj. negro’ |

6.4.3.2.3.2. Fluctuación /θ/ ~ /ʃ/

Se identificaron 36 casos para esta alternancia. En 23 casos, se seleccionó el lema que se realiza con la interdental /θ/. Algunos ejemplos son:

- | | | |
|-----------------|-------------------|--|
| 1. muday | /muθaj/ > /muʃaj/ | ‘s. cierta bebida’ |
| 2. pesha | /peʃa/ > /peθa/ | ‘v. encontrar una cosa perdida por otro’ |
| 3. weda | /weθa/ > /weɬa/ | ‘adj. mal, malo’ |

6.4.3.2.3.3. Fluctuación /θ/ ~ /f/ ~ /ɬ/

Finalmente, la fricativa interdental sorda /θ/ participa en relaciones de fluctuación fonemática con los segmentos /ɬ/ y /f/ en el postmargen, registrándose solo dos casos.

- | | | |
|-------------------|---------------------|-----------------|
| 1. müküf | /mækəf/ > /mækeθ/ | ‘v. apretar’ |
| 2. piwküll | /piwkəɬ/ > /piwkəθ/ | ‘v. dar vuelta’ |

6.4.3.2.4. Fluctuación /tʃ/ ~ /t/

Se registraron 6 casos que involucran estos dos fonemas. La africada se seleccionó en cuatro oportunidades. Algunos ejemplos de estas fluctuaciones son los siguientes:

- | | | |
|----------------|------------------------|-----------------|
| 1. chaw | /tʃaw/ > /tao ~ tʃsao/ | ‘s. padre’ |
| 2. che | /tʃe/ > /te/ | ‘s. gente’ |
| 3. chi | /tʃi/ > /ti/ | ‘art. el, la’ |
| 4. tayi | /taji/ > /tʃaj/ | ‘adv. denantes’ |

6.4.3.2.5. Otras fluctuaciones entre fonemas consonánticos

Se identificaron 21 casos de diversas fluctuaciones, respecto de los cuales se escogió un elemento siguiendo los criterios mencionados al principio de este apartado. A continuación se presentan algunos ejemplos:

- | | | |
|------------------|---------------------|------------------|
| 1. rümu | /ɾɐmu/ > /ɾɐlu/ | ‘v. zambullirse’ |
| 2. trupef | /tʃupef/ > /tʃunef/ | ‘v. espantarse’ |
| 3. mülfen | /mɐlfen/ > /mɐlfem/ | ‘s. el rocío’ |
| 4. ñaw | /ɲaw/ > /naw/ | ‘s. maullar’ |



6.4.3.2.6. Otras fluctuaciones con elisión o adición de uno o varios segmentos

Se registraron 26 casos de fluctuaciones donde uno de los lexemas presenta elisión de uno o dos segmentos, con diferentes selecciones. Ejemplos donde se seleccionó el lema que contenía un elemento adicionado son:

- | | | |
|-------------------|--------------------|----------------------|
| 1. deya | /θeja/ > /θea/ | ‘s. un familiar’ |
| 2. wengkol | /weŋkol/ > /enkol/ | ‘v. destroncar’ |
| 3. nüla | /nəla/ > /əla/ | ‘v. abrir o abrirse’ |

En cambio, los siguientes ejemplos seleccionan un lema con elisión:

- | | | |
|-----------------|--------------------|-------------------------|
| 1. koel | /koel/ > /kɔtʃeel/ | ‘mango’ |
| 2. latra | /latʃa/ > /lantʃa/ | ‘adj. abultado, grueso’ |

6.4.3.2.7. Fluctuaciones léxicas de patrones fonológicos mayores

Se identificaron 98 casos de variaciones léxicas, donde la fluctuación ocurre con estructuras fonológicas mayores. El listado completo se puede consultar en *Anexo: 3.24. Fluctuaciones léxicas de patrones fonológicos mayores*. A continuación algunos ejemplos:

- | | | |
|-------------------|--------------------------|--------------------------------|
| 1. nüwfü | /nəwfə/ > /nəwkəl/ | ‘v. pasar extendiendo la mano’ |
| 2. pelleñ | /peɬer/ > /peɬɛn/ | ‘s. verruga’ |
| 3. t’ue | /tue/ > /təwe/ | ‘s. tierra, champa’ |
| 4. üllngüd | /əlŋəθ/ > /wilŋəθ, əŋəθ/ | ‘s. coyuntura’ |

6.5. Metátesis

Se identificó una serie de pares de términos que sufrían metátesis. En dos casos, se trató de una transposición silábica con una fluctuación fonemática concomitante. Solo se seleccionó uno de los términos, siguiendo los mismos criterios generales ya aplicados.

6.5.1. Metátesis de /k/ ~ /θ/

- | | | |
|------------------|-----------------------|----------|
| kengdülla | /keŋθəɬa/ ~ /θeŋkəɬa/ | ‘s. eco’ |
|------------------|-----------------------|----------|

6.5.2. Metátesis de /ɬ/ ~ /w/, /u/

- | | | |
|-----------------|-------------------------|---------------------|
| charwa | /tʃaɬwa/ ~ /tʃawɬa/ | ‘s. un arbusto’ |
| cheru-fe | /tʃeɬu-fe/ ~ /tʃeuɬ-fe/ | ‘s. fenómeno ígneo’ |

6.5.3. Metátesis de /ə/ ~ /e/, /i/, /w/

trülef	/tʃəlef/ ~ /tʃeləf/	‘s. tiesto, pedazo de cántaro’
wichür	/wətʃi/ ~ /witʃə/	‘v. torcer la boca’
chiñüd	/tʃiɲəθ/ ~ /tʃəɲiθ/	‘v. cerner’
nüwfü	/nəwfə/ > /nəwəf/	‘v. pasar algo con la mano.’

6.5.4 Metátesis de /ɥ/ ~ /l/

regle	/ɥeɥle/ ~ /ɥelɥe/	‘num. siete’
--------------	-------------------	--------------

6.5.5. Metátesis silábicas con fluctuación fonemática

ngüti-w	/ti.ɲi-w/ ~ /ɲə.ti-w/	‘v. doblarse’
pultru	/pul.tʃu/ ~ /tul.pu/	‘s. amancay’

6.6. Reiteración fonológica y reduplicación

Durante el muestreo, se identificó una serie de unidades léxicas que presentaban una reiteración total o parcial de un patrón fonológico; por ejemplo, **chuchu** /tʃutʃu/ ‘s. abuela materna y sus nietos’, **chañchañ** /tʃaɲtʃaɲ/ ‘s. cierto pasto, **weluwelu** /weluwelu/ ‘adv. alternadamente’, **mungüllmungüll-nge-n** /muɲəɣmuɲəɣ/ ‘v. mecer el cuerpo’, etc.

Debido a que la reiteración de cualquier unidad segmental puede afectar los resultados de nuestro análisis, fue necesario controlar este fenómeno en las unidades donde era posible asociarlo con lo que la literatura ha denominado *construcción reduplicada*, fenómeno morfológico productivo en mapudungun que opera tanto en el dominio sintáctico, como en el semántico (Sandvig, 1986).

Siguiendo esta directriz, se controlaron los lexemas que presentaban reiteración de patrones fonológicos, reminiscentes del fenómeno de construcción reduplicada, seleccionando para nuestro corpus final solo la instancia ‘básica’ de la reiteración. Por ejemplo, en el adverbio **dagdag** /θauθau/ ‘en todas partes’, se seleccionó la sílaba **dag** /θau/ para nuestro conteo fonemático. Esta decisión opera principalmente en virtud de un criterio semántico, siguiendo las propiedades de significado identificadas por Sandvig (1986, p. 148) para estas estructuras, a saber: a) continuidad o repetición, b) intensidad y c) pluralidad individualizada.

Del total de 207 entradas seleccionadas que evidenciaban la reiteración de un patrón fonológico, se identificaron 160 reduplicaciones con el consiguiente control de su estructura fonemática, siguiendo los criterios ya señalados. Resulta interesante que en muchos de estos casos, la selección de la unidad lexemática considerada como construcción reduplicada coincidió con el hecho de que la estructura en cuestión tenía una extensión fonemática y silábica atípica para ser un lexema monomorfemático —por ejemplo, cuatro o más sílabas—, lo que respalda, con una base estructural, el hecho de considerar en el cómputo sólo la base de la forma reduplicada. En *Anexos*, es posible consultar el listado completo de formas consideradas reduplicaciones y de las formas que tienen una reiteración de un patrón fonológico.

6.7. Interjecciones

Durante el muestreo de las entradas, se identificaron lemas que correspondían a interjecciones y onomatopeyas. Se decidió controlar el efecto de distorsión de algunas de estas estructuras fonológicas, especialmente cuando presentaban distribuciones fonológicas anómalas, por lo que no se incluyeron en el corpus de nuestra investigación. Es el caso, por ejemplo, de la entrada *fotr* (/foʃs/) ‘interjección para expresar impaciencia’ (p. 49), la cual contiene una consonante africada en coda.

6.8. Falta de definición

En tres casos, no fue posible acceder a una definición, por no estar impresa en la obra original. Concretamente, hubo tres casos en que cierta entrada remitía a una segunda entrada que no aparecía en el diccionario. Se prefirió no incluir dichos términos.

1. cholkün. V. sholkün. (p. 25)
2. pelüŋ|n*, n., = Chelün. (p. 169)
3. ñəŋkül, s., V. *ñüŋkül* (p. 156)



7. Análisis y Resultados

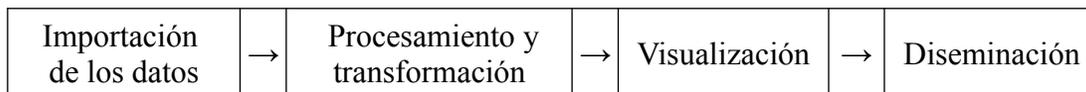
7.1. Introducción

Una vez realizado el proceso de muestreo, que consistió en el análisis caso a caso de las 9703 entradas del *Diccionario mapudungún-español* de Augusta (1916/2017), procedemos al análisis y reporte de nuestros resultados. Bajo la orientación de las preguntas de investigación, que procuran, entre otros objetivos, determinar la frecuencia de fonemas del mapudungun en un listado léxico de unidades monomorfémicas, se procesaron los datos recabados durante el muestreo.

Para el proceso de análisis y reporte de resultados, fue esencial la utilización de R, lenguaje de programación que permite realizar computación estadística (R Core Team, 2021). En específico, utilizamos un enfoque que se ha denominado *Ciencia de Datos*, definido como un conjunto de técnicas que permiten extraer información significativa a partir del procesamiento de datos (Kotu y Deshpande, 2019). Para su despliegue, se aplican una serie de pasos sucesivos, que comienzan con el *importe de los datos* al ambiente de R, para luego pasar a la *limpieza y organización* de los mismos siguiendo una estructura que es compatible con la filosofía de la propuesta, luego se continúa con la *exploración* de estos datos ya procesados, para terminar con su *comunicación o diseminación* (Wickham y Grolemund, 2017).

Dentro de estos cuatro pasos, la *exploración* se puede subdividir en: 1) el procesamiento y transformación de los datos por medio de la selección de variables de interés, 2) la creación de nuevas variables, cálculos estadísticos, etc., 3) la visualización de esta información, y, en el caso que exista planteamiento de una hipótesis, 4) un modelamiento (Wickham y Grolemund, 2017). Este es un proceso iterativo, que permite la reiteración de los pasos de manera sucesiva, de acuerdo con los resultados y la información proporcionada por la visualización de los mismos.

En nuestro estudio, esta serie de etapas adoptadas a partir de los principios de la *ciencia de datos* se traduce, en términos concretos, en cuatro etapas principales de análisis y reporte de nuestros resultados. Estas son:



A continuación, describiremos la forma en que desarrollamos estos pasos. Dado que R es un lenguaje de programación, expondremos, a modo de ilustración, el procedimiento y las rutinas programáticas necesarias para importar, procesar, transformar y visualizar nuestros datos.

7.2. Carga de datos y preparación preliminar

Una vez iniciada la sesión de trabajo en R, se deben cargar los recursos que permiten operar con distintas funciones y rutinas:

```
library(tidyverse)
library(here)
library(wirintukufe)
```



El recurso o librería **Tidyverse** (Wickham et al., 2019) es un conjunto de paquetes o subprogramas para R que permite unificar y facilitar la ejecución de la *ciencia de datos* antes descrita (Wickham y Grolemund, 2017). El recurso **here()** (Müller, 2020) es un paquete que facilita el acceso a los archivos y carpetas que forman parte de nuestro estudio, asegurando su replicabilidad. El recurso **Wirintukufe** (‘escritor/a’, en mapudungun) es un paquete creado especialmente para este proyecto, que permite la fonetización o conversión automática a la escritura del Alfabeto Fonético Internacional, a partir de un texto escrito en el Alfabeto Mapuche Unificado.

Posteriormente, se cargó el archivo .csv proveniente de nuestra etapa de muestreo en el ambiente de R, y se le asignó a una tabla de datos, denominada “Augusta”:

```
Augusta <- read_csv(here("03_data", "Augusta.csv"))
```

La tabla de datos así creada fue transformada por medio del filtrado de sus filas, cuya codificación correspondía a “aceptadas” (código “1”) y la selección de las columnas de interés; a saber: la unidades lexemáticas (“lexeme”), la clase de palabra (“pos”, del inglés *part of speech*) y un resumen de 25 caracteres de la definición original (“summary”):

```
corpus_lexicografico <- Augusta %>%  
  filter(code == 1) %>%  
  mutate(summary = str_trunc(as.character(def), 35)) %>%  
  select(lexeme, pos, summary) %>%  
  mutate(IPA = phonemizer(lexeme, simple = TRUE)) %>%  
  mutate(length = str_count(IPA, "\\w"))
```

Este código también crea de manera automatizada una nueva variable que contiene las entradas escritas en el Alfabeto Fonético Internacional (“IPA”), utilizando la función *phonemizer()*, creada especialmente para este estudio, y que se encuentra en el paquete **Wirintukufe**. Finalmente, se incluye una nueva variable (“length”), que contiene el número de fonemas por lexema, es decir, la longitud fonológica.

A modo de ejemplo de la estructura de datos resultante, en la **Tabla 7.1** replicamos 10 entradas seleccionadas al azar del corpus lexicográfico:

Tabla 7.1: Extracto de 10 entradas seleccionadas al azar del corpus lexicográfico, creado por medio de la transformación de la planilla de datos proveniente de nuestra etapa de muestreo.

Lexeme	IPA	POS	Summary	Length
yellu	/jeɬu/	verbo	desatarse, disolverse	4
l'aku	/ɭaku/	sustantivo	cierto familiar	4
cheñfill	/tʃeɲfiɬ/	sustantivo	Estera	6
kukaw	/kukaw/	sustantivo	hueco en los árboles arriba	5
eko	/eko/	verbo	Gatear	3
chikū	/tʃikə/	sustantivo	Anzuelo	4
ayl'en	/ajɬen/	sustantivo	Brasa	5
iwüd	/iwəθ/	verbo	Envolver	4
rale	/ɭale/	verbo	Parchar	4
tayil	/tajil/	verbo	interpretar un tipo de canto	5

7.3. Procesamiento y transformación

7.3.1. Fonemización automática

La creación de una columna con las mismas unidades lexemáticas, pero registradas en una escritura fonológica especialmente diseñada para este estudio, se realizó por medio de un algoritmo de reemplazo automático, que se sustenta en la equivalencia exacta entre grafema y fonema que tiene el Alfabeto Mapuche Unificado (Catrileo, 1984; Hernández, 1986), grafemario en el cual está escrito el *Diccionario mapudungún-español* de Augusta, en su versión moderna (1916/2017).

La Escritura Fonológica Simplificada (EFS) que utilizamos en esta etapa consiste en una serie de grafías especialmente seleccionadas para poder realizar el conteo fonológico. Se trata de una serie de símbolos discretos que representan de manera unívoca cada fonema del mapudungun, lo que tiene una utilidad específica para el conteo de aquellos fonemas que en AFI requieren la combinación de símbolos para su codificación.

La elección del símbolo EFS particular está motivada por dos requerimientos. En primer lugar, que su número sea relativamente bajo en el mapa de caracteres de Unicode, lo que asegura su correcta interpretación en cualquier plataforma; y en segundo lugar, que tenga un simbolismo visual relacionado con su símbolo AFI. La elección de este símbolo es meramente funcional y restringido a la etapa de análisis. Esto quiere decir que una vez concluida la etapa de transformación de datos es posible revertir esta EFS a su símbolo AFI correspondiente, y así poder hacer el reporte de los resultados con la nomenclatura adecuada a un estudio de esta índole.

La **Tabla 7.2**, a continuación, indica las equivalencias entre estos sistemas:

Tabla 7.2: Tabla de equivalencia de caracteres.

AMU	AFI	EFS	Unicode Hexadecimal
<ch>	/tʃ/	<tʃ>	U+02A7
<tr>	/tʂ/	<tʂ>	U+0167
<t'>	/t̥/	<t̥>	U+0163
<n'>	/n̥/	<n̥>	U+0146
<l'>	/l̥/	<l̥>	U+013C

La función de fonemización automática se creó por medio de un algoritmo que de manera sucesiva identifica los caracteres en Unificado y los convierte a EFS, utilizando una función de reemplazo en R, denominada **gsub()**, que permite reemplazar un patrón de caracteres al interior de un vector de datos. De igual modo, esta función se integró a la rutina para poder obtener el texto en escritura IPA, una vez realizado el conteo de las unidades.

7.3.2. Creación de tabla de fonemas

Establecido el corpus lexemático, se procedió a la creación de un corpus fonemático por medio de la extracción de cada una de las unidades segmentales, manteniendo información como el lexema de origen, su clase de palabra, definición y longitud fonológica. El código necesario para realizar esta transformación es el siguiente:

```
corpus_fonologico <- corpus_lexicografico %>%  
  mutate(phoneme = strsplit(as.character(IPA), "")) %>%  
  unnest(cols = phoneme) %>%  
  group_by(IPA) %>%  
  mutate(position = 1:n())
```

La primera línea especifica que la nueva tabla de datos se llamará “corpus fonológico”. La segunda línea crea una nueva variable con el nombre de “phoneme”, poblada con cada carácter que se encuentra en la columna “IPA”, es decir, con las unidades lexemáticas. La siguiente función **unnest()** distribuye la información de cada fila correspondiente a cada una de las variables. Finalmente, al agrupar la información por medio de la variable “IPA”, es posible crear una nueva variable denominada “position”, que especifica el lugar que ocupa el fonema en la estructura fonológica del lexema de donde fue extraído. Esta variable nos permitirá estudiar las frecuencias en distintas posiciones dentro del lexema; concretamente, posiciones inicial, media y final.

La **Tabla 7.3**, a continuación, se construyó a modo de ilustración de esta estructura de datos, escogiendo cuatro lexemas al azar. Se ha oscurecido la columna que contiene cada uno de los fonemas, para su mejor identificación:

Tabla 7.3: Extracto de 15 entradas del corpus fonológico, a partir de cuatro entradas léxicas seleccionadas al azar.

Lexeme	IPA	POS	Summary	phoneme	position
atru	/aʈsu/	verbo	fatigarse	/a/	1
atru	/aʈsu/	verbo	fatigarse	/ʈs/	2
atru	/aʈsu/	verbo	fatigarse	/u/	3
chakañ	/ʈʃakəŋ/	sustantivo	un marisco	/ʈʃ/	1
chakañ	/ʈʃakəŋ/	sustantivo	un marisco	/a/	2
chakañ	/ʈʃakəŋ/	sustantivo	un marisco	/k/	3
chakañ	/ʈʃakəŋ/	sustantivo	un marisco	/a/	4
chakañ	/ʈʃakəŋ/	sustantivo	un marisco	/ŋ/	5
trüküfü	/ʈʂəkəfə/	verbo	contraerse, acalambrarse	/ʈʂ/	1
trüküfü	/ʈʂəkəfə/	verbo	contraerse, acalambrarse	/ə/	2
trüküfü	/ʈʂəkəfə/	verbo	contraerse, acalambrarse	/k/	3
trüküfü	/ʈʂəkəfə/	verbo	contraerse, acalambrarse	/ə/	4
trüküfü	/ʈʂəkəfə/	verbo	contraerse, acalambrarse	/f/	5
trüküfü	/ʈʂəkəfə/	verbo	contraerse, acalambrarse	/ə/	6
wiwi	/wiwi/	sustantivo	cierto animalito	/w/	1
wiwi	/wiwi/	sustantivo	cierto animalito	/i/	2
wiwi	/wiwi/	sustantivo	cierto animalito	/w/	3
wiwi	/wiwi/	sustantivo	cierto animalito	/i/	4

7.3.3. Conteo de los fonemas

Una vez establecida la tabla de fonemas, se procedió a contar las instancias de estos, por medio de la tabulación de sus ocurrencias. Esto se puede realizar con el siguiente código:

```

tabla_fonemas <- corpus_fonologico %>%
  group_by(phoneme) %>%
  summarise(count = n()) %>%
  mutate(Fon. = phonemizer(Fon.)) %>%
  arrange(desc(Frec.))

```

La primera línea guarda el resultado en una nueva tabla de datos denominada “tabla_fonemas”. La segunda y tercera líneas permiten contabilizar las instancias de la variable “phoneme”, creando una tabulación de fonema más su conteo de frecuencia.

La penúltima función convierte la escritura EFS en símbolos AFI, para poder reportar los resultados. Finalmente, se ordena la información de mayor a menor frecuencia.

Posteriormente, también se incluye información como frecuencia relativa, bajo el nombre de “Perc.”, y ratio de segmentos observados / esperados, computado por medio de la división de la frecuencia relativa y una frecuencia relativa esperada de 3.6%, que considera la misma probabilidad de ocurrencia relativa para cada uno de los 28 fonemas del mapudungun normalizado a 100 segmentos ($100 / 28 = 3.6\%$). Este último indicador, como se verá en su sección respectiva, permite hacer comparaciones entre sistemas fonológicos que asumen distinto número de segmentos (Gordon, 2016). El código para estas nuevas variables es el siguiente:

```
tabla_fonemas <- tabla_fonemas %>%  
  rename(Fon. = phoneme, Frec. = count) %>%  
  mutate(Perc. = round((Frec. / sum(Frec.))*100, 2)) %>%  
  mutate(Ratio = Perc. / 3.571429) %>%  
  mutate(Ratio = round(Ratio, 2)) %>%  
  mutate(Fon. = phonemizer(Fon.)) %>%  
  arrange(desc(Frec.))
```

Finalmente, y con el objetivo de crear subconjuntos de datos de acuerdo con parámetros articulatorios y fonológicos, se procedió a agregar tres variables relacionadas con la naturaleza de cada segmento fonológico: el tipo de articulación –vocal o consonante–, el modo, y el lugar de articulación. Con esta información, es posible calcular otras proporciones e indicadores, como la ratio de segmentos sonoros y áfonos, resonantes y obstruyentes, etc.

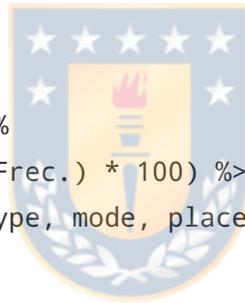
En la subsección *Resultados*, se reportarán estos últimos análisis, por corresponder a las respuestas de nuestras preguntas de investigación.

7.4. Visualización

Para la visualización de los datos procesados, se utilizan diversas técnicas gráficas, como el uso de tablas, gráficos de barra, etc. Estos gráficos y tablas serán reproducidos en la sección correspondiente a los resultados; en el apartado en curso registramos su codificación como ejemplo.

El código que genera cada gráfico consiste, en primer lugar, en la selección de las unidades de interés por medio del filtraje de las variables atendiendo a consideraciones como modo de articulación, dirección de la lengua, etc.

```
vocales <- results %>%  
  filter(type == "vocal") %>%  
  mutate(perc = Frec. / sum(Frec.) * 100) %>%  
  select(Fon., Frec., AMU, type, mode, place, perc)
```



El subconjunto creado por esta rutina puede cargarse en la función **ggplot()** para crear un gráfico:

```
graph <- ggplot(voc, aes(x = reorder(fonos, -perc), y = perc)) +  
  geom_bar(stat = "identity") +  
  scale_y_continuous(labels = function(x){paste0(x, "%")})
```

La mayoría de nuestros gráficos se crearon con códigos similares a este.

7.5. Resultados

7.5.1. Cuadro fonológico

Antes de proceder al reporte de las frecuencias absolutas y porcentuales de nuestro estudio, se plasman en la **Tabla 7.4** y la **Tabla 7.5** —con símbolos AFI y descripciones articulatorias— los 28 fonemas del mapudungun que se asumen en este estudio.

Tabla 7.4: Vocales del mapudungun.

	ANTERIOR	CENTRAL	POSTERIOR
ALTA	/i/		/u/
MEDIA	/e/	/ə/	/o/
BAJA		/a/	

Como se observa, las vocales del mapudungun son 6 y se releva una realización medio-centralizada para la denominada *sexta vocal*, en virtud de estudios que atribuyen esos rasgos articulatorios para este fonema (Catrileo, 1972, 1984, 1986; Henríquez y Salamanca, 2012; Soto-Barba, Lara y Salamanca 2016, Mena, Figueroa, Rogers y Salamanca, 2019). Debido a esta decisión, se grafica el fonema con el símbolo <ə>, también denominado *schwa*.

Tabla 7.5: Consonantes del mapudungun.

	BILABIAL	LABIO-DENTAL	INTER-DENTAL	ALVEOLAR	ALVEOPAL	RETRO-FLEJA	PALATAL	VELAR	LABIO-VELAR
OCCLUSIVA	/p/		/t̪/	/t/				/k/	
AFRICADA					/tʃ/	/tʃ̺/			
NASAL	/m/		/n̪/	/n/			/ɲ/	/ŋ/	
FRICATIVA		/f/	/θ/	/s/	/ʃ/				
APROXIMANTE						/ɻ/	/j/	/w̥/	/w/
LATERAL			/l̪/	/l/			/ʎ/		

Respecto de las consonantes, existen 22 unidades funcionales. Asumimos una realización aproximante para el fonema velar oral sonoro y lo transcribimos con el símbolo AFI /w̥/. Esto es

en virtud de un principio de simetría sistémica, puesto que /ɥ/⁷, junto con /j/ y /w/, es parte de una serie de fonemas que alternan con sonidos que son articulatoriamente vocales, a saber: [ɨ], [i] y [u], respectivamente (Salas, 1976).

7.5.2. Preparación preliminar

Las 9703 entradas del *Diccionario mapudungún-español* de Augusta (1916/2017) fueron procesadas siguiendo los criterios establecidos en la sección correspondiente a los criterios metodológicos. Tras ese exhaustivo proceso de muestreo, se arribó al siguiente número de entradas aceptadas:

Tabla 7.6: Número de entradas según su estatus de aceptadas y no pertinentes.

Estatus de entradas	n	porcentaje (%)
entradas no pertinentes	6789	69.97%
lexemas aceptados	2914	30.03%
Total	9703	100%

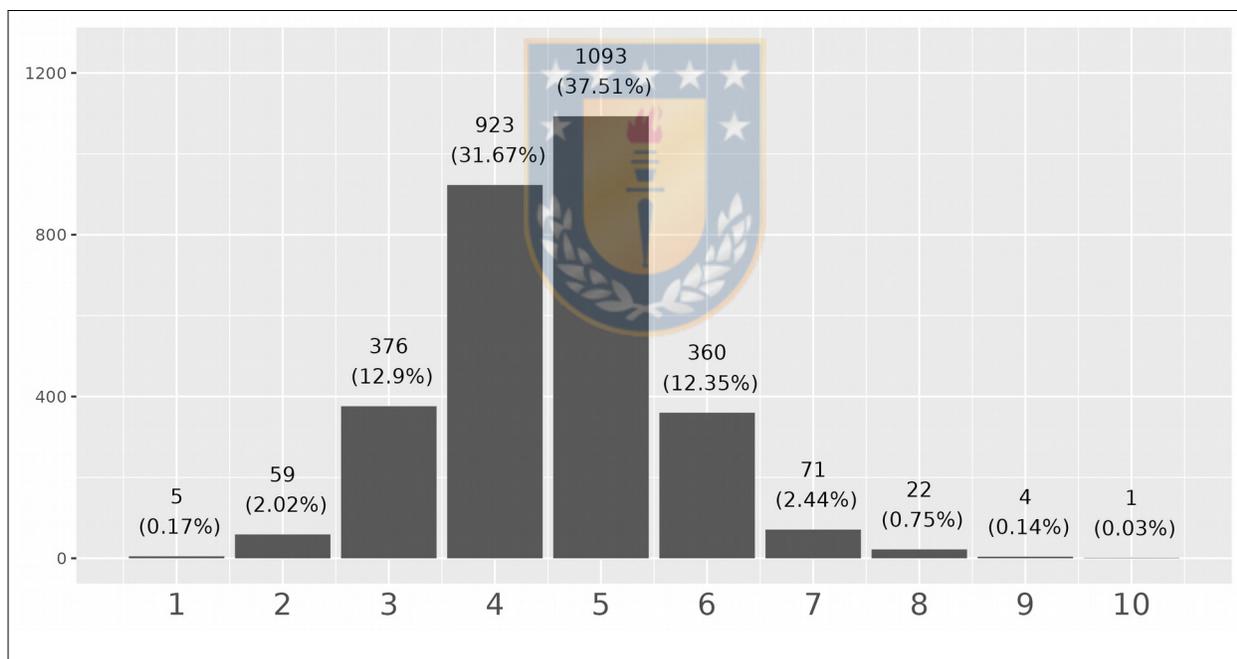
Del total de lemas, se aceptaron **2914** lexemas, que corresponde al 30% de las entradas del diccionario; el 70% restante no formó parte de nuestro corpus lexicográfico, pues fue filtrado de acuerdo con los procedimientos señalados en la sección *Criterios para el Filtrado de las Unidades*.

7 En otro estudio, cuya versión final está en proceso, hemos representado este segmento a través de la combinación del símbolo cardinal de la fricativa velar sonora /ɣ/, más un diacrítico de descenso /ɣ̣/, esto con el objetivo de enfatizar su origen fonológico diacrónico, debido a un proceso de reducción fonológica. Asimismo, ese símbolo estaba más en línea con la noción de que en varios estudios este fonema se ha descrito como fricativo (véanse Salinas y Salamanca, 2016; Pérez y Salamanca, 2017; Saldivia y Salamanca, 2020; Urrea y Salamanca, 2021, entre otros).

7.6. Corpus lexicográfico

Como hemos señalado, el corpus se compone de **2914** unidades lexemáticas monomorfémicas. Con el fin de describir la naturaleza de este corpus, entregaremos a continuación algunas de sus características principales. En primer lugar, debemos señalar que la extensión fonológica de los lexemas aceptados tiene un rango que va desde 1 fonema a 10 fonemas. En el **Gráfico 7.1**, se indican las frecuencias de los lexemas según su longitud fonológica. Se observa que los lexemas con una extensión de 4 y 5 fonemas son los más frecuentes (más del 69% de todos los lexemas), seguidos de aquellos con 3 y 6 fonemas (más del 25% de todos los lexemas). Valores extremos, como lexemas con dos o menos fonemas, por un lado, y lexemas con 7 o más fonemas, por el otro, constituyen menos del 6% del corpus.

Gráfico 7.1: Frecuencia de las distintas longitudes fonológicas de los lexemas.



Esta distribución de la extensión fonológica se comprueba al observar algunos estadígrafos, como la mediana ($M = 5$) y el promedio ($\bar{x} = 4,6$). Los lexemas con un solo fonema, en un extremo, y los de más de ocho fonemas, por el otro, se consideran valores atípicos por la baja frecuencia que manifiestan.

Tabla 7.7: Estadígrafos de la longitud fonológica de los lexemas aceptados.

MIN.	1 ST QU.	MEDIAN	MEAN	3 RD QU.	MAX.
1,00	4,00	5,00	4,56	5,00	10,00

Otra variable posible de medir respecto de los lexemas que forman parte de nuestro corpus es su pertenencia a una clase de palabra. La **Tabla 7.8** contiene las frecuencias absolutas y relativas de dichos lexemas, según este parámetro.

A grandes rasgos, un gran número y porcentaje de los lexemas son sustantivos (45,7%), seguido de cerca por los lexemas verbales (42,6%). Los lexemas que son modificadores —como adjetivos y adverbios— corresponden al 10%. Finalmente, otras clases de palabras, como interjecciones, partículas y numerales, etc., corresponden al 1,6%.

Tabla 7.8: Distribución de los lexemas de acuerdo a la clase de palabra a la que pertenecen.

POS	n	porcentaje
1. sustantivo	1331	45,68%
2. verbo	1240	42,55%
3. modificadores	296	10,16%
4. otros	47	1,61%
Total	2914	100%

7.7. Corpus fonológico

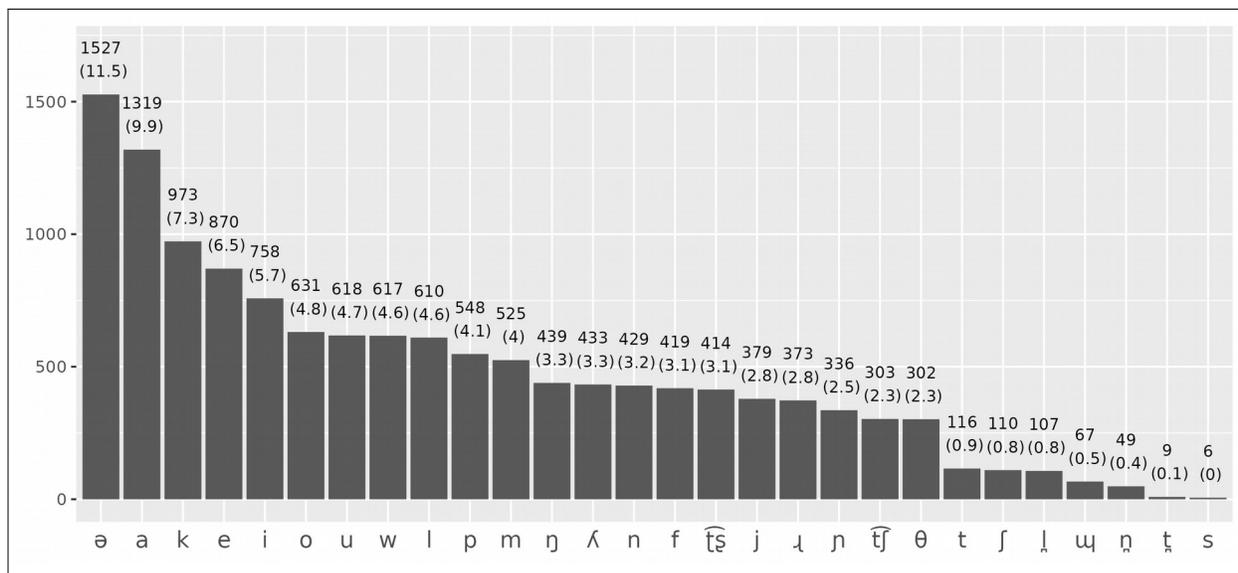
Descrito el corpus lexemático, se identifica el corpus fonológico. Este se compone de **13287** fonemas, cuyas frecuencias absolutas y relativas se distribuyen según la **Tabla 7.9**:

Tabla 7.9: Jerarquía de los fonemas del mapudungun según sus frecuencias absolutas y relativas.

RANK.	FON.	FREC.	PERC.	RANK.	FON.	FREC.	PERC.	RANK.	FON.	FREC.	PERC.
1	/ə/	1527	11,49	11	/m/	525	3,95	21	/θ/	302	2,27
2	/a/	1319	9,93	12	/ŋ/	439	3,3	22	/t/	116	0,87
3	/k/	973	7,32	13	/ʎ/	433	3,26	23	/ʃ/	110	0,83
4	/e/	870	6,55	14	/n/	429	3,23	24	/ɭ/	107	0,81
5	/i/	758	5,7	15	/f/	419	3,15	25	/ɥ/	67	0,5
6	/o/	631	4,75	16	/tʂ/	414	3,12	26	/ŋ̃/	49	0,37
7	/u/	618	4,65	17	/j/	379	2,85	27	/t̃/	9	0,07
8	/w/	617	4,64	18	/ɬ/	373	2,81	28	/s/	6	0,05
9	/l/	610	4,59	19	/ɲ/	336	2,53				
10	/p/	548	4,12	20	/tʃ/	303	2,28	Total 13287			

En el **Gráfico 7.2**, a continuación, se visualiza la jerarquía de segmentos del mapudungun según su frecuencia de ocurrencia. Es posible visualizar la alta frecuencia de las vocales centrales, tanto de /ə/ (**11,49%**) como de /a/ (**9,93%**), al compararlos con todos los otros segmentos. De igual modo, los fonemas vocálicos se ubican entre los primeros siete segmentos más frecuentes (sobre **617** instancias de cada caso), siendo /k/ la única consonante que tiene una frecuencia comparativamente muy alta (**7,32%**). La alta frecuencia de la consonante velar sorprende, por cuanto tiene una fonotaxis restringida al premargen (salvo en algunas interjecciones que no son parte de nuestro muestreo), en comparación con otras consonantes de alta frecuencia como /w/ o /l/, que pueden ocurrir tanto en el premargen, como en el postmargen.

Gráfico 7.2: Frecuencias absolutas y relativas (entre paréntesis) de los fonemas del mapudungun.



Entre los segmentos menos frecuentes se encuentran las oclusivas alveolar e interdental, la lateral y nasal interdental, las fricativas alveolar y alveopalatal, y la aproximante velar. En términos tipológicos, llama la atención que la oclusiva alveolar sorda /t/ (**0,87%**) tenga una frecuencia tan baja, característica que comparte con otras obstruyentes sordas como /t͡/ (**0,07%**) y /s/ (**0,05%**), pero no así con las resonantes /l/ y /n/ que manifiestan frecuencias más altas (**4,59%** y **3,23%**, respectivamente). Esta asociación del lugar y modo de articulación se profundiza en la sección *Recuento estadístico por clase natural*.

La distribución general de las frecuencias fonológicas sigue un patrón típico de estas distribuciones, donde algunos segmentos tienen una muy alta frecuencia (/ə/ > /a/ > /k/, etc.), seguidos por otros segmentos con una frecuencia media (/o/ > /u/ > /w/, etc.) y algunos con una frecuencia muy baja (/t/ > /ʃ/ > /ɬ/, etc). Volveremos sobre este punto en la sección donde abordamos una categorización de frecuencias *ad hoc* para los fonemas del mapudungun.

7.8. Determinación de proporciones

En esta sección, se reporta la ratio consonante / vocal (ratio C/V), la proporción de segmentos resonantes / obstruyentes, segmentos sonoros / áfonos, y, finalmente, una ratio por fonemas según su frecuencia esperada versus la observada.

