



Universitat Autònoma de Barcelona

**ADVERTIMENT.** L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons:  [http://cat.creativecommons.org/?page\\_id=184](http://cat.creativecommons.org/?page_id=184)

**ADVERTENCIA.** El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons:  <http://es.creativecommons.org/blog/licencias/>

**WARNING.** The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license:  <https://creativecommons.org/licenses/?lang=en>



**Universitat Autònoma  
de Barcelona**

**EL COMPORTAMIENTO FÓNICO DE LOS  
SINOHABLANTES ANTE LAS VOCALES DEL ESPAÑOL:  
EFECTOS DE LA DISTANCIA LINGÜÍSTICA SOBRE EL  
PROCESO DE ADQUISICIÓN**

**ALBA IGARRETA FERNÁNDEZ**

TESIS DOCTORAL DIRIGIDA POR LA  
DRA. DOLORS POCH OLIVÉ

Programa de Doctorado en Filología Española  
Departamento de Filología Española  
Facultad de Filosofía y Letras

Universitat Autònoma de Barcelona  
Bellaterra, junio de 2019



*A mis aïtas.*

*Por enseñarme que se hace camino al andar.*

*Nunca dejéis de bailar.*



## AGRADECIMIENTOS

Cuando pienso en todos estos años me doy cuenta de que una hacer una tesis doctoral es como prepararse para una actuación de *ballet*. La constancia, el esfuerzo y la dedicación son absolutamente necesarias, de la misma manera que es indispensable tener una buena maestra. Pero, además, son muchas las personas que, de una manera u otra, han formado parte de este camino que nos trae hasta el día de hoy.

En primer lugar, me gustaría mostrar mi más sincero agradecimiento a mi directora: Dolors Poch Olivé. Has sido una gran maestra y mentora. Me has guiado, aconsejado y enseñado mucho a lo largo de estos años. Muchas gracias por todo. Esta tesis no habría sido posible de no ser por ti.

Gracias al Departamento de Filología Española de la Universidad Autónoma de Barcelona por concederme una beca predoctoral que me ha permitido ampliar mi labor investigadora y mi experiencia docente, la cual ha sido sumamente enriquecedora.

Querría mostrar mi más sincero agradecimiento a la Dra. Beatriz Ferrús, ya que su generosidad y su apoyo durante todo este tiempo han sido fundamentales. Ha sido un gran placer trabajar estos años teniéndole como guía y referente.

Querría agradecer a la Dra. Gloria Clavería haberme permitido formar parte del *Grup de Lexicografia i Diacronia*, lo que me ha permitido ampliar mi formación como investigadora, y también me gustaría mostrar mi agradecimiento a todos los miembros del grupo de investigación “El español en contacto con el catalán” por todo lo que he aprendido gracias a ellos. Merece una mención especial la Dra. María Machuca, cuyos consejos y sugerencias han resultado muy valiosos.

Gracias a la Dra. Pilar Wang por acogerme como lectora en la Universidad de Ciencia y Tecnología de Macao, lo que me permitió tener mi primer contacto con alumnos orientales y detectar los problemas objeto de estudio de esta investigación. Esta tesis no habría sido posible sin informantes, por lo que quiero dar las gracias a Xiaoqi Zheng, Changchang Liu, Yixin Peng, Yaxuan Liu, Hufanyi Yan y Ruofei Ma por acceder a ser grabadas durante su estancia en Barcelona, y a Ane, Celia, Idoia, Cintya, Sandra y Sara haberme prestado su tiempo y sus voces tan desinteresadamente.

No puedo olvidarme, como es obvio, de los miembros del Seminario de Filología e Informática. Gracias a la Dra. Marta Prat, a la Dra. Cristina Buenafuentes, al Dr. Joan Torruella, a la Dra. Carolina Julià y a la Dra. Laura Muñoz por todos los buenos

momentos, la amistad y el apoyo. En especial, gracias a Natalia Terrón por las charlas al sol, los cafés y las confidencias. Nada hubiera sido lo mismo sin vosotros.

Merece una mención especial Anabel Blasco, del Servicio de Estadística Aplicada de la Universidad Autónoma de Barcelona, por su ayuda y paciencia con los aspectos relacionados con el estudio estadístico.

Gracias a Amaia, Arantza, Clara y Silvia por ser y estar.

Quisiera darle las gracias a Sandra porque, a pesar de la distancia, siempre está presente. Gracias por compartir toda esta experiencia conmigo.

Joseph, no voy a decir nada que no sepas ya. Una de las mejores partes de hacer esta tesis ha sido todo lo que hemos vivido juntos, y creo que todos esos momentos hablan por sí solos. Has sido una parte fundamental durante toda esta etapa. Gracias, de todo corazón, por ser un gran amigo.

Querría dar las gracias a mi familia por los momentos de alegría compartidos, por los ánimos, y por toda la ayuda prestada. Gracias, también, a mi querida Pilar, quien es una más de la familia.

Gracias a mis abuelos y a mi *amuña* porque, aunque no os he podido ver todo lo que me hubiera gustado, siempre os he sentido muy cerca. Os quiero mucho y os debo mucho. A ti también, *aitona*. Sé que me hubieras apoyado como todos.

Quisiera hacer una mención especial a la persona que, desde que apareció, no me ha soltado la mano en ningún momento. Ya sabes que todo lo que escriba se queda corto. Gracias por ser mi compañero, Uri.

A Ane. Además de ser hermana, amiga, apoyo incondicional y parte fundamental de mi vida, siempre estás presente, y haces que todo mejore. Darte las gracias se queda corto para todo lo que me gustaría agradecerte.

Esta tesis se la dedico a mis *aitas*, Bernardo y M<sup>a</sup> Lourdes, porque todo ha sido posible gracias a vosotros. Nunca os perdísteis ninguna actuación, y no os habéis perdido ni un solo momento de este camino que he recorrido.

Gracias al Dr. José Manuel Blecua, la Dra. Vicenta González Argüello y la Dra. María García Antuña por acceder a formar parte del tribunal de esta tesis. Es un gran honor para mí contar con profesionales que son un referente en el ámbito de la Fonética Experimental y la enseñanza de español para extranjeros.

# ÍNDICE

<b>I. INTRODUCCIÓN .....</b>	<b>3</b>
<b>II. MARCO TEÓRICO.....</b>	<b>7</b>
1. La adquisición del sistema fonológico de segundas lenguas.....	7
2. La adquisición del sistema vocálico español por sinohablantes.....	21
3. Recapitulación .....	43
<b>III. METODOLOGÍA Y DISEÑO EXPERIMENTAL .....</b>	<b>46</b>
1. Selección de las informantes .....	46
1.1 Grupo de control.....	47
1.2 Informantes sinohablantes .....	47
2. Diseño del corpus .....	49
2.1 Criterios seguidos para el diseño del corpus .....	49
2.2 Composición del corpus .....	50
2.3 Grabación.....	55
3. Análisis auditivo .....	55
3.1 Problemas preliminares .....	55
4. Análisis acústico .....	57
4.1 Parámetros acústicos analizados.....	57
4.2 Almacenamiento de los datos .....	59
5. Tratamiento estadístico de los datos.....	61
5.1 Validación estadística del análisis auditivo .....	61
5.2 Análisis de la duración vocálica en función de distintas variables.....	62
5.3 El análisis del timbre .....	62
5.4 El análisis del timbre: comparación entre el grupo de control y el grupo de informantes sinohablantes .....	63
<b>IV. RESULTADOS.....</b>	<b>64</b>
<b>Primera parte: Influencia de las variables independientes sobre la realización de las vocales. Procesos de resilabación.....</b>	<b>64</b>
1. La acentuación.....	64
2. Entorno consonántico .....	67
2.1 Entorno anterior.....	67
2.2 Entorno posterior .....	79
2.3 Realización de la vocal en posición final de palabra aislada.....	91



2.4	Recapitulación .....	92
3.	Estructura silábica.....	95
3.1	Estructura CV-CV .....	97
3.2	Estructura CV-CVC.....	99
3.3	Estructura CVC-CV.....	100
3.4	Recapitulación .....	101
<b>Segunda parte: La duración de las vocales .....</b>		<b>104</b>
1.	Grupo de control.....	104
1.1	Duración general.....	104
1.2	Acentuación de la vocal.....	105
1.3	Entorno consonántico .....	107
1.4	Estructura silábica.....	125
2.	Informantes sinohablantes .....	134
2.1	Duración general.....	135
2.2	Acentuación .....	136
2.3	Entorno .....	137
2.4	Estructura silábica.....	155
3.	Comparación: grupo de control e informantes sinohablantes .....	177
3.1	Duración general.....	177
3.2	Acentuación .....	178
3.3	Entorno .....	181
3.4	Estructura silábica.....	203
4	Recapitulación .....	211
<b>Tercera parte: El timbre vocálico .....</b>		<b>213</b>
1.	Grupo de control.....	213
1.1	Características de las vocales .....	213
1.2	Áreas de dispersión.....	219
1.3	Recapitulación .....	241
2.	Informantes sinohablantes .....	242
2.1	Vocal /i/ .....	243
2.2	Vocal /e/.....	247
2.3	Vocal /a/.....	252
2.4	Vocal /o/ .....	256

2.5 Vocal /u/ .....	261
2.6 Recapitulación .....	265
3. Diferencias entre el grupo de control y el de informantes sinohablantes .....	268
3.1 Vocal /i/ .....	268
3.2 Vocal /e/ .....	272
3.3 Vocal /a/ .....	275
3.4 Vocal /o/ .....	279
3.5 Vocal /u/ .....	283
3.6 Recapitulación .....	287
<b>V. DISCUSION DE LOS RESULTADOS .....</b>	<b>290</b>
1. Influencia de las variables independientes sobre la realización de las vocales y procesos de resilabación .....	290
1.1 Acentuación .....	290
1.2 Entorno consonántico .....	290
1.3 Estructura silábica .....	295
2. Duración de las vocales .....	298
2.1. Duración de las vocales en función del acento .....	298
2.2. Duración de las vocales en función del entorno .....	298
2.3. Duración de las vocales en función de la estructura silábica .....	299
3. Timbre de las vocales .....	300
<b>VI. CONCLUSIONES GENERALES .....</b>	<b>305</b>
<b>BIBLIOGRAFÍA .....</b>	<b>309</b>



**CAPÍTULO I**  
**INTRODUCCIÓN**



# I. INTRODUCCIÓN

La presente investigación, enmarcada en el ámbito de la Fonética Experimental aplicada a la adquisición de segundas lenguas, se plantea como objetivo proporcionar datos precisos sobre la adquisición de las vocales del español por sinohablantes y sobre los fenómenos que dicho proceso provoca en el conjunto de la cadena fónica emitida por los estudiantes. Su objetivo último es constituir una contribución al complejo proceso fonético-fonológico que consiste en adquirir la pronunciación de una lengua extranjera.

La enseñanza del español en China se inicia en el año 1953 con la creación del primer Departamento de Español en la Universidad de Estudios Extranjeros de Pekín (Instituto Cervantes, 2000). Dada la inexistencia de recursos bibliográficos, durante los primeros años los profesores emplean un manual de español publicado en Moscú para enseñar español a sinohablantes, y las únicas asignaturas que se imparten son lectura intensiva y gramática (Jingsheng, 2014). En la década de 1990 el número de centros que imparten docencia superior de español en China, de modo consolidado, es ya muy numeroso (Fisac, 2000, Jingsheng, 2014), por lo que es a finales de esta década cuando comienzan a proliferar los estudios que analizan la adquisición del español por chinos.

A lo largo de la historia de la enseñanza de lenguas extranjeras se ha producido una evolución de las teorías sobre la adquisición del sistema fonológico de segundas lenguas que ha corrido paralela a la aparición de las distintas orientaciones sobre el estudio del lenguaje que se han desarrollado durante el siglo XX. Los trabajos de Polivanov y Trubetzkoy han sido los pioneros en este campo. A ellos les ha seguido el Análisis Contrastivo tradicional y, durante la segunda mitad del siglo XX, deben destacarse los modelos elaborados por E. Flege. Sin embargo, tal y como se muestra en esta investigación, los trabajos que analizan la adquisición del español por sinohablantes han adaptado un punto de vista tradicional y han tomado como base el análisis contrastivo para predecir las posibles dificultades de los estudiantes.

En relación con la adquisición de las vocales del español, la mayoría de trabajos indican que, pese a que el chino mandarín posee más unidades segmentales que el español, cinco de ellas coinciden o son muy similares a las de esta lengua. Dado que estas observaciones se realizan desde la perspectiva de la inteligibilidad de los estudiantes, la mayoría de estos trabajos consideran que la correcta realización de las vocales del español

constituye un problema menor habida cuenta de que las vocales del español estarían “contenidas” en el sistema vocálico del chino. Por tanto, son muy pocos los trabajos que se centran en la adquisición del sistema vocálico del español por sinohablantes.