7.8.1. Ratio consonante/vocal

La ratio consonante/vocal puede calcularse respecto del inventario de una lengua particular —lo que nos entrega la capacidad potencial de su sistema fonológico—, pero también es posible realizarla respecto de la frecuencia de ocurrencia de los fonemas en un listado léxico. En este último caso, nos entrega la carga funcional observada según el tipo de segmento fonológico. A la primera, se le denomina ratio C/V del inventario fonológico, y a la segunda, ratio C/V en el léxico.

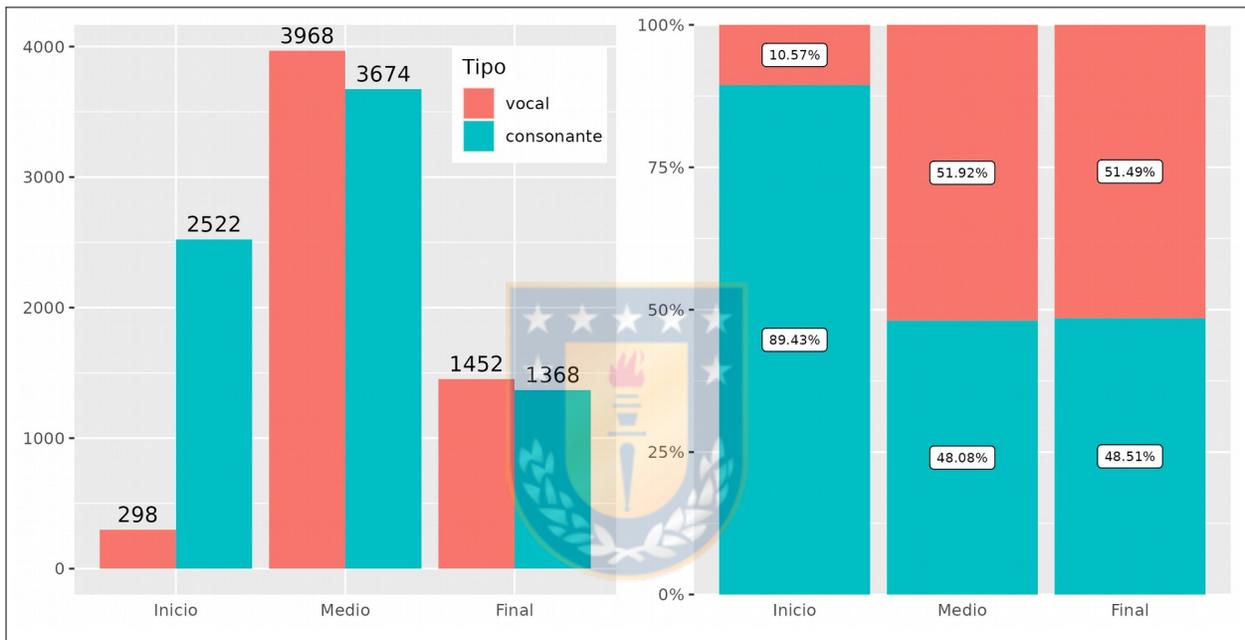
Por un lado, la ratio C/V del inventario fonológico ha sido una forma de evaluar la complejidad de los sistemas fonemáticos de las lenguas en el mundo, y se calcula por medio de la división del número de consonantes (C) por el número de vocales (V) en el inventario fonológico (Maddieson, 2013b). Como valor intralingüístico, permite describir y caracterizar los fonemas de una lengua. En el caso del mapudungun, al contar con 22 consonantes y 6 vocales, la ratio consonante/vocal del inventario fonológico es de 3,7:1, vale decir casi 4 consonantes por cada vocal. Comparado con otras lenguas, esta ratio es muy cercana a la mediana de 3.5:1 y se clasifica como “promedio”, por encontrarse entre los valores 2,75 y 4,5, junto con otras 234 lenguas, de un total de 564 (Maddieson, 2013).

Por otro lado, creemos que reportar la ratio C/V en el léxico es una forma de profundizar en la caracterización de la carga funcional de los segmentos fonémicos del mapudungun. Respecto de nuestro estudio, de los **13287** fonemas identificados, **7564** corresponden a consonantes y **5723** a vocales, lo que nos entrega una ratio C/V en el léxico de **1,32:1**. Si extrapolamos este

coeficiente, podemos señalar que por cada 100 vocales del mapudungun ocurren 132 consonantes en el listado léxico monomorfémico.

Además de reportar esta relación, creemos pertinente profundizar respecto de este dato por medio del reporte de la distribución de las vocales y consonantes en tres posiciones en la estructura fonológica de las unidades monomorfémicas puestas en foco (inicial, media y final).

Gráfico 7.3: Frecuencia de los fonemas vocálicos y consonánticos en las posiciones inicial, media y final.



Es posible visualizar, a partir del **Gráfico 7.3**, que la distribución de las consonantes y vocales en las posiciones media y final son muy similares, siendo las vocales levemente más frecuentes en estas posiciones (52% vs. 48%, y 51% vs. 49%, respectivamente). En cambio en la posición inicial, predominan con creces las realizaciones consonánticas (89%) por sobre las vocálicas (11%).

A continuación, con el fin de explorar con mayor detalle estas distribuciones y tendencias, se presentan en la **Tabla 7.10** distintas tabulaciones de los segmentos según tipo y rasgo articulatorio en las tres posiciones en el lexema.

Tabla 7.10: Distribución de frecuencias absolutas, según modo de articulación, y posición en el lexema.

	/p/	/t̥/	/t/	/tʃ/	/t͡ʂ/	/k/	/m/	/ɱ/	/n/	/ɲ/	/ŋ/
<i>inicio</i>	305	2	49	158	194	468	219	9	68	54	126
<i>medio</i>	243	7	67	145	220	505	209	20	221	121	233
<i>final</i>	N/A	N/A	N/A	N/A	N/A	N/A	97	20	140	161	80

	/f/	/θ/	/s/	/ʃ/	/ɬ/	/j/	/ɰ/	/w/
<i>inicio</i>	78	82	6	51	119	58	0	246
<i>medio</i>	208	127	0	46	172	208	29	252
<i>final</i>	133	93	0	13	82	113	38	119

	/ɭ/	/ɮ/	/ʎ/	/i/	/e/	/ə/	/a/	/o/	/u/
<i>inicio</i>	24	103	103	44	34	103	91	6	20
<i>medio</i>	65	334	242	538	603	1069	912	474	372
<i>final</i>	18	173	88	175	232	354	315	151	225

Gráfico 7.4: Distribución de las frecuencias fonemáticas en tres posiciones del lexema (valores relativos).

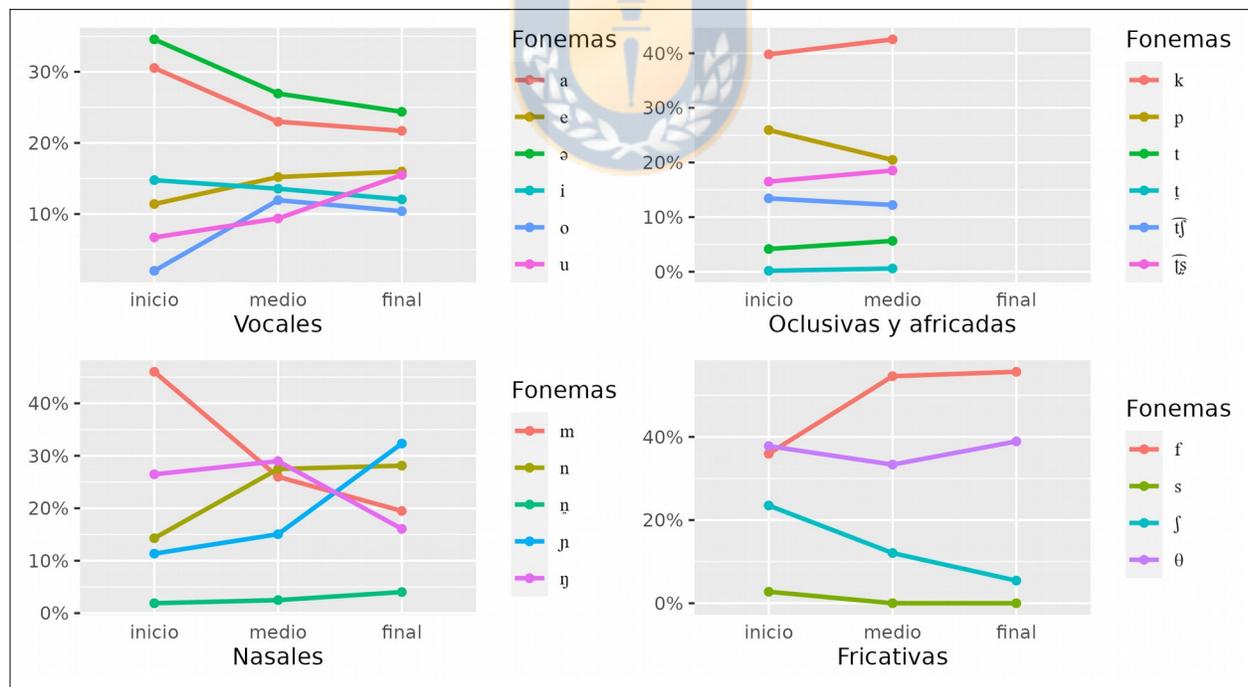
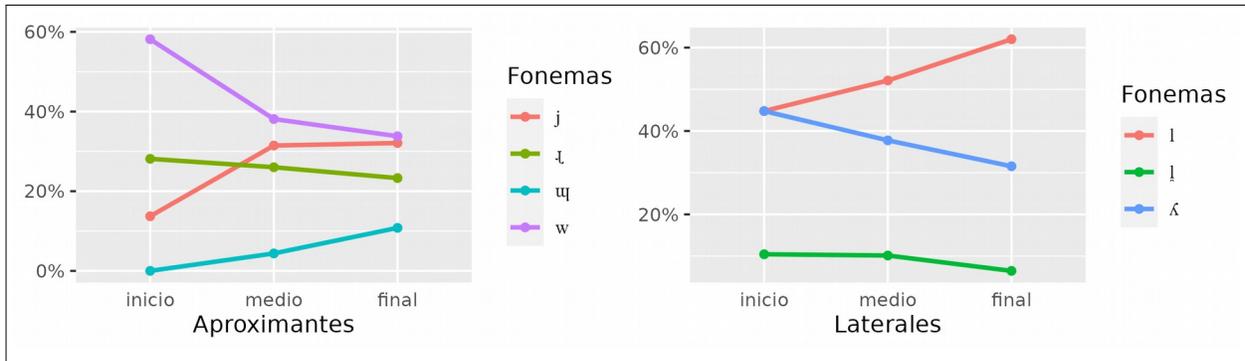


Gráfico 7.4: (continuación)



En el gráfico precedente se pueden visualizar de manera porcentual las distribuciones en las tres posiciones en el lexema. En el caso de las oclusivas y africadas, no existe una alta variación relativa entre las posiciones iniciales y media (cabe señalar que estos fonemas no ocurren en el postmargen silábico, por ende tampoco se registran al final del lexema).

Respecto de las fricativas, existe un poco más de variación de las frecuencias en las distintas posiciones. Por ejemplo, /f/ y /θ/ tienen porcentajes similares al inicio, pero no así en las otras posiciones.

En el caso de las vocales, existe una mayor variación en los porcentajes de distribución según la posición en el lexema. En este caso, las vocales más frecuentes tienen una mayor frecuencia relativa al inicio, en comparación con los otros fonemas vocálicos. No obstante, esta diferencia se estabiliza en las otras dos posiciones.

Respecto de las nasales, llama la atención que la nasal velar no tenga una mayor frecuencia relativa al final del lexema, a pesar de ser una posición fonotáctica tipológicamente preferida para este segmento. En esta clase de fonemas, es la nasal palatal la que tiene mayor ocurrencia en esa posición.

Respecto de las aproximantes, sus distribuciones son relativamente equitativas, salvo por /w/, que tiene una mayor frecuencia en la posición inicial, y /j/, que tiene una menor frecuencia en ese contexto.

Finalmente, respecto de las interdentes, en los distintos gráficos, llama la atención que sus distribuciones no sufren gran variación en las distintas posiciones.

Estas visualizaciones son una aproximación a la distribución fonemática en estructuras mayores, y creemos prudente y necesario profundizar en este asunto, por ejemplo, estudiando la posición en otros patrones fonológicos, como la de grupos vocálicos y consonánticos inter-léxicos y sílabas, como proyección para futuros estudios.

7.8.2. Ratio segmentos resonantes/no-resonantes

De acuerdo con Tambovtsev (2010), la frecuencia de los segmentos resonantes en la cadena del habla tiene un efecto en el grado de eufonía de una lengua particular. De acuerdo con el autor, la eufonía o “melodicidad” es una propiedad universal, al ser un rasgo básico en todas las lenguas (2010, p. 100).

La forma de comparar entre lenguas el grado de mayor o menor eufonía, entonces, es por medio del cálculo de la proporción de segmentos resonantes respecto de los no-resonantes. Asimismo, es posible establecer una ratio entre los segmentos resonantes observados respecto de los esperados en un corpus lingüístico; en el caso de nuestro estudio, respecto de un listado léxico monomorfémico.

El rasgo de resonancia de un segmento tiene que ver con el grado de constricción en su articulación y la posibilidad de que el sonido sea sonoro. Si la cerrazón es muy pronunciada, se imposibilita la sonoridad espontánea (Chomsky y Halle, 1968, p. 302; Burquest, 1993, p. 17). Los sonidos cuya modalidad es oclusiva, africada y fricativa suelen clasificarse como no

resonantes u obstruyentes; en cambio las vocales, aproximantes, semivocales, líquidas y nasales suelen clasificarse como sonantes o resonantes.

En primer lugar, identificamos en mapudungun aquellos segmentos que contienen el rasgo de resonancia. En este grupo de segmentos incluimos los segmentos vocálicos, más las consonantes nasales, laterales y aproximantes. En los segmentos obstruyentes, en cambio, incluimos las consonantes oclusivas, africadas y fricativas. Entonces, en resumen, en el inventario fonológico del mapudungun, se identifican 18 segmentos resonantes y 10 obstruyentes, lo que corresponde a una ratio de 1,8.

Tabla 7.11: Segmentos fonemáticos resonantes y no resonantes en mapudungun.

[+resonantes]		[-resonante]	
Vocales	/i/, /e/, /ə/, /a/, /o/, /u/	Oclusivas	/p/, /t̪/, /t/, /k/
Nasales	/m/, /ɲ/, /n/, /ɲ/, /ŋ/	Africadas	/tʃ/, /tʂ/
Laterales	/l̪/, /l/, /ɬ/	Fricativas	/f/, /θ/, /s/, /ʃ/
Aproximantes	/ɹ/, /j/, /w/, /w/		

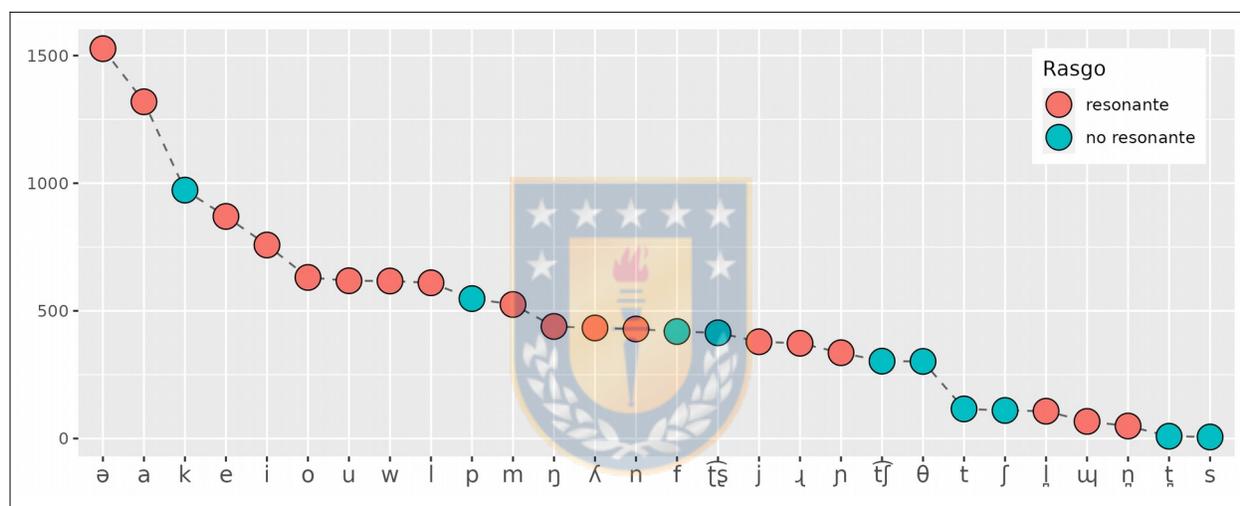
En cuanto a la ratio segmentos resonantes / obstruyentes identificados a partir de nuestro corpus lexicográfico, del total de **13287** fonemas, **10087** registran el rasgo de resonancia, mientras que solo **3200** son de tipo no resonante u obstruyente. Con esta información, se calcula una ratio de 3,15 para los segmentos resonantes. Si extrapolamos esta ratio, consideradas 100 unidades segmentales del mapudungun en el léxico, 76 de ellas serían resonantes y 24, obstruyentes.

Tabla 7.12: Frecuencias absolutas y relativas de los segmentos fonemáticos resonantes y no resonantes en mapudungun.

	n	porcentaje
resonante	10087	76
no resonante	3200	24
total	13287	100

No es sorprendente que los segmentos resonantes tengan una mayor frecuencia que los obstruyentes, pues los fonemas no se distribuyen equitativamente respecto de este rasgo en el inventario: 18 son segmentos resonantes y 10 son obstruyentes, con una ratio de 1,8. Además, es notorio el hecho de que, en nuestro corpus, los segmentos más frecuentes (con un porcentaje mayor a 4%) son mayoritariamente segmentos resonantes, incluyendo todas las vocales, más los segmentos /w/ y /l/; la excepción la constituye el segmento oclusivo velar áfono /k/. Esto se puede visualizar en el **Gráfico 7.5**.

Gráfico 7.5: Distribución de los segmentos resonantes y no resonantes en la frecuencia fonemática.



Respecto de la proporción de segmentos resonantes observados respecto de los esperados, Tambovtsev (2010) establece que el grado de eufonía esperable en cualquier lengua es de 66,7% (50% para los segmentos vocálicos + 16,7% para los segmentos consonánticos resonantes). Bajo este supuesto, el mapudungun se encuentra en el rango superior de los valores de eufonía, con una proporción de 76% de sonidos resonantes.

7.8.3. Ratio segmentos esperados y observados

Para este estudio, calculamos también la ratio entre fonemas observados y esperados (Gordon, 2016), una forma de entender la alta y baja frecuencia de los fonemas estudiados. Este indicador, además, permite hacer comparaciones con otras lenguas, al controlar la diferencia de segmentos en los inventarios fonológicos comparados (Gordon, 2016). Por último, nos permite comparar nuestros resultados con otros estudios que han reportado la frecuencia fonemática del mapudungun, pero que consignan un número distinto de fonemas (por ejemplo, Álvarez-Santullano, 1986 y Lagos, 2001)⁸.

Esta ratio se computó mediante la obtención del porcentaje esperado de cada fonema del mapudungun ($100 / 28 = 3,6\%$). Posteriormente, cada porcentaje observado se dividió por este porcentaje esperado, de tal modo que se obtuvo una ratio particular para cada fonema. En este contexto, en primer lugar se calculó este indicador de manera general para todos los segmentos; y luego separadamente según el tipo de fonema (vocálico o consonántico), con el fin de representar mejor su valor, controlando el tipo de segmento, puesto que —como vimos en la sección sobre la ratio consonante/vocal—, las consonantes tienen una mayor frecuencia que las vocales, posiblemente debido a tendencias fonotácticas de la lengua.

Eso sí, cabe señalar que este resultado puede ser objeto de mayores refinamientos, pues, por ejemplo, debería complementarse con un cálculo que controlara algunas restricciones fonotácticas, como la imposibilidad de algunas obstruyentes de ocurrir en coda silábica, y así arribar a una frecuencia esperada más precisa aún.

La **Tabla 7.13**, a continuación, grafica las ratio de frecuencia observada / esperada, tanto de todos los segmentos, como de las vocales, por un lado, y de las consonantes, por otro.

8 En *Anexos*, hemos desplegado un ejercicio que explica la ventaja del uso de este tipo de indicador de Ratio Obs./Esp., respecto del uso de porcentajes o valores absolutos.

Tabla 7.13: Ratio de frecuencias observadas y esperadas en el conteo general (“gen.”) de todos los segmentos, en las vocales (“voc.”), y en las consonantes (“cons.”).

Ratio Obs/Esp				Ratio Obs/Esp			Ratio Obs/Esp		
Fon.	gen.	voc.	cons.	Fon.	gen.	cons.	Fon.	gen.	cons.
/ə/	3.22:1	1.6:1		/m/	1.11:1	1.53:1	/θ/	0.64:1	0.88:1
/a/	2.78:1	1.38:1		/ŋ/	0.93:1	1.28:1	/t/	0.24:1	0.34:1
/k/	2.05:1		2.83:1	/ʎ/	0.91:1	1.26:1	/ʃ/	0.23:1	0.32:1
/e/	1.83:1	0.91:1		/n/	0.9:1	1.25:1	/ɺ/	0.23:1	0.31:1
/i/	1.6:1	0.79:1		/f/	0.88:1	1.22:1	/ɰ/	0.14:1	0.19:1
/o/	1.33:1	0.66:1		/t͡s/	0.87:1	1.2:1	/ɲ/	0.1:1	0.14:1
/u/	1.3:1	0.65:1		/j/	0.8:1	1.1:1	/t͡/	0.02:1	0.03:1
/w/	1.3:1		1.79:1	/ɻ/	0.79:1	1.08:1	/s/	0.01:1	0.02:1
/l/	1.29:1		1.77:1	/ɲ/	0.71:1	0.98:1			
/p/	1.15:1		1.59:1	/t͡f/	0.64:1	0.88:1			

7.8.4. Ratio segmentos sonoros y áfonos

En mapudungun, como en toda lengua, existen segmentos que se realizan en concomitancia con la activación de las cuerdas vocales o vibración glotática, los cuales se han denominado en la literatura como *segmento sonoros* (Burquest, 1993), diferenciándose de aquellos que se producen sin tal vibración, denominados *áfonos* (Burquest, 1993). Muchos idiomas incluso llegan a diferenciar dos pares de sonidos en virtud de este rasgo articulatorio (Ladefoged y Ferrari, 2012, p. 137), lo que, al menos en lo que respecta a la variante asumida en este estudio, no es el caso para el mapudungun central (Sadowsky et al., 2013).

De todas maneras, en virtud de la importancia del rasgo de sonoridad en lo que respecta a las variantes dialectales del idioma mapuche, se considera pertinente reportar la ratio de segmentos sonoros y áfonos en el corpus léxico bajo análisis.

En variantes nortinas del cordón cordillerano —grupos II y IV, de acuerdo con Croese (1980)— y en variantes nortinas de la costa (grupo I croesiano), existen realizaciones sonoras de los

fonemas fricativos /f/ y /θ/ del mapudungun. Así las cosas, palabras que en la zona central tienen /f/ y /θ/, se realizan con /v/ y /ð/, respectivamente, en esas variantes dialectales (Sánchez Cabezas, 1989; Salamanca, 1997; Salamanca y Quintrileo, 2009; Sánchez Pérez y Salamanca, 2015; Pérez y Salamanca, 2017; Mena, 2020; Saldivia y Salamanca, 2020; Urrea y Salamanca, 2021).

En lo que respecta al inventario fonológico asumido en nuestro estudio, son sonoros los cinco segmentos nasales (/m/, /ɲ/, /n/, /ɲ/, /ŋ/), las cuatro aproximantes (/ɹ/, /j/, /ɥ/, /w/), las tres laterales (/l/, /l/, /ʎ/) y los seis fonemas vocálicos (/i/, /e/, /ə/, /a/, /o/, /u/); en cambio, son áfonas, las cuatro oclusivas (/p/, /t/, /t/, /k/), las dos africadas (/tʃ/, /tʃ/) y las cuatro fricativas (/f/, /θ/, /s/, /f/). En virtud de estas dos categorías, la ratio de sonidos sonoros y áfonos es de 1,8, en el inventario fonológico. Esta configuración coincide plenamente con la reportada en la subsección sobre segmentos resonantes / no-resonantes, por lo que, de igual modo, se calcula una ratio de 3,15 para los segmentos sonoros en el corpus bajo estudio.

Ahora, en el contexto de la variación dialectal señalada, resulta interesante preguntarse por la proporción de segmentos sonoros y áfonos en el corpus léxico, de haberse considerado una variante diatópica que relevara como unidades segmentales las fricativas sonoras /v/ y /ð/. En virtud de esta interrogante, se arriba a una configuración distinta, que asume 20 fonemas sonoros y 6 fonemas sordos, con una ratio de **3,3**. En el corpus de unidades lexemáticas, en tanto, de los **13287** fonemas identificados, **10808** serían sonoros y **2479** áfonos, si asumimos una variante “sonora” de la lengua mapuche. Estas cifras nos otorgan una ratio específica para las variantes sonoras de **4,36**, bastante más elevada que la ratio de **3,15**, de la variante central asumida en este estudio.

7.9. Recuento estadístico por clase natural

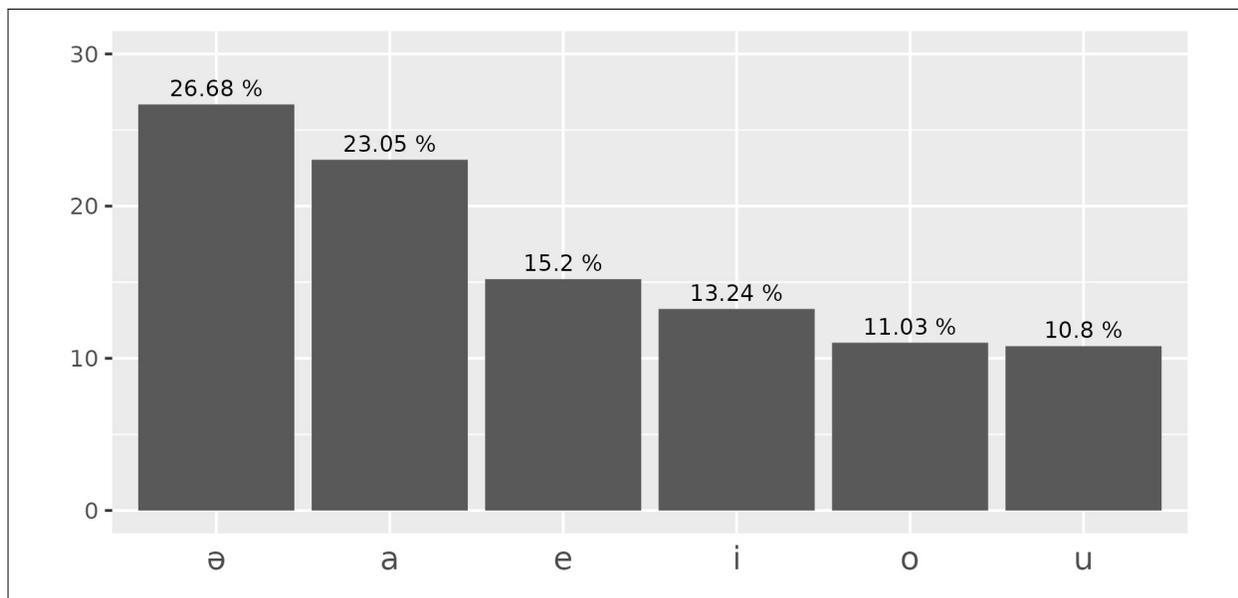
7.9.1. Vocales

Se identificaron **5723** vocales, cuyas distribuciones de frecuencia tienen la siguiente jerarquía: /ə/ > /a/ > /e/ > /i/ > /o/ > /u/. Respecto de los valores absolutos y relativos, la vocal más frecuente es la media central no-redondeada /ə/, con **1527** fonemas y un porcentaje de 26,68%; seguida de la baja central no-redondeada /a/, con **1319** unidades y un porcentaje de 23,05%. A estas dos, le siguen las vocales anterior media /e/ y anterior alta /i/, ambas no-redondeadas, con **870** y **758** unidades cada una, que corresponden al 15,2% y 13,24%, respectivamente. En el último rango, se encuentran las vocales posterior media /o/ y posterior alta /u/, ambas con rasgos de redondez labial, con **631** y **618** unidades cada una, correspondiente al 11,03% y 10,8%, respectivamente.

Tabla 7.14: Jerarquía de los fonemas vocálicos.

Fonema	Ranking	Frecuencia	Porcentaje %
ə	1	1527	26,68
a	2	1319	23,05
e	3	870	15,2
i	4	758	13,24
o	5	631	11,03
u	6	618	10,8
Total:		5723	100

Gráfico 7.6: Visualización de las frecuencias relativas de los seis fonemas del mapudungun.



A partir de una inspección visual del **Gráfico 7.6**, es interesante apreciar que las vocales /ə/ y /a/ —ambas vocales centrales— se distancian de manera importante del resto de las otras vocales. Volveremos a este punto en el apartado *Discusión*.

Como se puede apreciar en la **Tabla 7.4** —al inicio de esta sub-sección de resultados—, las vocales en mapudungun pueden ser descritas fonológicamente por dos parámetros articulatorios relacionados con la posición de la lengua: su altura, en el eje vertical; y su dirección, en el eje horizontal. Estos son los parámetros que analizamos a continuación.

7.9.1.1. Dirección de la lengua

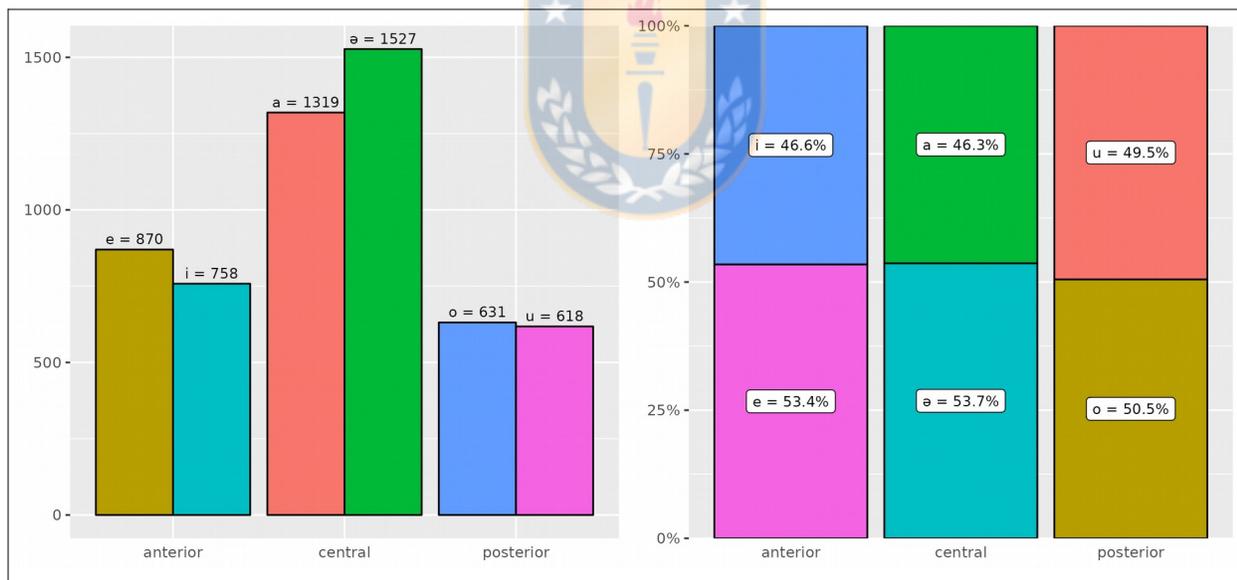
El rasgo de dirección de la lengua refiere, en líneas generales, a la posición de la lengua en la cavidad oral en un eje horizontal (Ladefoged y Johnson, 2001). En el caso de sonidos anteriores, como los fonemas /i/ y /e/, la lengua adopta una posición frontal. Desde esta posición, la lengua se desplaza hacia la zona central para la realización de sonidos adscritos a los fonemas /ə/ y /a/. Por último, a la zona posterior se adscriben los fonemas /o/ y /u/. En términos acústicos, la dirección de la lengua está relacionada con la frecuencia del segundo formante vocálico, donde

valores más altos de este indicador se corresponden con una articulación más frontal o anterior, y viceversa.

Respecto de este rasgo, los tipos de vocales se ordenan en nuestros datos de la siguiente manera: central > anterior > posterior. Las vocales centrales /ə/ y /a/ son las más frecuentes, con **2846** realizaciones, correspondientes al 49,73%. Luego, las vocales anteriores /e/ e /i/ se realizan **1628** veces, lo que corresponde al 28,45% de las realizaciones vocálicas. Finalmente, las vocales posteriores /o/ y /u/ se realizan **1249** veces, lo que equivale al 21,82%.

Como se ve, existe una asociación entre el rasgo de dirección de la lengua y la frecuencia de realización de las vocales, donde las vocales centrales tienen una mayor frecuencia que las anteriores y estas respecto de las posteriores.

Gráfico 7.7: Frecuencias absolutas y relativas de las vocales del mapudungun según el parámetro articulatorio de dirección de la lengua.



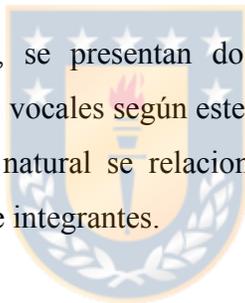
El aporte que realiza cada fonema a su propia clase según el rasgo dirección de la lengua es similar, como es posible apreciar en el **Gráfico 7.7**. Específicamente, en el caso de las vocales anteriores y centrales, la diferencia entre los dos pares de fonemas corresponde a alrededor de 7 puntos porcentuales; y en el caso de las vocales posteriores, esa diferencia es cercana al 1%.

7.9.1.2. Altura de la lengua

El segundo parámetro articulatorio que es analizado respecto de las vocales es el de altura de la lengua. En términos acústicos, la altura de la lengua está inversamente relacionada con la frecuencia del primer formante vocálico, donde valores más altos de este indicador se corresponden con una articulación más baja y viceversa.

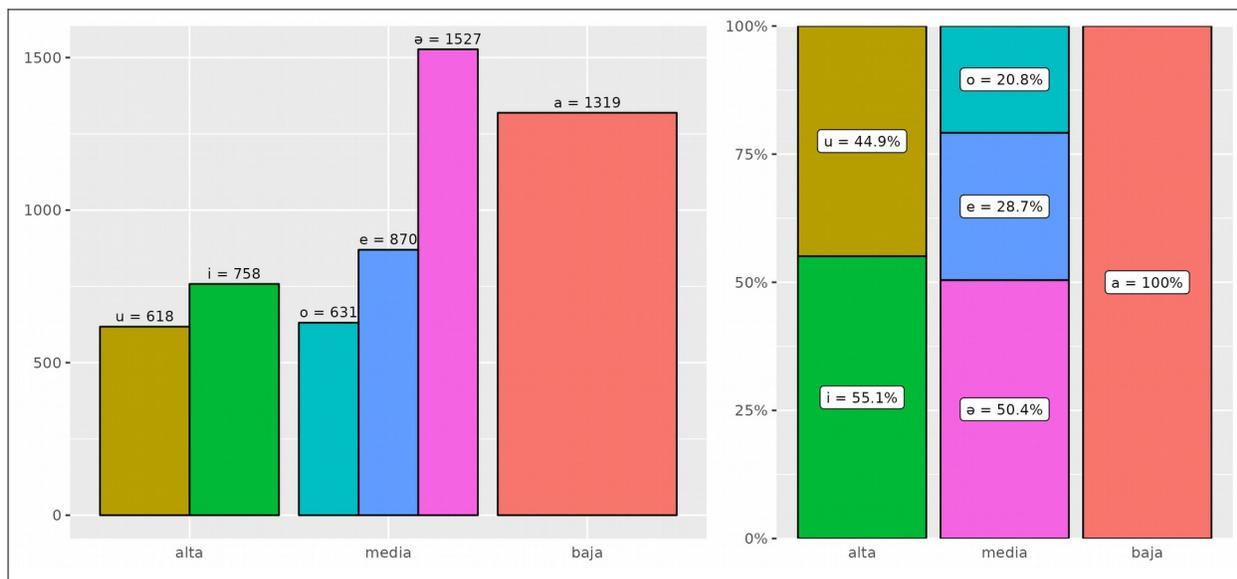
Respecto de este rasgo, los tipos de vocales en nuestros datos se ordenan de la siguiente manera: media > alta > baja. Las vocales medias, /ə/, /e/ y /o/, son las más frecuentes, con **3028** realizaciones, equivalente a casi el 53% de todas las realizaciones vocálicas bajo este rasgo articulatorio. Son seguidas por las vocales altas, representadas por /i/ y /u/, con **1376** realizaciones en total, correspondiente al 24%. Finalmente, la vocal baja, con un solo representante (/a/), es la de menor frecuencia, con **1319** realizaciones, equivalente al 23%.

A continuación, en el **Gráfico 7.8**, se presentan dos visualizaciones que despliegan las frecuencias absolutas y relativas de las vocales según este parámetro. Como se observa, el aporte que realiza cada fonema a su clase natural se relaciona con el hecho de que dichas zonas articulatorias tienen distinto número de integrantes.



En el caso de las vocales medias, la /ə/ es la que tiene una frecuencia mayor, correspondiente a un poco más de la mitad de los casos (50,4%); seguida de /e/, con un 28,7%; y /o/, con un 20,8%. En cambio, en el caso de las vocales altas, la diferencia entre ambos aportes es levemente superior a los 10 puntos porcentuales, donde /i/ es la más frecuente, con más del 55% de los casos. La vocal /a/, por último, al ser la única integrante de la clase natural “baja”, instancia el 100% de los casos.

Gráfico 7.8: Distribución de las frecuencias absolutas y relativas de las vocales, según altura.



Finalmente, la **Tabla 7.15**, a continuación, compara el aporte de cada segmento vocálico de manera general y de acuerdo a ambos rasgos articulatorios ya revisados. A modo de ejemplificación de la forma de leer la tabla, podemos señalar que la vocal media central /ə/ tiene una frecuencia absoluta de **1527** realizaciones, valor que comprende el **50,4%** de las **3028** realizaciones de segmentos según *altura media*, y, a su vez, abarca un **53,7%** de las **2846** realizaciones de la dirección *central*. Finalmente, la última cifra corresponde al porcentaje de /ə/ respecto de las vocales en general.

De igual modo, la **Tabla 7.15** informa que las distribuciones según los parámetros de altura y dirección de la lengua tienen sus clases predominantes, cercanas al 50% de los casos: **49,7%** para las vocales centrales, según la dirección de la lengua, y casi **53%** para las vocales medias, según altura de la lengua. Las otras clases comparten valores porcentuales similares que van desde el **21,8%** al **28,4%**.