Los estudios existentes toman como referencia la lengua común de China, el *putonghua*, también conocido como *chino estándar*, *chino*, *chino mandarín* o *mandarín*. Investigadores como Ramsey (1987), Pan (2000), Chaoju y Van Heuven (2007, 2008, 2009) y Gong, Chow y Ahlstrom (2010) coinciden en que la lengua china se encuentra dividida entre norte y sur, y que existen siete grandes grupos dialectales, de los cuales el conocido como *mandarín* engloba todas las variedades del norte (Ramsey, 1987: 87-88). Así, el chino toma como estándar de pronunciación el habla de Pekín, la cual se engloba dentro del grupo dialectal mandarín. El chino mandarín es la variedad que se enseña en escuelas y universidades, es la prestigiosa y la que se considera “de cultura” y, por tanto, es la que dominan los sinohablantes. Además, muchos hablantes emplean dialectos propios de su región que tienen un uso restringido al ámbito familiar y no se consideran “de cultura”. Sería impensable que en un país como China todos sus ciudadanos hablaran una única lengua homogénea. Es necesario señalar que, dado el amplio espectro de familias lingüísticas y variedades dentro de los grandes dialectos de China, no hay acuerdo entre los sinólogos en lo que respecta al sistema vocálico del chino estándar, como se puede comprobar en los trabajos de Lin (2007) y, Duanmu (2007). Este autor señala que, mientras que existe un estándar desde el punto de vista gramatical, no ocurre lo mismo en el ámbito de la pronunciación, ya que, tal y como señala,

many speakers in the Mandarin dialect family see little need to modify their pronunciation, because their dialect is often closer to Standard Chinese than what is attempted by people from other dialect families. As a result, most Standard Chinese speakers, or most of those who think they are speaking Standard Chinese, do not have a perfect pronunciation. (Duanmu, 2007: 5)

Este problema es mencionado en los trabajos que analizan la adquisición de las vocales del español por sinohablantes, aunque, como se ha indicado anteriormente, no se considera una cuestión importante porque hay acuerdo, entre todos los estudiosos, en considerar que las vocales del español están “contenidas” dentro del sistema vocálico chino.

Desde los primeros años del siglo XXI, la enseñanza del español a sinohablantes no está circunscrita únicamente a los cursos impartidos en China. Los alumnos de dicho país, por razones socioeconómicas, se interesan por el español sea porque quieren establecer relaciones industriales o comerciales con países de habla hispana, sea porque

quieren dedicarse a la enseñanza del español. Ello ha conducido a que dichos estudiantes estén muy interesados en aprender el español en los países hispanohablantes, especialmente en España, por el prestigio que supone, para un estudiante chino, estar en posesión de una Licenciatura o de un Máster otorgados por una universidad española (Real Instituto Elcano, 2018: 111). En el año 2006, la administración española cifra en 908 los chinos con permiso de estudios en España, mientras que 10 años más tarde, dicha cifra aumenta a 7500 personas con dicho permiso (Real Instituto Elcano, 2018: 111) Así, la llegada a estas universidades de importantes contingentes de estudiantes chinos constituye un importante reto para los profesionales de la enseñanza de la lengua y la cultura hispanas. Los autores de los trabajos de investigación sobre la adquisición del español a los que se ha aludido en los párrafos anteriores son, generalmente, profesores que trabajan en China utilizando la metodología y los materiales de trabajo tradicionales en dicho país. El manual con el que han estudiado español la mayor parte de los alumnos que llegan a las aulas universitarias españolas es el titulado *Español Moderno* (Dong y Liu, 2006), que sigue un enfoque tradicional en la enseñanza de la Gramática y también en la enseñanza de la pronunciación. Los estudiantes sinohablantes, pues, no han estado en contacto con el Enfoque Comunicativo, la orientación metodológica más desarrollada actualmente en Europa en el dominio de la enseñanza de lenguas extranjeras. Esta situación ha puesto de manifiesto la necesidad de desarrollar material didáctico adaptado a la metodología habitualmente empleada en la enseñanza del inglés, del francés, del alemán, etc. Pero, para poder hacerlo, es necesario disponer de información sobre cómo se produce el acercamiento a la lengua española por parte de los sinohablantes y sobre cuáles son las principales dificultades con las que se encuentran tanto en el plano gramatical como en el del aprendizaje del léxico o de la pronunciación.

La presente tesis se propone aportar datos sobre el proceso de aprendizaje de las vocales del español por sinohablantes con el fin contribuir, en primer lugar, al desarrollo de modelos de adquisición adaptados al comportamiento de hablantes que aprenden lenguas cuya estructura está muy alejada de la propia de su lengua base y, en segundo lugar, a la preparación de actividades didácticas que faciliten la adquisición de la pronunciación.

Con el objetivo de alcanzar estos fines, la presente tesis doctoral está estructurada en varios capítulos que se presentan a continuación (en cada uno de ellos, la numeración de figuras, tablas y espectrogramas es independiente de los otros para facilitar la lectura):



En primer lugar, el capítulo dedicado al marco teórico (capítulo II) consta de dos apartados diferenciados. En el primero de ellos se presentan, desde una perspectiva diacrónica, los modelos que ofrecen una explicación al proceso de adquisición general de segundas lenguas o lenguas extranjeras. En el segundo se ofrece una síntesis de los estudios dedicados al análisis de la adquisición del sistema fonético del español por sinohablantes.

El capítulo III está dedicado a la metodología y el diseño experimental. Se describen, en primer lugar, las características de las informantes pues se han estudiado dos grupos de hablantes: el de las hispanohablantes nativas que han funcionado como “grupo de control” y el de las sinohablantes; en segundo lugar, se indican los requisitos establecidos para el diseño del corpus y cuál es la composición del mismo y a continuación se describe cómo se han llevado a cabo las grabaciones. Posteriormente se establecen los dos tipos de análisis llevados a cabo, el auditivo y el acústico y, en último lugar, se explica cuál es el tratamiento estadístico que se ha aplicado a los datos.

Los resultados obtenidos constituyen el capítulo IV, el cual está dividido en tres partes. La primera de ellas muestra la influencia de las variables independientes (la acentuación, el entorno consonántico y la estructura silábica) sobre la realización de las vocales del español por parte de las informantes sinohablantes. La segunda parte recoge los datos obtenidos referentes a la duración de las vocales según las distintas variables señaladas anteriormente. Para ello, se analizan las vocales producidas por el grupo de control en primer lugar, las de las informantes sinohablantes en segundo lugar, y finalmente se comparan ambos grupos. La tercera parte muestra los resultados del análisis acústico del timbre de las vocales del español producidas, en primer lugar, por las hablantes nativas de español, en segundo lugar, por las estudiantes chinas, y en último lugar, se comparan ambos grupos.

El capítulo V está dedicado a la discusión de los resultados obtenidos. En él se contrastan los datos obtenidos en el trabajo con los obtenidos por los otros autores que han estudiado la cuestión y se ponen de relieve las aportaciones realizadas al conocimiento del proceso de adquisición de las vocales del español por estudiantes sinohablantes.

Finalmente, el capítulo VI contiene las conclusiones generales de esta investigación.

**CAPÍTULO II**  
**MARCO TEÓRICO**



## II. MARCO TEÓRICO

### 1. LA ADQUISICIÓN DEL SISTEMA FONOLÓGICO DE SEGUNDAS LENGUAS

Desde comienzos del siglo XX y hasta la actualidad la adquisición del sistema fonológico de segundas lenguas ha despertado el interés de numerosos investigadores, y son muchos los trabajos realizados en torno a esta temática. Desde los estudios pioneros de Polivanov y Trubetzkoy, las investigaciones han pasado por distintas etapas y la evolución teórica ha permitido llegar a los modelos más actuales, surgidos, principalmente, de los trabajos de Flege.

En 1931 Polivanov, el precursor de estos estudios, indica que las representaciones fonémicas de una segunda lengua (L2) se perciben de acuerdo con el sistema de la primera lengua (L1), es decir, que las dificultades en la producción de sonidos de una L2 surgen de la influencia de la estructura fonológica de la L1. Unos años más tarde, Trubetzkoy (1939), otro de los pioneros en este campo, señala que

(e)l sistema fonológico de una lengua es comparable a una criba a través de la cual pasa todo lo que se dice. En la criba quedan únicamente las marcas fónicas que son pertinentes para la individualidad de los fonemas. Todo lo demás cae en otra criba donde a su vez quedan las particularidades fónicas pertinentes en el plano apelativo; más abajo aún se encuentra otra criba en la cual se seleccionan los rasgos fónicos característicos de la expresión del locutor, etc. Las personas se apropian del sistema de su lengua materna, y cuando oyen hablar otra lengua emplean involuntariamente, para el análisis de lo que oyen, la “criba” fonológica que les es habitual, es decir, la de su lengua materna. Pero como esta “criba” no se adapta a la lengua extranjera, surgen numerosos errores e incomprensiones. Los sonidos de la lengua extranjera reciben una interpretación fonológica inexacta debido a que se los ha hecho pasar por la “criba” fonológica de la propia lengua. (Trubetzkoy, 1939: 46)

Por este motivo, el investigador indica que

el llamado “acento extranjero” no depende exclusivamente de que el extranjero no pueda pronunciar un sonido determinado, sino más bien de que no interpreta con corrección dicho sonido. Y esta interpretación errónea está condicionada por la diferencia entre la estructura fonológica de la lengua extranjera y la de la lengua materna del locutor. (Trubetzkoy, 1939: 48)

Esta es la razón por la que la percepción y la producción en la L2 difieren de la de los hablantes nativos de dicha lengua.

Siguiendo a sus predecesores y dentro del ámbito de la Escuela Lingüística de Praga, Jakobson (1941) comienza a hablar e investigar en torno a la importancia de la marcación

en la adquisición de la L2. Esto implica que los sonidos de la lengua son clasificados como no marcados o como marcados. Los primeros se consideran más simples, básicos y naturales que los segundos, los cuales son menos extendidos. Weinreich (1957), por su parte, analiza la transferencia de características de la L1 a la L2 y describe diversos tipos: 1) la sustitución de sonidos, según la cual el estudiante de la L2 emplea el equivalente de la L1 más cercano al sonido de la L2 en cuestión, por ejemplo, el empleo de la vibrante del inglés [ɹ] para la [r] del español en el caso en el que la L1 es el inglés y la L2 el español; 2) los procesos fonológicos, es decir, cuando se emplea un alófono de la L1 que no sucede en el mismo entorno consonántico o vocálico en la L2, como en el caso de la [l] en posición de coda para la [l] velarizada cuando la L1 es el francés y la L2 el inglés; 3) la baja diferenciación (*underdifferentiation*), la cual tiene lugar cuando la L2 tiene distinciones que la L1 no tiene, como en el caso de [d] y [ð] cuando la L1 es el español y la L2 el inglés; 4) la reinterpretación de distinciones, o lo que es lo mismo, reinterpretar características secundarias como si fueran principales o distintivas, como en el caso en el que la L1 es el alemán y la L2 el inglés, y el estudiante interpreta las distinciones tensas/laxas del inglés como largas/cortas; 5) la interferencia fonotáctica (*phonotactic interference*), es decir, realizar la estructura silábica de la L2 conforme con la de la L1, por ejemplo, articular *pic[i]nic[i]* cuando la L1 es el portugués y la L2 el inglés; 6) y, por último, la interferencia prosódica, como en el caso de los hablantes de inglés como L1 y de mandarín como L2 que producen entonación decreciente en vocales en posición final, independientemente del tono en mandarín. Estos trabajos pioneros condujeron al nacimiento de la lingüística contrastiva.

Una de las corrientes más seguidas por los investigadores que han analizado la adquisición del sistema fonético-fonológico de la L2 es la Hipótesis del Análisis Contrastivo (CAH). Esta indica que las diferencias entre la L1 y la L2 y la transferencia de rasgos de la primera a la segunda son fundamentales a la hora de explicar las características del habla en la L2. Lado (1957) es uno de los primeros en hablar de esta hipótesis. El investigador da por sentado que para los estudiantes de una lengua extranjera habrá aspectos de la misma que resulten fáciles de adquirir o aprender, mientras que otros resultarán extremadamente difíciles:

In learning the sound system of a foreign language one finds sounds that are physically similar to those of the native language, that structure similarly to them, and that are similarly distributed. "Learning" of such phonemes occurs by simple transfer without difficulty. On the other hand, one also finds sounds that are not part of the sound system of the native language, that structure differently, or that are differently distributed. Learning of these occurs more slowly, and difficulty with them is more persistent. In fact,

learning of the latter actually means learning the sounds of the language. (Lado, 1957: 12)

Ejemplo de ello es que, de acuerdo con el autor, en español [d] y [ð] son alófonos del mismo fonema, mientras que en inglés son distintos. Según el investigador, esta es la razón por la que percibir y producir el contraste de los mismos en inglés representa un alto grado de dificultad para los hispanohablantes nativos que aprenden inglés.