También es posible señalar que la predominancia de la clase *central* en el rasgo dirección de la lengua se debe a la presencia de las vocales /ə/ y /a/, que tienen las mayores frecuencias fonemáticas de la lengua en general. En el caso de la altura de la lengua, la mayor frecuencia de

las vocales medias se relaciona, sin duda, con el hecho de que, como hemos señalado, ocurren tres integrantes, donde uno de ellos, /ə/, es la vocal más frecuente del idioma.

Tabla 7.15: Frecuencias absolutas y relativas de vocales según rasgos articulatorios. Las cifras en negrita refieren a valores absolutos. El primer porcentaje (en azul) refiere a la proporción del segmento según clase *altura de la lengua*, el segundo porcentaje (en rojo), a su proporción según *dirección de la lengua*; y el tercer porcentaje, a su proporción general respecto de todas las vocales.

<i>altura</i>	<i>dirección</i>			<i>totales</i>
	<i>anterior</i>	<i>central</i>	<i>posterior</i>	
<i>alta</i>	758		618	1376
	55,1%		44,9%	24%
	46,6%		49,5%	
	13,2%		10,8%	
<i>media</i>	870	1579	631	3028
	28,7%	50,4%	20,8%	52,9%
	53,4%	53,7%	50,5%	
	15,2%	26,7%	11,0%	
<i>baja</i>		1319		1319
		100,0%		23,0%
		46,3%		
		23,0%		
<i>totales</i>	1628	2846	1249	5723
	28,4%	49,7%	21,8%	100%

7.9.2. Consonantes

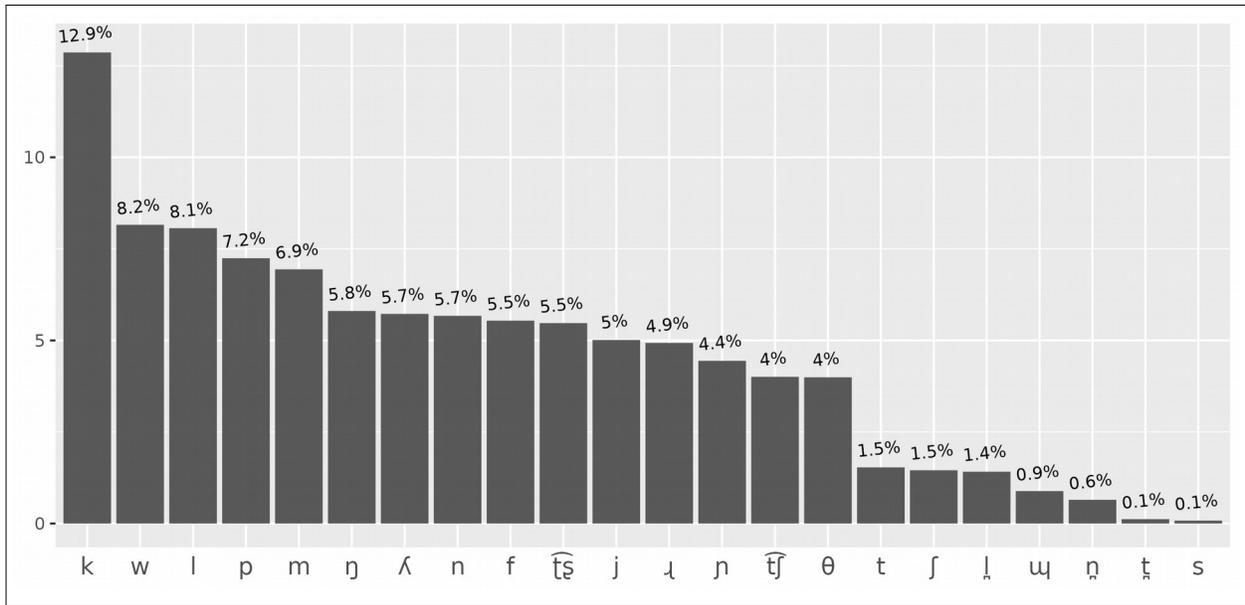
En el caso de las consonantes, se aborda su frecuencia según los parámetros articulatorios de modo de articulación y lugar de articulación. Estos parámetros han sido mencionados en la subsección 7.5.1. *Cuadro fonológico*, que contiene los fonemas del mapudungun dispuestos en una tabla según su pertenencia a estas distintas clases articulatorias.

En primer lugar, en la **Tabla 7.16** reiteramos las frecuencias absolutas de los fonemas consonánticos del mapudungun reportados por nuestro análisis y añadimos sus frecuencias relativas porcentuales y la posición que ocupan en un orden decreciente (considerado sólo este tipo de segmentos y no los fonemas en general). El **Gráfico 7.9**, por su parte, se focaliza en este último aspecto.

Tabla 7.16: Jerarquía según las frecuencias absolutas y relativas de las consonantes mapuche.

Rank.	Fon.	Frec.	Perc.	Rank.	Fon.	Frec.	Perc.	Rank.	Fon.	Frec.	Perc.
1	/k/	973	12,86	9	/f/	419	5,54	17	/ʃ/	110	1,45
2	/w/	617	8,16	10	/t͡s/	414	5,47	18	/l̥/	107	1,41
3	/l/	610	8,06	11	/j/	379	5,01	19	/ɥ/	67	0,89
4	/p/	548	7,24	12	/ɬ/	373	4,93	20	/ŋ̥/	49	0,65
5	/m/	525	6,94	13	/ɲ/	336	4,44	21	/t͡ʃ/	9	0,12
6	/ŋ/	439	5,8	14	/t͡ʃ̥/	303	4,01	22	/s/	6	0,08
7	/ʎ/	433	5,72	15	/θ/	302	3,99				
8	/n/	429	5,67	16	/t/	116	1,53				
									Total:	7564	

Gráfico 7.9: Frecuencia relativa de las consonantes del mapudungun.

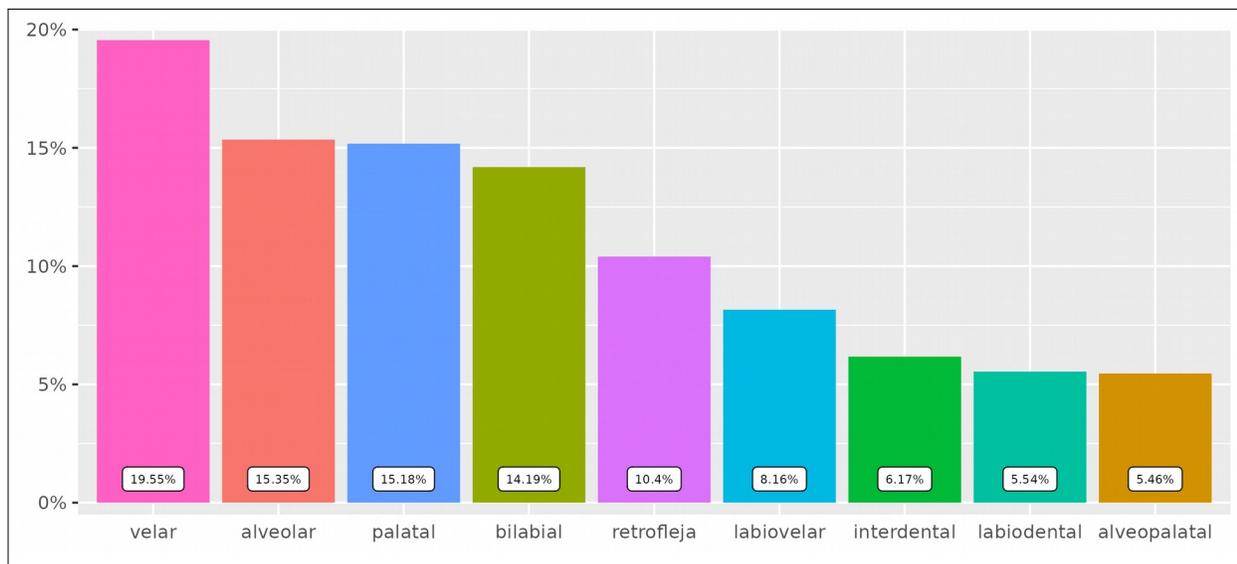


Como se observa, a diferencia de los fonemas vocálicos, existe una consonante que tienen una frecuencia ostensiblemente mayor que el resto de los sonidos consonánticos: la oclusiva velar /k/.

7.9.2.1. Consonantes por lugar de articulación

Las consonantes se pueden agrupar por su lugar de articulación. La jerarquía, en este caso, corresponde a velares > alveolares > palatales > bilabiales > retroflejas > labiovelar > interdental > labiodental > alveopalatales, tal como se muestra en el **Gráfico 7.10**.

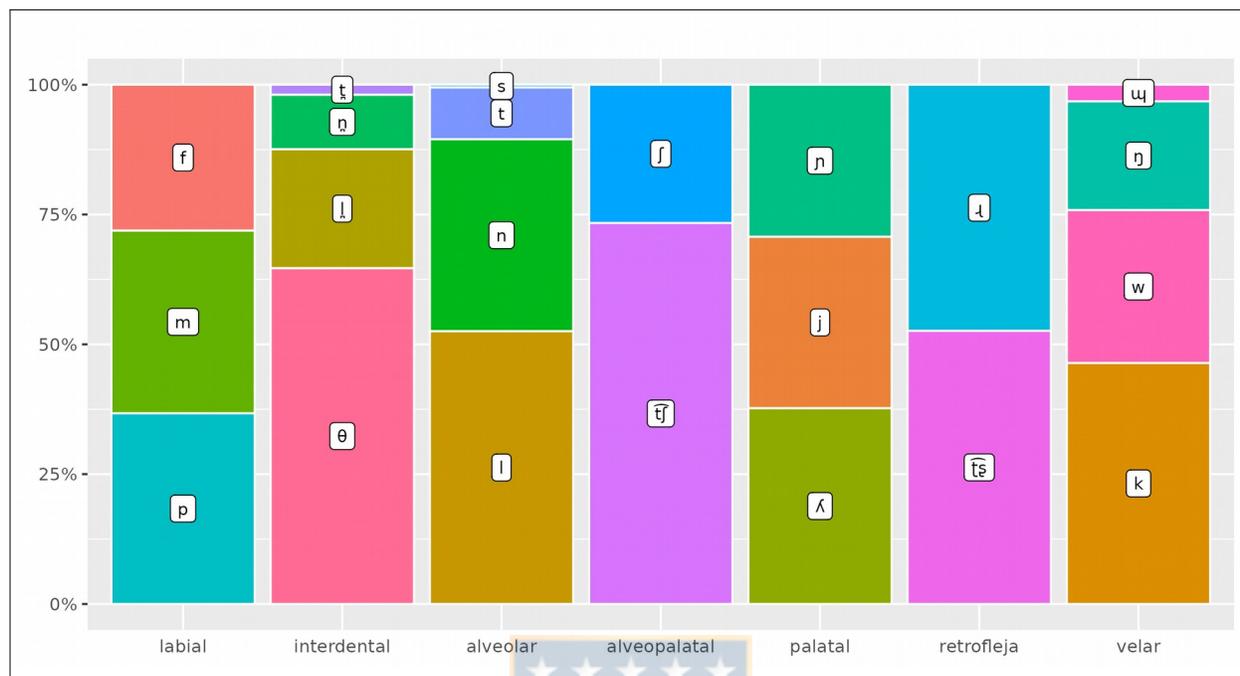
Gráfico 7.10: Frecuencias relativas de las consonantes del mapudungun según lugar de articulación.



El predominio del punto de articulación velar, superior al 19%, corresponde a las frecuencias combinadas de las consonantes /k/, /ŋ/ y /ŋ̃/. Este lugar de articulación es seguido por el alveolar, el cual tiene cuatro integrantes: /l/, /n/, /t/ y /s/; el palatal, con tres: /ʎ/, /j/ y /ɲ/; el bilabial, con dos: /p/ y /m/; y el retroflejo, también con dos: /tʂ/ y /ɬ/. Más abajo en la jerarquía de frecuencia, se encuentra la zona labiovelar, con un integrante: /w/; la zona interdental, con cuatro: /θ/, /l̪/, /n̪/ y /t̪/; la zona labiodental, con uno: /f/; y la zona alveopalatal, con dos: /ʃ/ y /tʃ/.

En el **Gráfico 7.11**, a continuación, es posible apreciar el aporte de cada uno de los fonemas a su clase natural, según el punto de articulación. En este gráfico, se han acumulado los valores de las consonantes bilabiales y labiodental en una sola zona, denominada labial, con tres integrantes. Asimismo, la consonante labiovelar /w/ se ha incluido dentro de las velares.

Gráfico 7.11: Distribución de las frecuencias relativas de las consonantes del mapudungun según lugar de articulación.



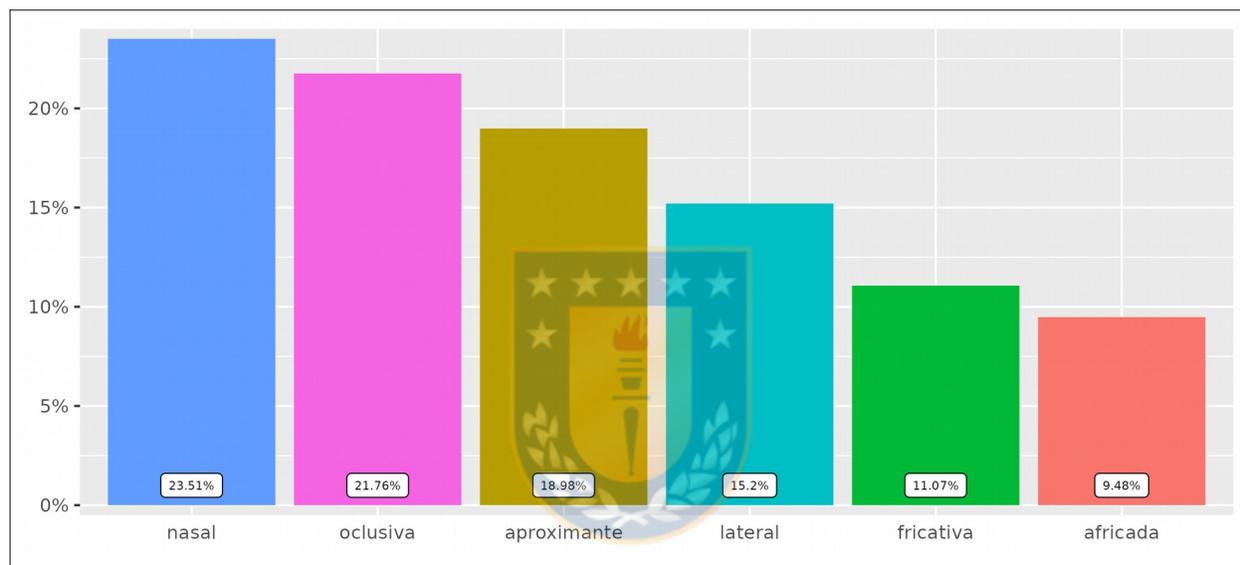
A partir del **Gráfico 7.11**, es posible apreciar el aporte de cada segmento a su clase natural respectiva. En el caso del lugar de articulación labial, que comprende tanto a las bilabiales /p/ (36,7%) y /m/ (35,2%), como a la labiodental /f/ (28,1%), se visualiza una distribución relativamente equitativa en sus porcentajes; caso similar al lugar de articulación palatal: /ʌ/ (37,7%), /j/ (33%) y /ɲ/ (29,3%); y al retroflejo: /t͡ʂ/ (52,6%) y /ɕ/ (47,4%). Por otro lado, los lugares de articulación interdental, alveolar, alveopalatal y velar exhiben distribuciones asimétricas. En efecto, en el caso de las interdental, la fricativa /θ/ (64,7%) tiene un porcentaje mucho mayor que los segmentos /l/ (22,9%), /ɲ/ (10,5%) y /t/ (1,9%). En la zona alveolar, se despliega una situación similar, donde la lateral /l/ (52,5%) y la nasal /n/ (37%) tienen porcentajes considerablemente más altos que /t/ (10%) y /s/ (0,5%). Mismo caso para la zona alveopalatal, donde la africada /tʃ/ (73,4%) tiene una frecuencia mucho mayor que su contraparte fricativa /ʃ/ (26,6%). Por último, en la zona velar —que comprende también una aproximante velar con redondeamiento labial /w/—, la oclusiva /k/ aporta casi la mitad de los casos (46,4%), seguida de la aproximante labiovelar /w/ (29,4%) y la nasal /ɲ/ (20,9%); en cambio, la fricativa /w̥/ presenta la frecuencia más baja de esta zona articulatoria, con solo un 3,2%. En este contexto,

sería interesante investigar si en otras lenguas donde ocurren los mismos segmentos en (algunas de) estas zonas articulatorias, sus aportes son similares o se producen diferencias importantes.

7.9.2.2. Consonantes por modo de articulación

Respecto del modo de articulación, la jerarquía corresponde a nasales > oclusivas > aproximantes > laterales > fricativas > africadas, tal como se aprecia en el siguiente gráfico.

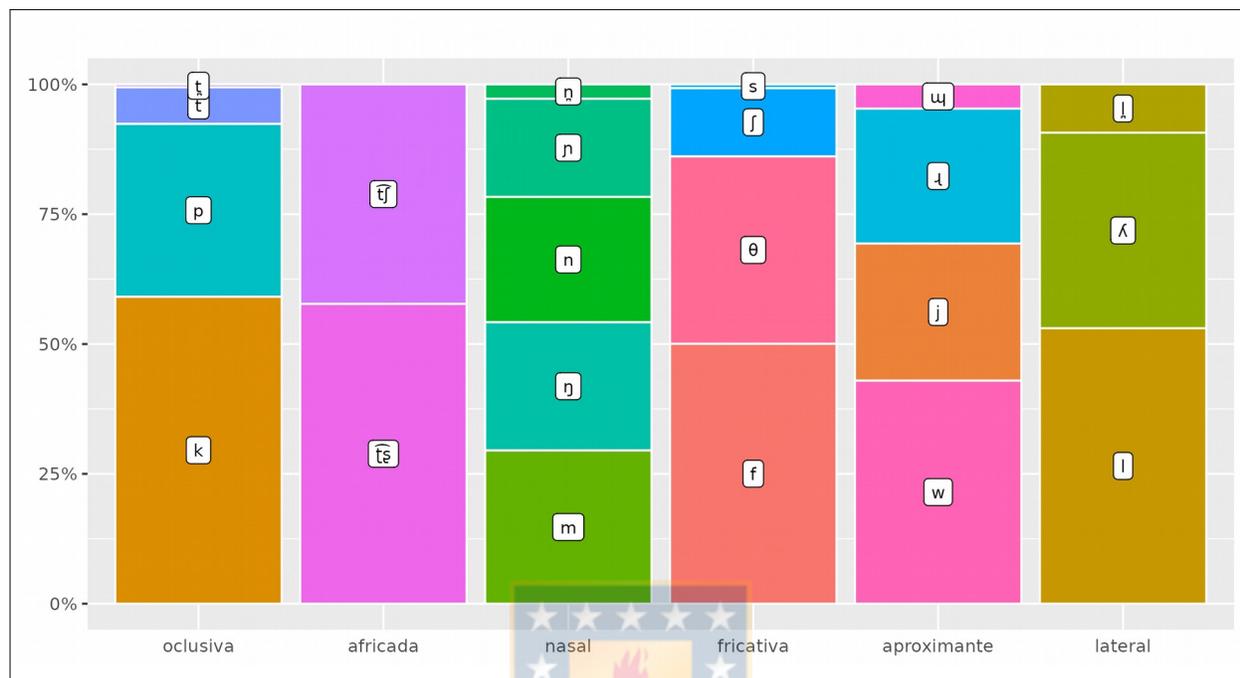
Gráfico 7.12: Frecuencias relativas de las consonantes del mapudungun según modo de articulación.



Las consonantes nasales (/m/ > /ŋ/ > /n/ > /ɲ/ > /ɳ/) son las más frecuentes, con más del 23,5% de las realizaciones; luego, siguen las consonantes oclusivas (/k/ > /p/ > /t/ > /tʰ/), con casi un 22%; las aproximantes (/w/ > /j/ > /ɹ/ > /ɻ/), con casi un 19%; las laterales (/l/ > /ʎ/ > /ɭ/), con más de 15%; las fricativas (/f/ > /θ/ > /s/ > /ʃ/), con un 11%; y, finalmente, las africadas (/tʃ/ > /tʂ/), con más de un 9%.

En términos generales, en la mayoría de los casos se cumple la relación entre el número de integrantes de la clase y la mayor frecuencia relativa del parámetro; salvo en el caso de las consonantes fricativas, las cuales, a pesar de contar con cuatro integrantes, se sitúan en el penúltimo lugar de la jerarquía.

Gráfico 7.13: Distribución de las frecuencias relativas de las consonantes del mapudungun según modo de articulación.



Al igual que en la visualización de las consonantes por lugar de articulación, al analizar los respectivos aportes de los segmentos a su categoría según modo de articulación, es posible identificar al menos una clase donde se produce una distribución equitativa; a saber, la de las africadas, donde /t̥s̥/ (57,7%) y /t̥ʃ/ (42,3%) registran porcentaje relativamente cercanos.

En cambio, en las otras categorías, se produce el escenario donde dos o más fonemas realizan un aporte relativo mucho mayor que los otros segmentos, configurándose categorías cuyas distribuciones son asimétricas. Tal es el caso del modo oclusivo, donde /k/ (59,1%) tiene el porcentaje mayor —más de la mitad de los casos—, seguido de /p/ (33,3%) y /t/ (7%); en cambio, la interdental /t̥/ registra un porcentaje extremadamente bajo (0,5%). En el caso del modo nasal, se aprecian porcentajes similares para la bilabial /m/ (29,5%), velar /ŋ/ (24,7%), alveolar /n/ (24,1%) y palatal /ɲ/ (18,9%), los cuales son considerablemente mayores que el porcentaje de la interdental /ɲ/ (2,8%). Para las fricativas, la labiodental /f/ (50,1%) registra la mitad de los casos; seguida de la interdental /θ/ (36,1%); luego, la alveopalatal /ʃ/ (13,1%); y, por

último, la alveolar /s/ (0,7%), que registra un porcentaje considerablemente más bajo. Caso similar para las las aproximantes, donde la labiovelar /w/ (43%) tiene el mayor porcentaje; seguida por /j/ (26,4%) y /ɲ/ (26%), con porcentajes similares; y, por último, /ɥ/ (4,7%), con el valor más bajo de esta serie. Finalmente, en las laterales, existe una situación similar, donde más de la mitad de los casos los registra la alveolar /l/ (53%); seguida de la palatal /ʎ/ (37,7%); y, más abajo, la interdental /ʎ/ (9,3%).

A continuación, en la **Tabla 7.17**, desplegamos los valores de las frecuencias absolutas y relativas de las consonantes mediante una matriz de contingencia que permite comparar los diversos aportes de los segmentos fonológicos según las variables articulatorias de *modo* y *lugar de articulación*. También se incluye la proporción porcentual general respecto de todos los segmentos consonánticos. En este caso, a diferencia de las descripciones del **Gráfico 7.11**, ya mencionado, se configuran las consonantes labiodental /f/ y labiovelar /w/ en sus respectivas categorías según *lugar de articulación*, por lo que los porcentajes específicos de sus variables articulatorias varían respecto a los ya reportados en esos gráficos.

A modo de ejemplo de la lectura de la **Tabla 7.17**, si atendemos a la oclusiva velar /k/, podemos señalar que tiene un valor absoluto de **973** ocurrencias en el corpus, que corresponde al **59,1%** de las **1646** realizaciones oclusivas, y al **65,8%** de las **1479** realizaciones velares. Finalmente, el porcentaje de **12,9%** refiere al aporte de este segmento a la totalidad de fonemas consonánticos (7564). De este modo, gracias a la **Tabla 7.17**, es posible visualizar valores y porcentajes de todos los segmentos consonánticos, en virtud de dos rasgos articulatorios basales.

Tabla 7.17: Frecuencias absolutas y relativas de las consonantes según rasgos articulatorios. La cifra en negrita refiere a valores absolutos. El primer porcentaje (en azul) refiere a la proporción del segmento según *modo de articulación*, el segundo porcentaje (en rojo) a su proporción según *lugar de articulación*, y el tercer porcentaje a su proporción general respecto de todas las consonantes.

	BILAB.	LABDEN.	INTERD.	ALVEOL.	ALVPAL.	RETROF.	PALAT.	VELAR	LABVEL.	TOTAL
OCLUSIVA	548 33,3% 51,1% 7,2%		9 0,5% 1,9% 0,1%	116 7,0% 10,0% 1,5%				973 59,1% 65,8% 12,9%		1646 21,8%
AFRICADA					303 42,3% 73,4% 4,0%	414 57,7% 52,6% 5,5%				717 9,5%
NASAL	525 29,5% 48,9% 6,9%		49 2,8% 10,5% 0,6%	429 24,1% 37,0% 5,7%			336 18,9% 29,3% 4,4%	439 24,7% 29,7% 5,8%		1778 23,5%
FRICATIVA		419 50,1% 100,0% 5,5%	302 36,1% 64,7% 4,0%	6 0,7% 0,5% 0,1%	110 13,1% 26,6% 1,5%					837 11,1%
APROX.						373 26,0% 47,4% 4,9%	379 26,4% 33,0% 5,0%	67 4,7% 4,5% 0,9%	617 43,0% 100,0% 8,2%	1436 19%
LATERAL			107 9,3% 22,9% 1,4%	610 53,0% 52,5% 8,1%			433 37,7% 37,7% 5,7%			1150 15,2%
TOTAL	1073 14,2%	419 5,5%	467 6,2%	1161 15,3%	413 5,5%	787 10,4%	1148 15,2%	1479 19,6%	617 8,2%	7564 100%

7.10. Clasificación en categorías

Para la clasificación de las frecuencias fonemáticas en categorías, se utilizó un algoritmo de agrupamiento implementado en el paquete de R `Ckmeans.1d.dp` (Wang y Song, 2011), que permite realizar un análisis de *cluster* o agrupamiento con datos unidimensionales; es decir, con una sola variable (en este caso, *frecuencia de ocurrencia*).

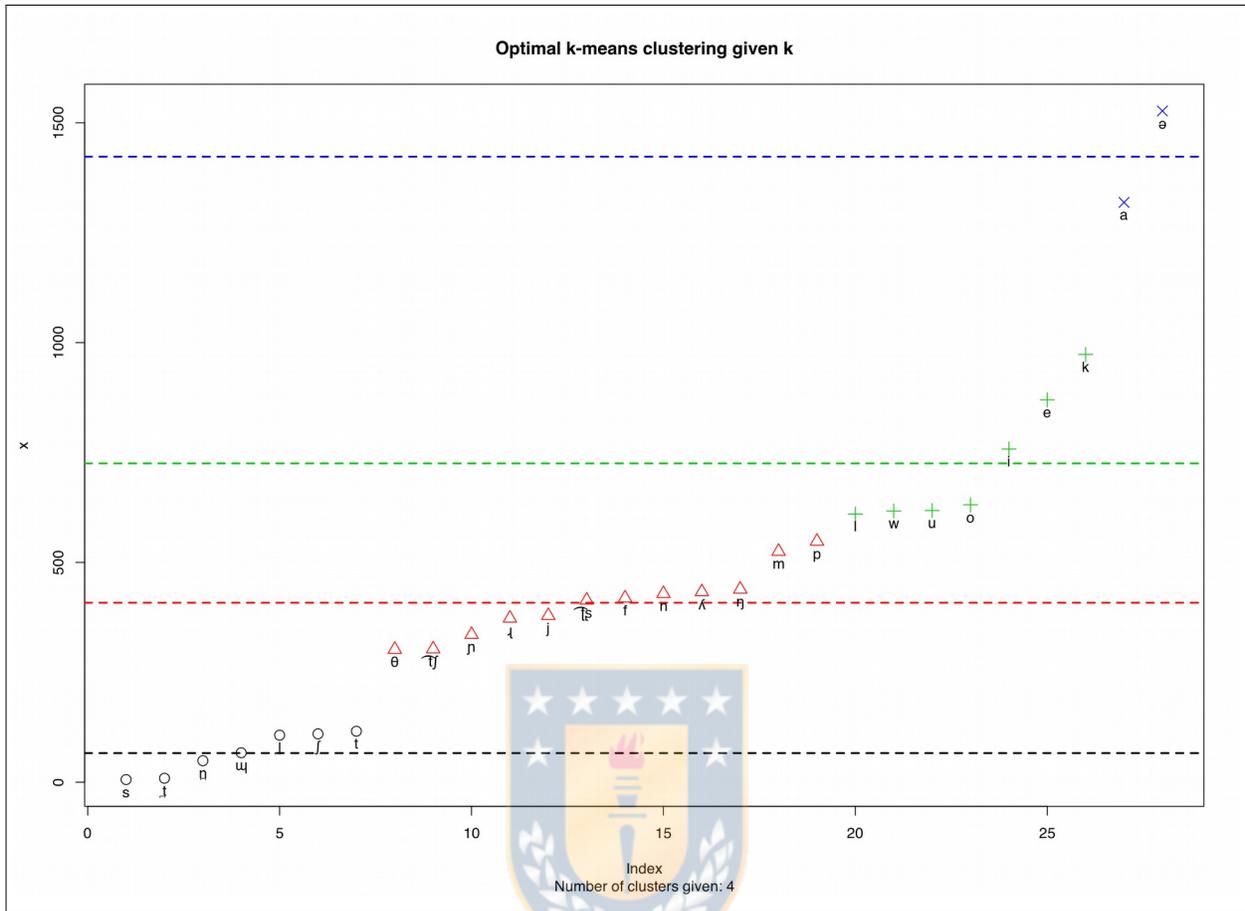
El algoritmo calcula de manera iterativa el centro de los agrupamientos y las distancias óptimas de los valores respecto de esos centros (Wang y Song, 2011)⁹. En el caso de nuestras frecuencias, nos permite identificar, dados cuatro o cinco grupos, el centro de cada agrupamiento y de ese modo identificar estadísticamente cuáles son los integrantes óptimos de esas mediciones.

Esta distinción en cuatro o cinco grupos viene motivada por el número de categorías que se han reportado en la literatura sobre frecuencia fonemática; particularmente, para el castellano (por ejemplo, Alarcos (1965) y Rojo (1991) reportan cinco categorías; mientras que Navarro (1946) y Guirao y García (1990) dividen las frecuencias fonemáticas en cuatro grupos).

En primer lugar, calculamos los agrupamientos con cuatro centros. Al cargar nuestros valores de frecuencia fonemática en la función `Ckmeans.1d.dp` (Wang y Song, 2011), se identificaron los siguientes valores para cuatro centros de agrupamiento: 66,3, 408,3, 725,3 y 1423, identificados en el **Gráfico 7.14**.

9 “The algorithm iteratively calculates the within-cluster sum of squared distances, modifies group membership of each point to reduce the within-cluster sum of squared distances, and computes new cluster centers until local convergence is achieved.” (Wang y Song, 2011, p. 1)

Gráfico 7.14: Análisis de agrupamiento con cuatro centros.



En este sentido, las frecuencias fonemáticas del mapudungun, al dividirse en cuatro categorías, forman los agrupamientos que se presentan en la **Tabla 7.18**:

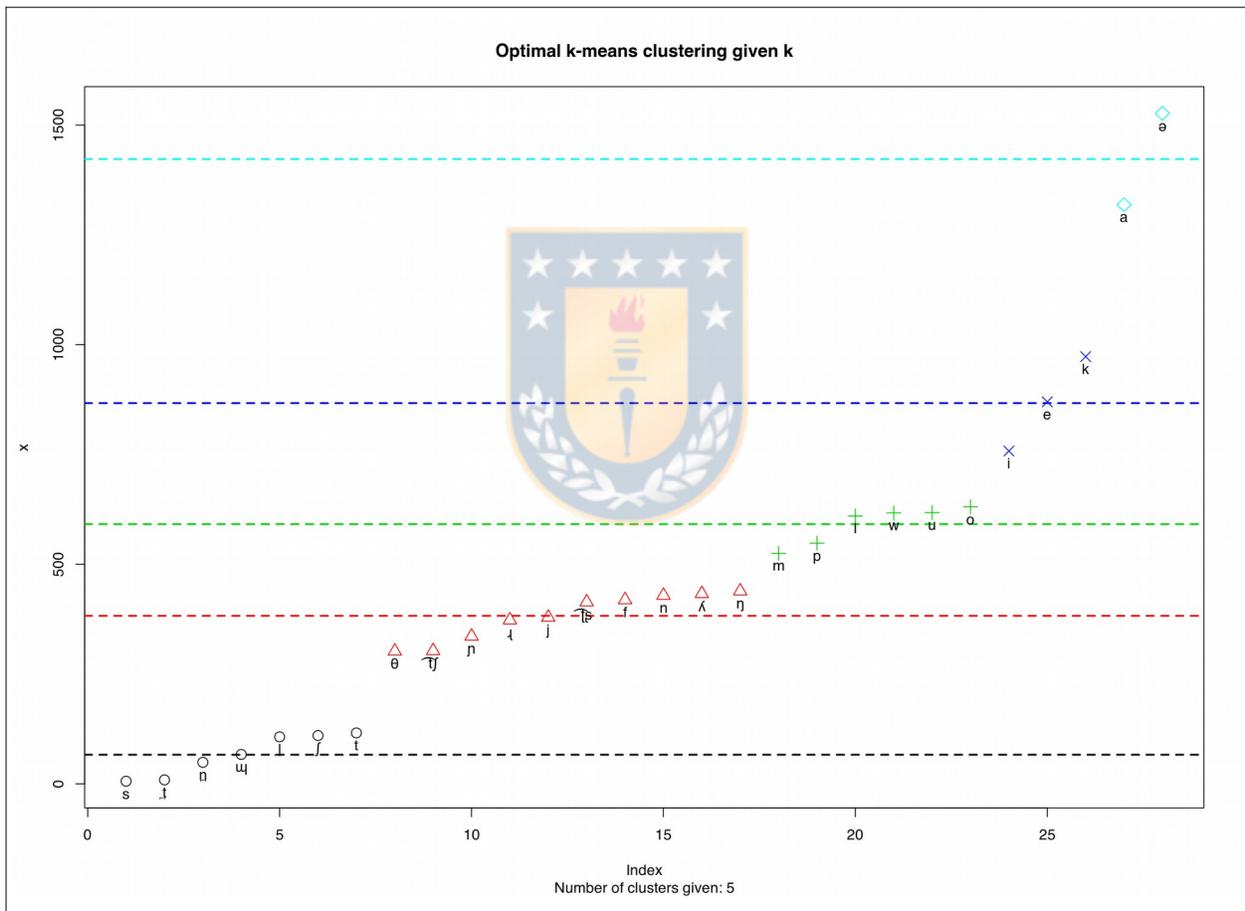
Tabla 7.18: Categorías de agrupamiento de los fonemas mapuche según su frecuencia (cuatro grupos).

Categorías	Fonemas
I) Fonemas de frecuencia mayor o 'muy frecuentes'	/ə/ > /a/
II) fonemas de alta frecuencia o 'bastante frecuentes'	/k/ > /e/ > /i/ > /o/ > /u/ > /w/ > /l/
III) fonemas de frecuencia moderada / normal o 'frecuentes'	/p/ > /m/ > /η/ > /l/ > /n/ > /f/ > /ts/ > /j/ > /ʎ/ > /ɲ/ > /tf/ > /θ/
IV) fonemas de baja frecuencia o 'poco frecuentes'	/t/ > /f/ > /l/ > /w/ > /ɲ/ > /t/ > /s/

De igual modo, calculamos los agrupamientos respecto de cinco centros, con el fin de identificar ese número de clases de frecuencias. Los centros de agrupamiento en este caso se producen en torno a los siguientes valores absolutos: 66,3, 382,7, 591,5, 867 y 1423.

En el **Gráfico 7.15**, se visualizan las distribuciones de la frecuencia fonemática y, por medio de líneas horizontales, los cinco centros. Los distintos colores y formas identifican las cinco clases de fonemas según su frecuencia.

Gráfico 7.15: Análisis de agrupamiento con cinco centros.



Los cinco agrupamientos que resultan del cálculo realizado por la función Ckmeans.1d.dp (Wang y Song, 2011) se despliegan en la **Tabla 7.19**. En la primera columna, se explicitan las categorías propuestas; en la columna central, se despliegan los cinco rangos de la ratio de frecuencia observada y esperada que se reportó en la sección 7.8.3, para los fonemas del mapudungun; y en la última columna se señalan los fonemas que se adscriben a cada categoría.

Tabla 7.19: Cinco categorías de agrupamiento de los fonemas mapuche según su frecuencia, comparado con la ratio de frecuencia observada/esperada.

Categorías	Ratio Obs/Esp	Fonemas
I) fonemas de muy alta frecuencia	2,1 - 3,2	/ə/ > /a/
II) fonemas de alta frecuencia	1,51 – 2,0	/k/ > /e/ > /i/
III) fonemas de frecuencia moderada	1,1 – 1,5	/o/ > /u/ > /w/ > /l/ > /p/ > /m/
IV) fonemas de baja frecuencia	0,51 – 1	/ŋ/ > /ʎ/ > /n/ > /f/ > /tʂ/ > /j/ > /ɬ/ > /ɲ/ > /tʃ/ > /θ/
V) fonemas de muy baja frecuencia	0,01 – 0,5	/t/ > /ʃ/ > /ɭ/ > /ɰ/ > /ɳ/ > /t̪/ > /s/

Así, entonces, los fonemas del mapudungun también se pueden categorizar en cinco clases, definibles de acuerdo con la ratio de frecuencia observada y esperada. Si se comienza por los siete segmentos de “muy baja frecuencia” (/s/, /t̪/, /ɳ/, /ɰ/, /ɭ/, /ʃ/ y /t/), se observa que estos se encuentran en el rango que va desde 0 a 0,5, en esta ratio. La siguiente categoría, la más poblada, en tanto, comprende diez fonemas “de baja frecuencia”, con una ratio Obs./Esp. que va desde 0,51 a 1 –vale decir, menor a la frecuencia esperada, que es igual a 1–, e incluye los fonemas /θ/, /tʃ/, /ɲ/, /ɬ/, /j/, /tʂ/, /f/, /n/, /ʎ/ y /ŋ/. La tercera categoría –que incluye a los fonemas de “frecuencia moderada”– agrupa los fonemas con una frecuencia observada superior a la esperada, con un rango de 1,1 a 1,5, e incluye a seis fonemas: /m/, /p/, /l/, /w/, /u/, /o/. La siguiente categoría agrupa los fonemas de “alta frecuencia”, con un rango Obs./Esp. de 1,51 a 2,0, e incluye a tres fonemas (/i/, /e/, /k/), donde cabe señalar que /k/, con una ratio de 2,05, se establece en el límite alto de esta categoría. Finalmente, la categoría de fonemas de “muy alta frecuencia”, con un valor de ratio de frecuencia observada / esperada superior a 2,1, incluye solamente los fonemas vocálicos /a/ y /ə/.