Lado (1957) indica que durante el proceso de adquisición de una lengua extranjera hay una tendencia a transferir el sistema de la lengua nativa a la lengua que se está aprendiendo. Según el autor, “(w)e tend to transfer to that language our phonemes and their variants, our stress and rhythm pattern, our transitions, our intonation patterns and their interaction with other phonemes” (Lado, 1957: 11). Esta explicación del autor está relacionada con la criba fonológica de Trubetzkoy y con la percepción, ya que, según el investigador,

the speaker of one language listening to another does not actually hear the foreign language sound units - phonemes. He hears his own. Phonemic differences in the foreign language will be consistently missed by him if there is no similar phonemic difference in his native language. (Lado, 1957: 11)

Lado (1957) considera necesario realizar una comparación estructural de los sistemas de ambas lenguas con el fin de identificar los distintos problemas y, así, poder predecir qué aspectos de la L2 serán más difíciles y cuáles serán más fáciles para los estudiantes de esta lengua. Además, el investigador describe cuáles son las etapas a seguir durante el proceso de análisis comparativo de dos lenguas:

(1) Analysis of sound systems. (2) Comparison of units. (3) Location and description of segmental problems. (4) Sequences of sounds. (5) The analysis of sequence problems. (6) Difficulty in pronouncing a phoneme versus difficulty in pronouncing a sequence. (7) Difficulty in pronouncing a word versus difficulty in pronouncing a phoneme or a phoneme sequence. (8) Problems of spelling pronunciation. (9) Pronunciation problems with words that show similarity in two languages. (10) Perception of phonemic contrast through nonphonemic sound features. (11) Perception and production of a phonemic difference through different structural interpretation. (12) Dialect differences and pronunciation problems. (13) Unpredictable alternation between two potential substitutions. (Lado, 1957: 12-27)

Como se puede observar, en esta primera etapa de estudio de la adquisición de segundas lenguas los trabajos de Polivanov (1931), Trubetzkoy (1939), Jakobson (1941), Weinreich (1957) y Lado (1957) sientan las bases sobre las que se sostienen las investigaciones posteriores.

En la década de los 60, los trabajos siguen, generalmente, la Hipótesis del Análisis Contrastivo, incidiendo en que las diferencias entre la lengua materna y la segunda lengua

son las que suponen un escollo para los hablantes al adquirir el sistema fonológico de la segunda. No obstante, tal y como se muestra a continuación, algunos autores comienzan a otorgar mayor relevancia a la adquisición de la L1 para explicar los fenómenos que influyen en la adquisición de la L2.

En esta década, el Centro de Lingüística Aplicada de EEUU, del que forman parte Stockwell y Bowen (1965), presenta una serie de estudios contrastivos que describen las similitudes y diferencias entre el inglés y cada una de las cinco lenguas extranjeras más comúnmente enseñadas en Estados Unidos: francés, alemán, italiano, ruso y español. Los investigadores de dicho centro indican que uno de los mayores problemas que surge en el proceso de aprendizaje de una lengua extranjera es la interferencia causada por las diferencias estructurales entre la lengua nativa del estudiante y la segunda lengua. Es por ello que los autores apuntan que es necesario realizar un análisis contrastivo en detalle de las dos lenguas con el fin de poder dar una mejor respuesta a aquellos problemas que presentan los alumnos. Stockwell y Bowen (1965) indican que la lengua debe considerarse como un conjunto de reglas que especifican si un aspecto de la misma es opcional, obligatorio o nulo, y cuáles son las restricciones de dicha lengua. En relación con el sistema fonético-fonológico de una lengua, un sonido nulo es aquel que no aparece en esta lengua, aquel que no depende del contexto fonológico en el que se encuentra es opcional, y el que está condicionado por el entorno es obligatorio, como es el caso de los alófonos de un mismo fonema. Los investigadores indican que “(i)n a strict sense, these optional and obligatory choices and the restrictions which govern them are unique to each language, but probably no two languages are so completely different that some choices and restrictions are not alike, or at least similar, in them” (Stockwell y Bowe, 1965: 2). En su trabajo estudian la adquisición del español por hablantes nativos de inglés e indican que

(t)he fact that there is substantial similarity in the choices and restrictions of related languages makes Spanish easier for a speaker of English to learn than, say, for a speaker of Chinese. But more to the point, the fact that a large number of the restrictions in English and Spanish are different creates most of the problems in the classroom. (Stockwell y Bowen, 1965: 2)

Es decir, los investigadores apuntan que las diferencias entre lenguas son una mayor fuente de dificultad a la hora de aprender una lengua extranjera, mientras que las similitudes pueden ayudar a los estudiantes a adquirir el sistema fonético de dicha lengua.

Cabe señalar que a partir de la década de los 60 los estudios en torno a la adquisición de la L2 otorgan mayor relevancia a trabajos que analizan la adquisición de

la L1, como las investigaciones de Penfield (1965), Lenneberg (1967) o Chomsky y Halle (1968). Lenneberg (1967: 374 y 404) indica que la función cognitiva subyacente al lenguaje consiste en un proceso de categorización y extracción de similitudes. Además, el autor señala que una teoría de fonética universal puede permitir formular ciertas leyes que rigen las secuencias permitidas en un idioma en concreto. Dichas leyes pueden ser propuestas como leyes de la fonética universal. En cuanto a Chomsky y Halle (1968), los investigadores señalan que

the phonetic representation of an utterance in a given language is a matrix with rows labelled by features of universal phonetics. The grammar of the language assigns to this phonetic representation a “structural description” that indicates how it is to be interpreted, ideally, in this language. (Chomsky y Halle, 1968: 5)

Se observa, pues, que tanto Lenneberg (1967), como Chomsky y Halle (1968) consideran que todas las lenguas están marcadas por la fonética universal.

Weinreich (1968) indica que aquellas desviaciones de la norma que aparecen en el habla de los bilingües son el resultado del contacto de lenguas, y estas las trata como fenómenos de interferencia. De acuerdo con el investigador,

(t)he problem of phonic interference concerns the manner in which a speaker perceives and reproduces the sounds of one language, which might be designated secondary, in terms of another, to be called primary. Interference arises when a bilingual identifies a phoneme of the secondary system with one in the primary system and, in reproducing it, subjects it to the phonetic rules of the primary language. Phenomena of this type were described traditionally under the heading of “sound substitution”. With the development of phonemic theory, investigators advanced from a mere description of the “mispronounced” sounds to the search for exact and verifiable causes of mispronunciation inherent in the speakers’ primary systems. Since the actual sounds produced by the bilingual lie, as it were, in the structural no man’s land between two phonemic systems, their interpretation in functional, i.e. phonemic, terms is subject to special difficulties. (Weinreich, 1968: 14)

Es decir, la interferencia concierne tanto a la percepción como a la reproducción de los sonidos, los cuales se sitúan en un espacio indefinido entre la lengua nativa y la segunda lengua. En su trabajo, el investigador compara los sistemas del *romansh* y el *schwyzertütsch*, dialectos hablados en Suiza, lo que le permite redefinir los diferentes tipos básicos de interferencias:

(1) UNDER-DIFFERENTIATION OF PHONEMES occurs when two sounds of the secondary system whose counterparts are not distinguished in the primary system are confused. (...) (2) OVER-DIFFERENTIATION OF PHONEMES involves the imposition of phonemic distinctions from the primary system on the sounds of the secondary system, where they are not required. The process can be inferred from a comparison of the sound systems in contact even if it is not always noticeable. (...) (3) REINTERPRETATION OF DISTINCTIONS occurs when the bilingual distinguishes phonemes of the secondary system by features which in that system are merely concomitant or redundant, but which are relevant in his primary system. (...) (4) Actual PHONE SUBSTITUTION, in the narrow sense of the term,



applies to phonemes that are identically defined in two languages but whose normal pronunciation differs. (Weinreich, 1968: 18-19)

Como se ha visto, durante la década de los 60 la mayoría de trabajos realizados en torno a la pronunciación de la L2 se desarrollan dentro del marco de la Hipótesis del Análisis Contrastivo (CAH), la cual intenta explicar las dificultades fonológicas de la L2 en base a las diferencias entre la L1 y la L2. Pero a partir de la década de los 70, los estudios entran en una nueva fase, y si bien también siguen la Hipótesis del Análisis Contrastivo, los investigadores consideran que son aquellos sonidos similares en la L1 y la L2 los que crean más dificultades a los estudiantes en el proceso de adquisición de la segunda lengua. Además, tal y como se muestra a continuación, en esta década se acuña, por primera vez, el término *interlengua*.

El trabajo de de Oller y Ziahosseiny (1970) es uno de los primeros que, estando dentro del marco de la CAH, sigue la idea de que las similitudes entre la L1 y la L2 son la causa de las dificultades que surgen al pronunciar una lengua extranjera.

A comienzos de esta década, autores como Corder (1971), Nemser (1971) y Selinker (1972) analizan las características de la lengua producida por los hablantes que se sitúa, de acuerdo con Weinreich (1968), en un espacio indefinido entre la lengua nativa y la segunda lengua. Según los investigadores, la adquisición de una segunda lengua se convierte en la construcción de una gramática mental basada en el *input* recibido en dicha L2. Este lenguaje del estudiante es denominado por Corder (1971) *dialecto idiosincrásico*, *sistema aproximativo* por Nemser (1971), e *interlengua* por Selinker (1972), siendo esta última denominación la que ha perdurado.

Siguiendo la propuesta de Oller y Ziahosseiny (1970), Wode (1976, 1978, 1983a, 1983b, 1994) propone el *Crucial Similarity Measure*, un modelo que, al contrario que las propuestas anteriores, se centra en analizar las similitudes entre la L1 y la L2 con el fin de poder anticipar las dificultades de los estudiantes en su producción en la L2.

Eckman (1977), por su parte, señala la necesidad de incorporar el concepto de marcado tipológico a la Hipótesis del Análisis Contrastivo de Lado (1957) como otro aspecto más a tener en cuenta a la hora de definir las dificultades a las que se enfrenta el estudiante durante el proceso de adquisición de una segunda lengua. El autor propone la Hipótesis de Marcado Diferencial (*Markedness Differential Hypothesis*) ya que mientras que la hipótesis de Lado intenta explicar las dificultades en el aprendizaje de la L2 únicamente basándose en las diferencias entre la L1 y la L2, esta nueva hipótesis señala que esta explicación es necesaria, pero no suficiente, por lo que cree necesario incorporar

el concepto de marcado tipológico. Esta hipótesis afirma que dentro de las diferencias entre la L1 y la L2 existen estructuras marcadas, las cuales son más difíciles de percibir y adquirir por parte de los estudiantes, mientras que las no marcadas son más fáciles.

Como se ha visto, hasta finales de la década de los 70 la evolución ha sido teórica. Sin embargo, en la década de los 80 hay una nueva reorientación en los estudios que analizan la adquisición del sistema fonológico de segundas lenguas, y los investigadores comienzan a publicar los primeros estudios experimentales, como se indica a continuación.

En 1980, Flege publica un trabajo de fonética experimental, en el cual estudia la producción de hablantes nativos de árabe cuya segunda lengua es el inglés, y siguiendo los trabajos de Corder (1971), Nemser (1971a) y Selinker (1972), indica que

(t)he term “interlanguage” has been used to describe a learner’s foreign language competence. (...) The existence of an “interlanguage” seems to depend on the assumption that a learner’s L2 output reflects an evolving linguistic competence which is reorganized at successive stages in the learning process, only gradually approximating the linguistic competence of native speakers of L2. (Flege, 1980: 118-119)

Más adelante diversos investigadores tratan de probar la afirmación de Flege, según la cual la interlengua no es estática.

Hammerly (1982) también sigue la Hipótesis del Análisis Contrastivo. Su trabajo muestra que, de las seis áreas de pronunciación que define como más problemáticas para los estudiantes de una L2, las tres de mayor dificultad son aquellas relacionadas con los alófonos de ambas lenguas. De acuerdo con el investigador, la mayor dificultad reside en la supresión de alófonos de la L1 al pronunciar la L2, es decir, suprimir alófonos de la L1 que no existen en la L2 supone un gran reto para los hablantes. La segunda causa de problemas reside en la dificultad para producir alófonos de la L1 con diferente distribución en la L2, como es el caso de la [d] y [ð] cuando el español es la L1, y el inglés la L2. En tercer lugar, otra de las mayores dificultades que, según Hammerly, presentan los estudiantes es la pronunciación de un alófono de la L2 que no existe en la L1.

Strange y Dittmann (1984), realizan uno de los primeros estudios experimentales que trata de mostrar si se puede mejorar la percepción en inglés de un grupo de estudiantes cuya L1 es el japonés. Los resultados de este trabajo muestran que no hay una mejora significativa en la capacidad de los alumnos de percibir las diferencias entre /r/ y /l/.

Young-Scholten (1985), por su parte, siguiendo la estela del trabajo de Oller y Ziahosseiny (1970), lleva a cabo un estudio para probar el *Crucial Similarity Measure* propuesto por Wode (1976, 1978, 1983a, 1983b), y observa que gracias a este modelo

puede predecir y detectar un mayor número de dificultades y errores que siguiendo la Hipótesis del Análisis Contrastivo.

Tal y como se ha señalado anteriormente, los trabajos desarrollados en torno a la adquisición de la L1 resultan relevantes en el avance de los estudios que analizan la adquisición de segundas lenguas. Sagey (1986) y Archangeli (1988), cuyos trabajos se engloban dentro de la teoría de la *Feature Geometry*, indican que el sistema fonológico de una lengua se estructura jerárquicamente, dividida en distintas categorías, de manera que unos aspectos de la misma dependen de otros. Esto implica que en la adquisición de segundas lenguas los estudiantes perciben los sonidos de la L2 a través de las categorías de la estructura fonológica de la L1. Estas, a su vez, limitan los sonidos de la L2 que pueden percibirse correctamente y los sonidos que el alumno puede producir con éxito.

Henly y Sheldon (1986) analizan a un grupo de estudiantes cuya lengua materna es el japonés o el cantonés y que estudian el inglés como L2. Los autores señalan que la fonología de la L1 filtra las características acústicas de la /r/ y la /l/ inglesas, y lo hará de una manera u otra en función de la existencia y/o distribución de dichos fonemas en la L1 del estudiante.

Flege (1987), al igual que Lado (1957), Stockwell y Bowen (1965), Oller y Ziahosseiny (1970), Hammerly (1982) o Young-Scholten (1985) compara la L1 y la L2 y realiza una clasificación de fonemas de la L2 desde el punto de vista de la L1:

From the standpoint of L2, the phones in an L2 may be taxonomized acoustically as “identical”, “similar” or “new” (Flege, 1968a). “New” L2 phones have no counterpart in the L1 and so, by definition, differ acoustically from phones found in L1. For example, realizations of French /y/ are “new” phones for native speakers of English. (...) “Similar” L2 phones, on the other hand, differ systematically from an easily identifiable counterpart in L1. For example, /t/ is found in both French and English, but it is implemented as a short-lag stop with dental place of articulation in French, and as a long-lag stop with alveolar place of articulation in English. (Flege, 1987a: 48)

En esta línea de comparación de la L1 y la L2, el trabajo de Best *et al* (1988) señala que la pronunciación de los estudiantes de la L2 depende de las similitudes entre los sonidos de esta lengua y los de la L1, aunque también hace referencia a la percepción, ya que, según los investigadores, aquellos sonidos que son distintos de los de la L1 son más fáciles de percibir y, por lo tanto, más fáciles producir. Sin embargo, aquellos que se asemejan a los de la L1 son más difíciles de percibir, y por lo tanto, más difíciles de producir.

Lehiste (1988), centra su trabajo en el análisis de las lenguas en contacto e indica que las interferencias son desviaciones de las normas de cualquiera de las dos lenguas

que suceden en el habla de los bilingües como resultado de su familiaridad con más de una lengua. En cuanto al concepto de interferencia fonética, la investigadora indica que

*(p)honic interference* occurs when a bilingual perceives and reproduces the sounds of one language (the secondary language, language B) in terms of his primary language (language A). Interference arises when the bilingual identifies a phoneme of the secondary language with a phoneme of the primary language and, in reproducing it, subjects it to the phonic rules of the primary language. (Lehiste, 1988: 2)

Al igual que Weinreich (1953, 1968), Lehiste realiza una clasificación de las interferencias fonéticas: 1) Sustitución de sonidos, ocurre cuando la pronunciación de la L1 se transmite a la L2, cuando la L2 tiene un fonema que la L1 no tiene, o cuando la L1 tiene un fonema y la L2 tiene dos, los cuales presentan cierta similitud; 2) transferencia de reglas, la cual tiene lugar cuando el estudiante no puede aplicar una regla de la L2 que no existe en la L1; 3) baja diferenciación, la cual sucede cuando la L1 no tiene un contraste que sí existe en la L2, es decir, la L1 posee dos alófonos de un mismo fonema que en la L2 son diferentes fonemas; 4) sobre-diferenciación (*overdifferentiation*), es la imposición de distinciones fonéticas de la L1 en la L2, es decir, los alófonos de la L2 son tratados como fonemas debido a que así lo son en la L1; 5) reinterpretación de las distinciones, el hablante distingue fonemas de la L2 que son distintivos en la L1, pero superfluos en la L2. Tal y como se puede observar, la autora indica que la pronunciación de los hablantes depende tanto de las diferencias y similitudes entre las dos lenguas, como de la percepción de los sonidos de la L2.

Como se ha visto, en la década de los años 80 los investigadores comienzan a realizar estudios experimentales con el fin de probar las teorías propuestas en etapas anteriores, y dado que estas no les permiten explicar todos los fenómenos que detectan en el proceso de adquisición del sistema fonológico de segundas lenguas, en la década de los 90 comienzan a proliferar nuevos modelos y teorías, aunque el *Speech Learning Model* propuesto por Flege es el que mayor aceptación tiene entre los investigadores, como se muestra a continuación.

En 1991 Eckman propone la Hipótesis de Conformidad Estructural (*Structural Conformity Hypothesis*). Según el autor, esta teoría no está formulada dentro de ninguna escuela particular y se puede seguir en cualquier programa de investigación que siga los universales lingüísticos. Y es que esta hipótesis afirma que tanto la lengua nativa como la interlengua obedecen a un conjunto de generalizaciones universales.

Tal y como se ha señalado anteriormente, en el año 1984 Strange y Dittmann realizan un estudio del que se infiere que los estudiantes de inglés cuya lengua materna

es el japonés no mejoran su capacidad de percepción de los sonidos de la L2. Sin embargo, estudios como los de Logan, Lively y Pisony (1991) o Strange (1992) muestran que el hecho de que los hablantes no nativos perciban algunos sonidos de la L2 de forma diferente a como lo hacen los nativos, no significa que su percepción permanezca constante. Es decir, los investigadores sugieren que la percepción de los sonidos puede cambiar durante el aprendizaje de la L2. El objetivo del trabajo de Logan, Lively y Pisony (1991) es determinar bajo qué condiciones un grupo de hablantes nativos de japonés puede aprender a identificar los fonemas /r/ y /l/ del inglés. Los resultados de este experimento demuestran que, si bien se pueden emplear técnicas para modificar la percepción, que el estudiante mejore su capacidad perceptiva depende, generalmente, del entorno fonético en el que se ubica el sonido en cuestión. Strange (1992) señala que aunque los estudios sobre la percepción del habla muestran que los hablantes pueden diferenciar fácil y automáticamente las categorías fonéticas que son distintas a las de su L1, la percepción de los contrastes fonéticos de la L2 es, en general, más difícil. Esto refleja la relación de los procesos perceptivos con la información acústica y articulatoria que es relevante en la L1. La investigadora apunta que aprender una lengua extranjera en la edad adulta implica la reacomodación de los procesos perceptivos y la reorganización de las categorías fonéticas.

También Flege (1991) indica que son muchos los investigadores que sugieren que los acentos extranjeros son causados por una percepción inexacta de los sonidos de la L2, y añade:

The substitutions seen in nonnative speakers' production of sounds in a second language (L2) vary systematically according to their L1 background. For those who learn an L2 after the L1 has been thoroughly established, the substitutions do not derive from immaturity of the speech production mechanism or inability to fully control its functioning (...), but rather from the perceptual misinterpretation of phonetic contrasts in the L2 and/or the inappropriate use of patterns of production used for L1 sounds in the production of L2 words. (Flege, 1991: 249)

Anteriormente se ha señalado que a partir de la década de los 60 son muchos los investigadores que tienen en cuenta los estudios realizados en torno a la adquisición de la L1 para realizar sus investigaciones en torno a la L2. Prince y Smolensky (1993, 2004), por su parte, proponen la Teoría de la Optimidad (*Optimality Theory*). De acuerdo con este modelo, las expresiones lingüísticas admiten cierta variabilidad, y las formas que presenta una lengua pasan una serie de restricciones jerarquizadas. Según esta teoría, durante el proceso de adquisición de la L1 se aprende el orden relativo de las restricciones.

Este modelo tiene gran relevancia en estudios posteriores, ya que son muchos los autores que se basan en el mismo para explicar la adquisición de la L2.

Se observa que el número de trabajos que analizan la percepción en el proceso de adquisición de segundas lenguas ha aumentado considerablemente, y como consecuencia surgen diferentes modelos que tratan de explicar aquellos aspectos a los que la Hipótesis de Análisis Contrastivo no logra dar una explicación. Uno de estos modelos es el *Perceptual Assimilation Model* (PAM), propuesto por Best (1993, 1994, 1995, 1996). Según este modelo, durante la adquisición de la L2 la percepción del habla se sintoniza con los elementos fónicos contrastivos de la L1. Por esta razón, es posible que los estudiantes de la L2 no puedan discernir las diferencias fonéticas entre distintos pares de sonidos de esta lengua, o entre sonidos de la L1 y la L2, ya sea porque son fonéticamente distintos en la L2, o porque son asimilados a una sola categoría de la L1.

Hancin-Bhatt (1994) propone el *Feature Competition Model* (FCM), el cual basa sus predicciones en la idea de que en el sistema fonético de las lenguas existen aspectos del mismo más relevantes que otros. Esto significa que aquellos rasgos que se emplean con más frecuencia en la L1, aquellos que son más prominentes, tendrán mayor influencia en la percepción que tendrá el estudiante de los nuevos sonidos de la L2, y serán aquellas características distintivas de L1 más relevantes las que determinen cómo se clasifican los sonidos de la L2 en las categorías ya existentes de la L1.

El modelo de percepción y producción del habla que más aceptación ha generado es el *Speech Learning Model* (SLM) propuesto por Flege (1995). Siguiendo a Trubetzkoy (1939), y al igual que muchos de los trabajos citados anteriormente, Flege (1995: 237) indica que los bilingües tienden a interpretar los sonidos de la L2 a través de la criba fonológica de su L1. Como resultado de las investigaciones desarrolladas hasta el momento, Flege propone el SLM, el cual tiene como objetivo analizar los límites relacionados con la edad de los hablantes y su capacidad para producir sonidos consonánticos y vocálicos en la L2 igual que los nativos. Este modelo se centra en bilingües de nivel medio-avanzado, dejando de lado a aquellos estudiantes que se encuentran en una fase inicial del aprendizaje de la L2. Según Flege, durante el proceso de adquisición de la L2 la percepción del habla se sintoniza con los elementos fónicos contrastivos de la L1. Este modelo afirma que, aunque muchos errores de producción de la L2 tienen una base perceptiva, no todos están motivados perceptualmente. El SLM de Flege (1995) propone dos mecanismos específicos a través de los cuales las categorías fonéticas que comprenden los subsistemas fonéticos de la L1 y la L2 interactúan: la

asimilación y la disimilación de categorías. Según Flege (2002), operará uno u otro en función de si se ha formado una nueva categoría o no. Los sistemas fonéticos de la L1 y la L2 interactúan a través del mecanismo de asimilación de categorías cuando se ha bloqueado la formación de las mismas. En tal caso, el alumno produce un sonido de la L2 como si fuera uno de la L1, es decir, sin ninguna modificación, debido a que lo equipara con este. Sin embargo, si el sonido de la L2 difiere de forma audible del sonido de L1 con el que se ha equiparado, se espera una modificación en la producción del mismo. En cuanto al mecanismo de disimilación, el investigador señala que los sistemas fonéticos de la L1 y la L2 interactúan a través del mismo cuando se establece una nueva categoría para un sonido que se encuentra en la L2, pero no en la L1.

Hancin-Bhatt y Bhatt (1997), por su parte, se centran en el estudio de la estructura silábica del inglés por estudiantes cuya lengua materna es el japonés o el español. El objetivo de su trabajo es mostrar evidencias empíricas de que tanto la transferencia de rasgos entre la L1 y la L2, como las restricciones presentes en ambas lenguas tienen influencia en la realización de las estructuras silábicas del inglés producidas por hablantes de japonés o español. Según los autores, la Teoría de la Optimidad es el modelo más adecuado para explicar los resultados de su estudio, dado que, según ellos, abarca y explica más aspectos de la lengua que el resto de modelos. De acuerdo con los investigadores, las clasificaciones de los distintos elementos de la L2 que se realizan durante la adquisición de esta están basados en los que el hablante ya tiene de la L1, pero según avanza el proceso de aprendizaje, estos ránkines se vuelven más parecidos a los propios de la L2. Además, Hancin-Bhatt y Bhatt (1997) recalcan la importancia de asumir que la fonología de la interlengua es dinámica para poder explicar la variación en el proceso de adquisición por parte de los distintos hablantes.

El trabajo de Eckman e Iverson (1997) indica que existen, por lo menos, dos aspectos de la L1 y la L2 que pueden causar dificultad en la pronunciación de la segunda. Por un lado, se encuentran las restricciones relacionadas con el inventario de sonidos de ambas lenguas, en cuyo caso el estudiante tiene que ser capaz de superarlas y ser capaz de pronunciar la L2 sin que nada de la L1 le sirva de apoyo para ello. Por otro lado, están las restricciones relacionadas con la posición, las cuales tienen lugar cuando la L2 posee un contraste entre sonidos, mientras que en la L1 dichos sonidos son alófonos, como es el caso [d] y [ð] cuando la L1 es el español y la L2 el inglés. Los autores señalan que cuando se transfiere una regla alofónica de la L1 a la interlengua, esta causa errores dependiendo de si el estudiante toma conciencia de los contrastes de la L2. Si estos no se

aprenden, la regla transferida continuará aplicándose, pero cuando el alumno adquiere dicho contraste, la pronunciación mejora y se vuelve más parecida a la propia de la L2.