A la luz de todo lo señalado, esto es, al poder establecer una asociación entre los rangos de frecuencia observada y esperada para los fonemas del mapudungun estudiados a partir del léxico y el cálculo establecido por medio de la función de $Ckmeans.1d.dp$ (Wang y Song, 2011), es posible señalar que cinco grupos o categorías son adecuadas para clasificar los fonemas del mapudungun, según su frecuencia de ocurrencia en un corpus lexemático.

La **Tabla 7.20** indica los porcentajes de distribución de cada una de las cinco categorías ya identificadas, además del número de fonemas y la frecuencia acumulada por grupo. El primer grupo, con solo dos representantes de muy alta frecuencia, comprende el 21,4% de las realizaciones. El segundo grupo, con tres fonemas de alta frecuencia, representa el 19,6%. Estos dos grupos, con cinco fonemas en total, abarcan casi el 41% de todas las realizaciones segmentales del mapudungun. El tercer grupo, con seis segmentos de frecuencia moderada, comprende un 26,71% de todas las realizaciones. El cuarto grupo, con diez fonemas de baja frecuencia, es el grupo con más realizaciones y aporta un 28,8% del total. Finalmente, los siete fonemas restantes forman parte del grupo con la menor frecuencia (aportan solo el 3,5% del total).

Tabla 7.20: Número de realizaciones y sus porcentajes por categoría de frecuencia de fonemas.

Categorías	N.º de Fon.	Freq. Abs.	Freq. Rel.	Freq. Acum.
I) Muy alta frecuencia	2	2846	21,42%	21,42%
II) Alta frecuencia	3	2601	19,58%	40,99%
III) Frecuencia moderada	6	3549	26,71%	67,71%
IV) Baja frecuencia	10	3827	28,80%	96,51%
V) Muy baja frecuencia	7	464	3,49%	100,00%
Totales:	28	13280	100%	

8. Discusión

Ya señalados los resultados de nuestro estudio, y atendidos los objetivos y preguntas de investigación, en esta sección procederemos a discutir algunos aspectos de los mismos, a la luz de la literatura que existe sobre frecuencia fonemática, en general, y sobre la frecuencia de los fonemas del mapudungun, en particular.

En primer lugar, se resumen los principales resultados de nuestro estudio. Posteriormente, se comparan dichos resultados con los contenidos en Márquez (2016) y Lagos (2001), poniendo atención en las divergencias y congruencias entre estos tres estudios que consideran una variedad central del idioma mapuche. Posteriormente, comparamos nuestros resultados con los reportados en Álvarez-Santullano (1986), estudio que se focalizó en una variedad meridional del mapudungun, denominada *williche* ('sureña'). Para realizar estas comparaciones, utilizamos la ratio de frecuencia de segmentos observados, en relación con la de segmentos esperados (ratio Obs./Esp.), razón que nos permite comparar resultados de sistemas fonológicos disímiles (Gordon, 2016); esto, en nuestro caso, debido a diferencias dialectales y/o de análisis fonémico¹⁰.

Una vez comparados los estudios de frecuencia fonemática del mapudungun, nos abocamos a un foco de interés que surgió de nuestro estudio e intentamos establecer una ruta que permita dar explicación a su manifestación. En específico, nos referiremos a la frecuencia más alta de la vocal /ə/ en nuestro estudio, resultado que no se condice con los reportes intralingüísticos de frecuencia fonemática sobre esta lengua, como se verá a continuación, y tampoco con la tendencia “universal” de que /a/ sea la vocal intralingüísticamente e interlingüísticamente más frecuente (Crothers, 1978; Ladefoged y Ferrari, 2012; Gordon, 2016).

Finalmente, evaluamos la pertinencia de nuestros criterios de filtrado, detallados en el *Capítulo 6*, y el mérito que a nuestro juicio poseen para su posible utilización en estudios de frecuencia fonemática en otras lenguas.

10 En *Anexos*, hemos desplegado un ejercicio que explica la ventaja del uso de este tipo de indicador de ratio Obs./Esp., por sobre el uso de porcentajes o valores absolutos.

8.1. Resumen de los hallazgos

Nuestro estudio tuvo como objetivo general de investigación determinar la frecuencia de los 28 fonemas del mapudungun que ocurren en unidades lexemáticas monomorfémicas. Nos abocamos a este propósito preparando un corpus de **2914** lexemas, a partir de un proceso exhaustivo de filtrado de cada una de las 9703 entradas de la versión modernizada del *Diccionario mapudungún-español* de Augusta (1916/2017), considerada la obra lexicográfica de mayor envergadura de la lengua mapuche (Salas, 1985, 1992b; Molineaux, 2016). El proceso de muestreo del diccionario consideró una serie de criterios de índole lingüística —tanto en el dominio de la fonología como de la gramática—, además de criterios sociolingüísticos. Respecto de los criterios lingüísticos, se controlaron fenómenos morfológicos, como la derivación y flexión afijal, la reduplicación y la fonotaxis anómala en algunas interjecciones. En el caso de *fluctuaciones fonemáticas*, se seleccionó un lexema por medio de un criterio de tipo lexicográfico, sustentado en la relevancia que otorga el autor a ciertas entradas por sobre otras. En cuanto a los criterios sociolingüísticos, se filtraron aquellas entradas que correspondían a préstamos lingüísticos del castellano y el quechua, y se seleccionó una entrada en el caso de *alternancias fonemáticas*, siguiendo un criterio que privilegiaba la Realización Indicadora de Vitalidad. Creemos que el tipo de muestreo desarrollado en este estudio puede ser replicado en otras lenguas, con el fin de desarrollar investigaciones de tipo comparativas que permitan aumentar el caudal de información fonostadística intralingüística y extralingüística.

Una vez obtenido el corpus de unidades lexemáticas monomorfémicas, por medio de diversas funciones del programa **R** (R Core Team, 2021) y la *suite* de paquetes **Tidyverse** (Wickham et al., 2019) —procesos documentados en las secciones respectivas, con el fin de asegurar la replicabilidad y fiabilidad de nuestra investigación—, se extrajeron **13287** unidades fonológicas. Estas unidades se transcribieron siguiendo una notación fonológica *ad hoc* para su conteo fonostadístico y se tabularon con el fin de reportar las frecuencias absolutas y relativas de los 28 fonemas del mapudungun, asumidos en este estudio siguiendo la literatura sobre este tema (por ejemplo, Salas, 1976; Smeets, 2008; Sadowsky et al., 2013, entre otros). Cada fonema, entonces, se contabilizó tomando en consideración las variables pertinentes, como el lexema de origen, su

posición en la estructura fonológica, su pertenencia al tipo vocálico o consonántico, sus rasgos articulatorios mínimos para distinguirlos de otros segmentos, etc. Esta información nos permitió dar cuenta de la jerarquía general de los fonemas del mapudungun en virtud de su frecuencia, que corresponde al primer objetivo específico de nuestro estudio. A modo de resumen, se obtuvo una distribución similar para frecuencias fonemática según se reporta de manera visual en la literatura (por ejemplo, González, 2016), donde unos pocos fonemas aportan la mayor cantidad de instancias, configurando de este modo una distribución muy desigual (Martin, 2007, p. 4)¹¹.

Respecto de nuestro segundo objetivo específico, indicamos la *ratio consonante/vocal*. Sobre este aspecto, cabe señalar que los segmentos consonánticos ocurren con una mayor frecuencia que los vocálicos, con una ratio de 1,32:1. En términos porcentuales, esta ratio significa que en un corpus de 100 unidades fonemáticas, alrededor de 57 de ellas son consonantes. Al investigar con mayor profundidad esta relación, se estableció que el lugar donde se manifiesta esta asimetría en las distribuciones vocal-consonante es al inicio de la unidad lexemática monomorfémica (10,6% para las vocales; 89,4% para las consonantes), puesto que en las otras dos posiciones (media y final), la distribución tiende, en términos comparativos, a ser casi simétrica, con una ligera ventaja de las vocales.

Siempre dentro del segundo objetivo específico sobre ratio de tipos de segmentos, reportamos que los segmentos *resonantes* en mapudungun ocurren con una ratio de 3,16 respecto de los *obstruyentes*. La misma ratio se obtuvo para los segmentos *sonoros* respecto de los *áfons*, debido a la congruencia entre los rasgos articulatorios que dan cuenta de la sonoridad, por un lado, y la resonancia, por otro, en la variedad dialectal asumida en este estudio. Ahora, con el fin de incorporar la variación dialectal que manifiesta la fonología del mapudungun, se reportó una ratio de segmentos sonoros/áfons para las variedades dialectales denominada “sonoras” —‘rama norte’, siguiendo la nomenclatura del estudio dialectológico de Croese (1980)—, debido a la activación de las cuerdas vocales de dos de sus fricativas: /v/ y /ð/ (algunos estudios que

11 Véase Macklin-Cordes y Round (2020) para una discusión respecto de la dificultad de evaluar modelos matemáticos que expliquen las distribuciones de las frecuencias fonemáticas intralingüísticas.

tratan sobre estas variedades son Fernández Garay, 1988; Sánchez, 1989; Salamanca, 1997; Viegas Barros, 1999; Viegas Barros, 2005; Salamanca y Quintrileo, 2009; Sánchez y Salamanca, 2015; Fuentes y Salamanca, 2016; Henríquez, 2017; Saldivia y Salamanca, 2020; Mena, Figueroa y Salamanca, 2020; Urrea y Salamanca, 2021, entre otros). En este caso, la ratio sonoridad/no sonoridad aumentó a un 4,37 para los segmentos sonoros, respecto de los áfonos.

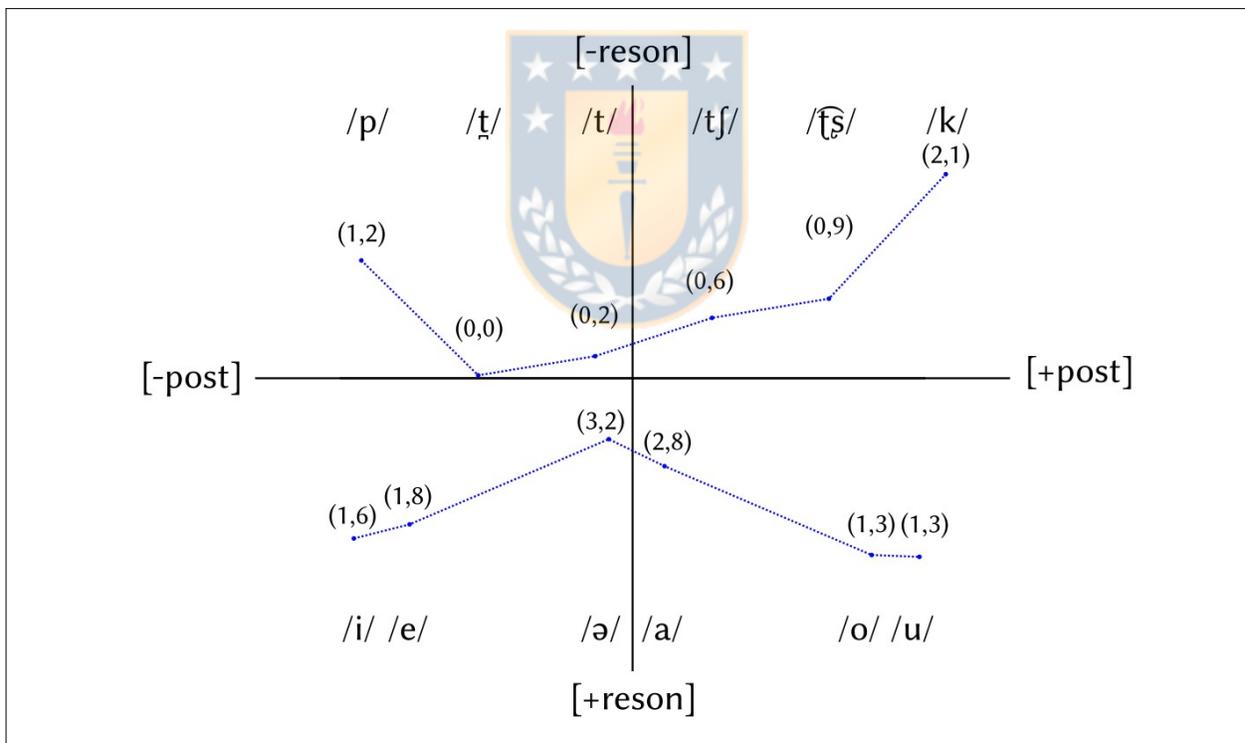
Por último, respecto de estas relaciones entre tipos de segmentos, consideramos pertinente indicar una *ratio de frecuencia observada / esperada* para cada fonema del mapudungun, calculada como la razón entre el número de ocurrencias fonemáticas y el número de instancias de cada segmento, si ocurrieran con la misma frecuencia (Gordon, 2016). En este sentido, según nuestra investigación, de los 28 fonemas del mapudungun, 11 tienen una frecuencia observada que es superior a la esperada, y 17 tienen una frecuencia menor. Es interesante que los intervalos de la ratio Obs./Esp. coinciden con las cinco categorías propuestas para la clasificación de los fonemas del mapudungun según su frecuencia.

En cuanto al tercer objetivo específico, el cual propone realizar un recuento estadístico de las frecuencias de los fonemas del mapudungun en cuanto a clases naturales, se reportaron, en primer lugar, las frecuencias relativas para los segmentos vocálicos, encontrándose la siguiente jerarquía de frecuencia: /ə/ > /a/ > /e/ > /i/ > /o/ > /u/. Posteriormente, se reportaron las frecuencias relativas en virtud de dos parámetros articulatorios mínimos necesarios para describir los fonemas vocálicos del mapudungun, a saber: *altura* y *dirección de la lengua*. En el caso de la *dirección de la lengua*, se encontró una relación entre la dirección lingual y la frecuencia fonemática, donde las vocales centrales tienen las mayores frecuencias, seguidas de las vocales anteriores, y finalmente las vocales posteriores con las menores frecuencias. Es posible que esta asociación guarde relación con características tipológicas como son las de facilidad de articulación y mayor capacidad de percepción, especialmente al compararlo con la alta frecuencia de consonantes obstruyentes en la zona velar y bilabial, lo que podría ser una compensación articulatorio-perceptual de la dirección de la lengua y las zonas de constricción del flujo de aire.

En este sentido, ambos aspectos tributarían para construir y mantener lexemas con mayores características perceptuales, y una óptima configuración articulatoria. En este sentido, sería interesante comprobar si patrones del tipo /kə/ - /ka/ - /pə/ - /pa/, etc. son sobreutilizados, respecto de estructuras como /pu/ - /ki/ - /to/ - /te/, etc., para la expresión lexemática monomorfémica.

El **Gráfico 8.1**, a continuación, presenta la configuración de estos rasgos articulatorios. Esbozamos esta consideración, no obstante, con la cautela de que mayor investigación de esta eventual relación articulatoria y la frecuencia fonemática en el léxico del mapudungun es necesaria, especialmente con metodologías de estadística inferencial.

Gráfico 8.1: Representación de la configuración de los datos de frecuencias fonemática y rasgos de zona articulatoria junto con resonancia.



Siguiendo con este tercer objetivo específico, se reportaron las frecuencias fonemáticas absolutas y relativas de los segmentos consonánticos. En este caso, la consonante más frecuente, con creces, es /k/, con un 12,9% de las realizaciones. Luego, se encuentran las resonantes /w/ y /l/,

con valores cercanos al 8% para cada una. El cuarto y quinto lugar lo ocupan dos consonantes bilabiales (/p/ y /m/), con aproximadamente 7% cada una. Entre el 4% y 6%, se encuentran diversos tipos de consonantes: /ŋ/, /ʎ/, /n/, /f/, /t͡ʃ/, /j/, /ɬ/, /ɲ/, /t͡ʃ/, y /θ/. Por último, con valores inferiores al 1,5%, se registran, en orden decreciente, /t/, /ʃ/, /l/, /w/, /ɲ/, /t/ y /s/.

En cuanto a la frecuencia fonémica de acuerdo con las clases naturales, seguimos dos parámetros articulatorios, los cuales se consideran suficientes para describir todos estos tipos de segmentos en la lengua mapuche; a saber, *lugar y modo de articulación*. Así, la jerarquía en el caso de las consonantes por lugar de articulación es: velares > alveolares > palatales > bilabiales > retroflejas > labiovelar > interdental > labiodental > alveopalatal. La zona velar es la de mayor frecuencia relativa (19,6%), debido a que incluye la alta frecuencia de /k/. Al acumular los valores de otra articulación relacionada —los de la velar con redondamiento labial /w—, esta zona de articulación asciende a una frecuencia de casi 28%. Las zonas alveolar y palatal se encuentran en el segundo y tercer lugar, respectivamente, con valores superiores al 15%. Si bien la zona bilabial ocupa el cuarto lugar, con 14,2%, si se considerara entre sus valores los de la zona labiodental, en conjunto ocuparía el segundo lugar, con 19,7%. La zona retrofleja ocupa el quinto lugar, con 10%, seguida de la interdental y alveopalatal, con valores cercanos al 5,5%.

Finalmente, y en atención al cuarto objetivo específico de investigación, se clasificaron los fonemas del mapudungun en cinco categorías, siguiendo en parte lo adoptado por la literatura (por ejemplo, Alarcos, 1965; Rojo, 1991) y en parte por lo arrojado por un algoritmo que nos permitió clasificar los fonemas en torno a centros de agrupamiento. Las cinco categorías, es interesante consignar, coincidieron con intervalos de la ratio de frecuencia observada/esperada. En consecuencia, los fonemas del mapudungun se pueden clasificar en fonemas de muy alta frecuencia (/ə/ y /a/), de alta frecuencia (/k/, /e/, /i/), de frecuencia moderada (/o/, /u/, /w/, /l/, /p/, /m/), de baja frecuencia (/ŋ/, /ʎ/, /n/, /f/, /t͡ʃ/, /j/, /ɬ/, /ɲ/, /t͡ʃ/, y /θ/), y de muy baja frecuencia (/t/, /ʃ/, /l/, /w/, /ɲ/, /t/ y /s/).

8.2. Comparación con otros estudios de frecuencia fonemática

Con el fin de comparar nuestros resultados con los reportados en la literatura sobre frecuencia fonemática del mapudungun, nos concentramos en tres reportes. En primer lugar, contrastamos nuestras frecuencias con las contenidas en Márquez (2016) y Lagos (2001), que también se realizaron a partir de un dialecto central del mapudungun (la comparación con Márquez (2016), además de coincidir con esta variedad dialectal, es interesante por cuanto emplea criterios e insumos similares a los utilizados para nuestro estudio). Finalmente, comparamos nuestros resultados con los de Álvarez-Santullano (1986), por ser una variedad dialectal distinta y por contar, incluso, con unidades funcionales no representadas en las otras variedades. Todos estos datos se despliegan en la **Tabla 8.1**, a continuación:

Tabla 8.1: Frecuencias fonemáticas de tres corpus de mapudungun, expresadas como la ratio de la frecuencia observada / frecuencia esperada (ratio Obs./Esp.). Se destaca la cifra comparativamente más alta.

FON.	BERR.	MÁR.	LAG.	FON.	BERR.	MÁR.	LAG.	FON.	BERR.	MÁR.	LAG.
ə, u	3,22	2,30	1,01	m	1,11	1,13	1,47	θ, ð	0,64	0,55	0,30
a	2,78	3,18	2,92	ŋ	0,93	0,70	0,14	t	0,24	0,41	1,51
k	2,05	2,23	1,35	ʎ	0,91	0,83	0,20	ʃ	0,23	0,03	N/A
e	1,83	2,23	4,18	n	0,90	0,96	1,77	ɭ	0,23	0,22	0,08
i	1,60	1,77	2,92	f	0,88	0,74	0,76	ɰ, ɣ	0,14	0,07	N/A
o	1,33	1,26	0,98	ʧ	0,87	0,76	0,59	ɲ	0,10	0,30	0,03
u	1,30	1,61	1,09	j	0,80	0,69	0,52	ʈ	0,02	0,05	0,02
w	1,30	1,43	0,64	ɬ	0,79	0,81	0,41	s	0,01	0,16	0,13
l	1,29	1,01	0,95	ɳ	0,71	0,76	0,78				
p	1,16	1,13	0,61	ʧ̂	0,64	0,67	0,64				

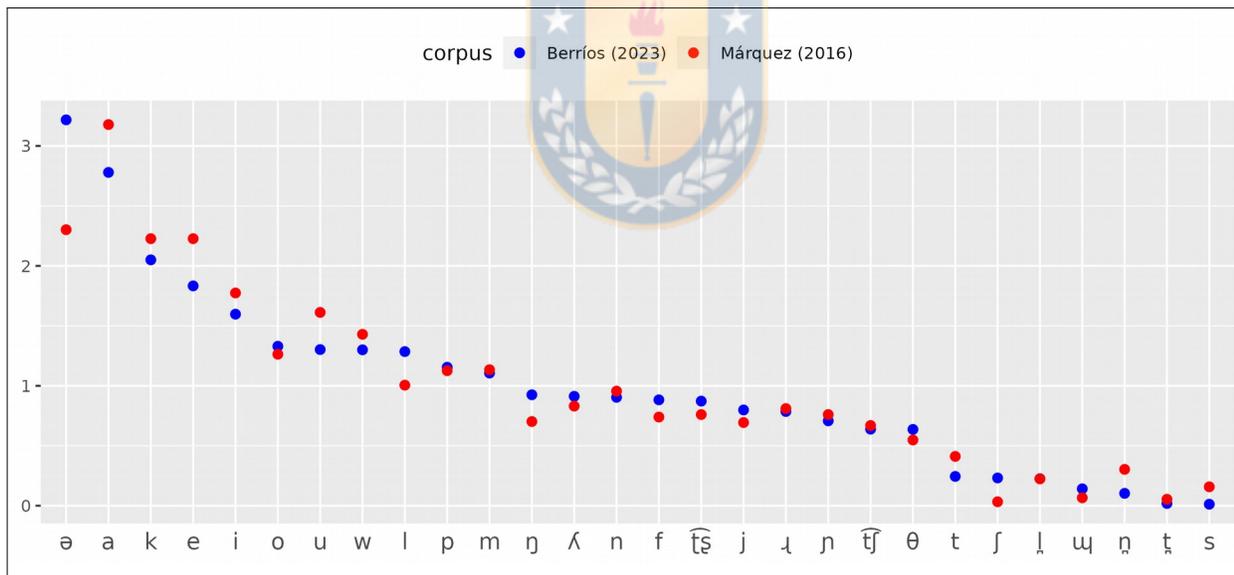
Esta tabla resume las frecuencias fonemáticas reportadas en nuestro estudio, titulado “BERR.”, y que utiliza como insumo el *Diccionario mapudungún-español* de Augusta (1916/2017), además de las señaladas en Márquez (2016), titulado “MÁR.”, y que corresponde a los datos acumulados de los diccionarios de Catrileo (1996) y Hernández y Ramos (1997); finalmente, se incluyen los resultados contenidos en Lagos (2001), abreviado como “LAG.”, en dicha tabla. Con el fin de

hacer posible la comparación, se normalizaron los símbolos fonemáticos asumidos para cada estudio. Esto quiere decir, por ejemplo, que los fonemas /ð/ y /ɣ/ plasmados en Márquez (2016) se expresan en este trabajo con los fonemas /θ/ y /ɥ/, respectivamente; y el fonema /w/, de Lagos (2001), se representa aquí junto con /ə/. En este caso, no se mencionan los datos reportados en Álvarez-Santullano (1986), por ser otra variedad dialectal. Dichos resultados serán indicados en su propia sub-sección, más adelante.

8.2.1. Comparación con Márquez (2016)

Con el fin de visualizar las principales diferencias, presentamos los resultados de nuestra investigación y los reportados de manera acumulada en Márquez (2016), en la forma de un gráfico de dispersión.

Gráfico 8.2: Comparación de las frecuencias, por medio de la ratio Obs./Esp., de nuestro estudio y las reportadas en Márquez (2016).



A partir de la visualización del **Gráfico 8.2**, es posible apreciar que nuestros resultados y los reportados en Márquez (2016) presentan valores similares de ratio de frecuencias observadas / esperadas, cuando se consideran desde un punto de vista general; es decir, existen valores muy congruentes en la forma de la curva que se genera a partir de la frecuencia fonemática.

La **Tabla 8.2**, a continuación, permite la comparación de las frecuencias de los fonemas en ambos estudios, por medio de la indicación de la diferencia de los valores de la ratio Obs./Esp. Hay algunas discrepancias asociadas a segmentos clave; por ejemplo, el segmento /ə/ en nuestro estudio tiene una frecuencia observada que es tres veces superior a la esperable; en cambio, en el caso de Márquez (2016), este fonema ocurre un poco más del doble de la frecuencia esperada. De igual modo, si bien el fonema /o/ también ocurre con una mayor frecuencia en nuestro corpus, esta diferencia es bastante menor (0,1, como se puede apreciar en la tabla). Por otro lado, las vocales /a/, /e/, /i/, y /u/, en Márquez (2016), se registran con una frecuencia mayor en comparación con nuestro estudio, aunque para /i/, la diferencia es menor (-0,17).

Tabla 8.2: Comparación de las frecuencias de fonemas en el corpus del presente estudio y en el reportado en Márquez (2016), utilizando la ratio de frecuencia observada / esperada.

Fon.	/ə/	/a/	/k/	/e/	/i/	/o/	/u/	/w/	/l/	/p/	/m/	/ŋ/	/ʎ/	/n/
Berr.	3,22	2,78	2,05	1,83	1,6	1,33	1,3	1,3	1,29	1,15	1,11	0,93	0,91	0,9
Már.	2,3	3,18	2,23	2,23	1,77	1,26	1,61	1,43	1,01	1,13	1,13	0,7	0,83	0,96
Diff.	0,92	-0,4	-0,18	-0,4	-0,17	0,07	-0,31	-0,13	0,28	0,02	-0,02	0,23	0,08	-0,06

Fon.	/f/	/t͡s/	/j/	/ɹ/	/p/	/t͡ʃ/	/θ/	/t/	/ʃ/	/l/	/w/	/ŋ/	/t/	/s/
Berr.	0,88	0,87	0,8	0,79	0,71	0,64	0,64	0,24	0,23	0,23	0,14	0,1	0,02	0,01
Már.	0,74	0,76	0,69	0,81	0,76	0,67	0,55	0,41	0,03	0,22	0,07	0,3	0,05	0,16
Diff.	0,14	0,11	0,11	-0,02	-0,05	-0,03	0,09	-0,17	0,2	0,01	0,07	-0,2	-0,03	-0,15

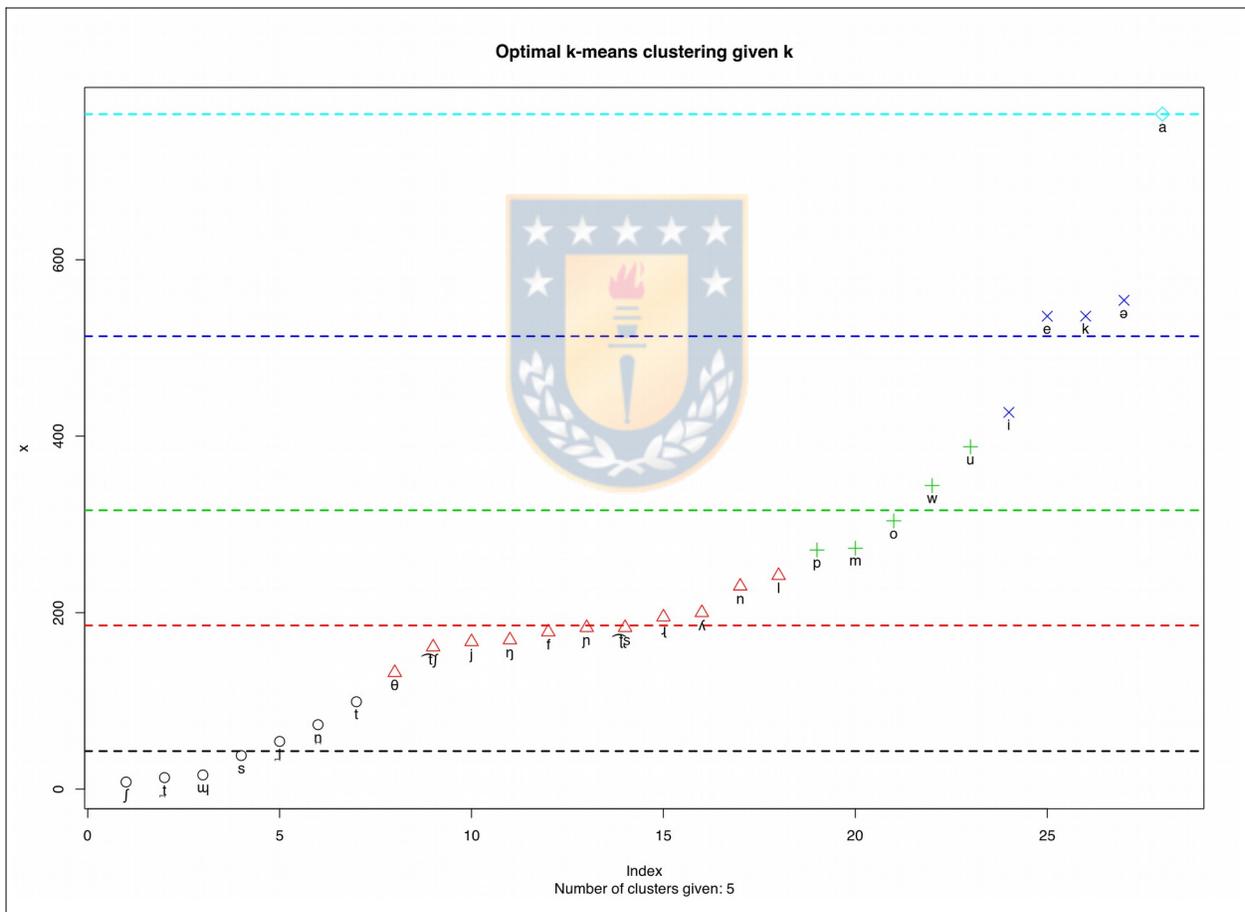
En el caso de las consonantes, nuestro estudio reporta frecuencias observadas mayores respecto del estudio de Márquez (2016) para los fonemas consonánticos /l/, /p/ /ŋ/, /ʎ/, /f/, /t͡s/, /j/, /θ/, /ʃ/ y /w/, pero en general las diferencias no son muy altas, si se compara con lo que se observa respecto de algunas vocales. En resumen, solo para cinco segmentos /ə/, /a/, /e/, /u/ y /l/ encontramos diferencias entre ambos estudios que son mayores a 0,25.

Asimismo, a modo de comparación, también agrupamos los resultados de frecuencia fonemática de Márquez (2016) en cinco centros de agrupamiento, siguiendo el algoritmo implementado en el paquete de R Ckmeans.1d.dp (Wang y Song, 2011), para así poder contrastar los resultados de

esta operación con los de nuestro estudio, los cuales fueron expuestos en la sub-sección 7.10 *Clasificación en categorías*.

El **Gráfico 8.3**, a continuación, identifica los valores para cinco centros de agrupamiento de las frecuencias reportadas en Márquez (2016), los cuales gravitan en torno a los valores: 43, 185, 316, 513 y 765. Alrededor de estos centros, se ordenan las frecuencias de los distintos fonemas, configurando de ese modo cinco categorías.

Gráfico 8.3: Análisis de agrupamiento con cinco centros, según reporte de frecuencias en Márquez (2016).



La **Tabla 8.3**, a continuación, ordena los resultados de Márquez (2016) —además de los resultados de nuestro estudio— en cinco categorías de frecuencia.

Tabla 8.3: Cinco categorías de agrupamiento de los fonemas mapuche según su frecuencia fonemática reportada en este estudio y en el de Márquez (2016). En color rojo, se relevan dos segmentos que varían en cuanto a su pertenencia a una categoría específica en nuestro estudio.

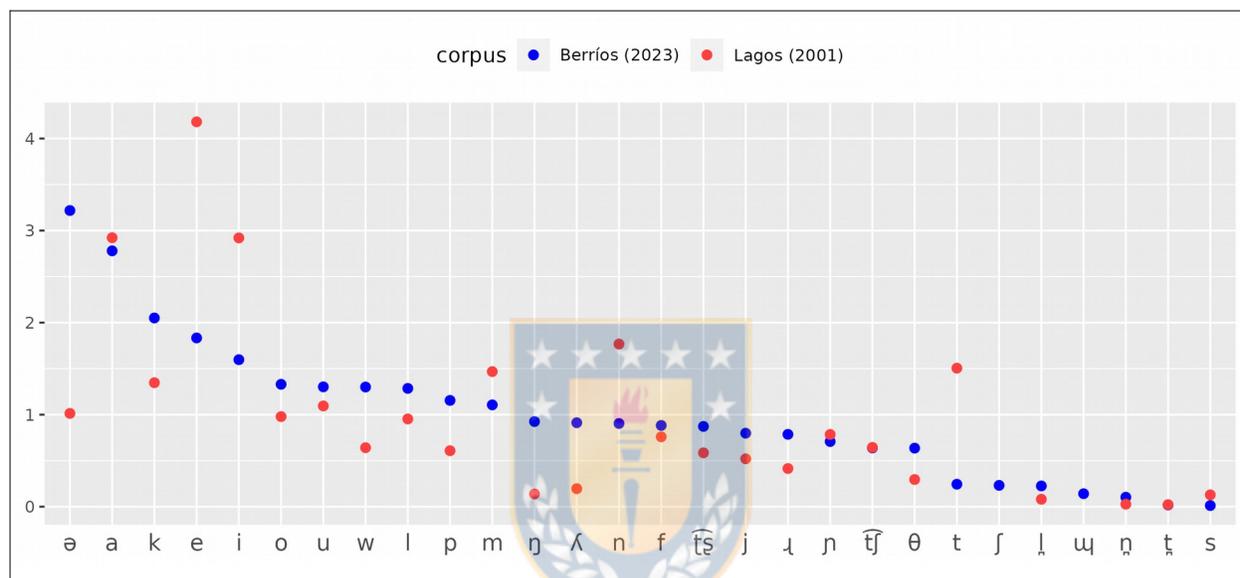
Categorías	Fonemas (Márquez, 2016)	Fonemas (Berríos, 2023)
I) Muy alta frecuencia	/a/	/ə/ > /a/
II) Alta frecuencia	/ə/ > /k/ > /e/ > /i/	/k/ > /e/ > /i/
III) Frecuencia moderada	/u/ > /w/ > /o/ > /m/ > /p/	/o/ > /u/ > /w/ > /l/ > /p/ > /m/
IV) Baja frecuencia	/l/ > /n/ > /ʎ/ > /ɲ/ > /tʂ/ > /ɲ/ > /f/ > /ɲ/ > /j/ > /tʃ/ > /θ/	/ɲ/ > /ʎ/ > /n/ > /f/ > /tʂ/ > /j/ > /ɲ/ > /ɲ/ > /tʃ/ > /θ/
V) Muy baja frecuencia	/t/ > /ɲ/ > /l/ > /s/ > /w/ > /t/ > /ʃ/	/t/ > /ʃ/ > /l/ > /w/ > /ɲ/ > /t/ > /s/

En términos comparativos, si bien existen diferencias en el orden o jerarquía fonemática para varios segmentos, al momento de considerar ambos resultados en términos categóricos, las asimetrías más importantes ocurren con el desplazamiento de dos segmentos entre clases de fonemas adyacentes. En primer lugar, el segmento /ə/ en Márquez (2016) ocurre en la segunda categoría, en lugar de ser el segmento de más alta frecuencia como se ha reportado en nuestro estudio. En el caso del segmento /l/, que en nuestro estudio ocurre en la III categoría (de frecuencia moderada), en Márquez (2016) aparece en la IV (de baja frecuencia). Esto refuerza la noción de que ambos estudios se diferencian solo en segmentos muy específicos, como la frecuencia de /ə/, asunto que retomaremos más adelante en esta sección de discusión.

8.2.2. Comparación con Lagos (2001)

En el caso del estudio de Lagos (2001), se presentan asimetrías mayores al comparar sus resultados con los nuestros, tanto en el número de fonemas que son distintos, como en la magnitud de esa diferencia. De manera visual, es posible apreciar este hecho en el **Gráfico 8.4**.

Gráfico 8.4: Comparación de las frecuencias, por medio de la ratio Obs./Esp., de nuestro estudio y las reportadas en Lagos (2001).



En primer lugar, en las vocales, se registran importantes diferencias; especialmente, en la llamada sexta vocal y en la vocal /e/. En nuestro estudio, en efecto, /ə/ tiene la más alta jerarquía (Ratio Obs./Esp. = 3,2); en cambio, en Lagos, apenas sobrepasa la frecuencia esperada (1,01), muy cercana al valor de /o/ (0,98), que en nuestro estudio corresponde a una de las vocales menos frecuente, junto con /u/. Respecto de /e/ e /i/, ocurre la situación inversa: Lagos (2001) reporta frecuencias muy altas para estos fonemas, cuando los comparamos con nuestros resultados. En Lagos (2001), además, /e/ ocurre con una frecuencia observada que es más de cuatro veces la frecuencia esperada, /i/ ocurre con una frecuencia observada que es tres veces más que la esperada, al igual que /a/. Ahora bien, a pesar de estas diferencias, las frecuencias observadas de /a/ en ambos estudios son bastante similares, con una diferencia de -0,14 de la ratio obs./esp., respecto de nuestro estudio.

Entre las consonantes, las mayores asimetrías se producen con el fonema /ŋ/ (con una diferencia de 0,8), seguida de /k/, /k/ y /w/ (con diferencias cercanas a 0,7) y /p/ (diferencia de 0,6). Todos estos fonemas en nuestro estudio tienen frecuencias observadas mayores. Por otro lado, las consonantes /m/, /n/, /t/, en el caso de los valores reportados por Lagos (2001), son considerablemente mayores. Es probable que parte del material gramatical —que es muy frecuente en texto continuo, y que suponemos es el suministro de tal estudio— esté elevando la frecuencia de estas consonantes, puesto que el sufijo **-m** forma verbos transitivos, /t/ es parte de un sufijo muy común en mapudungun (**-tu**), y **-n** es el sufijo nominalizador que se usa como forma de cita para los verbos. En el caso del fonema /e/, es probable que su alta ocurrencia en sufijos muy frecuentes en textos continuos (por ejemplo, **-ke** HABITUAL, **-we** LOCATIVO, **-fe** AGENTE) explique su alta frecuencia en el reporte de Lagos (2001). En ese sentido, creemos que nuestros criterios logran controlar la alteración que pudiesen provocar los sufijos flexivos y derivativos del mapudungun.

En resumen, al comparar nuestro estudio con el de Lagos (2001), encontramos 16 fonemas que presentan diferencias superiores a 0,25, lo que difiere considerablemente de los 5 fonemas que tenían esta condición cuando comparamos nuestro estudio con los resultados en Márquez (2016).

Tabla 8.4: Comparación de las frecuencias de fonemas en el corpus del presente estudio y en el reportado en Lagos (2001). Los valores corresponden a la ratio Obs./Esp.

Fon.	/ə/	/a/	/k/	/e/	/i/	/o/	/u/	/w/	/l/	/p/	/m/	/ŋ/	/k/	/n/
Berr.	3,22	2,78	2,05	1,83	1,6	1,33	1,3	1,3	1,29	1,15	1,11	0,93	0,91	0,9
Lag.	1,01	2,92	1,35	4,18	2,92	0,98	1,09	0,64	0,95	0,61	1,47	0,14	0,2	1,77
Diff.	2,21	-0,14	0,7	-2,35	-1,32	0,35	0,21	0,66	0,34	0,54	-0,36	0,79	0,71	-0,87
Fon.	/f/	/tʃ/	/j/	/ɬ/	/ɲ/	/tʃ/	/θ/	/t/	/ʃ/	/l̥/	/w/	/n̥/	/t̥/	/s/
Berr.	0,88	0,87	0,8	0,79	0,71	0,64	0,64	0,24	0,23	0,23	0,14	0,1	0,02	0,01
Lag.	0,76	0,59	0,52	0,41	0,78	0,64	0,3	1,51	N/A	0,08	N/A	0,03	0,02	0,13
Diff.	0,12	0,28	0,28	0,38	-0,07	0,00	0,34	-1,27	N/A	0,15	N/A	0,07	0,00	-0,12

8.2.3. Comparación con Álvarez-Santullano (1986)

Dadas las diferencias entre los inventarios fonológicos reportados por Álvarez-Santullano (1986) y el nuestro, especialmente debido a que los estudios se basan en variedades dialectales diferentes, la comparación de los resultados entre estas dos investigaciones fonostadísticas se torna un poco más difícil.

En primer lugar, reportaremos por medio de la *ratio de frecuencia observada y esperada* el aporte de cada fonema al propio inventario fonológico asumido. En nuestro caso, 28 fonemas; en el caso de Álvarez-Santullano (1986), 25 fonemas (6 vocálicos, 19 consonánticos). De igual modo, cuando fue posible, se normalizaron las frecuencias de algunos fonemas, atribuyéndolas al fonema correspondiente en la variable dialectal central (por ejemplo, el fonema fricativo retroflejo sordo /ʃ/ de la variante *williche* se despliega junto al aproximante retroflejo sonoro /ɟ/, de la variante central). No obstante, el fonema lateral retroflejo sonoro /ɮ/, unidad funcional de acuerdo con el estudio de Álvarez-Santullano (1986), no fue posible atribuirlo a un segmento de la variedad central de la lengua mapuche. Por eso, preferimos reportar su valor intralingüístico de manera separada, por medio del indicador de la ratio Obs./Esp.

Tabla 8.5: Comparación de las frecuencias de fonemas en el corpus de este estudio y en el reportado en Álvarez-Santullano (1986). (El fonema /ɮ/, que no aparece en la tabla, tiene una ratio Obs./Esp. de **0,52**.)

Fon.	/ə/	/a/	/k/	/e/	/i/	/o/	/u/	/w/	/l/	/p/	/m/	/ŋ/	/ʎ/	/n/
Berr.	3,22	2,78	2,05	1,83	1,6	1,33	1,3	1,3	1,29	1,15	1,11	0,93	0,91	0,9
Ál.S.	2,07	2,24	1,9	1,98	1,77	0,99	1,94	1,16	0,82	1,12	1,51	0,39	0,22	1,94
Diff.	1,15	0,54	0,15	-0,15	-0,17	0,34	-0,64	0,14	0,47	0,03	-0,4	0,54	0,69	-1,04

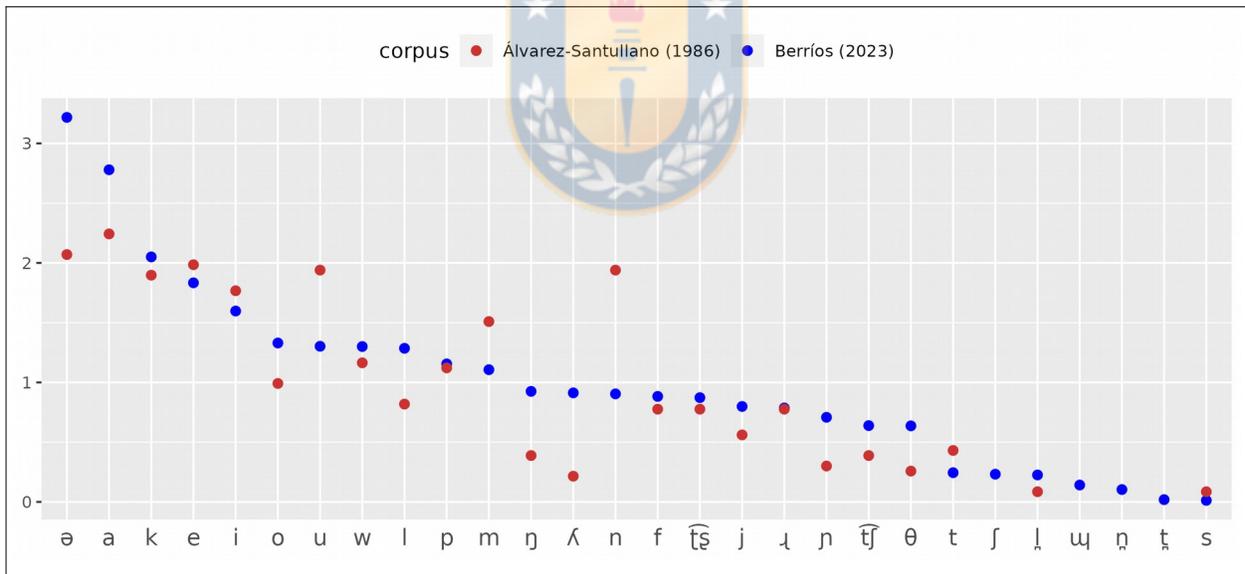
Fon.	/f/	/tʃ/	/j/	/ɟ/	/p/	/tʃ/	/θ/	/t/	/ʃ/	/ɬ/	/w/	/ɲ/	/t/	/s/
Berr.	0,88	0,87	0,8	0,79	0,71	0,64	0,64	0,24	0,23	0,23	0,14	0,1	0,02	0,01
Ál.S.	0,78	0,78	0,56	0,78	0,3	0,39	0,26	0,43	N/A	0,09	N/A	N/A	N/A	0,09
Diff.	0,1	0,09	0,24	0,01	0,41	0,25	0,38	-0,19	N/A	0,14	N/A	N/A	N/A	-0,08

Las mayores diferencias, nuevamente, se producen al comparar el segmento vocálico /ə/, que en el caso del estudio de Álvarez-Santullano (1986) registra una ratio dos veces más frecuente que

la frecuencia esperada. Otras vocales que en nuestro estudio tienen una frecuencia moderadamente mayor son /a/ y /o/. Igualmente, dentro de las vocales, /u/ es más frecuente en el inventario del *williche*, al igual que /e/ e /i/, aunque con una menor diferencia respecto de nuestro estudio.

Asimismo, otras asimetrías importantes entre ambos estudios se producen respecto de las frecuencias de algunas consonantes resonantes, como es el caso de /ʎ/, /ŋ/, /ɲ/ y /l/, todas con frecuencias mayores en nuestro estudio. En el caso de /n/ y /m/, éstas ocurren con mayor frecuencia en el estudio de Álvarez-Santullano (1986). El resto de las consonantes no tienen diferencias importantes entre ambos estudios, salvo /θ/, /j/ y /tʃ/, que ocurren en nuestro estudio con una mayor frecuencia observada, si se compara con la esperada. El **Gráfico 8.5** permite visualizar y comparar estas tendencias entre ambos reportes.

Gráfico 8.5: Comparación de las frecuencias, por medio de la ratio Obs./Esp., de nuestro estudio y las reportadas en Álvarez-Santullano (1986).



En resumen, las frecuencias de fonemas son similares entre las cuatro investigaciones reportadas; es decir, se evidencia una alta frecuencia de las vocales y algunas consonantes resonantes, además de una alta frecuencia de una consonante obstruyente. En el caso de nuestro estudio, en el de Márquez (2016) y Álvarez-Santullano (1986), esta consonante es /k/; en Lagos, en cambio,

es la consonante /t/. De igual modo, es notoria la similitud de nuestro estudio con el de Márquez (2016) respecto de varios puntos, lo que se podría vincular con el hecho de que los insumos, la metodología de muestreo y los criterios de filtrado son, en términos generales, similares. Un efecto positivo de ello es que nos permite hacer comparaciones y establecer, por ejemplo, que a pesar de existir casi 80 años entre la recolección de ambos insumos lexicográficos, existe una persistencia general en cuanto a las distribuciones fonostadísticas (Martin, 2007)¹². Las diferencias específicas encontradas, y que se profundizarán a continuación, pueden atribuirse a otras motivaciones; específicamente, motivaciones de índole extralingüística.

8.3. Frecuencia de /ə/

La vocal central media no redondeada /ə/ ha sido foco de atención en estudios de diversa índole; por ejemplo, fonético-acústicos (Alvar López et al., 1997; Soto-Barba et al., 2016; Mena et al., 2019) y articulatorios (Álvarez et al., 2020). Esto evidencia su estatus de fonema de relevancia en los estudios de las ciencias del habla. En nuestro caso, se confirma este estatus de segmento de atención, debido a su primera posición en la jerarquía de las frecuencias fonemáticas del mapudungun.

Si bien la literatura sobre frecuencia fonemática del mapudungun ha establecido la alta ocurrencia de este fonema en diversos corpus, su jerarquía específica ocupaba, en dichos estudios, el segundo o tercer lugar, tras otras vocales más frecuentes como /a/ o /e/; en nuestro reporte, en cambio, esta ocupa la posición más alta, superando a la vocal /a/ en dos puntos porcentuales. Esta posición dentro de la jerarquía de frecuencia fonemática amerita, a nuestro juicio, la formulación de alguna hipótesis para explicar este resultado, que no solo no es coincidente con la literatura que se ocupa del tema fonostadístico para el mapudungun, sino que con las tendencias tipológicas en general, las cuales suelen reportar vocales periféricas, como /a/, como las más frecuentes en términos intralingüísticos (Gordon, 2016).

12 “Words with phonotactic patterns that facilitate lexical access will tend to be used more than words without those patterns, leading over time to a lexicon in which these “good” patterns predominate” (Martin, 2007, p. xiv)

En este contexto, en primer lugar, descartamos que nuestra decisión metodológica de operar con un criterio que escoge realizaciones indicadoras de vitalidad (asunto desarrollado en la sección 6.4.2. *Alternancias fonemáticas*), frente a pares de lexemas con significado idéntico o similar y que se diferencian solo por la alternancia de /ə/ con cualquier otra vocal, tuviera alguna incidencia en ubicarla en la primera posición. En efecto, como se detalló en la sección 6.4.2.2. *Exploración del efecto del criterio RIV y alternativas*, incluso con la exclusión de esa vocal en esos pares, la posición de /ə/ continuaba siendo la más alta (el efecto concomitante de esa omisión era aproximar su valor relativo al de /a/, pero sin llegar a ser menor que esta última).

Por otro lado, desde el punto de vista morfológico, es posible que el minucioso control del material gramatical en las entradas del diccionario pueda explicar, en parte, la alta presencia del fonema /ə/, pues sufijos muy comunes ocurren con otras vocales. Es el caso, por ejemplo, de **-tu**, **-le**, **-nge**, **-ma**, **-ka**, etc., los cuales, hasta donde sabemos, no han sido controlados en la literatura que estudia la frecuencia fonemática del mapudungun —salvo por los sufijos **-we** y **-fe**, que sí fueron controlados en Márquez (2016). No obstante, esta decisión de controlar el material gramatical también afecta la frecuencia de /ə/, puesto que, en algunos casos, esta vocal es realizada como epéntesis para solucionar grupos consonánticos anómalos en verbos derivados, por lo que su filtrado también debería tributar en una baja de su frecuencia no controlada. De igual modo, dentro del criterio gramatical, si se hubiesen incluido los seudosufijos **-dü** y **-kü**, con una muy alta frecuencia, también habría aumentado de manera importante el valor de ocurrencia de /ə/. Entonces, por lo señalado, no creemos que los criterios desplegados para analizar el material gramatical hayan generado un sesgo metodológico que explique la alta frecuencia de esta vocal.

Es posible pensar que la muestra utilizada como insumo para nuestro estudio, el *Diccionario mapudungún-español* de Augusta (1916/2017), presente un sesgo diatópico que en términos fonológicos explique la mayor frecuencia de lexemas con sexta vocal, debido a preferencias dialectales o idiolectales (Catrileo, 1986, p. 5). No obstante, la complejidad de esta argumentación está en explicar cómo es posible que coincidan en este supuesto sesgo fonológico

por /ə/ dos variantes geográficamente distantes, como son las que se incluyen en el diccionario utilizado. Debemos recordar que los hablantes que colaboraron en la creación del diccionario son, por un lado, de la zona de Huapi, mapudungun hablado en la zona costera de la Araucanía, y que es la variante hablada por Pascual Segundo Painemilla Ñamcucheu y José Francisco Coluñ, y, por otro lado, de la zona de Panguipulli, mapudungun hablado en el sur hacia la cordillera, representado por Domingo Segundo Huenuñamco (Augusta, 1916/2017, p. XVI). Al ser una muestra representativa de más de una zona dialectal —además de ser zonas geográficamente distantes—, se reduce la posibilidad de estar ante una explicación de tipo diatópica.

Con todo, estimamos que es posible apuntar a fenómenos diacrónicos y sociolingüísticos como posibles explicaciones para la alta frecuencia de /ə/ en nuestro estudio respecto de otros anteriores; en específico, por tener, posiblemente, el mapudungun una mayor vitalidad a principios del siglo XX, fecha en que se recogen los vocablos que forman parte del *Diccionario mapudungún-español* de Augusta (1916/2017). Una idea que podría profundizarse es la posibilidad de que las unidades léxicas con /ə/ no sean transmitidas con el devenir generacional, y que hayan sido reemplazadas por neologismos producidos por medio de préstamos lingüísticos, por un lado, o por medio de derivaciones gramaticales que no contengan esta sexta vocal, por otro.

Esta noción se ha denominado *empobrecimiento lexical* (Chiodi y Loncon, 1999) y consiste en la contracción del caudal léxico del idioma. De acuerdo con Chiodi y Loncon (1999), la pérdida de vigencia de las palabras mapuche se debe al concomitante abandono de “ideas, hábitos y costumbres y prácticas culturales mapuche” y también a la sustitución de las mismas por medio de la adopción “innecesaria” de vocablos del castellano (1999, p. 11, las comillas son nuestras). Según los autores, este fenómeno conduce, por un lado, a la pérdida de las palabras y, por otro, impide que el vocabulario sea recreado a la par con la realidad de nuevas ideas y materialidad (1999, p. 11).

En este sentido, podemos especular que vocabulario de urdimbre mapuche con una alta frecuencia de /ə/ en su estructura fonológica fue cayendo en desuso y perdiendo su transmisión conforme se adquiría la lengua entre las generaciones. Las nuevas palabras que se incorporaron

al idioma, entonces, tienen la particularidad de ser préstamos del castellano o de ser sustantivos derivados por medio de recursos gramaticales intralingüísticos (Villena, 2014) —que en la lengua mapuche es “uno de los mecanismos de mayor efectividad” (Chiodi y Loncon, 1999)—, pero es poco frecuente que surjan neologismos espontáneos por medio de la invención de palabras combinando los sonidos de la lengua en forma inédita (Chiodi y Loncon, 1999¹³).

La reutilización de recursos morfológicos como la derivación y la composición agudiza la reducción del léxico con /ə/, porque permite un mecanismo de expresión alternativo que va explotando la disponibilidad de los fonemas en un estadio diacrónico determinado, en desmedro de las frecuencias fonemáticas del estadio anterior; de esa manera se reducen las frecuencias en este caso de la vocal más frecuente según nuestro reporte, llegando a un nuevo estadio sincrónico donde /ə/ adquiere una frecuencia que es menor a /a/, como se reporta en Márquez (2016).

Tabla 8.6: Propuesta de representación del reemplazo del léxico por medio de préstamos y derivados gramaticales, y cómo podría afectar las distribuciones de las vocales en las unidades lexemáticas.

Tiempo ₁	Tiempo ₂	Tiempo ₃
lexema 1 con /ə/	lexema 1 con /ə/	lexema 1 con /ə/
lexema 2 con /ə/	préstamo con /a/	préstamo con /a/
lexema 3 con /ə/	lexema 3 con /ə/	derivado de lexema 5 con /a/
lexema 4 con /a/	lexema 4 con /a/	derivado de lexema 1 con /ə/
lexema 5 con /a/	lexema 5 con /a/	lexema 5 con /a/
etc.	etc.	etc.

En la **Tabla 8.6**, visualizamos una propuesta de esquema temporal de este proceso. Así, dado un *tiempo*₁, esto es, la época donde se recogen los lexemas con la frecuencia reportada en este estudio, se da un proceso de cambio lingüístico conforme nuevas generaciones aprenden y transmiten la lengua, por lo que de manera sucesiva el léxico va siendo reemplazado por otros recursos de creación léxica, como préstamos y estructuras derivadas, arribando al *tiempo*₃, más

13 Si bien Chiodi y Loncon (1999) mencionan la validez del recurso de invención para ampliación del léxico, señalan que es poco usado y poco práctico, y por lo mismo lo desaconsejan. Curiosamente, uno de los ejemplos que proveen para ilustrar este recurso neológico —*ükü* ‘pediatra’ (p. 157)— contiene dos instancias de /ə/, además de una de /k/, dos de los fonemas con mayores frecuencias en nuestro reporte.

contemporáneo, donde existe una menor diferencia entre las frecuencias relativas de dos de las vocales más frecuentes.

Un efecto similar, aunque no tan marcado, ocurre con otros fonemas que evidencian vitalidad, como /ɲ/, /ʎ/, /t͡s/, /θ/, /ʃ/, /ɻ/ y /ɰ/, aunque, de manera interesante, no para las interdentes /ɳ/ y /t͡/, que registran ratios Obs./Esp. inferiores en nuestro estudio, al compararlo con el de Márquez (2016). Es posible que la baja carga funcional de estos dos fonemas, que en nuestro estudio se clasifican dentro de los fonemas de más baja frecuencia (menor al 0,37%), contribuyan a la dificultad de comparar, en este caso, los resultados de ambos estudios.

En el mismo sentido, es posible que en algunos casos las diferencias se deban a fluctuación fonemática que no quedó registrada en el caso del *Diccionario mapudungún-español* de Augusta (1916/2017). A modo de ejemplo, la palabra /antə/ —que en Augusta se escribe <antü>, remitiendo a dos consonantes alveolares—, en Hernández y Ramos (1997) se simboliza <antü>, remitiendo a consonantes interdentes, posiblemente debido a un fenómeno de *fluctuación fonemática*. Es posible que algunas diferencias de este tipo, además del bajo número de casos, puedan explicar un aparente incremento en las realizaciones de /t͡/ y /ɳ/, al comparar ambos estudios. De todas maneras, mayor información es necesaria para poder esclarecer esta eventual relación entre el *empobrecimiento léxico* y la reducción de las frecuencias fonemáticas de fonemas que evidencian vitalidad.

Una posibilidad de investigación de esta noción de disponibilidad fonemática para la invención de nuevo léxico podría consistir en un estudio de creación de *logotomas* o *pseudopalabras* en mapudungun, que permita establecer el repertorio de fonemas disponibles, como se ha hecho con variedades del español de Chile en Concepción (Mora y Sanhueza, 2017; Jiménez et al., 2019). La hipótesis, en ese caso, es que la disponibilidad fonemática tiende a una permanencia mayor que la de los lexemas observados, producto de este proceso de reducción léxica que sufre una lengua como el mapudungun, la cual se encuentra en peligro ante el mayor uso de una lengua mayoritaria como el castellano.

8.4. Criterios de filtrado

Nuestro estudio contó con un exhaustivo proceso de filtrado, donde se analizaron cada una de las 9703 entradas de la versión modernizada del *Diccionario mapudungún-español* de Augusta (1916/2017), con el objetivo de acceder a las unidades lexemáticas monomorfémicas. En virtud de este objetivo, se siguieron criterios de índole lingüística que atendían a las características morfológicas de la lengua mapuche; en especial, la relación prácticamente transparente entre forma y significado, dada por su carácter de lengua aglutinante, lo que nos permitió identificar morfemas de manera discreta, tanto para aislar las unidades lexemáticas, como para controlar la frecuencia de morfemas gramaticales derivativos y flexivos, identificados en las entradas del diccionario.

En otros casos, debido a que en mapudungun la composición nominal y verbal, además de la incorporación nominal al complejo morfémico verbal, son procesos morfológicos productivos para la creación de palabras (Zuñiga, 2014), fue necesario identificar los bloques elementales que conforman estos compuestos, anotando la forma tipo y controlando cualquier instancia de los más de 160 compuestos identificados.

Un fenómeno que evidenciamos tempranamente durante este proceso de control de las unidades fue un tipo de variación muy frecuente en mapudungun, denominado en la literatura *fluctuación fonemática* (Key, 1968; Clairis, 1981; Salas, 2006), el cual se manifiesta con dos formas léxicas, de igual o muy similar significado, que solo se diferencian en un fonema. Respecto de estas variaciones, identificamos y enumeramos, en virtud de la literatura, dos tipos: la *fluctuación fonemática* (propriadamente dicha) y la *alternancia fonemática*. En el primer caso, la intercambiabilidad de los fonemas se considera un fenómeno propio de la fonología del mapudungun, aunque su alcance y posible motivación todavía no han sido elucidados (Smeets, 2008, p. 31). El segundo caso, la *alternancia fonemática*, en cambio, se considera un fenómeno atribuible a la situación de contacto del mapudungun con una lengua de comunicación general: el castellano (Henríquez y Salamanca, 2015).

Decidimos, entonces, controlar estos dos fenómenos de variabilidad léxica. Así, respecto de la *alternancia*, se escogió el lexema que tenía el segmento denominado en la literatura como Realización Indicadora de Vitalidad (RIV), cuando los pares de lemas en alternancia así lo permitían (Henríquez y Salamanca, 2015). En el caso de las *fluctuaciones fonemáticas* entre segmentos que no corresponden a una realización indicadora de vitalidad o transferencia, operamos bajo un criterio de relevancia léxica, que se manifestaba en la disposición lexicográfica escogida por el autor de la obra. Este relevamiento lexical opera del siguiente modo: ante un par de entradas con términos de similar significado y estructura fonológica, escogimos aquel lema que ostenta el despliegue de una definición y ejemplos, y consideramos de orden secundario aquella entrada que remitía de manera abreviada al primer término. Esta decisión se sustenta en que este relevamiento léxico no era predecible a partir de criterios como el ordenamiento alfabético o la variedad dialectal, por lo que entendemos que su elección se debía a una mayor importancia o relevancia de ese término por parte del autor. Concretamente, se identificaron alrededor de 300 casos de *alternancias fonemáticas* y un número cercano a 160 casos de *fluctuaciones fonemáticas*, en los que se operó con estos criterios.

De igual modo, se controló el efecto que podía tener la inclusión de léxico proveniente de sistemas fonológicos foráneos. Fue así como se excluyeron quechuismos, hispanismos e interjecciones con fonotaxis atípica. De igual modo, fue necesario filtrar parte de la estructura fonológica de lexemas que claramente manifestaban *reduplicación*, un fenómeno que ocurre en el mapudungun, y que fue identificado por medio de la concomitancia de sufijos que funcionan con este recurso morfológico, o bien por sus significados, los cuales evocan los campos semánticos típicos de este recurso lingüístico, tales como intensidad, pluralidad, etc. (Sandvig, 1986).

Estos diversos criterios de índole lingüística y extralingüística permiten sistematizar un filtrado de unidades lexemáticas monomorfémicas que atiende a las características propias de la lengua mapuche. Ahora bien, independiente de la ocurrencia de algunos criterios ad hoc para el mapudungun, creemos que es posible replicar este tipo de estudio en otras lenguas y de ese modo proporcionar y comparar la frecuencia fonemática en el léxico, controlando alteraciones en la

distribución de frecuencias debido a procesos morfosintácticos. De este modo, estimamos, es posible elucidar el caudal fonemático efectivamente disponible en un listado de lexemas en una lengua en particular.

Así, entonces, es posible que algunos de estos criterios, ajustados al carácter aglutinativo de la morfología mapuche, no puedan ser replicados en lenguas con una morfología fusional, donde se torna más complejo identificar los componentes semántico-funcionales de los léxicos en una determinada palabra. Por este motivo, es posible que el objetivo de poder controlar la alta frecuencia de fonemas que aparezcan en la estructura fonológica del material gramatical en lenguas como el castellano o el romané, por mencionar algunas, requiera caminos analíticos algo diferentes, como la analogía con otras estructuras, y siempre valiéndose de las últimas nociones e investigaciones en el campo de la morfosintaxis de la lengua en particular.

Por último, creemos que las ventajas de este *modus operandi* no se agotan en el estudio de la frecuencia fonemática, sino que también pueden permitir el análisis de otros indicadores fonestadísticos, como la *frecuencia léxica* en corpus de textos continuos, ya sea orales o escritos, puesto que permite contar con una herramienta de conteo de las unidades lexemáticas, que es independiente de su realización como formas derivadas o flexivas, y de ese modo acceder a las unidades lexicales como unidades *tipo*. Un conteo de la frecuencia lexemática permitiría estudiar la relación de la frecuencia de estructuras fonológicas mayores con otras áreas del quehacer de las ciencias del lenguaje como la psicolingüística; específicamente, respecto de mecanismos de procesamiento léxico, trabajos que investigan las interfaces fonológico-gramaticales y que hasta el momento no cuentan con este tipo de insumo (Véliz y Olate, 2008).

En consecuencia, en virtud de lo señalado con respecto a los criterios de filtrado establecidos en nuestra metodología y sus diversas aplicaciones, creemos que este aspecto de nuestra investigación se torna en una contribución importante no solo al área de la fonestadística, sino que a otros dominios de la ciencia del lenguaje.

9. Conclusiones y proyecciones

Consideramos que nuestro estudio es un aporte a la descripción fonológica del mapudungun, al reportar las frecuencias fonemáticas en las unidades monomorfémicas lexemáticas, utilizando como insumo la obra lexicográfica más importante del acervo de esta lengua: el *Diccionario mapudungún-español* de Augusta (1916/2017). Asimismo, los resultados indicados en esta investigación no solo contribuyen a profundizar la descripción fonostadística de esta lengua en particular, sino que también potencian el estudio comparativo de las tendencias tipológicas de los segmentos fonemáticos, allanando el camino a estudios que incorporen patrones fonológicos mayores, y a su aplicación en otras áreas de la ciencia lingüística.

Un aspecto prominente de este trabajo es haber encontrado frecuencias que no habían sido reportadas en la literatura fonostadística sobre el mapudungun, como es la altísima frecuencia del segmento /ə/. La posible relación de este hallazgo con fenómenos sociolingüísticos como el denominado *empobrecimiento léxico* (Chiodi y Loncon, 1999) también contribuye con nuevas perspectivas al debate en torno a la vitalidad de la lengua mapuche, y a la urgencia concomitante de los procesos de revitalización.

Asimismo, y a manera de proyección respecto de este tema, es posible que un estudio de la disponibilidad fonemática de hablantes de la lengua para la creación de *logotomas* (también denominadas *pseudopalabras*) nos permitiría comprobar si existe una separación entre la distribución de frecuencias observadas en el léxico y aquellas disponibles para la conformación de nuevos vocabularios, entregando luces respecto de la relación entre el *empobrecimiento léxico* y la contracción fonológica. En términos prácticos, un estudio de esta índole sería de interés para la creación de neologismos, que sigan patrones adecuados a las tendencias fonostadísticas de la lengua mapuche.

Por otro lado, el hecho de que nuestra investigación reporte que los segmentos interdientales evidencian una muy baja frecuencia en el léxico confirma, sin duda, las tendencias tipológicas de las lenguas del mundo, donde el contraste entre dos articulaciones coronales es muy poco común

(Molineaux, 2021). Además, concuerda con lo señalado por distintos autores (por ejemplo, Smeets 2008; Molineaux, 2021, entre otros), en cuanto a que la distinción entre la serie interdental /t̪/, /n̪/, /l̪/ y la alveolar /t/, /n/, /l/ está sufriendo un proceso de pérdida, a pesar de que en términos cualitativos el contraste entre estos fonemas es importante para los hablantes, especialmente desde un punto de vista identitario¹⁴, pues les permite distinguir entre variedades dialectales e incluso advertir grados de conocimiento y dominio de la lengua (Painequeo et al., 2018). En este sentido, creemos que una proyección de un estudio sobre frecuencias léxicas permitirá ponderar nuestros resultados con un nuevo indicador que permita establecer una posible asociación entre el mantenimiento de estos fonemas de muy baja frecuencia y su ocurrencia en lexemas de alta frecuencia.

Entre las limitaciones de nuestra investigación, debemos señalar que nuestro estudio se concentró en las unidades lexemáticas monomorfémicas y no abordó la frecuencia fonemática en otras estructuras. Como se evidenció en las secciones correspondientes de filtrado y criterios metodológicos, el dominio lexemático entraña sus propios desafíos, por lo que la inclusión de un área como la gramatical requiere del diseño e implementación de un nuevo estudio que acometa este trabajo de manera focalizada. En el caso de un estudio que aborde las unidades gramaticales, es necesario proveer un nuevo insumo que permita maximizar la posibilidad de encontrar instancias de este tipo. En este sentido, es fundamental generar un corpus de textos continuos, puesto que en una obra lexicográfica se tienden a agrupar las realizaciones en torno a formas de cita. De todas maneras, el conteo debe realizarse en unidades gramaticales tipo, considerando criterios para fenómenos propios de la gramática del mapudungun, como la morfofonología (Smeets, 2008) y la alomorfía supletiva de algunos sufijos verbales (Fasola, 2015).

De igual modo, dado el alcance descriptivo de nuestra investigación, no operamos con una hipótesis cuantitativa, lo que hace necesario remitirnos y concentrarnos en la visualización de las tendencias encontradas por medio de este tipo de estadística. Nuestra decisión de tener un

14 Henríquez y Dinamarca (2018) mencionan cómo los hablantes tienen una actitud de valoración a quienes dominan y usan la lengua, y asumen como un proceso de pérdida de identidad las influencias que el castellano pueda tener en la estructura del idioma.

alcance descriptivo es congruente con el *status quaestionis* de la fonestadística en mapudungun, donde no se había acometido un estudio de esta envergadura, salvo por el trabajo de Márquez (2016), el cual, de todas formas, no problematiza en algunos de los focos relevados en nuestro estudio. El uso de estadística descriptiva para reportar los datos en fonestadística es una limitación reconocida en la literatura (Macklin-Cordes y Round, 2020), pero consideramos que es un primer paso a partir del cual es posible desarrollar nuevas investigaciones que busquen modelar matemáticamente nuestros resultados, y, de ese modo, comprobar asociaciones u otras relaciones, así como se han esbozado en este trabajo.

Por último, una limitación de nuestro estudio es la falta de comparación de nuestros resultados con distribuciones de frecuencia fonemática en otras lenguas. Decidimos no seguir esa ruta, porque esperamos reportar esta información a partir de la aplicación de la metodología y criterios desplegados en el presente estudio y así asegurar una comparación adecuada de los resultados interlingüísticos (aunque, por cierto, con las adecuaciones que la lengua del caso amerite). En este sentido, esta limitación se vuelve una oportunidad para la comprobación del tipo de metodología asumida en este trabajo y su insumo fuente (unidades léxicas monomorfemáticas).

Respecto de estas proyecciones y posibilidades que se abren a partir de nuestro estudio, creemos que, además de las direcciones ya señaladas, se deben realizar estudios de patrones fonológicos mayores, como grupos vocálicos y consonánticos, además de la sílaba. Este nuevo objetivo es posible de realizar gracias al corpus confeccionado en nuestro estudio y las herramientas tecnológicas propuestas en este documento.

10. Referencias

- Adelaar, W. F. H. (2004). *The Languages of the Andes*. Cambridge University Press.
- Alarcos, E. (1965). *Fonología española*. Editorial Gredos.
- Altmann, G. (2005). Phonic word structure. En R. Köhler, G. Altmann, R. Piotrowski (Eds.), *Quantitative Linguistics: An International Handbook* (pp. 191-208). Walter de Gruyter
- Alvar López, M., Moreno Fernández, F. y Ueda, H. (1997). La sexta vocal del mapuche. *Thesaurus: Boletín del instituto Caro y Cuervo*, 52(1), 183–189.
- Álvarez-Santullano, P. (1986). *Descripción fonológica del huilliche, un dialecto del mapuche o araucano del centro-sur de Chile* (Tesis de magister). Universidad de Concepción.
- Álvarez, G., Ruiz, M., Arias, A., Lezcano, M. F., & Fuentes, R. (2020). Description of Acoustic and Articulatory Parameters of Vowels in Mapudungun Speakers. Pilot Study. *International journal of odontostomatology*, 14(2), 205–212. <https://doi.org/10.4067/s0718-381x2020000200205>
- Andrews, E. (1990). *Markedness Theory: The Union of Asymmetry and Semiosis in Language*. Duke University Press.
- Antinao, C. (2014). *Diccionario. Ta iñ Mapun Dungun. Nuestra lengua mapuche*. Ediciones Comunidad de Historia Mapuche.
- Augusta, F. de. (1916). *Diccionario araucano-español y español-araucano*. Imprenta Universitaria.
- Augusta, F. de. (2017). *Diccionario mapudungún-español y español-mapudungún*. Villena, B. (ed.). Ediciones Universidad Católica de Temuco. (Obra original publicada en 1916)
- Basadi Palai, E. y O'Hanlon, L. (2004). Word and phoneme frequency of occurrence in conversational Setswana: A clinical linguistic application. *Southern African Linguistics and Applied Language Studies*, 22(3–4), 125–142.
- Burquest, D. (1993). *Phonological Analysis. A functional approach*. SIL International.
- Busa, R., Croatto-Martinolli, C., Croatto, L., Tagliavini, C. & Zampolli, A. (1962). Una Ricerca Statistica Sulla Composizione Fonologica Della Lingua Italiana Parlata, Eseguita Con Un Sistema IBM A Schede Perforate. En L. Croatto & C. Croatto- Martinolli

- (Eds.), XIIth International Speech and Voice Therapy Conference of the International Association of Logopedics and Phoniatics (pp. 542–562).
- Bybee, J. (2004). *Phonology and Language Use*. Cambridge University Press.
- Catrileo, M. (1972). *A Tagmemic Sketch of Mapuche Grammar*. University of Texas.
- Catrileo, M. (1984). Consideraciones lingüísticas en torno a un grafemario uniforme para el Mapudungu. *Actas de Lengua y Literatura Mapuche*, 1, pp. 29–40.
- Catrileo, M. (1986). La variación estilística en el nivel fonológico del mapudungu. *Actas de Lengua y Literatura Mapuche*, 1–19.
- Catrileo, M. (1996). *Diccionario lingüístico-etnográfico de la lengua mapuche*. Editorial Andrés Bello.
- Catrileo, M. (2010). *La lengua mapuche en el siglo XXI*. Universidad Austral de Chile.
- Chandía, A. et al. (2014). *Corpus lexicográfico del mapudüngun*. <http://corlexim.cl>
<https://web.archive.org/web/20220107113813/http://corlexim.cl/referencias>
- Chiodi, F. y Loncon, E. (1999). *Crear nuevas palabras: Innovación y expansión de los recursos lexicales de la lengua mapuche*. UFRO-CONADI.
- Chomsky, N. y Halle, M. (1968). *The sound patterns of English*. The MIT press.
- Clairis, C. (1981). La fluctuation de phonèmes. *Dilbilim Linguistique*, 6. 99-110
- Clairis, C. (2001) La lengua Qawasqar. En Ch. Clairis, A. M. Guerra, D. Lagos (ed.), A. Riffo y C. Villalón, *Fonologías de las lenguas indígenas de Chile* (pp. 55-110). Centro de Estudios Dialectológicos.
- Clairis, Ch., Guerra, A.M., Lagos, D. (ed.), Riffo, A. y Villalón, C. (2001). *Fonologías de las lenguas indígenas de Chile*. Centro de Estudios Dialectológicos.
- Cohen, E.-G. (2015). Phoneme complexity and frequency in the acquisition of Hebrew rhotics. *Journal of Child Language Acquisition and Development*, 3(1), 1–11.
- Coseriu, E. (1986). *Introducción a la lingüística*. Gredos.
- Creative Commons. (2022). *Attribution-NonCommercial-ShareAlike 4.0 International* (CC BY-NC-SA 4.0).
- Croese, R. (1980). Estudio dialectológico del mapuche. *Estudios Filológicos*, 15, 7–38.

- Croese, R., Salas, A., y Sepúlveda, G. (1978). Proposición de un sistema unificado de transcripción fonémica para el mapudungu. *RLA. Revista de Lingüística Teórica y Aplicada*, 16, 151–160.
- Crothers, J. (1978). Typology and universals of vowel systems. In *Volume 2 of Universals of Human Language* (pp. 95–152).
- Delattre, P. (1965). Comparing the Phonetic Features of English, French, German and Spanish. Julios Groos Verlag Heidelberg.
- Dewey, G. (1923). *Relative Frequency of English speech sounds*. Harvard University Press.
- Dwight, W. (1874). VIII. The elements of English Pronunciation. En *Oriental and Linguistics Studies* (pp. 202–276). Scribner, Armstrong and Company.
- Echeverría, M. (1964). Descripción fonológica del mapuche actual. *Boletín de Filología*, 16(16), 13–59.
- Echeverría, M. y Contreras, H. (1965). Araucanian Phonemics. *International Journal of American Linguistics*, 31(2), 132–135.
- Everett, C. (2018). The similar rates of occurrence of consonants across the world's languages: A quantitative analysis of phonetically transcribed word lists. *Language Sciences*, 69, 125–135. <https://doi.org/10.1016/j.langsci.2018.07.003>
- Fasola, C. (2015). *Topics in the Syntax of Mapudungun Subordinate Clauses*. The State University of New Jersey.
- Febrés, A. (1764). *Arte de la lengua general del reyno de Chile*. Calle de la Encarnación.
- Fernández Garay, A. (1988). Aspectos dialectales del Ranquel. *Actas de lengua y literatura mapuche*, 3, 73–90.
- Fernández Garay, A. (1998). *El tehuelche: una lengua en vías de extinción*. Estudios Filológicos.
- Frisch, S. A. (2015). A preliminary investigation of quantitative patterns in sonority sequencing. *Italian Journal of Linguistics*, 27(1), 9–27.
- Frisch, S.A. (2011). Frequency Effects. En M. van Oostendorp, C. Ewen, E. Hume, y K. Rice (eds), *The Blackwell Companion to Phonology, Volume IV* (pp. 2137-2163). John Wiley & Sons.
- Fuentes, D. y Salamanca, G. (2016). ¿Los fonemas / v / y / ð / presentan alófonos aproximantes en el Chedungun hablado en el Alto Bío-bío?: Evidencias desde una perspectiva

acústica. *Literatura y lingüística*, 33, 267–288. <https://doi.org/10.4067/s0716-58112016000100013>

González, G. (2016). *Determinación de las frecuencias fonémicas del idioma cabécar mediante un corpus electrónico*. Trabajo de fin de Máster. Universidad Nacional de Educación a Distancia.