Broselow, Chen y Wang (1998) realizan un estudio con estudiantes de inglés cuya lengua materna es el chino mandarín siguiendo los principios de la Teoría de la Optimidad. Esta, según los autores, difiere en gran medida de la fonología estándar basada en reglas generativas, ya que asume que las lenguas difieren en virtud de tener diferentes restricciones, mientras que según la Teoría de la Optimidad, las lenguas no difieren en su conjunto de restricciones, si no en las clasificaciones de las mismas, las cuales determinan la fuerza de una restricción en particular. Los investigadores concluyen su trabajo indicando que el modelo de adquisición de una L2 que proponen es aquel en el que los estudiantes construyen una interlengua cuyas clasificaciones restrictivas pueden diferir de las propias tanto de la L1 como de la L2.

Brown (1998, 2000), quien trabaja con estudiantes de inglés cuya lengua materna es el chino, el japonés o el coreano, desarrolla sus estudios según los principios de la *Feature Geometry*. La investigadora indica que la estructura geométrica adquirida por los niños en su L1 reduce la sensibilidad perceptiva a contrastes en otras lenguas, con lo que todos los sonidos del habla, tanto los nativos como los no nativos, se perciben según las características de la L1, aspecto que continúa influyendo en la percepción del habla adulta. De acuerdo con los principios de esta teoría, aunque la L2 es percibida en términos de las categorías de la L1, es la estructura geométrica de esta lengua la que asigna la señal acústica recibida a los diferentes niveles. La autora apunta que en aquellos casos en los que parece que la L1 no influye en la adquisición de la L2, lo que sucede es que la estructura geométrica de dicha lengua facilita la percepción y la adquisición de la misma. Por el contrario, en aquellos casos en los que la L1 sí parece tener influencia la estructura geométrica dificulta la percepción y, por lo tanto, impide la adquisición.

En el año 2000 Young-Scholten y Archibald publican un trabajo en el que realizan un análisis de las sílabas del inglés, el alemán y el holandés como L1. Los investigadores indican que intentan mostrar que la adquisición de la estructura silábica de la L2 está influida tanto por los principios universales de la estructura prosódica, como por las propiedades de la estructura silábica de la L1. También apunta que la cuestión de si la nueva estructura puede establecerse en la L2 sigue siendo incierta, ya que parece muy difícil superar la estructura de la L1.



Más recientemente, trabajos como los de Eckman (2004) Major (2008) o Hancin-Bhatt (2008) analizan la bibliografía relacionada con los estudios de segundas lenguas desarrollados desde Polivanov (1931) hasta los estudios más recientes.

Tal y como se ha expuesto en estas páginas, durante las últimas décadas han surgido diversas teorías, modelos y métodos cuyo objetivo es describir el proceso de adquisición de segundas lenguas (o lenguas extranjeras) y explicar el porqué de las características fonéticas de la L2 que presentan los estudiantes de la misma. La mayoría de investigadores han desarrollado sus teorías y modelos siguiendo la Hipótesis del Análisis Contrastivo, según la cual las diferencias entre la L1 y la L2, y la transferencia de rasgos de la primera a la segunda son fundamentales para explicar los rasgos propios del habla en la L2.

Planteamientos más recientes como el *Speech Learning Model* propuesto por Flege (1995) afirma, contrariamente al análisis contrastivo, que los sonidos similares son difíciles de percibir y producir, mientras que los sonidos nuevos resultan más fáciles. Best (1995), por su parte, propone el *Perceptual Assimilation Model* que se asemeja al SLM de Flege, ya que también indica que la L2 será percibida de acuerdo con las similitudes entre esta y la L1.

De acuerdo con la fonología generativa clásica, la adquisición de la L2 comienza siguiendo las reglas de la L1, pasa por un estado en el que las reglas siguen las de la fonética universal, y finalmente se siguen las reglas de la L2. La Teoría de la Optimidad define la adquisición de la L2 de manera similar, pero reemplaza las reglas por ránkines de restricciones. Según este punto de vista, la adquisición de la L2 atravesaría las siguientes fases: en primer lugar, los ránkines de restricciones seguidos por el estudiante serían las de la L1, después pasarían a situarse entre los de la L1 y los de la L2, y finalmente serían ya los de la L2. De estos trabajos se concluye que la transferencia no debería estudiarse como un fenómeno aislado, sino en relación con otros factores, incluyendo la marcación, las similitudes y diferencias, y los principios universales de la lengua.

Son muchos los estudios realizados en torno a la adquisición del inglés por hablantes de distintas lenguas maternas como el español, el francés, el japonés, el coreano, o el chino, entre muchas otras. También abundan los trabajos que analizan la adquisición del español como lengua extranjera por hablantes de otras lenguas, especialmente aquellas que pertenecen a la familia de las lenguas románicas. Sin

**embargo, la cantidad de estudios relacionados con la adquisición del español por sinohablantes es reducida comparada con la de otras lenguas. En el siguiente apartado se analiza la bibliografía que trata de describir las características del proceso de adquisición de la pronunciación del español por hablantes de chino mandarín y que analiza los problemas que surgen durante el aprendizaje del español como lengua extranjera (ELE).**

## **2. LA ADQUISICIÓN DEL SISTEMA VOCÁLICO ESPAÑOL POR SINOHABLANTES**

En el apartado anterior se ha podido observar que la mayoría de los estudios realizados en torno a la adquisición de segundas lenguas o lenguas extranjeras se ha servido, generalmente, de la fonética contrastiva para predecir los posibles errores o problemas a los que se enfrentan los estudiantes de dichas lenguas. Ejemplo de ello son los estudios que comparan el chino mandarín con lenguas como el inglés, el francés o el alemán y analizan la adquisición del sistema fonético de dichas lenguas por estudiantes sinohablantes. Sin embargo, no existen trabajos previos a la década de los 90 que comparen el chino y el español o que analicen las características del español hablado por estudiantes cuya lengua materna es el mandarín. Por este motivo, se puede considerar que la tesis doctoral de Ortí Mateu (1990) es pionera en este campo de estudio. Esta investigación ofrece un análisis detallado de las características articulatorias y acústicas de los sonidos vocálicos y consonánticos del chino y del español, y un estudio contrastivo del sistema fonético-fonológico de ambas lenguas. Los resultados de una serie de pruebas de contrastes fónicos y de pronunciación realizadas a 120 estudiantes sinohablantes permiten predecir o diagnosticar cuáles son los elementos segmentales del español que plantean más dificultades a los alumnos sinohablantes. El trabajo concluye con una serie de observaciones y recomendaciones didácticas que pretenden ser de ayuda para los profesores que enseñan la pronunciación de dicha lengua.

La autora registra todas las dificultades detectadas en el conjunto de alumnos estudiados que, como se verá más adelante, serán también detectadas y analizadas por otros investigadores. Presentamos a continuación las dificultades detectadas por Ortí Mateu (1990: 13-14) por orden de importancia según la autora:

- Dificultades en la pronunciación de sonidos
  - Vibrantes [r, r]
  - Oclusivas sonoras [b, d, g]
  - Lateral [ʎ]
  - Interdental [θ]
  - Vocales medias [e, o]
  - Silbante [s]
- Dificultades en los contrastes fónicos
  - Lateral – vibrante [l – r]
  - Sorda – sonora: [k – g], [t – d], [p – b]
  - Lateral – fricativa [ʎ – ʝ]
  - Labiodental – interdental [f – θ]
- Dificultades en las estructuras silábicas
  - Consonante trabada: C1 C2 V; C2
  - Consonante final, excepto si es [n]: C1 V C3; C3
  - Otras estructuras menos frecuentes: V C3 C4; C1 V C3 C4; C1 C2 V C3 C4; C1 V1 V C3; C1 C2 V1 V
- Dificultades en la entonación
  - Pronunciar unidas las sílabas del grupo fónico
  - Colocar el acento oracional en el lugar correspondiente de la cadena hablada.

Según la autora, la complejidad del sistema vocálico chino contrasta con la simplicidad del sistema vocálico español (Ortí Mateu, 1990: 2). El sistema vocálico chino consta, de acuerdo con Ortí Mateu, de seis fonemas /i, y, u, e, o, a/ y once alófonos, mientras que el español presenta cinco fonemas y cinco alófonos. Cinco de estas vocales existen en ambos idiomas, aunque se dan notables diferencias fónicas entre las mismas. De acuerdo con la autora, también el porcentaje de frecuencia de aparición en la cadena hablada es diferente, ya que [i] y [u] son las vocales más frecuentes en chino, mientras que las más frecuentes en español son [a] y [e] (Ortí Mateu, 1990: 3). Así, según esta autora, la adquisición del sistema vocálico del español no representaría una gran dificultad para los alumnos sinohablantes.

Unos años más tarde, Zhou (1995), en su tesis doctoral, realiza un estudio comparativo de aspectos lingüísticos y culturales del chino y del español. Su trabajo se centra, principalmente, en los problemas que presentan los estudiantes hispanohablantes que estudian el chino como lengua extranjera, aunque en algunas ocasiones también hace referencia a las dificultades de los sinohablantes cuando estudian español. Su trabajo resulta interesante para esta investigación ya que, tal y como se podrá observar, la descripción que realiza del sistema vocálico del chino mandarín difiere en algunos aspectos de la realizada por otros investigadores que se citan en este apartado. Según el autor, en la lengua china existen seis vocales simples (*i, ü, a, o, e, u*), mientras que el español consta de cinco (*i, e, a, o, u*). De estas, el investigador indica que *a, i, o* y *u* se pronuncian igual en ambas lenguas y que solamente dos de ellas son diferentes: *e* y *ü*

(Zhou, 1995: 50). Un aspecto reseñable que menciona el autor en relación con las vocales del chino es que cuatro (*a, o, e, i*) de los seis sonidos que conforman el sistema vocálico del mandarín experimentan cambios en la pronunciación según el entorno fónico en el que se encuentren. Zhou comenta algunos de estos cambios en relación con la “e”: cuando aparece delante de [r], [n] o [ŋ] se pronuncia [ə], y si se encuentra delante de [i], se articula [e]. Cuando va precedida por [i] o [ü], la pronunciación de la misma es [ɛ]. Y en el resto de entornos se pronuncia [ɤ] (Zhou, 1995: 56-57). En cuanto al sistema consonántico de ambas lenguas, el autor indica que una de las mayores dificultades con la que se encuentran los estudiantes sinohablantes de ELE, así como los hispanohablantes que aprenden chino mandarín es la distinción en la pronunciación de las consonantes oclusivas puesto que mientras en español el contraste se establece mediante una oposición sonoridad, en chino el rasgo que establece la diferencia es la aspiración (Zhou, 1995: 56-57).

A partir de finales de los años 1990, uno de los investigadores que más trabajos ha publicado en torno a la adquisición del sistema fonético del español por estudiantes sinohablantes es Cortés Moreno. Sus estudios de 1998, 2001a y 2001b se centran en distintos aspectos de la entonación. En el primero de ellos el investigador analiza el orden de adquisición de la entonación del español en el plano perceptivo, en el segundo se centra en las dificultades concretas que presentan los estudiantes sinohablantes en la percepción de la entonación del español, y en el tercero analiza “en qué medida se corresponden la producción de patrones entonativos en español por parte de sinohablantes y la interpretación de dichos patrones por parte de hablantes/oyentes nativos de castellano y catalán” (Cortés Moreno, 2001b: 193). Estos tres artículos coinciden en que el orden de adquisición de la entonación y las dificultades en la percepción de este rasgo suprasegmental están íntimamente relacionados. En primer lugar, los alumnos sinohablantes adquieren la entonación de los enunciados declarativos, por lo que esta estructura es la que menos problemas plantea; en segundo lugar, adquieren la entonación correspondiente a las preguntas nominales; más adelante, la de las preguntas absolutas y, finalmente, la de los enunciados enfáticos, siendo esta entonación la que mayores dificultades plantea a los estudiantes. (Cortés Moreno 1998, 2001a, 2001b)

En sus trabajos de 2003 (a y b), Cortés Moreno se centra en la acentuación. En la primera publicación analiza la percepción de esta, y en la segunda, la producción de la misma por estudiantes sinohablantes. El investigador señala que la dificultad principal se encuentra en la percepción del acento en las palabras agudas, mientras que su percepción en palabras

llanas presenta menos dificultades para los alumnos (Cortés Moreno, 2003a). A pesar de ello, y siempre según el punto de vista del autor, la acentuación del español no plantea dificultades reseñables a estos estudiantes (Cortés Moreno, 2003b).

En 2004 Cortés Moreno publica un nuevo trabajo cuyo objetivo es caracterizar los contornos entonativos que producen los estudiantes sinohablantes de ELE y comprobar si existe una transferencia de los mismos desde su lengua materna al español. Los resultados de esta investigación muestran que sí existe una transferencia del chino al español y corroboran los resultados de sus trabajos anteriores (Cortés Moreno 1998, 2001a, 2001b), ya que indica que la dificultad principal para estos estudiantes es la entonación enfática. En 2005 Cortés Moreno se vuelve a centrar en las dificultades que los estudiantes sinohablantes encuentran en la adquisición y aprendizaje de la acentuación y la entonación del español y el investigador detecta que se transfieren algunos patrones acentuales y entonativos del chino al español, tanto en la percepción como en la producción. El autor afirma, de nuevo, que el aprendizaje de la acentuación no plantea dificultades notables (Cortés Moreno, 2003b), y los resultados obtenidos en el análisis de la entonación son iguales que los señalados en anteriores publicaciones (Cortés Moreno 1998, 2001a, 2001b). En su trabajo de 2006 el autor apunta que existe una transferencia de contornos tonales del chino al español por parte de los sinohablantes, y que existen importantes oscilaciones melódicas intrasilábicas, principalmente en sílabas tónicas y en el primer pico de la curva melódica.

En 2009 Cortés Moreno publica dos artículos en los que realiza un análisis contrastivo chino-español. En 2009a expone los errores segmentales, detectados en el aula de ELE, producidos por estudiantes sinohablantes. El autor indica que no todos los errores se deben a procesos de interferencia de la L1, sino que también los hay de desarrollo o intralingüísticos, es decir, que se tratan de errores parecidos a los que cometen los niños en el proceso de adquisición de su L1 (Cortés Moreno, 2009a, 2009b). El investigador señala que “una comparación entre las vocales y las consonantes de las lenguas en cuestión nos permite predecir ciertos problemas de percepción y de producción de los sinohablantes nativos que estudian ELE” (Cortés Moreno, 2009a: 2). Cortés Moreno no se pronuncia sobre el grado de dificultad que pueden presentar los elementos segmentales durante el proceso de adquisición del español pues apunta que, aunque determinados elementos sean parecidos en las dos lenguas, eso no significa que vayan a resultar fáciles para los estudiantes, o que aquellos que son distintos sean difíciles (Cortés Moreno, 2009b: 46).

En relación con el sistema vocálico del chino y del español, el investigador indica que

del contraste entre las vocales de una y otra lengua se desprende que, a priori, los aprendientes chinos de ELE no tendrán dificultades ni para percibir ni para producir las vocales españolas, supuesto que estas coinciden con algunas de su L1 o, dicho sea con otros términos, porque el vocalismo español está contenido en el sistema vocálico de la lengua china. (Cortés Moreno, 2009a: 4)

Y señala, además, el autor (2009a, 2009b) que en chino mandarín existen dos fonemas que no existen en español:

(t) tanto en chino como en español existen estos cinco fonemas vocálicos: /i/, /e/, /a/, /o/ y /u/. Además, el sistema vocálico chino cuenta con otros dos fonemas: /y/ -comparable a la *u* francesa de *une* o a la *y* danesa de *dyne*- y /ə/ -comparable a la *e* en la palabra catalana *pare* o en la inglesa *father*. En chino existen cuatro diptongos decrecientes -/ai/, /ei/, /au/ y /ou/- y cinco diptongos crecientes -/ja/, /je/, /wa/, /wo/ y /ye/-. Existen asimismo cuatro triptongos: /jau/, /jou/, /wai/ y /wei/. (Cortés Moreno, 2009b: 183-184)

Como puede apreciarse, Cortés Moreno (2009a, 2009b) coincide con Ortí Mateu (1990) y con Zhou (1995) en señalar que el chino mandarín posee más sonidos vocálicos que el español, aunque no coincide con estos investigadores en la descripción de aquellos sonidos que no existen en castellano. Los tres autores coinciden, además, en la creencia de que aprender el sistema vocálico de una lengua extranjera (español) que posee menos unidades que la materna (chino mandarín) no plantea ningún problema porque lo difícil es, exclusivamente, el aprendizaje de sonidos nuevos.

Cortés Moreno (2009b: 185) también realiza un análisis contrastivo de las diferencias y similitudes entre los sistemas consonánticos del español y del chino mandarín del cual extrae una lista de los principales problemas que presentan los estudiantes sinohablantes en su producción oral en español. En primer lugar, apunta a la confusión entre /s/ y /θ/, dado que este último no existe en chino. En segundo lugar, el autor indica que otra dificultad a la que se enfrentan los estudiantes es la discriminación de /l/, /r/ y /r/, dado que la lateral y la vibrante simple son alófonos en chino y la vibrante múltiple no existe en dicha lengua. En tercer lugar, señala como problema importante la discriminación entre las oclusivas sonoras y sordas, es decir, entre /b/ y /p/, entre /d/ y /t/, y entre /g/ y /k/. Cortés Moreno (2009b), de la misma forma que Ortí Mateu (1990) y Zhou (1995), apunta que las oclusivas chinas son sordas, y son clasificadas en aspiradas y no aspiradas, mientras que en español unas son sordas y otras sonoras. El investigador señala que, debido a las dificultades en la percepción y la pronunciación, surgen problemas de escritura entre los estudiantes sinohablantes relacionados con la confusión entre /s/ y /θ/, entre /l/, /r/ y /r/, así como entre los fonemas oclusivos sonoros /b, d, g/ y sus correlatos oclusivos sordos /p, t, k/ (Cortés Moreno, 2009b: 188). Todos estos aspectos son

corroborados por el estudio que realiza en 2009, lo que significa que, según su hipótesis, y al igual que señala Ortí Mateu (1990) en su tesis, los estudiantes sinohablantes de ELE presentarán dificultades para discriminar [s] y [θ], la lateral [l] y las vibrantes [r] y [r̄], y las oclusivas sordas [p], [t] y [k] de las sonoras [b], [d] y [g].

En relación con la sílaba y la estructura silábica, Cortés Moreno (2009a: 7) apunta que “la sílaba china es más simple que la española”. Esto se debe a que la sílaba china consta de tres componentes, el inicial, el final y el tonema, y la parte inicial, por lo general, está formada por una sola consonante. Es decir, no hay grupos consonánticos iniciales, lo que plantea dificultades a los estudiantes sinohablantes de ELE, aunque no solo en esta posición, si no en cualquier otra. En cuanto a la coda silábica, generalmente consiste en chino en una sola vocal, un diptongo o una vocal seguida de una nasal (Cortés Moreno, 2009b: 186).

El investigador también describe la acentuación y la entonación de estas dos lenguas, y señala que

(1) La diferencia principal entre ambas lenguas en materia de prosodia reside en que el chino es una lengua tonal y el español, una lengua no tonal. Se aprecia en el habla de los alumnos chinos una cierta tendencia a pronunciar sílaba por sílaba en lugar de agruparlas en palabras fónicas y grupos fónicos, muy en especial, cuando leen y, sobre todo, en los niveles iniciales. Ello impregna su discurso de un característico ritmo staccato, que unido a la entonación propia de su lengua, llega a plantearles dificultades de comunicación en ELE. (Cortés Moreno, 2009b: 186)

En cuanto a las oscilaciones melódicas dentro de las sílabas producidas por los estudiantes, Cortés Moreno (2009b: 187) indica que no se trata de una transferencia de tonemas del chino al español, sino de pequeñas alteraciones tonales, es decir, pequeñas inflexiones melódicas. En relación con la acentuación, cabe señalar que, de acuerdo con el autor, las palabras agudas son las más frecuentes en chino y las llanas son las más usuales en español. Sin embargo, a pesar de las diferencias existentes entre ambas lenguas en este aspecto, la acentuación del español no plantea grandes dificultades a los estudiantes sinohablantes de ELE según Cortés Moreno (2009b). Y en lo que concierne a los contornos entonativos, la diferencia entre estas dos lenguas en los contornos interrogativos estriba en que los valores tonales altos en español aparecen, generalmente, en la última sílaba del enunciado en una inflexión ascendente, mientras que en chino tienen lugar al principio del enunciado y el registro es semialto. (Cortés Moreno, 2009b: 187-188)

Tal y como se puede observar, si bien Cortés Moreno ha tratado numerosos aspectos relacionados con la adquisición del sistema fonético del español por sinohablantes, se

observa que el grueso de sus trabajos analiza distintos aspectos suprasegmentales, siendo menor y más reciente el número de trabajos que se centran en el análisis de las características del sistema segmental. Además, se desprende también del análisis de sus publicaciones que el autor no considera que el aprendizaje de las vocales constituya un problema grave debido a que las vocales del español estarían “contenidas” en el sistema vocálico del chino.

Otro trabajo de relevancia para el objeto de estudio de esta investigación es el de Wang (2001) que describe y compara los sistemas fonéticos español y chino. Al igual que Ortí Mateu (1990), Zhou (1995) y, posteriormente, Cortés Moreno (2009b), el autor indica que tanto el chino como el español poseen los fonemas vocálicos /a/, /e/, /o/, /i/ y /u/. Además, añade dos sonidos más que no existen en español, y que difieren de los señalados por los autores anteriormente citados: [w], [ɣ] y [ə] (Wang, 2001: 100-101). En cuanto al sistema consonántico de estas lenguas, Wang (2001) señala una serie de diferencias entre ambas. La primera de ellas que también es señalada por los trabajos de Ortí Mateu (1990), Zhou (1995) y Cortés Moreno (2009a, 2009b), es que a pesar de que tanto el chino mandarín como el español poseen seis fonemas oclusivos, todos ellos son sordos en chino, lengua en la cual la distinción se hace entre aspirado-no aspirado, mientras que en español se dividen en sordos y sonoros. En cuanto a las consonantes fricativas, el chino mandarín conoce seis y el español cinco, y de estas, la interdental sorda [θ] no existe en chino. Otra diferencia que se detecta en el análisis contrastivo de estas dos lenguas es que el español posee dos fonemas laterales, [l] y [λ], y el chino mandarín solo conoce el primero. El español posee las vibrantes simple y múltiple, mientras que, según el autor, estos sonidos no existen en chino mandarín, pero en esta lengua sí existen cuatro fonemas retroflejos que no existen en español: /tʂ/, /tʂʰ/, /ʂ/, /ʐ/. (Wang, 2001: 102-109). En relación con la estructura silábica del chino mandarín, Wang (2001: 88) indica que se divide en dos partes: la inicial, que siempre es una consonante, y la final, que siempre es un elemento vocálico, y al igual que Cortés Moreno (2009a, 2009b), apunta que mientras que las consonantes trabadas abundan en español, no existen en chino.

Tal y como se puede observar a continuación, los trabajos desarrollados por otros autores siguen la estela de los hasta ahora citados, especialmente los realizados por Cortés Moreno, tanto corroborando lo expuesto por este investigador, como añadiendo algunas precisiones. Cabe señalar, sin embargo, que debido a que algunos de los estudios que se analizan a continuación utilizan diferentes grafías para representar los fonemas y los sonidos del chino y del español, en algunas ocasiones el estudio de las distintas



aportaciones de los autores se complica, dado que no resulta sencillo interpretar si se refieren a grafías del pinyin, fonemas o alófonos. Como se ha podido observar, en este apartado de estado de la cuestión se respeta la grafía empleada por los investigadores en sus trabajos.

Yufei (2007), realiza un estudio contrastivo de los fonemas oclusivos del español y del chino. La autora indica que en chino mandarín solo existen cinco consonantes sonoras: [l], [m], [n], [ŋ] y [ʐ], y no existen sonidos oclusivos sonoros, dado que los fonemas oclusivos chinos son sordos. Estos dos rasgos hacen que la pronunciación de los sonidos sonoros del español represente una gran dificultad para la mayoría de los estudiantes sinohablantes que los pronuncian como sordos. A esto, la autora añade que [β], [ð], [ɣ], sonidos muy característicos de la lengua española, son desconocidos en chino mandarín, lo que también conlleva dificultades enormes para los estudiantes sinohablantes (Yufei, 2007: 92-94).

Siguiendo la estela de Cortés Moreno (1998, 2001a, 2001b, 2004), Planas Morales y Villalba (2007) analizan las similitudes entre el contorno entonativo del chino mandarín y el de los grupos fónicos del español. Los autores establecen una serie de equivalencias entre el contorno entonativo de la primera lengua y los grupos fónicos de la segunda, dado que, según indican, existen similitudes entre ambos. Estas equivalencias, señalan, permiten que los alumnos sinohablantes apliquen la competencia fonética de su lengua materna a la L2.

Planas Morales (2008), por su parte, realiza un estudio contrastivo del sistema fonético del chino y del español. La investigadora se basa en la clasificación de Poch (1999) para realizar una previsión de las dificultades a las que se enfrenta el estudiante. Dichos obstáculos se dividen en cuatro grupos: los derivados de la existencia de sonidos que no son propios de la lengua del alumno, los derivados de la diferente distribución silábica de los sonidos, los derivados del sistema acentual y, finalmente, los derivados de la entonación oracional (Planas Morales, 2008: 498).