González, M. y Mejía, J. (2011). Frecuencia fonemática del español de Colombia. *Forma y Función*, 24(2), 69–102.

Gordon, M. (2016). *Phonological Typology*. Oxford University Press.

Greenberg, J. (2005). *Language Universals*. De Gruyter Mouton. (Obra original publicada en 1966)

Gregores, E. y Suárez, J. (1967). *A description of colloquial Guarani*. Mouton & Co.

Guerra, R. (1983). Estudio estadístico de la sílaba en español. En M. Esgueva y M. Cantarero (Eds.), *Estudios de Fonética* (pp. 11–112). Consejo Superior de Investigaciones Científicas.

Guirao, M., y García Jurado, M. A. (1990). *Frequency of Occurrence of Phonemes in American Spanish*. *Revue Québécoise de Linguistique*, 19(2), 135–149. <https://doi.org/10.7202/602680ar>

Haspelmath, M. y Sims, A. (2002). *Understanding Morphology*. Hodder Education. <https://doi.org/10.4324/9780203776506>

Henríquez, M. (2016). Vitalidad de la Fonología Segmental del Mapudungun de Escolares Mapuches de Tirúa. *Nueva Revista del Pacífico*, 64, 27–66. <https://doi.org/10.4067/s0719-51762016000100002>

Henríquez, M. (2017). Aproximaciones a la fonología del Chedungun: un estudio exploratorio en niños Pewenches del Alto Bío-Bío. *Nueva Revista del Pacífico*, 66, 99–114. <https://doi.org/10.4067/s0719-51762017000300099>

Henríquez, M. y Dinamarca, J. (2018). Actitudes lingüísticas hacia el mapudungun y el castellano: Estudio exploratorio en dos comunidades pewenche del Alto Biobío. *Nueva revista del Pacífico*, 69, 51–66. <https://doi.org/10.4067/s0719-51762018000200051>

- Henríquez, M. y Salamanca, G. (2012). Rasgos prominentes de la fonología segmental del chedungun hablado por escolares del Alto Bío-Bío. *Alpha*, 34, 153–171.
- Henríquez, M., y Salamanca, G. (2015). Vitalidad de la fonología segmental del chedungun hablado por escolares del Alto Biobío. *ALPHA: Revista de Artes, Letras y Filosofía*, 41, 207–231.
- Hernández-Sampieri, R., Fernández, C. y Baptista, M. del P. (2010). *Metodología de la investigación*. McGraw-Hill.
- Hernández, A. (1984). ¡Papeltañ Mapudungu Meo! Un Primer Ensayo de Ortografía Práctica en Mapudungu. *Cultura - Hombre - Sociedad CUHSO*, 1(1). <https://doi.org/10.7770/cuhso-v1n1-art136>
- Hernández, A. (1986). Encuentro para la unificación del alfabeto mapuche. Propositiones y acuerdos. *Cultura - Hombre - Sociedad CUHSO*, 3(1).
- Hernández, A. y Ramos, N. (1997). *Mapuche. Lengua y Cultura. Diccionario Mapudungun-español-inglés*. Maitén Editores.
- Hernández, J., Avendaño, A., y Rojas, J. (2012). Efecto de la frecuencia léxica y silábica en afasia anómica. *Revista Ciencias de La Salud*, 10(3), 337–346.
- Hume, E. (2011). Markedness. En M. van Oostendorp, C. Ewen, E. Hume y K. Rice (eds), *The Blackwell Companion to Phonology*, Volume 1 (pp. 79-106). John Wiley & Sons.
- Irwin, O. C. & Chen, H. P. (1946). Infant speech; vowel and consonant frequency. *The Journal of speech disorders*, 11, 123–125. <https://doi.org/10.1044/jshd.1102.123>
- Jiménez, M. J., Saavedra, C. & Sanhueza, M. (2019). *Frecuencia de fonemas en pseudopalabras del español de la Octava Región, Chile: un estudio comparativo*. Universidad de Concepción.
- Kalyani, N. y Sunitha, D. K. V. N. (2009). Syllable analysis to build a dictation system in Telugu language. *International Journal of Computer Science and Information Security*, 6(3), 171–176. <https://arxiv.org/abs/1001.2263v1>
- Key, M. R. (1968). Phonemic pattern and phoneme fluctuation in Bolivian Chama. *La Linguistique*, 2, 35–48.
- Kotu, V. y Deshpande, B. (2019). *Data Science: Concepts and Practice*. Morgan Kaufmann Publishers.

- Ladefoged, P. y Ferrari, S. (2012). *Vowels and consonants*. Wiley-Blackwell.
- Ladefoged, P. y Johnson, K. (2001). *A course in Phonetics*. Wadsworth Cengage Learning.
- Lagos, D. (1984). Fonología del mapuche hablado en Victoria. *Actas de Lengua y Literatura Mapuche*, 1, 41–50.
- Lagos, D. (2001). La lengua mapuche “Mapudungun”. En Ch. Clairis, A. M. Guerra, D. Lagos, (eds.), A. Riffo, y C. Villalón, *Fonologías de las lenguas indígenas de Chile* (pp. 224-267). Centro de Estudios Dialectológicos.
- Laime, T. (2007). *Diccionario Bilingüe. Iskay simipi yuyayk’ancha: Quechua - Castellano; Castellano-Quechua*. Auto-edición.
- Lenz, R. (1895-1897). *Estudios araucanos. Materiales para el estudio de la lengua, la literatura i las costumbres de los indios mapuche o araucanos*. Imprenta Cervantes.
- Lightner, T. M. (1978). Let’s hear it one more time on cluster frequencies. *Linguistics*, 16(211).
- Lloyd, P. y Schnitzer, R. (1967). A statistical study of the structure of the Spanish syllable. *Linguistics*, 5(37), 58-72.
- Loncon, E. (2017). Mapuzugun Mapuce Rakizuwam mew: Cumgefuy, Cumlen ka ñi Cumleam (El Mapuzugun desde el pensamiento mapuche: Pasado, Presente y Futuro). *Americania*, 204–219.
- Lotz, J. (1952). Vowel Frequency in Hungarian. *WORD*, 8(3), 227–235. <https://doi.org/10.1080/00437956.1952.11659435>
- Macklin-Cordes, J. L. y Round, E. R. (2020). Re-evaluating Phoneme Frequencies. *Frontiers in Psychology*, 11(November), 1–17. <https://doi.org/10.3389/fpsyg.2020.570895>
- Maddieson, I. (2013a). Consonant Inventories. En M.S. Dryer, y M. Haspelmath (eds.), *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology. Recuperado el 2022-04-20 de <http://wals.info/chapter/1>
- Maddieson, I. (2013b). Consonant-Vowel Ratio. En M.S. Dryer, y M. Haspelmath (eds.), *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology. Recuperado el 2022-04-27 de <http://wals.info/chapter/3>
- Malécot, A. (1974). Frequency of Occurrence of French Phonemes and Consonant Clusters. *Phonetica*, 29, 158–170.

- Márquez, C. (2016). *Frecuencia de fonemas y consideraciones teórico-metodológicas en un corpus del mapudungun*. Tesis para optar al grado de Licenciado en Educación. Universidad de Concepción.
- Martin, A. T. (2007). *The evolving lexicon*. A dissertation submitted in partial satisfaction of the requirements of the degree Doctor of Philosophy in Linguistics. [University of California].
- Martinet, A. (1964). *Economía de los cambios fonéticos*. Editorial Gredos.
- Mena, D. (2020). *Resolución de aspectos controversiales de la fonética y fonología del mapudungun mediante métodos de fonética acústica y estadística inferencial* (Tesis de doctorado). Universidad de Concepción, Facultad de Humanidades y Arte, Departamento de Español. <http://repositorio.udec.cl/jspui/handle/11594/620>
- Mena, D., Figueroa, M. y Salamanca, G. (2020). Análisis acústico-estadístico de las consonantes fricativas labiodentales e interdentes sonoras y áfonas del mapudungun hablado en Alto Biobío. *Logos: Revista de Lingüística, Filosofía y Literatura*, 30(2), 206–223. <https://doi.org/doi.org/10.15443/RL3017>
- Mena, D., Figueroa, M., Rogers, B. y Salamanca, G. (2019). Losing one allophone at a time: an acoustic and statistical study on mapudungun's sixth vowel. En S. Calhoun, P. Escudero, M. Tabain, y P. Warren (Eds.), *19th International Congress of Phonetic Sciences* (pp. 1426–1430).
- Mineduc. (2021). *Chawpin qichwata alli qillqanapaq maytu. Manual de escritura quechua central*. Edición digital.
- Molineaux, B. (2016). Vigencia del Diccionario Araucano de Félix de Augusta, a cien años de su publicación. *Boletín de Filología*, 51(1), 187–209. <https://doi.org/10.4067/S0718-93032016000100006>
- Molineaux, B. (2021). The dental-alveolar contrast in Mapudungun: Loss, preservation and extension. *Linguistics Vanguard* (En prensa).
- Mora, E. y Sanhueza, E. (2017). *Frecuencia de fonemas en pseudopalabras del español de la Octava Región, Chile: un estudio exploratorio*. Universidad de Concepción.

- Moulian, R., Catrileo, M. y Landeo, P. (2015). Afines quechua en el vocabulario mapuche de Luis de Valdivia. *RLA*, 53(2), 73–96. <https://doi.org/10.4067/S0718-48832015000200004>
- Müller, K. (2020). *'here': A Simpler Way to Find Your Files*. R package version 1.0.1. <https://CRAN.R-project.org/package=here>
- Munson, B. (2001). Phonological Pattern Frequency and Speech Production in Adults and Children. *Journal of Speech, Language, and Hearing Research*, 44(1–4), 778–792. [https://doi.org/10.1044/1092-4388\(2001/061\)](https://doi.org/10.1044/1092-4388(2001/061))
- Munthuli, A., Sirimujalin, P., Tantibundhit, C., Kosawat, K. y Onsuwan, C. (2013). A Corpus-Based Study of Phoneme Distribution in Thai. *Tenth Symposium on Natural Language Processing (SNLP2013)*.
- Munthuli, A., Tantibundhit, C., Onsuwan, C., Kosawat, K., y Wutiwiwatchai, C. (2015). Frequency of occurrence of phonemes and syllables in Thai: Analysis of spoken and written corpora. *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)*, August, 3–7.
- Nahar, K., Elshafei, M., Al-Khatib, W., Al-Muhtaseb, H., Nahar, K. M. O., Al-Khatib, W. G., y Alghamdi, M. M. (2012). Statistical Analysis of Arabic Phonemes for Continuous Arabic Speech Recognition. *International Journal of Computer and Information Technology*, 2279–0764
- Navarro, T. (1946). *Estudios de fonología española*. Syracuse University Press.
- Obediente, E. (2007). *Fonética y fonología*. Consejo de Publicaciones de la Universidad de Los Andes.
- Painequeo, J. H., Salamanca, G., y Jiménez, M. (2018). Estatus fonológico de los fonos interdentes [ɲ], [ʎ] y [tʃ] en el mapudungun hablado en el sector costa, Budi, Región de La Araucanía, Chile. *Revista Alpha*, 46, 111–128.
- Parodi, G. (2008). Lingüística de corpus: Una introducción al ámbito. *RLA - Revista de Lingüística Teórica y Aplicada*, 46(1), 93–119. <https://doi.org/10.4067/s0718-48832008000100006>
- Pavey, E. L. (2010). *The structure of language: An introduction to grammatical analysis*. Cambridge University Press.

- Peirce, C. S. (1906). Prolegomena to an apology for pragmaticism. *The Monist*, 16(4), 492-546.
- Pérez, C. y Salamanca, G. (2017). El mapuche hablado en Curarrehue: fonemas segmentales, fonotaxis y comparación con otras variedades. *Literatura y Lingüística*, 35, 315–335.
- Pérez, H. E. (2003). Frecuencia de fonemas. *Revista Electrónica En Tecnologías Del Habla* 1 (7).
- Pineda, L. A., Pineda, L. V., Cuétara, J., Castellanos, H., y López, I. (2004). DIMEx100: A new phonetic and speech corpus for Mexican Spanish. *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science)*, 3315(May 2014), 974–983.
- R Core Team (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Riffo, A. (2001). La lengua vanaña rapanui. En Ch. Clairis, A.M. Guerra, D. Lagos (eds.), A. Riffo, y C. Villalón, *Fonologías de las lenguas indígenas de Chile* (pp. 4-54). Centro de Estudios Dialectológicos.
- Robson, J., Pring, T., Marshall, J., y Chiat, S. (2003). Phoneme frequency effects in jargon aphasia: A phonological investigation of nonword errors. *Brain and Language*, 85(1), 109–124. [https://doi.org/10.1016/S0093-934X\(02\)00503-5](https://doi.org/10.1016/S0093-934X(02)00503-5)
- Rojo, G. (1991). Frecuencia de fonemas en español actual. En M. Brea y F. Fernández Rei (Eds.), *Homenaxe ó Profesor Constantino García* (pp. 451–467). Universidade de Santiago de Compostela.
- Sadowsky, S., Painequeo, J. H., Salamanca, G. & Avelino, H. (2013). Illustrations of the IPA: Mapudungun. *Journal of the International Phonetic Association*, 43(01), 87–96. <https://doi.org/10.1017/S0025100312000369>
- Salamanca, G. (1997). Fonología del pehuenche hablado en el Alto Bío-Bío. *RLA. Revista de Lingüística Teórica y Aplicada*, 35, 113–124.
- Salamanca, G. y Quintrileo, E. (2009). El mapuche hablado en Tirúa: Fonemas segmentales, fonotaxis y comparación con otras variedades. *RLA. Revista de lingüística teórica y aplicada*, 47(1), 13–35. <https://doi.org/10.4067/S0718-48832009000100002>
- Salas, A. (1976). Esbozo fonológico del mapuθuñu, lengua de los mapuče o araucanos de Chile central. *Estudios Filológicos*, 11, 143–154.

- Salas, A. (1985). Fray Feliz José de Augusta. Su aporte a los estudios de la lengua y la cultura de los mapuches o araucanos. *Cultura, Hombre, Sociedad*, 2(2), 197–272.
- Salas, A. (1992a). *El mapuche o araucano. Fonología, gramática y antología de cuentos*. Editorial MAPFRE S.A.
- Salas, A. (1992b). Lingüística mapuche. Guía bibliográfica. *Revista Andina*, 2, 473–537.
- Salas, A. (2006). *El mapuche o araucano*. Centro de Estudios Públicos.
- Saldivia, A. y Salamanca, G. (2020). Rasgos prominentes de la fonología segmental del lavkenche hablado en la comuna de Los Álamos, variante septentrional del mapudungun hablado en Chile. *Logos*, 30(1), 97–110.
- Salinas, C. y Salamanca, G. (2016). Reanálisis de aspectos controversiales de la fonología del chedungun hablado en alto Biobío: El estatus fonético-fonológico del fono fricativo alveopalatal áfono [χ]. *Nueva Revista del Pacífico*, 65, 137–159.
- Salminen, T. (2010). Europa y el Cáucaso. En Ch. Moseley (ed.), *Atlas de las Lenguas del Mundo en Peligro* (pp. 32-42). Ediciones UNESCO.
- Sampson, G. (2001). *Empirical Linguistics*. Continuum.
- Sánchez, G. (1989). Relatos orales en pewenche chileno. *Anales de la Universidad de Chile*, 17, 289-360. <https://doi.org/10.5354/0717-8883.1988.23915>
- Sánchez, M. y Salamanca, G. (2015). El mapuche hablado en Lonquimay: Fonemas segmentales, fonotaxis y comparación con otras variedades. *Literatura y Lingüística*, 31, 295–331. <https://doi.org/10.4067/S0718-48832009000100002>
- Sandvig, T. (1986). La reduplicación en mapudungun. *Actas de Lengua y Literatura Mapuche*.
- Saussure, F. (1945). *Curso de lingüística general*. Editorial Losada.
- Scholz, B., Pelletier, F., Pullum, G. y Nefdt R. (2022). *Philosophy of Linguistics*, *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.). <https://plato.stanford.edu/archives/spr2022/entries/linguistics/>
- Silber-Varod, V, Latin, M., y Moyal, A. (2017). Frequency of Hebrew phonemes and phoneme clusters in a data-driven approach Speech prosody of Hebrew. *Literacy and Language* (Oryanut Ve-Safa), 6, 22-36. (Artículo publicado en hebreo)
- Smeets, I. (2008). *A Grammar of Mapuche*. Mouton de Gruyter.

- Smirnova, N., y Chistikov, P. (2011). Statistics of Russian Monophones and Diphones. *Proceedings of Specom*, 218–223.
- Soto-Barba, J., Lara, I. y Salamanca, G. (2016). Descripción fonético-acústica de la sexta vocal en el chedungun hablado en Alto Bio-Bío. *Onomazein*, 34(2), 229–241.
- Stenneken, P., Hofmann, M., y Jacobs, A. M. (2005). Patterns of phoneme and syllable frequency in jargon aphasia. *Brain and Language*, 95(1 SPEC. ISS.), 221–222. <https://doi.org/10.1016/j.bandl.2005.07.114>
- Stevens, K. N. (2002). Toward a model for lexical access based on acoustic landmarks and distinctive features. *The Journal of the Acoustical Society of America*, 111(4), 1872–1891. <https://doi.org/10.1121/1.1458026>
- Suárez, J. A. (1959). The Phonemes of an Araucanian Dialect. *International Journal of American Linguistics*, 25(3), 177–181.
- Tamaoka, K. y Makioka, S. (2004). Frequency of occurrence for units of phonemes, morae, and syllables appearing in a lexical corpus of a Japanese newspaper. *Behavior Research Methods, Instruments, and Computers*, 36(3), 531–547.
- Tambovtsev, Y. (2003). Phonological similarity between Basque and other world languages based on the frequency of occurrence of certain typological consonantal features. *The Prague Bulletin of Mathematical Linguistics*, 79–80, 121–126.
- Tambovtsev, Y. (2010). Euphony in world languages. *Lingua Posnaniensis*, 52(1), 99–111. <https://doi.org/10.2478/V10122-010-0008-8>
- Trubetzkoy, N.S. (1939). *Principles of Phonology*. University of California Press.
- Trubetzkoy, N.S. (2001). Phonological systems considered in themselves and in relation to general language structure. En N.S. Trubetzkoy y A. Libermann (Ed.). *Studies in general linguistics and language structure*. (pp. 5-9), Duke University Press. (Obra original publicada en 1931)
- UNESCO. (2003). *Vitalidad y peligro de desaparición de las lenguas*. Documento adoptado por la Reunión Internacional de Expertos sobre el programa de la UNESCO “Salvaguardia de las Lenguas en Peligro”.

- Urrea, P. y Salamanca, G. (2021). Fonemas segmentales del mapudungun hablado en Icalma y configuración de un perfil fonético-fonológico del cordón cordillerano de habla mapuche-pewenche. *Logos* 31(2), 220–236.
- Valdivia, L. de. (1606). *Arte y gramática general de la lengua que corre en todo el Reyno de Chile, con un vocabulario, y confessorario*. Thomás López de Haro.
- Véliz, M., y Olate, A. (2008). Spoken word recognition in Mapudungu: A preliminary research. *Revista Signos*, 41(66), 107–125. <https://doi.org/10.4067/S0718-09342008000100005>
- Viegas Barros, J. P. (1999). Aspectos fonéticos y fonológicos de la dialectología del mapudungun en la argentina. *Actas de las III Jornadas de Etnolingüística*, 141–149.
- Viegas Barros, J. P. (2005). El mapudungun de la comunidad Millaín Kurikal (Neuquén, Argentina). Aspectos de la fonología y el léxico. En A. Tissera & J. Zigarán (Eds.), *Lenguas, educación y culturas* (pp. 77–89).
- Villena, B. (2014). *El mapudungún de Santiago de Chile: creación neológica y vitalidad interna*. Editorial Universitaria.
- Wang, H., y Song, M. (2011). Ckmeans.1d.dp: Optimal k-means clustering in one dimension by dynamic programming. *R Journal*, 3(2), 29–33. <https://doi.org/10.32614/rj-2011-015>
- Wickham, H. y Grolemund, G. (2017). *R for Data Science*. O’Reilly.
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K. y Yutani, H. (2019). “Welcome to the tidyverse.” *Journal of Open Source Software*, 4(43), 1686. doi: 10.21105/joss.01686
- Zamuner, T. (2003). *Input-Based Phonological Acquisition*. Routledge.
- Zipf, G.K. (1935). *The Psycho-Biology of language*. M.I.T.
- Zipf, G.K. y Rogers F. M. (1939). Phonemes and Variphones in four present-day Romance Languages and Classical Latin from the viewpoint of dynamic Philology. En *Archives Néerlandaises de Phonétique Expérimentale*, 15, 111-147. Citado en Alarcos, E. (1965)
- Zúñiga, F. (2000). *Mapudungun*. Lincon Europa.

Zúñiga, F. (2014). Nominal compounds in Mapudungun. En S. Danielsen, K. Hannss, y F. Zúñiga (Eds.), *Word Formation in South American Languages* (pp. 11–31). John Benjamins Publishing Company.



Anexos

1. Anexo sobre sufijos

1.1. Listado de sufijos identificados

A continuación, se identifican y glosan los sufijos identificados en nuestro análisis. Además, en la columna titulada ‘MDV’, se proponen las formas que podrían ser parte de un estudio para establecer las frecuencias de los fonemas en el dominio de la morfología derivativa y flexiva del mapudungun. En algunos casos, se considera que la forma gramatical se encuentra ya contabilizada al existir un lexema del cual deriva, y, en consecuencia, no aplica (N/A).

rank	n	glosa	Morfos	ejemplos	MDV
1	6671	INF	-n -ən	ko-n ‘v. traer agua’	-n
2	863	VBR ADV APL RST	-tu	müchay-tu-n ‘v. teñir con michay’	N/A
3	766	EST	-kəle -le	lef-küle-n ‘v. estar corriendo’	-le
4	636	CAUS	-l -əl	küpa-l-ün ‘v. hacer venir, traer’	-l
5	619	REFL	-w -uw -əw -wə	lif-w-ün ‘v. limpiarse’	-w
6	303	VBR	-ŋe	kure-nge-n ‘v. tener esposa, ser casado’	N/A
7	266	APL	-ma -ɲma -əɲma	piliñ-ma-n ‘v. helarse (las sementeras)’ kim-üñma-n ‘v. entender a alguien’	-ɲma
8	241	PERS	-kənu -kəno	nüla-künu-n ‘v. dejar abierto, abrir’	N/A
9	237	CAUS	-m -əm	narfü-m-ün ‘v. mojar, humedecer’	-m
10	216	COL	-ntəku -təku -əntəku -ɲtəku	malal-tüku-n ‘v. cercar’	N/A
11	214	CONT	-ka	nie-ka-n ‘v. retener’	-ka
12	214	LOC	-we	anü-m-ka-we ‘v. plantío, huerto’	-we

13	162	ABJ	-naɥ -nakəm	lof-nag-ün ‘v. desplomarse’ imül-naküm-ün ‘v. hacer rodar abajo’	N/A
14	158	CIRC	-jaw -kiaw -kəjaw	llallitu-yaw-ün ‘v. andar en visitas’ amu-yaw-ün ‘v. andar sin rumbo’	-jaw
15	144	CIS	-pa	amu-pa-n ‘v. venir acá’	-pa
16	135	EXTR	-nentu -entu	folil-entu-n ‘v. sacar de raíz’	N/A
17	131	AG	-fe -fə -fue	pütu-fe ‘s. bebedor’ wül’-fü ‘s. balsero’	-fe
18	118	APL	-el -lel	luyüf-el-ün ‘v. alisar’	-el
19	110	ADN	-kon	ütrüf-kon-ün ‘v. caer adentro’	N/A
20	106	PERS	-nie	el-ka-nie-n ‘v. tener guardado’	N/A
21	100	APL	-je	duam-ye-n ‘v. necesitar’	N/A
22	98	GRP	-ntu	kura-ntu ‘s. pedregal’	N/A
23	96	AND	-me	llallitu-me-n ‘v. ir a visitar’	-me
24	94	PL	-ke	af-ke-ntu ‘adv. sin cesar’	-ke
25	92	ARR	-pəɟa -ɲpəɟa -npəɟa	üwir-püra-m-ün ‘v. erguir’ witra-ñpüra-m-ün ‘v. alzar’	N/A
26	81	AFU	-tʃipa	lef-tripa-n ‘v. huir’	N/A
27	78	LOC	-pu	af-pu-n ‘v. llegar al fin’	-pu
28	67	RES	-j -ə -mu -pi -ken -nej -ɲeɬu -əɬ -ɲeləm -ɲen -aɬ -i -ləl -mun -ɲtʃə -ɲeɬ -ɲken -tʃa	chacha-y ‘s. papito’ kon-ü-we antü ‘s. el oeste’ llay-pi-n ‘v. escoger, reservar’ i-lül ‘s. cierto tipo de carne’ l’af-ken’ ‘s. mar’ epu-e ‘adv. hace dos días’ epu-ñchü-ke ‘adj. doble’ chang-üll ‘v. dedo’ lam-ngen ‘s. hermana’ few-all ‘adv. finalmente, al fin’ ngull-che-ñmaywe ‘s. cierto lugar’ chang-üll ‘s. dedos’ af-i ‘v. se acabó’	-j -ə -mu -pi -ken -nej -ɲeɬu -əɬ -ɲeləm -ɲen -aɬ -ləl -mun -ɲtʃə -ɲeɬ -ɲken -tʃa -tʃe

			-tʃe -tʃil -θel -θə -θi		-tʃil -θel -θə -θi
29	62	ADJ	-fal -faɫ	ayü-fal ‘adj. amable’ illu-fal ‘adj. apetecible’	-fal
30	55	INTR	-ɹupa -ɹpa -ɹulpa -əɹpa -ɹəpa	mültü-rulpa-n ‘v. embestir’	N/A
31	48	ADJ	-tʃi	inaninan-chi pu che ‘s. súbditos’	-tʃi
32	42	NOM	-pejəm -peje -pejel -peɹ -pem	pütu-peye ‘s. vaso’ pütu-peyel ‘s. bebida’ pütu-peyüm ‘s. casa donde se bebe’	N/A
33	40	NOM	-el -uɛl -eɫ	ñuke-ye-el ‘s. tía materna’ allkü-pe-no-el ‘adj. inaudito’	-el
34	36	VBR	-kə	düching-kü-n ‘v. ralear’	-kə
35	35	INTR	-ɹpu -əɹpu -ɹəpu	fe-le-rpu-n ‘v. ser así en adelante’	N/A
36	35	RECR	-kantu	kude-kantu-n ‘v. jugar a apostar’	-kantu
37	33	SOC	-wen	kellu-wen ‘s. colaborador’	-wen
38	32	NOM	-lu	mün’a-rume-lu ‘adj. de poca circunferencia’	-lu
39	31	INTR	-ɹume -wme	chaykad-rume-n ‘v. pasar agachado’ ka-ke-wme ‘adj. de varias clases’	N/A
40	30	INTER	-ɹu -ɹul	lef-ru-tu-n ‘v. embestir’	N/A
41	26	ADV	-ketʃi	nor-kechi ‘adv. rectamente’	N/A
42	25	NOM	-ɹi	raki-ñ ‘s. cuenta’	-ɹi
43	23	NEG	-no -nu -ki -la	pelo-nu-lu ‘s. ciego’	-no -ki -la
44	22	BEN	-mpe -npe -ɹpe	weycha-ñpe-n ‘v. defender (a alguien)’ illku-mpe-n ‘v. reprender’	-npe

45	19	DIM	-kip -kin -kił -kəł	pang-küll ‘s. cachorro de puma’ traf-kin ‘s. cierta relación social’ müllo-kiñ ‘s. especie de albóndigas’ chang-kill ‘s. ramas de árbol’	-kip
46	17	FUT	-a -ja	now-ün kewa-ya-lu ‘exp. trabar una pelea’	-a
47	16	AND	-amu	ipün-amu-n ‘v. alejarse volando’	N/A
48	13	NOM	-nte -ntje -ten -nte	yewe-nte-n ‘adj. vergonzoso’ afü-nche-n ‘adj. blando al cocerse’ trepe-w-ten ‘adj. espantadizo’	-nte
49	12	INTS	-kətuje -kətie	wima-kütie-n ‘v. azotar con muchos varillazos’	-kətuje
50	11	PROX	-pe	allkü-pe-no-el ‘s. inaudito’	-pe
51	10	INTR	-wəl	kecha-wül-ün ‘v. arrear (de atrás)’	-wəl
52	10	PROG	-meke	tüng-meke-n ‘v. quedarse tranquilo en un lugar’	N/A
53	9	INTS	-kepi	püd-üm-ün-kepi-n ‘v. dispersar en todas direcciones’	-kepi
54	9	SIML	-wfaluw -faluw	pe-wfaluw-ün ‘v. hacerse visible’	-faluw
55	9	VAN	-pəθa	pi-püda-n ‘v. decir sin motivo’	N/A
56	9	RPR	-i -n -iŋ -u -e -m -ŋ -ən -tje	-i-n-che-iñ ‘pers. nosotros PL’ -e-i-m-ün ‘pers. ustedes DL’ -e-ng-u ‘pers. DL’ -i-u ‘poss. nuestro DL’ -i-iñ ‘poss. nuestro PL’	-i -iŋ -u -e -m -ŋ -ən
57	7	DES	-fel	nge-no-a-fel ‘adv. a todo trance’	N/A
58	7	MOV	-jekəme -jekowme -jekəpa -jekuwme	al’ü-yeküme-n ‘v. ir en aumento’	N/A
59	6	NOM	-mo -mom -mum	rupa-mom ‘adv. después de una fecha’	N/A
60	6	VEN	-kəpa	witra-küpa-l-ün ‘v. atraer’	N/A
61	5	AFR	-ła -ləe	awü-lla-n ‘v. hacer sufrir (mucho)’	-ləe

62	5	BEN	-ɲmu	i-ñmu-n ‘v. comerse algo solo’	-ɲmu
63	5	TEMP	-jəm	kiñe-yüm ‘adv. alternativamente’	-jəm
64	5	NOM	-kən -kəm	chem-kün ‘s. objeto’ karü-n-kün ‘s. el verdor’	-kən
65	5	REP	-ə.ɬke -ɬka -ɬke -ɬkə	trewa rke! ‘exp. ¡oh, un perro!’	-ɬke
66	3	CIRC	-tiaw	wallo-tiaw-ün ‘v. girar’	-tiaw
67	3	TRAV	-ɲo	weyel-n'o-n ‘v. pasar a nado’	N/A
68	3	NOM	-am	pi-am ‘exp. dizque’	-am
69	3	PL	-küf	yall-küf-el ‘s. los hijos nacidos’	-küf
70	3	VBR	-ku	ngaw-ku-n ‘v. lavar las manos’	-ku
71	2	INM	-fem	rupa-fem-ün ‘v. pasar de inmediato’	N/A
72	2	NOM	-en	palpal-en ‘s. cierto arbusto’	-en
73	2	NOM	-tra	püna-tra ‘v. cierto hongo’	-tra
74	2	NOM	-wma	wadkü-wma ko ‘s. agua ya cocida’	-wma
75	1	PLPF	-uje	allkü-uye-el ‘s. una noticia’	-uje
76	1	ADV	-əlke	fam-ülke ‘adv. en aquel tiempo’	-əlke



1.2. Formas alternantes con la vocal /ə/ en el radical verbal

Las siguientes tablas identifican los pares de formas que evidencian alternancia respecto de la realización de una vocal /ə/ al final del radical.

1.2.1. Formas seleccionadas sin la vocal /ə/

#	Selección	Alternancia	Definición
1	along	alongü	‘v. relumbrar’
2	añid	añidü	‘v. secar un palo’
3	ayong	ayongü	‘v. estar muy claro (el sol)’
4	chaykad	chaykadü	‘v. filtrarse, absorberse’
5	chidküy	chidküyü	‘v. amarotarse’
6	chongof	chongofü	‘v. caerse las hojas del haba’
7	chüngar	chüngarü	‘v. herir con cuchillo o puñal’
8	demay	demayü	‘v. envanecerse’
9	dum·dum	dum·dumü	‘v. ofuscarse’
10	efkeñ	efkeñü	‘v. respirar rápidamente’
11	far·far	far·farü	‘v. sudar mucho’
12	fey	feyü	‘v. bastar, caber’
13	foltray	foltrayü	‘v. matar con los dientes’
14	iwüd	iwüdü	‘v. envolver’
15	kalpud	kalpudü	‘v. arrancar, tirar (pasto)’
16	kengkür	kengkürü	‘igualar, emparejar’
17	kuf·kuf	kuf·kufü	‘v. caminar con pasos oíbles’
18	konkül	konkülü	‘v. gruñir’
19	küd·küd	küd·küdü	‘v. silbar (ratones, culebras)’
20	künay	künayü	‘v. flotar’
21	küntray	küntrayü	‘v. flotar, boyar’
22	künül	künülü	‘v. sonar las flemas’
23	küyküy	küyküyü	‘v. estar muy ocupado’
24	lar	larü	‘v. caerse’
25	lutrung	lutrungü	‘v. pisar fuerte, con ruido’
26	luyüf	luyüfü	‘v. liso, pulido’
27	llay	llayü	‘v. quedarse atrás’

28	llayllay	llayllayü	‘v. aparecer los rayos del sol’
29	llumed	llumedü	‘v. ser sabroso.’
30	may·may	may·mayü	‘v. separar (trigo)’
31	melkay	melkayü	‘v. resbalar’
32	mül’·mül’	mül’·mül’ü	‘v. tiritar, temblar’
33	mütrong	mütrongü	‘v. dar golpes’
34	mütrüm	mütrümü	‘v. llamar a gritos’
35	mütrur	mütrurü	‘v. tropezar’
36	müyüd	müyüdü	‘v. balancear, tambalearse, vacilar’
37	mongof	mongofü	‘v. pacer’
38	nükür	nükürü	‘v. sollozar’
39	nüngkül	ngünkülü	‘v. arrancar’
40	nüy·nüy	nüy·nüyü	‘v. esparcir’
41	ngeyküll	ngeyküllü	‘v. mecer’
42	ngülof	ngülofü	‘v. abollarse’
43	ngülör	ngülörü	‘v. cosechar (el maíz)’
44	ngüllud	ngülludü	‘v. enjugar’
45	pingüd	pingüdü	‘v. refregar’
46	pitrong	pitrongü	‘v. (h.) picotear, (p.) agobiarse’
47	punul’	punul’ü	‘v. hincharse’
48	püner	pünerü	‘v. cauterizar’
49	püntrüf	püntrüfü	‘v. ablandarse (al cocinar)’
50	pütaf	pütafü	‘v. quitar a papirotes’
51	shilüm	shilümü	‘adormecerse (un miembro).’
52	toftof	toftofü	‘v. hacer ruido’
53	tülang	tülangü	‘v. talonear (al caballo)’
54	tralof	tralofü	‘v. golpear’
55	trelol	trelolü	‘v. levantar (las asentaderas)’
56	trelpong	trelpongü	‘v. trotar’
57	trentil	trentilü	‘v. pelarse (el cutis), descostrarse’
58	trolol	trololü	‘v. ahuecar’
59	tromir	tromirü	‘v. tener parálisis (ojo)’
60	trongkoy	trongkoyü	‘v. volcarse, darse vuelta, rotar’
61	trong·trong	trong·trongü	‘v. hincharse’

62	truf·truf	truf·trufü	‘v. hervir a borbotones’
63	truy·truy	truy·truyü	‘v. colmar’
64	trupef	trupefü	‘v. espantarse’
65	trüker	trükerü	‘v. torcer los ojos con rabia’
66	trüf	trüfü	‘v. sofocarse’
67	trümpay	trümpayü	‘v. ladearse’
68	trümüm	trümümü	‘v. ensordecer, zumbar’
69	trünüy	trünüyü	‘v. vibrar’
70	ungul	ungulü	‘v. deslizarse’
71	üngar	üngarü	‘v. mascar con ruido’
72	üngür	üngürü	‘v. gastarse’
73	üngül	üngülü	‘v. roer’
74	ütrir	ütrirü	‘v. envidiar’
75	wangil	wangilü	‘v. abundar cosas agudas paradas’
76	wechod	wechodü	‘v. agujerearse’
77	weñang	weñangü	‘v. tener deseo por algo’
78	wetraf	wetrafü	‘v. doblarse para atrás (cuerpo)’
79	wichor	wichorü	‘v. tener mucha corriente (agua)’
80	wingüd	wingüdü	‘v. arrastrar’
81	wingul	wingulü	‘v. deslizarse’
82	wilüf	wilüfü	‘v. brillar’
83	wilfod	wilfodü	‘v. empaparse, mojarse mucho’
84	wilper	wilperü	‘v. faltar a un objeto un pedazo en las orillas.’
85	wiltraf	wiltrafü	‘v. pegar’
86	wimpüll	wimpüllü	‘v. estar prendido de sangre’
87	wiraf	wirafü	‘v. pegar (cosas flexibles)’
88	witüf	witüfü	‘v. salirse solo’
89	würwür	würwürü	‘v. gruñir, bramar’
90	yifüd	yifüdü	‘v. quitar raspando’
91	yof·yof	yof·yofü	‘v. elevarse borbollando’
92	yufül	yufülü	‘v. hacer vueltas y tornos’