En cuanto a la descripción de dichas dificultades, la autora indica que los estudiantes sinohablantes, “(e)n general, son alumnos que no presentan demasiados problemas en la articulación de los sonidos en sílaba libre, pero sí en su percepción” (Planas Morales, 2008: 500). Ejemplo de ello son las oclusivas [p / b], [t / d], [k / g]. La autora señala que la confusión entre [l] y [r] se da en ambas direcciones. Otro aspecto destacable de la producción en español de los estudiantes sinohablantes es que los alumnos “tienden a omitir la -n final, nasalizando la vocal precedente” (Planas Morales, 2008: 501). La

investigadora también remarca la importancia de tener en cuenta las equivalencias entre sonidos y grafías, ya que pueden darse casos de errores derivados de la confusión entre ambas, en lugar de que la causa del problema esté relacionada con la discriminación de los sonidos.

En cuanto al sistema vocálico, Planas Morales (2008: 501), igual que Cortés Moreno en sus publicaciones posteriores Cortés Moreno (2009a, 2009b), señala que a pesar de que parece que las vocales no presentan problemas de pronunciación dado que el sistema del español es más reducido que el del chino, sí se encuentran dificultades, especialmente porque las equivalencias entre la grafía y el sonido se multiplican en las vocales chinas.

La autora explica que, por ejemplo:

la grafía *e* en pinyin tiene cuatro sonidos diferentes: [ɤ]; [ə], delante de [n, ŋ]; [e], delante de [i]; [ɛ], después de [i, y]. El sonido [ɤ], que es propiamente el del fonema, es parecido a la vocal neutra catalana [ə], pero un poco más posterior, con una [o] no labial. Por su parte, la grafía *a* puede ser [a] o [ɛ] (como la *e* abierta del catalán). Motivos por los cuales, los estudiantes chinos tienden a articular la *e* española de forma abierta y más posterior. (Planas Morales, 2008: 501)

En lo que se refiere a los errores en la acentuación, según Planas Morales (2008) los estudiantes sinohablantes conocen las reglas de acentuación del español, pero no saben agruparlas en grupos fónicos alrededor del acento primario, por lo que su producción oral es poco fluida y entrecortada. Y en cuanto a las funciones de la entonación, la autora indica que desconocen la función demarcativa de la misma puesto que en chino mandarín las funciones morfosintácticas se expresan a través del orden en el que aparecen las palabras en la oración. Por este motivo, los alumnos sinohablantes no distinguen entre la entonación de la interrogativa absoluta y la declarativa, y además no diferencian las distintas entonaciones que dan significado semántico y expresivo a las oraciones. Sin embargo, la investigadora indica que estas dificultades se superan a medida que los estudiantes avanzan en su estudio de la lengua, dejan de cometer errores de entonación y “desarrollan una gran destreza en imprimir matices entonativos” (Planas Morales, 2008: 501).

El trabajo de Méndez Marassa (2009) no difiere del de los autores anteriores, ya que indica que los problemas a los que se enfrentan los estudiantes sinohablantes de ELE son la distinción entre fonemas sordos y sonoros, la pronunciación de las vibrantes e interdentes que no existen en chino y la distinción entre la vibrante simple y la lateral /l/. Y, como puede verse, no menciona las dificultades que pudieran afectar a la pronunciación de las vocales.

Zhi (2011) desarrolla su trabajo con un grupo de 51 informantes sinohablantes estudiantes de ELE de primer año y analiza los errores articulatorios que presentan estos alumnos. En relación con los sonidos consonánticos del español, el autor, desde su experiencia como docente, indica que “las dificultades habituales consisten en la distinción de las oclusivas sonoras y sordas, el relajamiento de las oclusivas sonoras en posición intervocálica, así como de las vibrantes” (Zhi, 2011: 55). Los resultados de su trabajo corroboran su hipótesis, dado que casi la mitad del total de los problemas registrados en la producción de los informantes están relacionados, en primer lugar, con la articulación de la vibrante múltiple y, en segundo lugar, con la de las consonantes oclusivas. Resulta de gran interés el análisis en detalle que realiza el investigador de la producción oral de estos informantes. Por lo general, los estudiantes confunden las oclusivas sordas y las sonoras, y a pesar de que algunos son conscientes de la diferencia entre ambos tipos, no han desarrollado la capacidad de pronunciarlas bien. Se observa una tendencia a pronunciar las oclusivas sordas como sonoras, y, en menor medida, a pronunciar las sonoras como sordas. Otra tendencia detectada en lo que concierne a las oclusivas sonoras es que cuando estas están en posición intervocálica, los estudiantes no las relajan, es decir, no las realizan como aproximantes. En relación con las líquidas, el investigador señala dos problemas principales. Por un lado, la dificultad de los estudiantes para pronunciar la vibrante múltiple, ya que en la mayoría de los casos no la pronuncian o lo hacen mal. Esto se debe a, como indica Zhi, la laguna fonológica de los alumnos, ya que este es un sonido que no existe en el chino mandarín. Por otro lado, los estudiantes confunden la vibrante simple con la lateral, y articulan [l] en lugar de [r]. En lo que concierne a las consonantes nasales, el principal problema que observa Zhi (2011) es la escasa nasalización de las mismas. Y, por último, el investigador detecta otra serie de errores menos frecuentes, especialmente relacionados con el fonema /s/, como el alargamiento de s final, la pronunciación de la misma como [ʃ] o la realización [θ] de dicho sonido.

Liu (2012) también analiza los errores y la evolución de la interlengua de los estudiantes sinohablantes. La autora realiza una breve comparación entre los sistemas fonéticos del chino y del español basada en la caracterización de los sonidos del chino mandarín realizada por Cortés Moreno (2009a).

Lu (2012) estudia las dificultades que surgen cuando los estudiantes sinohablantes están aprendiendo español como lengua extranjera, y se centra en las relacionadas con la adquisición del sistema consonántico. Al igual que los anteriores trabajos, este está

realizado desde un enfoque contrastivo, y los aspectos que indica como problemáticos son, por lo general, los mismos que apuntan las investigaciones previas realizadas por otros autores: dificultad para discriminar y pronunciar la oposición sorda-sonora de las oclusivas, la pronunciación de las vibrantes y las interdental, ausentes en el idioma chino, la distinción entre las dos clases de vibrantes y la lateral, y la pronunciación de los grupos consonánticos, especialmente aquellos compuestos por una vibrante o una lateral.

La investigadora Muñoz Torres (2012) señala que la teoría propuesta por Flege (1987) no es plausible en lo que se refiere a los estudiantes sinohablantes de ELE. Como ya se ha apuntado anteriormente, según esta teoría los sonidos idénticos en la L1 y en la L2 y aquellos nuevos sin equivalencia en la L2 no supondrán dificultades en el aprendizaje, mientras que aquellos similares en la L1 y en la L2 son los que producirán el efecto de interferencia. Ejemplo de ello es que los alumnos sinófonos encuentran dificultades en la articulación de la vibrante múltiple pese a que este sonido no es propio de su L1 (Muñoz Torres, 2012: 212). Además, la autora presenta una síntesis de la bibliografía publicada hasta el momento en torno a las dificultades que presentan los estudiantes sinohablantes en la adquisición del español como lengua extranjera.

Resulta de gran interés el trabajo de López (2012), ya que realiza un estudio experimental en el que analiza los problemas de pronunciación del español de un grupo de sinohablantes que aprenden español en España. La investigadora compara la descripción del sistema vocálico del chino que hacen Ortí (1990), Wang (2001) y Cortés Moreno (2009a, 2009b), y señala que todos coinciden en la existencia de cinco fonemas vocálicos tanto en chino como en español: /i/, /e/, /a/, /o/ y /u/. Además, todos ellos añaden dos vocales, aunque, como se ha visto, existen matices diferenciadores entre las vocales descritas por los autores.

En relación a la vocal /u/ hay diferentes matices según los autores, para Wang (2009:91) tiene un alófono [w], que describe como *vocal alta*, redondeada y que se corresponde con el alófono [u] posterior, cerrada, redondeada, de Cortés. La vocal cerrada /i/, presenta dos alófonos, tanto para Wang como para Cortés, [i] e [y], esta segunda redondeada. Para la vocal /e/ Cortés (2009:4) presenta dos alófonos: [e] anterior, media semicerrada y, [ə] central, media semicerrada, y Wang presente los alófonos [e] y [ɤ]. [ɤ] está descrita como un sonido que se produce con los labios entreabiertos y con una pronunciación entre el sonido vocálico [e] y [o]. Su articulación es semejante a la del alófono de la consonante castellana /g/ [ɣ] fricativa, velar, sonora. (Wang, 2001:92) | Respecto a la vocal baja media [a], descrita como central anterior, abierta (RAE 2011:46), Ortí (1990) añade para /a/ otro alófono [ɑ], como posterior, abierta, sin especificar su grado de redondeamiento, si lo hubiere. (López, 2012: 15-16)

La investigadora, con el fin de unificar conceptos, sigue la propuesta de Cortés Moreno (2009), aunque también tiene en cuenta los matices expuestos por los otros expertos. Tras

haber realizado este análisis y dadas las diferencias entre ambos sistemas vocálicos, López (2012) realiza una previsión de problemas relacionados con el sistema vocálico:

La realización de la [i], al tener una gama de sonidos más amplia, [i], [y] e incluso [u] mucho más redondeada que la [u] del español, puede plantear problemas. Lo mismo sucede con los sonidos [e] y [o], pues en chino hay una gama que transcurre entre [e], [ə], [o] y [ɤ] y como se mencionó anteriormente, en chino las vocales /e/ y /o/ suelen ir acompañando a otra vocal, en los diptongos (Ortí, 1990) y no junto a consonante. (López, 2012: 16)

En cuanto a la previsión de errores derivados de los fonemas consonánticos, señala los siguientes: confusión entre la lateral [l] y las vibrantes simple [r] y múltiple [ʀ]; confusión entre las oclusivas sordas [p], [t] y [k] y las oclusivas sonoras [b], [d] y [g]; confusión entre la fricativa [s] y la [θ]; confusión entre la fricativa [f] y la [θ]; y confusión entre la lateral [ʎ] y la fricativa [j]. (López, 2012: 17-18)

En este trabajo, la investigadora realiza la caracterización acústica de los segmentos vocálicos /i/, /e/, /a/, /o/ y /u/, midiendo el punto medio de los formantes 1 y 2 de las mismas. También analiza los segmentos consonánticos oclusivos /p/-/b/, /t/-/d/; consonante interdental fricativo /θ/, así como las consonantes róticas /r/ y /ʀ/. Del análisis de las vocales producidas por los informantes, la autora extrae los siguientes datos que compara con los obtenidos por A. Quilis en sus estudios descriptivos sobre la lengua española:

	/i/	/e/	/a/	/o/	/u/
F1	370.2	532.9	964.7	542.3	394.9
F2	2631.7	2374.4	1614	1018.4	848.5

Tabla 1. Valores medios de las vocales tónicas producidas por las mujeres. (López, 2012: 43)

Por lo general, a pesar de que se observa cierto retroceso en la realización de /i/ que se ve reflejado en el descenso de las frecuencias de F2, las vocales cerradas presentan una mayor abertura en el caso de las mujeres, ya que las frecuencias de F1 son más elevadas; las vocales medias también presentan mayor abertura y adelantamiento que las propias del español; la realización de /e/ se acerca a [ɛ]; la /o/ adquiere una posición similar a la propia de [ɤ], y la vocal baja /a/ es más abierta que la española.

	/i/	/e/	/a/	/o/	/u/
F1	325.1	415.1	702.3	507.6	383
F2	2215	2089.8	1344.6	963.4	835.1

Tabla 2. Valores medios de las vocales tónicas producidas por los hombres. (López, 2012: 46)

En cuanto a la producción de los hombres, se observa que las vocales cerradas /i/ y /u/ presentan una elevación de las frecuencias del F1, por lo que la abertura bucal es mayor; sin embargo, F2 desciende en /i/, lo que se interpreta como un retroceso, y aumenta en /u/, lo que se traduce como un adelantamiento; las vocales medias /e/ y /o/ presentan elevación de frecuencias tanto en F1 como en F2; en cuanto a la vocal baja /a/, se observa una ligera elevación de frecuencias en ambos formantes. Tal y como señala la investigadora, tanto en el caso de las mujeres como en el de los hombres se observa una proporción entre los formantes de las vocales producidas por los informantes, aunque la producción de /i/ es la que presenta valores más alterados y diferenciados con respecto a los valores medios de la lengua española. (López, 2012: 46-47)

Tras realizar el análisis contrastivo de las dos lenguas, y el estudio de los errores, la investigadora hace el siguiente listado de problemas a los que se enfrentan los estudiantes sinohablantes de ELE: 1) la percusiva [ɾ] y la vibrante [r] son producidas como lateral [l]; 2) las consonantes oclusivas sonoras [b] [d] [g], son producidas como oclusivas sordas [p] [t] [k], y las aproximantes [β] [δ] [ɣ] son realizadas como oclusivas sordas; 3) la interdental [θ] es producida como dentoalveolar [s]; y 4) en cuanto a las vocales, la vocal baja [a] y las vocales medias [e] y [o] son producidas con mayor abertura, la vocal cerrada [u] es producida también con mayor abertura, y la [i] es producida con mayor abertura bucal, pero con mayor retraimiento lingual. (López, 2012: 49-50)

Alcoholado (2012) realiza una propuesta de modelos de actividades con la sílaba como unidad estructural que “puede servir de trampolín para un posterior control de otras unidades discursivas de mayor envergadura” (Alcoholado, 2012: 71). En 2013, el investigador centra sus trabajos en el análisis de las vocales en distintos contextos. El autor, al igual que Cortés Moreno (2009) y Wang (2001) indica que “la sílaba china es más simple que la española y carece de grupos consonánticos” (Alcoholado, 2013a: 34), lo que hace que la articulación de los distintos tipos de sílabas del español sea un reto para los sinohablantes. “(L)a pronunciación de combinaciones de consonantes en posición marginal (...) supone una dificultad particular para los alumnos chinos, sobre todo en su producción” (Alcoholado, 2013a: 34). Además, asume que aquellas combinaciones vocálicas inexistentes en chino también suponen un obstáculo, especialmente cuando dichas combinaciones aparecen entre palabras gráficamente separadas. Alcoholado también observa una tendencia por parte de los estudiantes sinófonos de ELE a pronunciar sílaba por sílaba.