1.2.2. Formas seleccionadas que contienen la vocal /ə/

#	Selección	Alternancia	Definición
1	keyü	key	‘adv. hasta, inclusive’ Cft. key-tu-n ~ keyü-tu-n ‘v. acertar’
2	patrüü	patrü	‘v. esparcir (semilla)’
3	pellpellü	pellpell	‘v. angustiarse’
4	pifarü	pifar	‘v. bufar (caballo)’
5	repülü	repül	‘v. cornear’
6	rükafü	rükaf	‘v. clavarse (el sujeto) algo como astillas, al pisar.’
7	trawü	traw	‘v. reunirse gente’
8	trükefü	trükef	‘v. encogerse, acalambrarse’
9	üfülü	üfül	‘v. sorber con los labios’
10	wicharü	wichar	‘v. rasgarse’
11	wileyü	wiley	‘v. despachurrar’
12	yillü	yill	‘v. multiplicar’
13	kallkoyü	kallkoy	‘v. enamorarse’
14	llamllamü	llallam	‘v. undular, formar bandada’
15	nüwfü	nüwüf	‘v. pasar algo extendiendo la mano.’
16	nüküfü	nüküf	‘v. callarse’
17	müpafü	müpaf	‘v. aletear’
18	piwü	piw	‘v. secarse’
19	tralüngü	tralüng	‘v. destaparse’
20	üfü	üf	‘v. apretar, atar’

2. Anexos sobre préstamos

2.1. Préstamos del castellano

#	lema	definición	< deriva de
1	akucha	s. aguja	< aguja
2	angkad	adv. en ancas	< ancas
3	añil	s. el añil	< añil
4	arimatu	adj. llegado de otra parte	< arrimar
5	arkolla	s. argolla	< argolla
6	asotin	v. azotar	< azote
7	aspa-we	s. el aspa	< aspa
8	asus	s. el ajo	< ajos
9	awar	s. el haba	< habas
10	cheda	s. cedazo	< cedazo
11	choapino	s. chañuntüku con las motas largas	< choapino
12	darma	s. salma	< jalma
13	defe	s. deuda	< deber, deuda
14	dengde	(P.) v. vender algo	< vender
15	echal	v. echar	< echar
16	echun	v. estar hecho, acostumbrado	< hecho
17	empeña	v. dar en prenda, empeñar	< empeñar
18	espejo-tu	v. mirarse en el espejo	< espejo
19	estipu	s. estribos	< estribo
20	fawül	(P.) s. baúl	< baúl
21	fela	s. vela para dar luz	< vela
22	finu	adj. fino	< fino
23	finu	s. el vino	< vino
24	fisaro	adj. perito, baquiano	< bizarro
25	fofo	adj. bobo	< bobo
26	foshal, fosal	s. el bozal	< bozal
27	fota	s. la bota	< bota
28	fota	s. odre en que se maza la leche	< bota
29	foya-gel ñeweñ,	s. los boyantes de la red	< boyar

	foy-el ñeweñ		
30	frasko	(P.) s. cuerno arreglado como frasco	< frasco
31	kalon	s. los dos palos que sirven para mantener estirada la red bajo el agua	< calar
32	kama–ruka	(P.) s. comp. los largueros de la casa	< cama
33	kansha-le	v. estar cansado	< cansar
34	karniru	s. el carnero	< carnero
35	karoti	s. el garrote	< garrote
36	kaso	v. hacer caso a alguno	< caso
37	kason	(P.) s. los calzones, pantalones	< calzón
38	kawchu	adj. soltero, soltera (es expresión ordinaria; otros no entienden soltero, sino el término gaucho)	< gaucho
39	kekaw	v. quejarse	< quejar
40	kontra	v. con pronunciación española: hacer oposición a alguno, declarar en su contra	< contra
41	koron	s. la correa	< cordón
42	koyotro	s. coyochó, bulbo de yuyo	< coyochó
43	kriollo püñeñ	(P.) cierta denominación	< criollo
44	krusel	s. la cruz	< cruz
45	kulpa	v. cometer culpa; hacer daño; ofender	< culpa
46	kupilka	s. cupilca, mezcla de harina tostada y de chicha	< cupilca, chupilca
47	küruselün	v. persignar	< cruz
48	kuyunta	s. coyunda	< coyunda
49	kuywintu	(P.) s. la coyunda	< coyunda
50	lastra tirador	s. cierta prenda de plata	< lastra, tirador
51	linu	(P.) s. linaza	< linaza
52	liñu	(H.) s. linaza	< linaza
53	llamada	s. ramada	< ramada
54	llongkañ	(P.) s. salchicha	< longaniza
55	lofo	adj. huidor	< lobo
56	madakal	s. borlón que llevan los caballos en la frente	< máscara
57	malkada	(P.) s. careta	< máscara
58	manshun	(P.) s. buey	< manso
59	manshun'	(H.) s. buey	< manso
60	maynan	v. manear (la bestia)	< manear

61	medi	v. medir, inquirir, examinar	< medir
62	moltero	s. mortero	< mortero
63	motro	adj. mocho, mutilado, obtuso	< mocho
64	napur	s. bulbo encima de la raíz, como en el rábano	< nabos
65	nempür	(P.) s. el membrillo	< membrillo
66	ñulluydu	(H.) adj. tullido	< tullido
67	olkita	s. la orqueta	< orqueta
68	pafu	s. el pavo	< pavo
69	pala	s. la pala	< pala
70	palenke	s. una estaca plantada en medio de la era	< palenque
71	pampino	(P.) adj. de lomo blanco	< pampino
72	pangku	s. el banco	< banco
73	pañu	s. el paño	< paño
74	paraka	s. los naipes	< baraja
75	paro	s. lance de cuatro caras blancas y cuatro negras en el juego de las habas	< paro
76	pashian	v. dar un paseo	< pasear
77	patia	s. la batea, artecilla	< batea
78	patiru	(P.) s. el padre misionero	< padre
79	pato	s. el pato (ave)	< pato
80	patron	s. el patrón	< patrón
81	peka	v. pecar	< pecar
82	penga	adv. indiscretamente	< pecar
83	pesheytu	(P.) v. mimar	< besito
84	peshid, peshis	(P.) s. perdiz	< perdiz
85	peyne	s. el peine	< peine
86	peyneta	s. el peine que tiene dientes a ambos lados	< peineta
87	pofu	adj. bobo	< bobo
88	pofre	adj. pobre	< pobre
89	ponshon	s. el punzón, alhaja de plata	< punzón
90	posho	s. pozo	< pozo
91	presu	s. prisión	< preso
92	pülawta	(H.) s. flauta	< flauta
93	re falta	adv. de balde, gratis, en vano	< falta, en balde

94	rey	s. rey	< rey
95	rilo	(H.) s. el hilo de fábrica	< hilo
96	saku	s. saco	< saco
97	salay, say	interj. para espantar a los perros	< sal ahí
98	salma	s. ciertos aparejos para el caballo	< jalma, enjalma
99	sampu	adj. crespo	< zambo
100	sanchu	s. chancho	< chancho
101	sechaw	s. dechado	< dechado
102	serfin	v. servir, cuidar	< servir
103	shofa	v. sobar	< sobar
104	oron, orron	s. zurrón	< zurrón
105	suran	(H.) v. jurar	< jurar
106	tesa	v. dar puñetazos a alguno; tesar	< tesar
107	timun	s. arado	< timón
108	tol'to	(P.) s. toldo	< toldo
109	topa	v. topar, caerle en suerte, lograr hacer algo	< topar
110	waka	s. el vacuno, la vaca	< vaca
111	were	(H.) interj. para estimular a los bueyes	< bueyes
112	yuku	s. yugo	< yugo
113	kaputira	(P.) s. aguja, alfiler pequeños	< agujita, puntero
114	kaputiru	(H.) s. aguja, alfiler pequeños	< agujita, puntero, agujero
115	piaw	(H.) adj. puyado, cierto color de caballo	< puyado, bayo
116	wachol	adj. (animal) criado en la casa, guacho	< guacho
117	waria	s. ciudad	< guardia
118	wella	s. bot. la huella (planta)	< huella
119	mushantu-fe	(H.) adj. músico	< música
120	palao	adj. bayo; claro	< pálido
121	pikuña makuñ	(P.) s. comp. mantas argentinas, que tienen flecos en toda la orilla	< vicuña
122	kachilla	s. trigo; trigal	< castilla
123	tapayu	s. negros (hombres y caballos)	< tapado
124	temlaria	(P.) s. vaso de plata para tomar chicha	Sin hipótesis.

2.2. Cognados del quechua

El número de página remite a las definiciones del diccionario de Laime (2007).

	Mapudungun		Quechua	
1	ampi	v. aplicar remedio de polvo o hierba quemada a la piel, a las llagas.	hampij	v. Curar. Sanar aplicando remedios. (p. 34)
2	atʃawəʎ ~atʃawaʎ	s. gallina, gallo.	wallpa	s. Gallina. Hembra del gallo, ave galliforme. (p. 128)
3	awka awkan	adj. Alzado, rebelde, muy travieso. s. Las yeguas (porque suelen ser chúcaras; pero ñom yegua 'una yegua mansa'). intr. Sublevarse. s. El alzamiento, la rebelión.	awqa awqaj	adj. y s. Enemigo, -ga. Contrario. intr. Guerrear, hacer guerra. (p. 16)
4	jana	s. mapuches ladinos criados entre los españoles.	jana	fig. adj. y s. Esclavo, -va. Persona que está bajo la absoluta dependencia de otra o que está bajo la dominación de alguien. (p. 137)
5	kaka pəpəp	s. Panguí. Sobrina política (expresión ridícula).	kaka ~kaki	s. Tío. Hermano de la madre respecto del sobrino o sobrina. (p. 42)
6	kamaɲ	s. Guardián, pastor.	kamaj	s. Ocupación. Trabajo, empleo, oficio o dignidad. (p.43)
7	kaɲka	tr. Asar. adj. Asado. s. El asado.	kankaj	tr. Asar. Someter un manjar al fuego. (p.43)
8	kaɭa	s. La ciudad.	pukara	s. Fortaleza. Castillo, recinto fortificado. (p. 79)
9	katʃu	s. El pasto, la hierba.	q'atʃu	s. Hierba verde que se da a los animales herbívoros. Pasto. Hierba que paca el ganado. (p. 96)
10	kawətu ~kawitu	s. Catre de cama.	kawitu	s. Hamaca. Mecedora, columpio. (p.43)
11	minka	intr. Panguí. Buscar gente para el trabajo y pagar con comida y bebida. Minkayan 'haré un mingaco, una minga'.	mink'a mink'aj	s. Servicio personal pagado. trabajo no remunerado, gratuito (Pot.). tr. Contratar personas. (p. 68)
12	mifki	adj. Dulce. s. La miel. mishki abeja. Miel de abejas.	misk'i	adj. Dulce. De sabor de la miel. (p. 68)
13	pataka	num. Ciento. epu pataka.	patʃak	adj. Cien. Apócope de

		Doscientos.	~patʃaq	ciento que determina noventa más diez unidades. (p. 75)
14	pətʃə ~pitʃi	adj. Huapi. Pequeño, chico, menudo. adv. Un poco. adj. Pangui. Chico, pequeño, poco.	pisi	adj. poco, -ca. Escaso, limitado. Adv. en pequeña cantidad, un poco. (p. 79)
15	poθkol ~poθkoł	s. Huapi. La levadura.	p'usqu ~p'utʃqu	s. Levadura. Fermento. (p. 86)
16	putʃu	intr. Sobrar. adj. Lo sobrante. Puchun kofke 'las sobras del pan'. Puchun cigarro 'la colilla de cigarro'.	putʃu	s. Residuo. Sobra, parte que queda de un todo (p. 79)
17	ɹaki	v contar, calcular; s. cuenta, cálculo	rak'i	s. Partición. Repartimiento, división. (p. 99)
18	tʃałwa	s. El pescado.	tʃałwa	s. Pescado. Pez sacado del agua. (p. 19)
19	tʃaɹki	s. Carne cortada en fajas y secada al sol o ahumada (cecina).	tʃ'arki	s. Cecina. Carne salada y secada al sol. (p. 27)
20	tʃilka ~tʃiłka	tr. Pangui. Marcar los animales, p. ej. cortándoles una oreja o de cualquier modo semejante. s. Libro, carta, noticia.	qiłqa	s. Escritura. Acto y arte de escribir. Documento (p. 89)
21	tʃutʃoka	s. Maíz tostado o cocido para conservarlo.	tʃutʃuqa	s. Maíz medio hervido y secado al sol, que suele guardarse en conserva. (p. 22)
22	wampo ~wampu	s. La canoa.	wamp'u	s. Barco. Embarcación para transportar mercancías o personas. (p. 129)
23	waraŋka	num. Mil	waranqa	adj. Mil. (p. 131)
24	jana	s. criado	jana	s. esclavo (p. 137)

3. Anexos sobre alternancia y fluctuación fonemática

A continuación, se incluyen las alternancias encontradas durante el muestreo. La primera forma es la forma de RIV, que indica la forma seleccionada. La segunda columna indica la alternancia con un segmento de RIT. Finalmente, se incluye una breve definición con la clase de palabra.

3.1. Alternancia /ə/ > /i/

	RIV: /ə/	RIT: /i/	Definición
1	afkə	afki	‘v. suspirar’
2	aɪəɲ	aɪiɲ	‘v. ahumarse la comida.’
3	atʃəθko	atʃiθko	‘v. tener acedia’
4	əle.ɔ	ileɪ	‘v. sobar, amasar’
5	əɣku	iɣku	‘v. enojarse, encolerizarse’
6	əpəɪ	əpiɪ	‘v. picar (las pulgas, etc.)’
7	əwəɪ	əwiɪ	‘v. erguirse, estirarse’
8	fəɪə	fiɪə	‘v. silbar el viento’
9	fəɪkə	fiɪkə	‘s. el fresco’
10	fotʃə	fotʃi	‘tr. extraer chupando’
11	jantʃə	jantʃi	‘v. tener escalofríos’
12	jəfkə	jiɪkə	‘v. desbastar’
13	jəwe	jiwe	‘s. taza de greda con asa’
14	kalofətʃa	kalofitʃa	‘s. sanguijuela’
15	kəɣə	kəɣi	‘v. acertar’
16	kətʃə	kətʃi	‘s. mirto, arrayàn’
17	kəθ	kiθ	‘s. sonido/silbido del gato o ratón’
18	koɲə	koɲi	‘v. parir’; ‘s. cría’
19	koθəw	koθiw	‘adj. apolillado’
20	kuɲə	kuɲi	‘s. el cuidado’
21	ləɰ	liɰ	‘adj. blanco’
22	ɣəfeθ	ɣifeθ	(H.) ‘v. flojear’; (P.) ‘v. vers como puntos’
23	ɣəɲəθ	ɣiɲəθ	‘adj. plano, llano’
24	ɣəpəθ	ɣipaθ	‘v. echarse (aves)’
25	ɣəw	ɣiw	‘v. vara grande’
26	ɣəwa	ɣiwa	‘v. notar’

27	majtʃə	majtʃi	‘v. hacer una seña con la mano.’
28	məki	miki	‘s. cierto arbusto’
29	məʎsən	məʎsɪn	‘s. catuto’
30	məɹəŋ	məɹiŋ	‘v. tener el cuerpo grueso y alto’
31	məʎsələ	məʎsilə	‘v. deshacer, destruir’
32	məʎsə	məʎsi	‘adj. gordo’
33	nəkəfə	ŋikəf	‘v. callarse’
34	nəkəm	ŋikəm	‘v. guarecerse de la lluvia’
35	nəkəɹ	nikəɹ	‘v. gemir’
36	nəni	nini	‘v. hormiguesar, hervir, abundar’
37	nəŋaj	niŋaj	‘v. impacientarse’
38	nəɹə	nəɹi	‘v. obstruirse (el vientre, la uretra).’
39	nəwəɹ	niwəɹ	‘v. pasar afirmándose’
40	ŋəkəɹ	ŋikəɹ	‘v. rizar, ajar telas’
41	ŋəmə	ŋəmi	‘tr. coger, alzar.’
42	ŋəpəθ	ŋipəθ	‘adj. pegadizo’
43	ŋəwme	ŋiwme	‘s. cierta planta’
44	ŋetʃə	ŋetʃi	‘v. evacuar el vientre’
45	ŋəju	ŋiju	‘v. endilgar’
46	pəjkəŋ	pəjkɪŋ	‘v. representarse la imagen de cosas o personas’
47	pəθwɪŋ	pilwɪŋ	‘cierto animal’
48	pəkɪŋ	pikɪŋ	‘v. piquinear el telar’
49	pəla	pila	‘s. una planta’
50	pəɹəf	piɹəf	‘adj. delgado, tierno, fino’
51	pɪʎəŋ	pɪʎiŋ	‘s. paico’
52	ɹəjə	ɹəji	‘v. revolver’
53	ɹəŋkə	ɹiŋkə	‘v. saltar’
54	ʃəŋəl	ʃiŋəl	‘v. pasar agachado, por agujero, etc.’
55	ʃəwmeŋ	θiwmeŋ	‘s. abeja’
56	ʃupə	ʃupɪ	‘v. matar piojos’
57	tʃəɹləm	tʃiɹləm	‘s. cierta planta’
58	tʃəɹləmə	tʃiɹlimə	‘v. adormecerse un miembro’
59	tʃəŋaj	tʃiŋaj	‘v. flecos, franjas’
60	tʃəŋkəθ	tʃiŋkəθ	‘adj. redondo’

61	ṭṭəŋko	ṭṭiŋko	‘v. atestar, rodear’
62	ṭṭəɽwa	ṭṭiɽwa	‘v. turbarse, enredarse’
63	ṭṭəθkə	ṭṭiθkə	‘v. chorrear’.
64	ṭṭiŋəθ	ṭṭiŋiθ	‘v. cerner’
65	ṭṭiɽəf	ṭṭiɽif	‘s. chirigüe, un pájaro’
66	ṭṭoʝə	ṭṭoʝi	‘v. nacer, brotar’
67	ṭṭaɽə	ṭṭaɽi	‘v. amarrar’
68	ṭṭəɽṭṭəw	ṭṭəɽṭṭiw	‘v. meter ruido (aves)’
69	ṭṭəɽfu	ṭṭiɽfu	‘v. fatigarse’
70	ṭṭəṭṭəɽə	ṭṭiṭṭəɽə	‘v. crujir’
71	ṭṭiŋṭṭə	ṭṭiŋṭṭi	‘adj. crespo’
72	ṭṭəkə	ṭṭoki	‘v. parecerle, hacer juicio’
73	wəɽəl	wəɽil	‘s. laringe con la tráquea’
74	wəkəlpe	wəkilpe	‘s. un pez, huaiquil’
75	wəf	wif	‘adj. largo, derecho’
76	wəfkə	wifkə	‘v. buscar la manutención’
77	wəlu	wilu	‘v. formar enjambres’
78	wəlwij	wiɽwij	‘v. tener diarrea’
79	wəɽŋo	wiɽŋo	‘s. cierta planta’
80	wəntu	wintu	‘adj. viejo’
81	wəfupəl	wifupəl	‘s. matapijos’
82	wəṭṭə	wiṭṭə	‘s. cuchara’
83	wəṭṭu	wiṭṭu	‘v. fluir (agua)’
84	wiɽə	wiɽi	‘v. rayar, dibujar, surcar’
85	wiɽəlkoŋ	wiɽiɽkoŋ	‘s. golondrina’
86	θiwəɽ	θiwil	‘v. revolver’
			Total: 86

3.2. Alternancia /ə/ > /u/

	RIV: /ə/	RIT: /u/	Definición
1	ṭṭəmiŋ	θumiŋ	‘s. oscuridad’; ‘adj. oscuro’
2	əwa	uwa ~ wa	‘s. maíz’
3	fəkup	fukup	‘s. una enfermedad; parte del cuerpo’

4	jəɲuf	juɲuf	‘v. deslizarse’
5	kawpə	kawpu	‘v. raer’
6	kəfəɫ	kufəɫ	‘v. tartamudear’
7	kəɫum	kuɫum	‘v. lavarse la cara’
8	kənuɲ	kunuɲ	(P.) ‘v. trotar’
9	kəw	kuw	‘v. ir, proseguir’
10	komə	komu	‘v. mirar con agrado’
11	kufə	kuθu	‘v. acostarse’
12	manpə	mampu	‘v. hacer caricias con la mano’
13	məfəɫ	mufəɫ	‘agitar con la mano’
14	məməl	muməl	‘v. formar, instruir’
15	məɲa	mɲa	‘intr. caber’.
16	məpi	mupi	‘v. cree’
17	mət̪sɯɫ	mut̪sɯɫə	‘v. tropezar’
18	nəkə	nəku	‘v. guarecerse de la lluvia’
19	pəf	puf	‘v. saltar’
20	pəlul	pulul	‘s. hueco’
21	pəɲku	puɲku	‘s. bulbo’
22	ɫəfə	ɫəfu	‘v. trasladar líquidos’
23	ɫəku	ɫuku	‘s. pecho’
24	ɫət̪sɛ	ɫut̪sɛ	‘v. pellizcar’
25	təmpaj	tumpaj	‘v. girar’
26	təmu	tumo	‘s. pata (animal), corvejón abajo’
27	təpu	tupu	‘s. adorno corporal’
28	t̪əɫ	t̪uɫ	‘s. colgajo’
29	t̪səɫmə	t̪səɫmu	‘v. filtrarse (agua)’
30	wafəm	wafum	‘s. colmillo’
31	wemə	wemu	‘v. repeler’
			Total: 31

3.3. Alternancia /ə/ > /e/

	RIV: /ə/	RIT: /e/	Definición
1	aɫkən	aɫken	‘s. baja (mar)’
2	ajməɲ	ajmeɲ	‘adv. algo, un poco’
3	eŋələlə	eŋələl	‘v. llamar (la yegua)’
4	kʌfə	kʌfe	‘v. rozar (bosques)’
5	kətʃu	ketʃu	‘s. cierta ave, el quetro’
6	l̥əm	lem	‘v. relumbrar de repente’
7	məkəf	məkeθ	‘v. apretar (el techo)’
8	məfəɫ	mefəɫ	‘v. agitar con la mano’
9	pəθkol	peθkol	‘s. cierto tipo de lluvia’
10	ɫəkən	ɫeken	‘v. hincharse’
11	ɫutʃə	ɫutʃe	‘v. impeler, empujar’
12	wəf	wef	‘v. aparecer, surgir’
			Total: 12



3.4. Alternancia /ə/ > /a/

	RIV: /ə/	RIT: /a/	Definición
1 ⁺¹	anənəj	ananaj	‘interjección de dolor’
2	ewən	ewan	‘v. tener arcadas’
3	ɫəpəθ	ɫipaθ	‘v. echarse (aves)’
4	ɲətʃu	ɲako	‘v. juntar, cerrar; apretar’
5	nəɫkə	nəɫka	(H.) ‘v. recalar (el agua) en el suelo’
6	pəkə	paka	(P.) ‘v. estar preñado el trigo’
7	pənə	panə	‘v. elevarse (un pájaro)’
8	tʃoŋkə	tʃoŋka	‘v. brincar, cabecear’
9	wilkəɫə	wilkəɫ	‘v. tocarse’; ‘adj. torcido’
10	θeɫkəntu	θeɫkantu	‘s. grasa del caldo’
			Total: 11

3.5. Alternancia /ə/ > /o/

	RIV: /ə/	RIT: /o/	Definición
1	əfələ	ofəl	‘v. sorber’
2	kəŋoj	koŋoj	‘s. filandro (mamífero)’
3	λaŋkə	λaŋko	‘v. caerse’
4	məlpə	molpə	‘v. defenderse, espantar insectos’
5	mətʃə	mətʃo	‘v. tr. Sacudir’
6	ɹenʃə	ɹenʃo	‘adj. contrapesado’
7	ʃəŋkaj	ʃoŋkaj	‘adj. circular’
8	wəλo	oλo	‘v. arrollar’
			Total: 8

3.6. Alternancia /ɬ/ > /l/, /ɲ/, /ŋ/ (en contexto asimilatorio), ∅

	RIV /ɬ/	RIT /l/	Definición
1	ʃiɬʃan	ʃiltʃan	‘v. macollarse’
2	iλhaw	ilhaw	‘adj. algo dulce’
3	kajɬi	ʃajli	‘v. dejar los huesos (sin la carne)’
4	koλoŋ	koloŋ	‘s. El disfrazado’
5	kowəλ	koəl	‘s. La fruta del nüpu foki, los cóguiles’
6	kəλeŋ	kəlen	‘v. cierta ave’
7	kəλpiŋ	kəlpin	‘v. amarrar un peso’
8	kujil	kujil	‘pref. sexto grado de ascendencia’
9	λaka	lakan	‘v. partir en el medio’
10	λam	lam	‘s. hermano, hermana’
11	law	law	‘v. pelar’; ‘v. perder el pelo’
12	lawtʃa	lawtʃa	‘s. ratoncillo, laucha’
13	λoŋo	loŋo	‘s. vaso tejido de quila’
14	λəfke	ləfke	‘s. Panguí. El relámpago’
15	ŋejkəλ	ŋejkəl	‘v. mecer, tabalear’
16	maɬtʃe	maɬtʃe	‘s. pescado cauque’
17	meλə	mel	‘v. hacerse cola (pasta)’
18	məλaf	mələf	‘v. aplastarse parte del cuerpo entre dos durezas’

19	məɫfi	məɫfi	‘s. varilla larga’
20	paɫ	paɫ	‘s. las siete cabrillas (constelación astral)’
21	piɫfə	piɫfə	‘v. un tipo de boqui’
22	piɫkaθ	piɫkaθ	‘v. descuartizar’
23	piɫkeɲ	piɫken	‘s. vestido’
24	piɫmajkeɲ	piɫmajken	‘s. golondrina’
25	poɫwiɲ	poɫwiɲ	‘s. renacuajo’
26	pəɹəɫ	pəɹəɫ	‘adj. negro y blanco’
27	ʃofəɫ	ɹofəɫ	‘v. abrazar’
28	ʃoɫki	θoɫki	‘v. desgranar’
29	əɫɲa	əɫɲa	‘v. cerner el pájaro’
30	əɫɲəθ	əɫɲəθ	‘s. coyunturas de los huesos’
31	əwəɫ	əwəɫ	‘v. jugar con juguete’
32	jiwəɫ	jiwəɫ	‘s. lanzadera’
33	paɫi	paɫi	‘s. tobillo, bola del palín’
	/ɫ/	/ɲ/	
34	ɫaɫi	ɲaɫi	‘v. visitar’
	/ɫ/	∅	
35	kufəɫ	kufə	‘v. caldear’
	/ɫ/	/ɲ/ (contexto asimilatorio)	
36	kuɫ	kuɲ(kuɫ)	‘s. cierto instrumento’
Total: 36			

3.7. Alternancia /ɫ/ > /l/, /n/

	RIV	RIT	Definición
	/ɫ/	/l/	
1 ⁺	ɫoɫon	lolon	‘v. provocar vómito con los dedos’
2	ɫoj	loj	‘adj. tonto.’
3	ɫəkan	ləkan	‘s. bazo, pajarilla’
4	məɫo	məlo	‘s. machorra, mujer estéril’

5	ə̃nel̃	ənel	‘s. sabor, aliño’
	/l̃/	/n/	
6	l̃el̃	nel̃	‘v. dejar ir’
			Total: 7

3.8. Alternancia /ɲ/ > /n/, /ɲ/, /l/

	RIV	RIT	Definición
	/ɲ/	/n/	
1	mə̃l̃pɯ̃ɲ	mə̃l̃pɯ̃n	‘s. hollín’
2	ənə̃ɲ	ənən	‘v. estar con las alas extendidas’
	ɲ/	/ɲ/	
3	ɲape	ɲape	‘adj. lerdo, perezoso’
4	wɛ̃ɲi	wɛ̃ɲi	‘s. muchacho’
	/ɲ/	/l/	
5	ə̃nel̃	ənel	‘s. sabor, aliño’
			Total: 5

3.9. Alternancia /ɲ/ > /f/, /m/, /n/, /ɲ/, /p/, /j/, ∅

	RIV	RIT	Definición
	/ɲ/	/f/	
1	alɔ̃ɲ(ə)	alɔf(ə)	‘v. relumbrar’
2	ʈsə̃ɲo	ʈsəfo	‘v. quebrar, quebrarse’
	/ɲ/	/m/	
3	ajɔ̃ɲ	ajom	‘v. estar muy claro (el sol)’
4	ləñfə	ləmfə	‘s. raíz del cochayuyo’
5	əl̃ʈsə̃ɲ	əl̃ʈsə̃m	‘v. atrapar con la boca y tragar o morder’
	/ɲ/	/n/	

6	kunajkijɨ	kunajkijɨ	‘adj. color de guanaco’
7 ⁺¹	ʈsɔŋʈsɔŋ	ʈsɔŋʈsɔŋ	‘s. bolsa’
	/ŋ/	/ɲ/	
8	ɲifəf	ɲifəf	‘v. coser’
	/ŋ/	/p/	
9	wiŋəɫ	wipəɫ	‘v. poner en fila’
	/ŋ/	/j/	
10	ŋəŋəw	ŋəjəw	‘v. pujar (al defecar, eruptar)’
	/ŋ/	∅	
11	piŋlijɨ	pilijɨ	‘s. helada’
<i>Contextos asimilatorios</i>			
12	mɨɲku	mɔŋku	‘adj. globo, entero, por todas partes’
13	mankə	maŋkə	‘v. patear’
14	ɲomke	ɲoŋke	‘s. sobrecubierta’
15	pɛnka	pɛŋka	‘v. el zapallo’
16	wɛnkə	wɛŋkə	‘s. romerillo (arbusto)’
			Total: 17

3.10. Alternancia /ʈs/ > /tʃ/, /t/, /m/, /l/, /f/, ∅

	RIV	RIT	Definición
	/ʈs/	/tʃ/	
1	kəlʈsɨf	kəlʈsɨf	‘v. mojarse mucho’
2	kəʈsəŋ	kəʈsəŋ	‘s. manojo, atado’
3	ɫafəʈsa	ɫafətʃa	‘s. bicho, sapo, insecto’
4	ɫotʃo	ɫotʃo	‘v. aflojarse’
5	miʈsiku	mitʃiku	‘v. fugarse’
6	məʈsəw	məʈsəw	‘v. andar allí y allá alumbrando con tizón’

7	patʃəə	patʃə	‘v. sembrar esparciendo’
8	ʃaltʃal	taltal	‘v. roncar, estar ronco’
9	ʃatʃakə	tʃatʃakə	‘v. cacarear’
10	ʃaf	tʃaf	‘adv. junto, unido, al lado’
11	ʃana	tʃana	‘v. golpear; desnudar’
12	ʃikəm	tʃikəm	‘s. punta de fierro’
13	ʃəko	tʃəko	‘s. ovillo’
14	ʃəʎfu	tʃəʎfu	‘v. estar fatigado’
15	ʃəmpol	tʃumpol	‘v. envolver, arrollar’
16	ətʃəf	ətʃəf	‘s. un pájaro’
17	wantʃo	wantʃo	‘v. caérse algo’
	/ʃ/	/t/	
18	keltʃa	kelta	‘v. cavar con azadón’
19	kotʃəɫ	kotə	‘v. tostar’
20	kəltʃo	kəltə	‘v. aflojarse’
21	kəltʃə	kəltu	‘v. elevarse en el aire’
22	ʃaj	taj	‘salto de agua, catarata, cascada’
23	ʃelponɔŋ	telponɔŋ	‘s. trote’
24	ʃeləŋ	teləŋ	‘v. estar vacío, aclarado’
25	ʃej	tej	‘v. derrumbarse’
26	ʃəŋkəl	təŋkəl	‘v. enrollar (cordel)’
27	ʃəɫku	təɫku	‘v. pelarse’
	/ʃ/	/m/	
28	ʃajlef	majlef	‘v. despejarse’
29	entʃif	empif	‘v. andar de puntillas’
	/ʃ/	/l/	
30	ʃəkeɫ	ləkeɫə	‘v. torcer los ojos con rabia’
	/ʃ/	/f/	
31	ʃeltʃonɔŋ	ʃelfonɔŋ	‘adj. gigantesco’

	/tʂ/	∅	
32	peltʂə	pelə	‘v. empujar algo’
			Total: 32

3.11. Alternancia /θ/ > /s/, /tʃ/, /n/, /j/, /l/, /f/, /p/, /p/, ∅

	RIV	RIT	Definición
	/θ/	/s/	
1	θetʃu	setʃu	‘v. apuntalar, apoyar’
2	θitʃo	sitʃu	‘v. indicar con el dedo’
3	θonə	sonə	‘s. La arruga’
4	θətʃiŋ	sətʃiŋ	‘s. rayo del sol’
	/θ/	/tʃ/	
5	θiwəl	tʃiwəl	‘v. revolver’
6	piŋθa	piŋtʃa	‘s. picaflor’
7	ŋəθal	ŋətʃal	‘v. instigar, incitar, azuzar’
	/θ/	/n/	
8	θonə	nonə	‘s. La arruga’
9	paθə	panə	‘v. cargar al hombro’
	/θ/	/j/	
10	θiwliŋ	jiwliŋ	‘s. moscardón, abejón’
11	weθa	weja	‘adj. malo’
	/θ/	/l/	
12	kujpəθ	kujpəl	‘v. chamuscar’
13	pəθwiŋ	pilwiŋ	‘s. cierto animal’
	/θ/	/f/	
14	θi	fi	‘v. alcanzar’
15	θew	few	‘v. concluirse (un trabajo)’

	/θ/	/ɲ/	
16	θəmi	ɲəmi	‘v. bordar’
	/θ/	/p/	
17	θawəl	pawəl	‘adj. estancado (agua)’
	/θ/	∅	
18	mentʃoθ	mentʃo	‘v. arrojar botones (planta)’
19	pənaθ	pəna	‘v. pegar’
20	ɲəɲkuθ	ɲəɲku	‘v. estar muy tupidas’
21	koɾəθ	koɾə	‘v. tenderse boca abajo’
			Total: 21

3.12. Otras alternancias

	RIV	RIT	Definición
	/ɹ/	/l/	
1	maɹaɲ	maləɲ	‘v. tener doblado para atrás (cabeza)’
2	əfəɹ	əfəl	‘v. sorber con los labios’
3	ʃsəɹa	ʃsəla	‘v. partirse, rajarse’
4	wəɹwan	wəlwan	‘s. vaho’
	/ɹ/	∅	
5	əɲaɹ	əɲa	‘v. comer con ruido’
6	ɹəɲkoɹ	əɲkoɹ	‘s. talón, calcañar’
	/ɹ/	/j/	
7	ɹawilma	jawilma	‘s. cierta ave’
	/ʃ/	/tʃ/	
8	ʃoɹpiw	tʃoɹpiwə	‘v. cantar la diuca’
	/t/	/tʃ/	
9	foɹəm	foɹtʃəm	‘s. hijo’
			Total: 9

3.13. Alternancia en el posmargen

3.13.1. Alternancia en el posmargen /ɲ/ > /n/, /m/, /ŋ/ (en contexto asimilatorio), ∅

	RIV	RIT	Definición
	/ɲ/	/n/	
1	θəkɔɲ	θəkɔn	‘s. coronta’
2	θəɲiɲ	θəɲin	‘s. cejas’
3	kawpəɲ	kawpən	‘s. cebada’
4	kəlo.ɲiɲ	kəlo.ɲin	‘v. roncar durmiendo’
5	kəɲke	kənke	‘s. cierta planta’
6	mɛɲki	menki	‘s. penacho (de aves)’
7	ɲaɲki	ɲaɲki, ɲanki	‘s. cierta planta’
8	pewɛɲ	pewen	‘s. araucaria’
9	piɭaɲ	piɭan	‘s. cualquier volcán’
10	piɭmajkɛɲ	piɭmajken	‘s. La golondrina’
11	piɲliɲ	piɲlin	‘s. helada’
12	ɹəʈʂiɲ	ɹəʈʂin	‘s. prendedor de palito o espino’
13	ʈʂaɲ	ʈʂan	‘v. caerse (quien está en pie)’
	/ɲ/	/ŋ/	
14	ətʃaɲ	ətʃaɲ(kə)	‘v. mascar y chupar el jugo’
15	pəɲku	pəɲku	‘s. bulbo’
	/ɲ/	/m/	
16	jaɲ	jam	‘v. tener miedo, temer’
	/ɲ/	∅	
17	awkiɲ	awki	‘v. hacer eco, resonar’
18	əkɔɲ	əku	‘v. plañir’; ‘v. producir ruido de repugnancia’
19	umaɲ	uma	‘v. alojarse’
20	ʈʃiɲɥe	ʈʃiɲe	‘s. chingue’
			Total: 20

3.13.2. Alternancia en el posmargen /f/ > /m/, /l/, ∅

	RIV	RIT	Definición
	/f/	/m/	
1	ləf	ləm	‘v. quemarse, arder’
2	iɫaf	iɫam	‘v. despreciar’
3	iʈsof	iʈsom	‘v. entrársele a uno polvo’
	/f/	∅	
4	lelif	leli	‘v. mirar, abrir los ojos’
5	ɹuef	ɹue	‘v. empujar algo’
	/f/	/l/	
6	əkəf	əkəl	v. estar encajado’
			Total: 6

3.13.3. Alternancia en el posmargen /ɥ/ > /w/, /l/

	RIV	RIT	Definición
	/ɥ/	/w/	
7	θaɥɫu	θawɫu	‘camarón de los esteros’
8	leɥpe	lewpe	‘s. callana’
9	ɫaɥ	ɫaw	‘v. brindar’
	/ɥ/	/l/	
10	kəɫeɥ	kəɫel	‘declive, cuesta’
			Total: 4

3.14. Fluctuación /o/ ~ /u/

	Selección	Alternativa	Definición
	/o/	/u/	
1	θit̃fo	θit̃fu	‘v. indicar’
2	koṭṣəɫ	kuṭṣəɫ	‘adj. tostado’
3	ɲoɫu	ɲuɫu	‘v. cambiar de cuero (las culebras)’
4	peθkol	peθkul	‘s. tempestad’
5	ɹəɲo	ɹəɲu	‘s. harina cruda’
6	foɲi	fuɲi	‘s. arveja’
7	t̃soɲa	t̃supa	‘s. pinta, mancha’
8	t̃səko	t̃səku	‘v. hacer ovillo’
9	əɲo	əɲu	‘v. cariarse’
10	weɲo	weɲu	‘s. sencillo (moneda)’
	/u/	/o/	
1	kua	koa	‘v. bramar’
2	mət̃suɫ	mət̃soɫə	‘v. tropezar’
3	ɲaθu	ɲaθo	‘s. cuñada de una mujer’
4	pəθu	pəθo	‘s. pudú, venado’
5	ɹut̃sa	ɹot̃sa	‘s. cierta medida’
6	təmu	təmo	‘s. pata (animal), corvejón abajo’
7	pətu	pəto	‘v. beber’; ‘v. tomar agua’
8	weku	weko	‘s. tío materno.’
			Total: 18

3.15. Otras fluctuaciones vocálicas

	Selección	Alternativa	Definición
	/a/	/e/	
1	aɫa	eɫa	‘adj. bonito’
2	apew	epew	‘s. cuento’
3	θefa	θefe	‘s. ballico, cierta planta’
4 ⁺¹	ɫawɫaw	ɫewɫew	‘s. fruto del roble’

	/a/	∅	
5	aɲaj	ɲaj	‘part. expresión enfática’
6	wəθa	wəθ	‘v. separarse’
	∅	/a/	
7	fən	fəna	‘s. tendón; nervio’
	/a/	/o/	
8	ətʃaɲ	ətʃoɲ	‘v. mascar y chupar jugo’
	/e/	/a/	
9	koλελα	κολαλα	‘s. hormiga’
10	ləɲeɹma	ləɲaɹma	‘s. muela del juicio’
	/e/	/i/	
11	ʃsajlef	ʃsajlif	‘v. despejarse’
	/e/	∅	
12	ʃsoɲe	ʃsoɲ	‘adj. denso, tupido’
	/i/	/e/	
13	θətʃiɲ	θətʃeɲ	‘s. rayo del sol’
14	witʃaɹə	wetʃaɹ	‘v. rasgarse; rasgar’
	∅	/i/	
15	pətiw	pətiwi	‘s. chispa’
	∅	/o/	
16	waɹ	waɹo	‘v. estar al rededor’
			Total: 17

3.16. Fluctuación /l/ ~ /ɮ/ ~ /θ/

	Selección	Alternativa	Definición
	/l/	/θ/	
1	kajpəl	kajpəθ	‘resbaloso’
	/l/	/ɮ/	
2	kəɫwi	kəɮwi	‘porotos’
3	ɫəm	ɮəm	‘relumbrar, producir reflejos’
4	mewɫen	mewɮen	‘s. torbellino’
5	məl	məɮ	‘v. tiritar, temblar’
6	pajɫa	pajɮa	‘adv. de espalda’
7	əɫɲu	əɮɲu	‘v. tener repugnancia a la comida’
			Total: 7

3.17. Fluctuación /ɟ/ ~ /ʃ/

	Selección	Alternativa	Definición
	/ɟ/	/ʃ/	
1	fu.ɟi	fufi	‘s. espalda; parte posterior’
2	fəɟkə	fəʃkə	‘adj. fresco’
3	ka.ɟə	kaʃə	‘adj. verde; crudo; no maduro’
4	ma.ɟi	maʃi	‘num. diez’
5	na.ɟki	naʃki	‘s. gato’
6	pe.ɟ	peʃ	‘s. borra de la chicha’
7	pe.ɟkiɲ	peʃkiɲ	‘s. barbas de ciertas plantas’
8	pə.ɟəɮ	pəʃəɮ	‘adj. negro y blanco’
9	ɟuka	ʃuka	‘s. casa’
10	ɟəmə	ʃəmə	‘s. una hierba’
11	ɟəɲa	ʃəɲa	‘v. cavar’
12	ɟəɲe	ʃəɲe	‘v. ser grueso’
13	ɟəɲə	ʃəɲə	‘adj. molido; v. moler’
14	wi.ɟka	wiʃka	‘s. pinta blanca’
	/ʃ/	/ɟ/	

15	piʃə	piɾə	‘v. piar’
16	ʃofəɫ	ɾofəl	‘v. abrazar’
			Total: 16

3.18. Fluctuación /θ/ ~ /ɾ/

	Selección	Alternativa	Definición
	/ɾ/	/θ/	
1	kəŋeɪ	kəŋeθ	‘v. restregarse (el chancho)’
2	kəɾpu	kəθpu	‘v. meter en el escote’
3	kuɪə	kuθə	‘adj. negro’
4	ɫufəɪ	ɫufəθ	‘s. bagre’
5	mukəɪ	mukəθ	‘adj. amargo’
6	napuɪ	napəθ	‘s. nieve’
7	ŋəɫko	ŋəθko	‘v. agolparse’
8	əɾəm	əθəm	‘s. encías’
9 ⁺¹	kuɪkuɪ	kuθkuθ	‘s. nido (cuadrúpedos)’
	/θ/	/ɾ/	
10	tʃəŋkəθ	tʃəŋkəɪ	‘adj. circular, redondo’
11	ɾaθal	ɾaɪal	‘s. radial, nogal’
12	wəθa	wəɪa	‘adj. mal, malo’
			Total: 13

3.19. Fluctuación /θ/ ~ /ʃ/

	Seleccionada	Alternativa	Definición
	/θ/	/ʃ/	
1	aθ	aʃ	‘s. exterior (forma, color, etc.)’
2	tʃaθi	tʃaʃi	‘s. sal’
3	tʃiθap	tʃiʃa	‘s. greda, lama’
4	θaɪɫu	ʃaɪɫu	‘s. camarón de estero’
5	θaŋe	ʃaŋe	‘s. nido’
6	θawe	ʃawe	‘s. quinua’

7	θitʃo	ʃitʃo	‘v. indicar’
8	θiwəλ	ʃiwəλ	‘v. revolver’
9	θoλəm	ʃoλəm	‘s. choritos’
10	θətʃo	ʃətʃo	‘s. una planta’
11	kiθu	kifu	‘adj. mismo’
12	kəθijɪ	kəʃijɪ	‘s. placenta’
13	kupeθ	kupeʃ	‘v. aplicar calor con las manos’
14	λumeθ	λumeʃ	‘v. ser sabroso’
15	maθi	maʃi	‘s. cierta planta’
16	maŋkaθ	maŋkaʃ	‘v. montar en ancas’
17	meθkeɪ	meʃken	‘adj. molido’
18	mejuθ	mejuʃ	‘adj. liso’
19	muθaj	muʃaj	‘s. una bebida’
20	ŋəfaθ	ŋəfaʃ	‘adj. blando (cama, postema)’
21	piθoj	piʃoj	‘s. tortera del huso’
22	weθa	wəʃa	‘adj. malo’
23	waθkə	waʃkə	‘v. hervir’
	/ʃ/	/θ/	
24	kuʃe	kuθe	‘adj. anciana’
25	λuʃu	λuθu	‘adj. recién nacido’
26	meʃa	meθa	‘v. regalar’
27	ŋiʃəʃ	ŋiθəʃ	‘v. coser’
28	paɪuʃ	paɪuθ	‘adj. suave, blando (de ropa)’
29	peʃa	peθa	‘v. encontrar una cosa perdida por otro’
30	piʃko	piθko	‘v. colchar riendas’
31	piʃku	piθku	‘adj. cocido en olla’
32	ʃiɲe	θiɲe	‘v. deslizarse, correrse’
33	ʃiweɪ	θiweɪ	‘s. compañero; hongo’
34	ʃoλki	θoλki	‘v. desgranar, descascarar’
35	ʃumpa	θumpa	‘v. andar a tientas’
36	ʃəwmeɪ	θiwmeɪ	‘s. cierto insecto’
			Total: 36

3.20. Fluctuación /θ/ ~ /f/ (en posmargen), /ɬ/

	Seleccionada	Alternativa	Definición
	/f/	/θ/	
1	məkəf	məkeθ	‘v. apretar (el techo)’
	/ɬ/	/θ/	
2	piwkəɬ	piwkəθ	‘v. dar vuelta (como a un tornillo)’
			Total: 2

3.21. Fluctuación /tʃ/ ~ /t/

	Selección	Alternativa	Definición
	/tʃ/	/t/	
1	tʃaw	tao ~ tʃao	‘s. padre’
2	tʃe	te	‘s. gente’
3	tʃi	ti	‘art. el, la’
4	tʃu	tu	‘demos. cómo’
	/t/	/tʃ/	
5	taji	tʃaj	‘adv. denantes’
6	wente	wentʃe	‘adv. encima’
			Total: 6

3.22. Otras fluctuaciones entre fonemas consonánticos

	Seleccionada	Alternativa	Definición
	/m/	/l/, /n/, /ɲ/	
1	ɬəmu	ɬəlu	‘v. zambullirse’
2	ɲomke	ɲonke	‘s. sobrecubierta de la cama’
3	mipoko	ɲipoko	‘s. salteador, bandido’
	/f/	/n/, /p/, /k/	
4	ɬəfe	ɬəne	‘v. hocicar (el chanco)’
5	kalofətʃa	kalopitʃa	‘s. sanguijuela’
6	nəwfəl	nəwkəl	‘v. pasar algo extendiendo la mano’

	/ɲ/	/n/	
7	ɲaw	naw	‘v. maullar’
8	ʈsajɲu	ʈsanu	‘s. abarca, ojota’
	/k/	/p/, /tʃ/	
9	kəlke	kəlpe	‘adj. medio despertar’
10	kajɫi	ʈfajli	‘v. dejar los huesos’
	/w/	/o/, /u/	
11	ʈfaw	tao	‘s. padre’
12	kujwan	kujuan	‘v. echar vaho’
	/n/	/m/, /ɲ/	
13	məɫfen	məɫfem	‘s. el rocío’
14	pinu	piɲu	‘v. escamarse, dar cuero, salir como caspa’
	/p/	/f/, /n/	
15	pəɫo	fəɫo	‘s. nudo’
16	ʈsupɛf	ʈsunɛf	‘v. espantarse’
	/j/	/m/	
17	əj	əm	‘s. nombre’
	/l/	/n/	
18	maɲɛl	maɲɛn	‘adj. convidado (a comer o beber)’
	/ŋ/	/ɫ/, /l/	
19	ɲɛɫi	ɫɛɫi	‘rogar, suplicar’
20	piɲko	pilko	‘s. tubito’
	ʈʂ/	/ɹ/	
21	ʈʂuna	ɹuna	‘s. puñado a dos manos’
			Total: 21

3.23. Otras fluctuaciones con elisión y adición

	Seleccionada	Alternativa	Definición
	/ɥ/	∅	
1	fojɥe	foje	‘s. canelo’
2	kalɥal	kalal	‘s. cierto hongo’
3	ʧsajɥen	ʧsajen	‘s. cascada’
	/j/	∅	
4	θeja	θea	‘s. un familiar’
5	əjnif	ənif	‘v. espigar; v. recoger lo que dejaron otros’
	/m/	∅	
6	maɾum	maɾu	‘s. recompensa, premio’
7	om	oom	‘interj. cierta expresión’
	/n/	∅	
8	nəla	əla	‘v. abrir o abrirse’
9	ənʧsiw	əʧsiw	‘s. quintral’
	/ɲ/	∅	
10	ɲiθo	iθo	‘s. parte donde empieza algo’
11	ɲiθo	iθo	‘v. masturbarse mutuamente’
	/w/	∅	
12	wɛɲkol	enkol	‘v. destroncar’
	/l/	∅	
13	pɯɲʧʊl	pɯɲʧʊ	‘v. ser engañoso, engañar’
	/k/	∅	
14	əwkeɲ	əweɲə	‘v. silbar’
	∅	/j/	
15	ifə	jifə	‘v. agrandarse (masa)’

16	iful	jiful	‘v. aporcar’
17	ifəlkoŋ	jəfəlko	‘s. cierta planta’
18	iwijŋ	jiwijŋ	‘s. grasa, gordura, manteca’
	∅	/n/	
19	entu	nentu	‘v. sacar, quitar’
20	kuʃsu	kuʃsun	‘s. lunar’
21	laʃsa	lanʃsa	‘adj. abultado, grueso’
	∅	/w/	
22	pu	Puw	‘v. llegar allá’
23	ənu	wənu	‘s. trago; v. tener asco’
	∅	/k/	
24	əkəŋ	kəkəŋ	‘s. fugitivo’
	∅	/l/	
25	tʃekəm	tʃelkəm	‘v. guarecerse’
	∅	/ɰ/	
26	maʎen	maʎɰen	‘s. mujer, niña’
			Total: 26

3.24. Fluctuaciones léxicas de patrones fonológicos mayores

	Selección	IPA	Controladas	Definición
1	achüdko	atʃəθko	atʃiɫko	‘v. tener acidez’
2	chedkeñ	tʃeθkeŋ	tʃeθkan	‘s. cierta ave’
3	chidañ	tʃiθaŋ	tʃiʃa	‘s. greda, lama’
4	chiñge	tʃiŋɰe	tʃiŋe	‘s. chingue’
5	chokif	tʃokif	tʃoke	‘v. ser o estar arqueado’
6	chollpiwü	tʃoʎpiwə	ʃolpiw	‘v. cantar la diuca’
7	chüñwad	tʃəŋwaθ	tʃənwaɫ	‘v. estar confuso, enredado’
8	chürwa	tʃəɫwa	tʃinwaɫ	‘v. enredarse, turbarse’

9	demoll	θemoł	moł	‘v. estar desnudo’
10	diwüll	θiwəł	tʃewəł	‘v. revolver’
11	duam	θuam	ʃum	‘s. intención’
12	dükoñ	θəkoŋ	θikon	‘s. coronta’
13	dümillko	θəmillko	ʃimiłko	‘s. La astilla’
14	focho	fotʃo	fɔʃko	‘v. mojarse’
15	fürkü	fəɫkə	fɪʃken	‘adj. fresco’
16	ifülkoñ	ifəłkoŋ	jəfəłkon	‘s. cierta planta’
17	iwüd	iwəθ	ijuwəθ	‘v. envolver’
18	kodwi	Koθwi	koθuj	‘s. tizón’
19	koel	Koel	koʃeel	‘s. mango’
20	kotrür	koʃsəɫ	kotə	‘adj. tostado’
21	kowüll	kowəł	koəl	‘s. cierta fruta’
22	küchü	kətʃə	kətʃoθə, kətʃo.ɪə, kətojə, kətʃoθ, kətʃo.ɪ	‘v. refregar remedios y sacar espuma’
23	kufüll	kufəł	kufəŋ, kofɪŋ, kofeŋ, kufə, kufɪŋ, kufił, kofi	‘v. caldear’
24	külilke	kəlilke	kəlɪkəli	‘s. cierta ave’
25	kultrafü	kultʃafə	kuʃsaf	‘v. sonar fuerte (galope)’
26	kümtrülü	kəmʃsələ	kəməjə	‘v. meter ruido’
27	küngüwma	kəŋəwma	kəŋojkəŋoj	‘s. cierto mamífero’
28	kürakuü	kəɫakuə	kəɫako	‘s. una enredadera’
29	kütrürü	kətʃsəɫə	kikəɫə	‘v. crujir’
30	küyaka	kəjaka	kiaka	‘s. cierto árbol’
31	l’awü	ɫawə	ɫawi	‘v. echarse (el ganado) a la sombra’
32	l’el’	ɫɛɫ	leθkəm	‘v. dejar ir’
33	llashka	ɫaʃka	laɫka	‘s. ciertos animales’
34	medkeñ	meθkeŋ	meʃken	‘adj. moler’
35	müküf	məkəf	məkəθ	‘v. apretar (el techo)’
36	mülleng	məɫeŋ	məɫoθ	‘v. caer de cara’
37	mültrün	məɫʃsən	məɫʃsɪŋ	‘s. catuto’
38	münchu	məntʃu	mentʃuɫ	‘v. caerse (las flores)’
39	muñku	mɪŋku	moŋko	‘adj. por todas partes’

40	n'ay	ɲaj	nejpi	‘v. atar’
41	nangka	naŋka	laŋka	‘v. tener celos el animal’
42	ngeyküll	ɲejkəɻ	ɲejku	‘v. mecer, tabalear’
43	ngüf	ɲəf	ɲətʃi, ɲətʃin, ɲəla- ɻə, ɲəɻin, ɲəɻin, ɲətʃif, ɲətʃaf, ɲətʃaɻ, ɲətʃaw, ɲətʃsi, ɲətʃsiɻ, ɲətʃsiθ, ɲətʃsə, ətʃsa, ətʃsaf	‘adj. obstruido, muy angosto’
44	ngüñkü	ɲəɲkə	ɲəlkeɻə	‘v. refregar, frotar’
45	ngünow	ɲənəw	ɲəno	‘s. buche (aves)’
46	ngüpu	ɲəpu	ɲəpif	‘v. encogerse (articulación)’
47	nüküfü	nəkəfə	ɲikəf	‘v. callarse’
48	nütrü	nətʃə	nətʃəɻ	‘v. forma de comer (animales)’
49	nüwfü	nəwfə	nəwkəl	‘v. pasar extendiendo la mano’
50	paf	paf	pafnu	‘s. postema’
51	paküf	pakəf	pakoθ, pakuɻ	‘adj. prominente’
52	palü	palə	paɻu	‘s. cierta araña’
53	pañüag	paɲəaɻ	paɲə	‘s. resolana’
54	pelleñ	peɻeɲ	peɻuɻen	‘s. verruga’
55	pelü	pelə	pelələ	‘v. empujar algo’
56	perimol	peɻimol	peɻimontu	‘s. mal agüero, presagio malo’
57	pewul’	pewuɻ	peul	‘s. remolino de agua’
58	pilcheñ	piltʃeɲ	piltʃaf, piltʃal	‘v. llevar entre dos’
59	pillkad	piɻkaθ	piltaɲ, pilkaθ	‘v. descuartizar’
60	pillkeñ	piɻkeɲ	pilken	‘s. cierto tipo de vestuario’
61	piltrü	piltʃə	piltʃaw	‘v. ampollarse’
62	pilwi	pilwi	piluj	‘s. la babosa (molusco)’
63	püdu	pəθu	pəʃo	‘s. pudú, venado’
64	püdwĩñ	pəθwiɲ	pilwiɲ	‘s. cierto animal’
65	püna	pəna	pənaθ	‘v. pegar’
66	puñchul	puɲtʃul	puɲtʃu	‘v. ser engañoso, engañar’
67	püñtrüf	pəɲtʃəf	pənəf	‘v. ablandarse’
68	pütiw	pətiw	pitʃəɻken	‘s. chispa’
69	pütrü	pətʃə	pətu	‘adv. en abundancia, mucho’

70	rekel	ɹekel	ɹeje	‘adv. seriamente, en realidad’
71	reyfü	ɹejfə	ɹejpu	‘v. entreverar, mezclar’
72	shifüñ	fəɲ	kuɫa	‘s. especie de fuente de greda’
73	shüñül	ʃəɲəl	ʃɪɲuɫ	‘v. agacharse, encajarse’
74	t’ue	t̪ue	t̪əwe	‘s. tierra, champa’
75	trawaw	t̪sawaw	t̪sajajə, t̪sajaj	‘v. aporrear, apalear’
76	triltrang	t̪sil̪t̪saɲ	t̪si̪t̪sa	‘adj. desnudo, en pelota’
77	trüfül	t̪səfəl	t̪səmfəl, t̪sumfəl, t̪soɲfəl, t̪somfəl	‘adj. entorcido, torcido’
78	trüllurü	t̪səɫuɹə	kəɫuθə	‘v. contraerse y desprende’
79	trümpol	t̪səmpol	t̪ʃumpol	‘v. arrollar (lazos, etc.)’
80	tüway	təwaj	tuaj	‘v. dar vuelta; girar’
81	tüye	təje	tie	‘demos. allá’
82	üchañ	ət̪ʃaɲ	ət̪ʃe	‘v. mascar y chupar el jugo’
83	üllngüd	əɫɲəθ	wilɲəθ, əɲəθ	‘s. coyuntura’
84	üngir	əɲiɹ	əjeɹə, əjeθə	‘v. mostrar los dientes’
85	ünün’	ənən̪	ənəl, ənəf, inəf, ənən, ənəl	‘v. estar con las alas extendidas’
86	üpür	əpəɹ	əpi	‘v. picar (las pulgas, etc.)’
87	ürüm	əɹəm	əθum	‘s. encías’
88	ürüwüll	əɹəwəl	əɹəl	‘s. especie de pato’
89	üwkeñ	əwkeɲ	əɹken	‘v. silbar’
90	uyü	ujə	wijə	‘v. marearse, aturdirse’
91	wangil	wəɲil	wajli	‘v. abundar cosas agudas’
92	wil’i	wiɹi	wilə	‘s. uña (persona, animales)’
93	wil’pü	wil’pü	wilpuθ, wilpəθ	‘v. recocerse (papas, carne, etc.)’
94	willad	wiɹaθ	wilaf	‘v. estar tendida en el suelo una persona muy flaca.’
95	wülü	wələ	wəlwi	‘v. (h.) guiar las arvejas; (p.) tener ya botones las habas.’
96	würwan	wəɹwan	oɹken, wəɹken	‘s el vaho’
97	wütru	wət̪s̪u	u̪t̪s̪u	‘v. derramarse; derramar’
98	yangül	jaɲəl	jaɲoθ	‘v. descompaginarse (huesos)’

4. Reduplicaciones y reiteraciones de patrones fonológicos

4.1. Reduplicaciones

	Entrada	Selección	Definición
1	chañchañ	chañ	‘s. un pasto’
2	charcharü-n	charü	‘v. sonar al inflamarse’
3	cheykülcheykül-nge-n	cheykül	‘v. tener la vista muy viva’
4	chillchill-kiaw-ün	chill	‘v. hacer tintín al andar’
5	chingdechingde-n	chingde	‘v. divergir en todas las direcciones’
6	chongchongü-n	chongü	‘v. tabalear, hacer ruido tocando’
7	chongchongü-n	chongü	‘v. hincharse por golpes, machucarse por bofetadas’
8	chukanchukan	chukan	‘s. melosa (una planta)’
9	chuluchulu-nge-n	chulu	‘v. borbollar’
10	dagdag	dag	‘adv. en todas partes’
11	demkodemko-le-n	demko	‘v. estar duro en parte, p. ej. un jergón mal arreglado’
12	dengkoll dengkoll-ün	dengkoll	‘v. deshacerse (una hinchazón que hay debajo de la piel)’
13	domdomü-tu-n	domü	‘v. rezongar’
14	dufdufü-n	dufü	‘v. enturbiarse’
15	dumdum-ün	dum	‘v. ofuscarse’
16	farfar-küle-n	far	‘v. sudar mucho’
17	fidafida	fida	‘s. paladar duro’
18	fotridfotrid	fotrid	‘s. cierta planta trepadora’
19	fudfud-amu-n	fud	‘v. avanzar el fuego de polvito a polvito en los aserrines’
20	fülfül-küle-n	fül	‘v. estar en efervescencia, espumar’
21	fünufünu-nge-n	fünu	‘v. chispear’
22	fürfür-trem-ün	für	‘v. quedarse atrasado en el crecimiento del cuerpo’
23	fürfürü-n	fürü	‘v. silbar (el viento, la varilla al partir el aire)’
24	füyfüy-ma-w-ün	füy	‘v. no poder sujetar el vientre (el que tiene mucha diarrea)’
25	kadkadeñ	kad	‘s. cierta ave’
26	kafakafa-nge-n	kafa	‘v. andar (el perro) muy cansado con la lengua afuera...’
27	kallekalle	kalle	‘s. cierta planta’
28	karekare	kare	‘s. gallina destinada al sacrificio’
29	karifikarif-künu-n	karif	‘v. hacer sonar algo suavemente’
30	kemukemu	kemu	‘s. palo del rewe’

31	kingkingü-n	kingü	‘v. ponersele colorados los ojos al que empieza a llorar’
32	korekore	kore	‘s. cierta hierba’
33	korkor	kor	‘s. cierta herramienta’
34	koyongkoyong-nge-n	koyong	‘v. girar continuamente’
35	küchüküchü	küchü	‘s. cierta ave’
36	kudallkudall	kudall	‘s. riñones’
37	küdküd-ün, küdküdü-n	küd	‘v. silbar (ratones, culebras)’
38	kufkuf-kiaw-ün	kuf	‘v. caminar con pasos oíbles’
39	kulawkulaw	kulaw	‘s. cierta araña’
40	külengküleng	küleng	‘s. cierta ave’
41	küllengkülleng	külleng	‘s. cierta ave’
42	kullkull	kull	‘s. cierto instrumento’
43	külolkülol	külol	‘s. laringe con tráquea y bronquios’
44	kümumkümum-nge-n	kümum	‘v. hacer mucho ruido’
45	kunawkunaw	kunaw	‘s. cierto alimento’
46	küngoyküngoy	küngoy	‘s. cierto animal’
47	kuningkuning	kuning	‘s. grillo’
48	küñküñ-küle-n	küñ	‘v. formar racimos’
49	kupükupü-tu-n	kupü	‘v. volquearse, revolcarse (las aves)’
50	küshküsh-ün	küsh	‘v. acallar (a la guagua diciéndole küshküsh)’
51	kütreküte	küte	‘s. jabalí’
52	küyküy-küle-n	küy	‘v. estar muy ocupado’
53	l'etegl'eteg-tu-n	l'eteg	‘v. hacer algo precipitadamente’
54	langlangü-n	langü	‘v. hacer estruendo’
55	lapilapi	lapi	‘s. falda vientre de la res’
56	lawlaw	law	‘s. gotera’
57	lengleng	leng	‘s. cráneo’
58	likuliku	liku	‘s. cierto insecto’
59	llafllaf	llaf	‘s. cierta parte del rehue’
60	llamkellamke	llamke	‘s. mariposa’
61	llamllamü-n	llamü	‘v. ondular’
62	llayllay	llay	‘s. rayos del sol’
63	llüfllüf	llüf	‘s. cierto adorno’
64	llüfollüfo	llüfo	‘s. cierta planta’

65	lükishlükish-nie-w-ün	lükish	‘v. coquetear, pololear’
66	lütrünglütrüng-nge-n	lütrüng	‘v. agitarse, moverse continuamente’
67	mafmafü-n	mafü	‘v. lamer’
68	malmal-küle-we-n	mal	‘v. haberse quedado sólo los huesos...’
69	mama-kü-n	ma	‘v. mugir la vaca’
70	marmar	mar	‘adj. peludo’
71	maymay-nentu-n	may	‘v. barrer con ramas el trigo cuando se avienta’
72	maytrumaytru-nge-n	maytru	‘v. patear con manos y pies como un epiléptico’
73	mellmellü-n	mellü	‘v. transformarse en cola (pasta)’
74	melludmellud	mellud	‘s. la manzanilla (del animal), que sirve para hacer queso’
75	meme-ka-n	me	‘v. balar’
76	mewmew-ün	mew	‘v. formar enjambres’
77	momoke-n	mo	‘v. mugir’
78	mongofmongof-nge-n	mongof	‘v. pacer’
79	mormor-ün	mor	‘v. pacer’
80	mül'mül'-nge-we-n	mül'	‘v. haberse quedado trémulo’
81	mulmul	mul	‘adj. papas hervidas con las cáscaras’
82	mungüllmungüll-nge-n	mungüll	‘v. mecer el cuerpo (quien está sentado sin ocupación)’
83	muñmuñ-küle-n	muñ	‘v. tener pelusa (como algunas frutas)’
84	münüymünüy-tu-n	münüy	‘v. menear (la cabeza, p. ej. en señal de consentimiento...)’
85	ñadñad-küle-n	ñad	‘v. tener sarpullido, sarpullirse’
86	ñawñaw-ün	ñaw	‘v. maullar’
87	newnew-ün	new	‘v. hervir, bullir, abundar’
88	ngal'ngal'ü-n	ngal'ü	‘v. crujir al quebrarse’
89	ngalawngalaw	ngalaw	‘s. cierta planta’
90	ngallngallü-n	ngallü	‘v. dolerle a uno como de cansancio (los huesos...)’
91	ngarngarü-n	ngarü	‘v. crujir’
92	ngawngaw-ün	ngaw	‘v. huapi. masticar algo con ruido, como el chanco las papas’
93	ngayngayü-n	ngayü	‘v. relinchar’
94	ngoyngoy-ün	ngoyü	‘v. hacer ruido confuso, mucho griterío (gente reunida, etc.)’
95	ngüchayngüchay	ngüchay	‘s. limpiaplata, cola de caballo (planta)’
96	ngüfüngüfü-nge-n	ngüfü	‘v. ser muy flexible’
97	ngünungünu	ngünu	‘s. intestino grueso’
98	ngushngush-ün	ngush	‘v. toser fuerte y de continuo...’

99	noynoyü-n	noyü	‘v. salir o entrar muchos de prisa por un paso estrecho...’
100	nufnufü-tu-n	nufü	‘v. olfatear’
101	nüminümi-tu-n	nümi	‘v. pestañear, parpadear’
102	nüynüy-el-ün	nüy	‘v. caer uno por uno; esparcir (polvos) como salando algo’
103	omiñomiñ-nge-n	omiñ	‘v. tener los ojos espasmódicamente cerrados’
104	pangpang-küle-n	pang	‘v. ser nueva y cargada de hojas la mata de papas’
105	pawpawweñ	pawweñ	‘s. cierta planta’
106	pellpellü-n	pellü	‘v. angustiarse, consternarse’
107	perper	per	‘s. borra’
108	pinpin-ün	pin	‘v. escamarse, salir como una caspa’
109	piongpiong-nge-n	piong	‘v. punzar continuamente, p. ej. la muela’
110	pishpishü-n	pishü	‘v. piar’
111	piwüllpiwüll	piwüll	‘s. aorta (probabl.)’
112	podüllpodüll-ün	podüll	‘v. deshacerse (una hinchazón subcutánea)’
113	pongpong-kü-n	pong	‘v. hincharse’
114	poñpoñ	poñ	‘s. frisos de la tela que se ven en su revés; (h.) esponja’
115	püllupüllu	püllu	‘s. cierta planta’
116	pünpünü-n	pünü	‘v. aletear’
117	puypuy-küle-n	puy	‘v. correr’
118	püypüy–trape-l-ün	püy	‘v. amarrar (una bestia) dándole todo el largo del lazo’
119	rara-kü-n	ra	‘v. hacer ruido (mar, gente, etc.)’
120	rüfrüf-kiaw-ün	rüf	‘v. roer allí y allá (los ratones)’
121	rümrüm-ün	rüm	‘v. susurrar, hablar (muchos) a media voz’
122	runrun-ün	run	‘v. dar silbidos’
123	shafshaf-amu-n-nge-n	shaf	‘v. arrastrar algo los pies al andar’
124	shayshay-küle-n	shay	‘v. ser bien overo’
125	shillishilli-tu-n	shilli	‘v. trinar (la diuca)’
126	shingshing	shing	‘s. chicharrón’
127	tapitapi	tapi	‘s. cierta hierba’
128	toftof-ün, toftofü-n	tof	‘v. tabalear, hacer ruido tocando’
129	traltralü-n	tralü	‘v. roncar’
130	trangatranga	tranga	‘s. quijada’
131	tratra-kü-n	tra	‘v. cacarear’
132	trawatrawa	trawa	‘s. cierta ave’

133	trawtraw-küle-n	traw	‘v. tener los ojos hinchados y colorados por llorar’
134	trawültrawül	trawül	‘s. cierto sapo’
135	trefrefü-n	trefü	‘v. palpitar, latir fuerte y seguido (el corazón)’
136	truftruf-ün	truf	‘v. hervir a borbotones, borbollar el agua’
137	trukantrukan	trukan	‘s. cierta planta’
138	trutru-ka	tru	‘s. cierto instrumento musical ’
139	trüytrüy-n	trüy	‘v. separar la paja de los granos’
140	truytruy-äpo-n	truy	‘v. colmar’
141	ülläüllä	üllä	‘s. sienes’
142	walwalü-n	walü	‘v. murmurar (corrientes de agua, cascadas, remolinos, tripas)’
143	wanwan-nge-n	wan	‘v. ser poco harinoso (el trigo), algo vacío el grano’
144	wawa-kü-n	wa	‘v. meter mucha bulla (gentío, ranas, etc.)’
145	wawa-n	wa	‘v. empezar a florecer el maíz’
146	wenwen-angka-le-n	wen	‘v. tener diarrea’
147	weyluweylu-n	weylu	‘v. ser disparejo, p. ej. un libro cuando aún no tiene corte’
148	widawida	wida	‘s. cierto pajarito’
149	wil'wil'ü-n	wil'ü	‘v. chorrear, salir goteando’
150	willwill-ün	will	‘v. rajar; marcar animales partiéndoles las orejas’
151	wilulwilul-nge-n	wilul	‘v. hacer mucho clamoreo’
152	win'win'	win'	‘adv. uno por uno, en fila’
153	wüchilwüchil	wüchil	‘s. cierta planta’
154	wulüwulü	wulü	‘s. cierto insecto volador’
155	wüño-llüllü-ngechi-n	llü	‘v. andar a reculadas, p. ej. la máquina del tren’
156	würwür-ün	wür	‘v. gruñir el chanco; silbar o bramar el viento’
157	yachalyachal-nge-n	yachal	‘v. agitarse como una ventana (abierta) al impulso del viento’
158	yafyaf-ün	yaf	‘v. sacudirse de escalofríos’
159	yagyag-küle-n	yag	‘v. estar despejado, aclarado un bosque...’
160	yofyof-püra-pa-n	yof	‘v. elevarse borbollando o espumando’

4.2. Entradas con un patrón fonológico que plasma reiteración de fonemas

	Entrada	Selección	Definición
1	molmol	mol·mol	‘adj. redondo, delgado (palo)’
2	wedwed	wed·wed	‘adj. travieso, desobediente, loco’
3	cheche	che·che	‘s. un familiar’
4	chimchim–ko	chim·chim	‘s. un animal’
5	chiñchiñ	chiñ·chiñ	‘s. un arbusto’
6	choñchoñ	choñ·choñ	‘s. un pájaro’
7	feñfeñ-el	feñ·feñ	‘s. boyante de la red’
8	fuyfuy	fuy·fuy	‘s. cierta planta’
9	ireire	ire·ire	‘s. cierta planta’
10	kadkad	kad·kad	‘s. cierto insecto’
11	kaykay	kay·kay	‘s. cierto animal mitológico’
12	kemkem	kem·kem	‘s. declive’
13	kewkew	kew·kew	‘s. cierta enredadera’
14	kilkil	kil·kil	‘s. cierta ave’
15	kodkod	kod·kod	‘s. cierto animal’
16	kongkong	kong·kong	‘s. buho’
17	kuku	ku·ku	‘s. cierto familiar’
18	külkül	kül·kül	‘s. cierto helecho’
19	kurkur	kur·kur	‘s. guarida de animales’
20	kuykuy	kuy·kuy	‘s. puente’
21	llalla	lla·lla	‘s. cierto familiar’
22	llangllang	llang·llang	‘s. murta’
23	llawllaw	llaw·llaw	‘s. cierta fruta’
24	llinglling	lling·lling	‘s. cierta planta’
25	ñaña	ña·ña	‘s. cierto familiar’
26	ñeñe	ñe·ñe	‘s. cierto familiar’
27	nefnef	nef·nef	‘s. bicho’
28	nganngan	ngan·ngan	‘s. cierta planta’
29	ngünngün	ngün·ngün	‘s. (h.) cierta ave; (p.) cierto personaje histórico’
30	ñoño	ño·ño	‘s. cierto familiar’
31	ñufñuf	ñuf·ñuf	‘s. cierto insecto del tamaño’

32	palpal	pal·pal	‘s. cierto arbusto’
33	pishpish	pish·pish	‘s. cierto pajarito ’
34	pülapüla	püla·püla	‘s. cierta planta’
35	püllpüll	püll·püll	‘s. cierta planta’
36	pülpül	pül·pül	‘s. bocha, bolita’
37	pülpül	pül·pül	‘s. cierta planta’
38	rere	re·re	‘s. cierta ave’
39	rürü	rü·rü	‘s. zancudo (insecto)’
40	tonton	ton·ton	‘s. mariposas nocturnas’
41	towtow	tow·tow	‘s. cierta ave nocturna’
42	traltral	tral·tral	‘s. cierta ave’
43	trengtreng	treng·treng	‘s. cierto cerro’
44	trotro	tro·tro	‘s. culo’
45	trongtrong	trong·trong	‘s. especie de bolsa’
46	wüdwüd	wüd·wüd	‘s. cierta ave’
47	toto-n	to·to	‘v. enterrar (al animal) el puñal en el cogote’



5. Ejercicio que explica la ratio frecuencia observada / frecuencia esperada

Se asumen dos lenguas: X e Y . La lengua X tiene un inventario fonológico de cuatro fonemas hipotéticos: A , B , C , D ; en cambio, la lengua Y tiene cinco fonemas: A , B , C , D , E . Ahora, se obtienen dos corpus de 100 unidades fonemáticas para cada lengua, con una distribución equitativa para efectos de esta explicación.

En el caso de la lengua X , cada uno de sus cuatro fonemas ocurre con una frecuencia absoluta de 25 unidades. En el caso de la lengua Y , cada fonema tiene una frecuencia absoluta de 20 unidades. Al calcularse las frecuencias relativas para cada lengua, estas corresponden —como es de esperarse— a un 25% para cada fonema de la lengua X y un 20% para cada fonema de la lengua Y , como se indica en la tabla a continuación:

Tabla A: Comparación de los fonemas hipotéticos de dos lenguas asumidas para este ejercicio.

Lengua X			Lengua Y		
Fonemas	Frec. Rel.	Ratio Obs./Esp.	Fonemas	Frec. Rel.	Ratio Obs./Esp.
A	25%	1:1	A	20%	1:1
B	25%	1:1	B	20%	1:1
C	25%	1:1	C	20%	1:1
D	25%	1:1	D	20%	1:1
			E	20%	1:1
Total	100%	4:4	Total	100%	5:5

Con el reporte de las frecuencias relativas, podemos comparar las frecuencias fonemáticas de manera interna en cada lengua y de ese manera dimensionar el aporte relativo de cada fonema; sin embargo, no nos permite comparar la frecuencia de A en la lengua X con la de A en la lengua Y , al tener porcentajes disímiles (25% vs. 20%), aun sabiendo que A tiene la misma probabilidad de ocurrir en ambos sistemas fonológicos.

Para eso, reportamos en nuestra investigación un indicador que sea sensible al número de fonemas del inventario en particular, denominado en la literatura, *ratio de frecuencia observada respecto de la frecuencia esperada* (Gordon, 2016). La forma de calcular esta ratio, abreviada *ratio Obs./Esp.*, es dividiendo la frecuencia de observaciones del fonema en cuestión por la frecuencia esperada, bajo el supuesto que cada fonema en la lengua tiene la misma frecuencia de ocurrencia.

En el caso de la lengua *X*, la frecuencia esperada de cada fonema es de 25 fonemas, obtenida al dividir el número total de observaciones por el número de fonemas en el inventario ($100 / 4 = 25$). Luego, dividimos la frecuencia observada de cada fonema (igual a 25, en nuestro ejercicio) por la frecuencia esperada calculada anteriormente ($25 / 25 = 1$), dándonos un ratio de 1:1 para cada fonema de la lengua *X*. En el caso de la lengua *Y*, la ratio también es 1:1, obtenido al dividir $20 / (100/5) = 1$.

Estos ratios de 1:1 nos permiten señalar que la frecuencia del fonema *A* en ambas lenguas es igual a la frecuencia esperada, permitiendo un canon comparativo en ambos sistemas, a pesar del número disímil de fonemas de cada inventario fonemático. De esta manera podemos especificar con mayor seguridad qué fonemas tienen una mayor frecuencia observada entre diferentes estudios, e incluso entre diferentes lenguas, permitiéndonos también comparaciones tipológicas.