Blanco y Nogueroles (2013) realizan una interesante descripción y categorización de los errores fónicos producidos por estudiantes de español como lengua extranjera. Las autoras clasifican los errores según cuatro parámetros: el primero es la naturaleza lingüística, es decir, si existe o no una equivalencia interlingüística entre la L1 y la L2; el segundo es el segmento al que afecta dicho error, es decir, si se trata de tipo segmental o no segmental, y además se especifica si es vocálico o consonántico, o si es silábico, de intensidad o de entonación; el tercer parámetro tiene que ver con la acción que se desencadena como consecuencia de la dificultad que experimenta el estudiante y que da lugar al error, como por ejemplo, la adición, omisión, sustitución, modificación o desplazamiento; finalmente, el cuarto parámetro está relacionado con el efecto que dicho error ocasiona en la comunicación (Blanco y Nogueroles, 2013: 206). Las investigadoras ofrecen una lista detallada de los problemas detectados en la producción oral de los informantes sinohablantes (Blanco y Nogueroles, 2013: 217-218):

- Diptongación de /o/ y /e/.
- Sustitución de los fonemas oclusivos sonoros por sus correspondientes sordos y, por tanto tendencia a no diferenciar entre /p/ y /b/, /t/ y /d/, /k/ y /g/.
- Ensondecimiento de las consonantes sonoras.
- Falta de ensondecimiento de las consonantes sordas.
- Pronunciación como oclusivas de las aproximantes [β] [ð] [ɣ].
- Sustitución de la vibrante simple /r/ por /l/, pues /r/ existe como alófono de /l/.
- Inserción de una [d] epentética delante de /r/.
- Pronunciación aproximante de las vibrantes [r] y [r], como el inglés [ɹ].
- Sustitución de la /θ/ por /s/.
- Simplificación de los grupos consonánticos en palabras como TRaCTor, PRaCTicar o iNSCRiPCión, en las que aparecen varias consonantes seguidas, ya sea en una misma sílaba o en sílabas contiguas.
- Sustitución de /x/ por [h].
- Pronunciación insuficientemente nasalizada de /n/.
- Pronunciación insuficientemente nasalizada de /m/.
- Desplazamiento de los patrones acentuales del chino al español, transformando, así, una palabra llana en una palabra aguda.
- Desplazamiento, menos generalizado, del acento hacia el principio de la palabra convirtiendo palabras llanas en esdrújulas.
- Modificación de los contornos entonativos por influencia del chino.
- Modificación del énfasis en los enunciados interrogativos por una excesiva atención a la zona inicial y desatención a la inflexión final, la zona entonativa más rica.
- Modificación de la zona tonal de los enunciados declarativos, con una realización más alta que lo habitual en español, aproximándose al entonema declarativo (.) en chino.
- Modificación del ritmo y pronunciación sílaba por sílaba, en lugar de agrupar estas sílabas en palabras y grupos fónicos.
- Oscilaciones melódicas intrasilábicas, principalmente, en las sílabas tónicas y en el primer pico de la curva melódica.
- Modificación de la entonación realizando una anacrusis ascendente, en lugar de llano.

- Modificación de la entonación realizando un primer pico en la zona semialta, en vez de en la zona semibaja, del campo tonal del discurso.
- Modificación de la entonación realizando inflexiones finales con pendientes insuficientes, tanto las descendentes como las ascendentes, e incluso inflexiones que siguen la dirección opuesta ascendente o descendente.

Tal y como se puede observar, exceptuando la realización de las vocales, en cuyo caso solo se ha detectado un problema, las dificultades señaladas por las investigadoras ya han sido detectadas anteriormente por otros autores.

Chih (2013), analiza los problemas de percepción de los estudiantes de ELE sinohablantes. Para ello, analiza la discriminación de distintos pares mínimos, y las dificultades, en orden de más a menos problemáticas son el contraste de [k] - [g]; [t]- [d]; [l] - [r] en posición final; [p] - [b]; [l] - [r] en posición inicial; [s] - [θ]; [r] - [d]; [r] - [d]; y contraste de [r] - [rr] (Chih, 2013: 24). Más adelante, el investigador realiza una segunda audición con el mismo grupo de informantes, y observa que las diferencias en la percepción son poco significativas, y que en casos como la discriminación de [p]-[b], [r]-[d] el porcentaje de errores aumenta, mientras que en relación con [k]-[g], [t]-[d], [l]-[r], y [s]-[θ] dicho porcentaje disminuye. (Chih, 2013: 27)

Planas Morales ya realiza en 2008 un estudio contrastivo entre los sistemas fonéticos del chino y del español. Sin embargo, en su trabajo de 2013 profundiza en el análisis contrastivo de estas dos lenguas, más concretamente de los sistemas vocálicos, lo que, tal y como indica la autora, permite encontrar claves para la enseñanza y corrección de errores de las vocales en la lengua extranjera. La investigadora señala que los errores y dificultades principales son el timbre inestable de [a], la sustitución de la vocal [e] por la vocal neutra y la pronunciación [u] de /o/. De acuerdo con la autora, quien sigue el trabajo de Lin (2007), el sistema vocálico del chino mandarín posee solamente cinco fonemas, aunque estos poseen un amplio número de alófonos que varían dependiendo de la distribución silábica del fonema en el núcleo silábico. Ejemplo de ellos es que los fonemas vocálicos /i/, /y/ y /u/, poseen los alófonos [j], [ç] y [w], respectivamente, que funcionan como consonante, cuando no son la vocal que nuclear de la sílaba. Planas Morales (2013) indica los problemas de timbre que presentan las vocales realizadas por los estudiantes sinohablantes al diferente número de alófonos del español y del mandarín ya que su distribución en el espacio vocálico es una diferencia muy importante entre los dos sistemas. Por otro lado, también hay que tener en cuenta (siempre según Planas Morales) que la transcripción de los caracteres chinos a la escritura alfabética romanizada o el paso del alfabeto a la transcripción fonética suponen un cambio conceptual para los



alumnos de español, lo que también es probable que sea causa de los problemas a los que se enfrentan en la adquisición del sistema fonético de esta lengua.

Miranda Márquez (2014), al igual que la mayoría de los investigadores anteriormente citados, realiza un análisis contrastivo del chino y el español e indica cuáles son las mayores dificultades a las que se enfrentan los estudiantes sinohablantes. Ya se ha señalado que autores como Cortés Moreno (2009a, 2009b) o Alcoholado (2013) indican que la estructura silábica del chino es más sencilla que la del español, y Miranda Márquez (2014) suscribe sus palabras indicando que mientras que en español hay nueve tipos de combinaciones silábicas posibles (si se tiene en consideración la parte vocálica), en chino solo hay cuatro: CV; CVC; V y VC (Miranda Márquez, 2014: 56-57). Según el investigador,

(e)sta compleja variedad de combinación silábica del español frente al chino provoca que, en ocasiones, resulta difícil para los sinohablantes distinguir algunas de las sílabas del español, a veces separando en dos o más partes una sola sílaba para adaptar el sonido original al sonido más cercano en su lengua materna, y, cuando sucede esto, pueden surgir ciertas alteraciones semánticas con respecto al enunciado original. (Miranda Márquez, 2014: 56-57)

En relación con las vocales del chino, cabe señalar que este autor, al igual que Ortí Mateu (1990), Wang (2001) y Cortés Moreno (2009a, 2009b), distingue siete vocales diferentes: /a/, /o/, /ə/, /i/, /u/, /y/, y /e/. Sin embargo, a pesar de que realiza una descripción de las mismas, no explica cuáles son las dificultades a las que se enfrentan los estudiantes sinohablantes a la hora de pronunciar las vocales del español.

En su trabajo de 2014, Poch e Igarreta, además de realizar una revisión de la bibliografía publicada hasta el momento en relación con los problemas de los estudiantes de ELE sinohablantes, realizan un análisis auditivo, de tipo cualitativo, de la producción de un grupo de estudiantes universitarios de primer año que estudian español como lengua extranjera. Este trabajo permite obtener una visión general de las dificultades que tienen estos hablantes cuando comienzan a aprender español. Contrariamente a otros autores, en este trabajo se señala que las vocales del español realizadas por estudiantes sinohablantes presentan importantes diferencias en sus características con respecto a las realizadas por hispanohablantes nativos. En relación con los grupos vocálicos, se observa que, generalmente, los diptongos son realizados como hiatos y no suelen pronunciar los triptongos. En cuanto al timbre de las vocales del español producidas por este grupo, las autoras señalan que se observa una gran inestabilidad en el mismo, y que es especialmente inestable en el caso de las vocales realizadas por el grupo de informantes oriundos de

Shanghái, ya que presentan un porcentaje de realizaciones adecuadas menor que el de los hablantes de otras regiones. De acuerdo con las autoras,

(l)as realizaciones vocálicas inadecuadas suelen ser claramente más abiertas que las que produciría un hispanohablante; en un número importante de casos los alumnos emiten el sonido [ə] en vez de [a] o de [e]; en lugar de la vocal [u] en la palabra *duda* aparece la vocal labializada [y] y, finalmente, en otras ocasiones los estudiantes producen una vocal de timbre distinto al esperado. (Poch e Igarreta, 2014: 150)

Además, las investigadoras apuntan que los datos apuntan a que el comportamiento diferenciado de los alumnos de Shanghái frente al de los de otras regiones se puede deber al perfil lingüístico de los hablantes o, especialmente, a las características del chino hablado en dicha zona.

En relación con la realización de los sonidos consonánticos, Poch e Igarreta (2014: 151) señalan que todos los alumnos presentan realizaciones sordas de las oclusivas sonoras, y son contadas las ocasiones en las que articulan sonidos aproximantes. En cuanto a las fricativas, no han detectado problemas con la consonante labiodental [f], y aunque tampoco los hay con los sonidos [θ] y [s], se ha observado que algunos hablantes han optado por la pronunciación seseante, otros por la distinguidora, y otros alternan ambas realizaciones. En cuanto a la fricativa velar [x], los alumnos siempre realizan la aspirada [h], lo cual puede deberse, según las autoras, a la influencia del inglés. Las consonantes africadas no suponen ningún problema para los informantes sinohablantes, lo mismo que las nasales [m] y [n], las cuales han sido realizadas correctamente. En relación con la consonante lateral alveolar, los hablantes no tienen problemas de pronunciación, aunque las investigadoras han detectado algunas realizaciones velarizadas. Además, los estudiantes procedentes de Anhui y Guizhou realizan [n] en lugar de [l], confusión que, de acuerdo con las autoras, “podría estar relacionada con la variedad de chino mandarín propia de esas zonas o con los dialectos de los estudiantes” (Poch e Igarreta, 2014: 153). Otro sonido que plantea grandes dificultades es la consonante lateral palatal [ʎ], la cual en pocas ocasiones es realizada como tal, y los estudiantes la realizan como [i], [ʒ] o incluso la eliden. En cuanto a las vibrantes, en la lista de palabras que constituye el corpus de análisis no aparecen ejemplos de vibrante múltiple, por lo que las investigadoras no han podido analizar la realización de dicho sonido. Sin embargo, sí aparecen casos de la vibrante simple. En este caso, las autoras señalan que “(e)l 89% de los estudiantes pronuncia la consonante lateral en lugar de la percusiva. Aquellos alumnos que emiten un sonido parecido a la percusiva realizan el sonido retroflejo [ɾ] propio del inglés” (Poch e Igarreta, 2014: 154).

En cuanto a la realización de los grupos consonánticos, las autoras señalan que aquellas palabras que contienen un grupo consonántico en su primera sílaba

han constituido un escollo casi insalvable por la mayor parte de los alumnos analizados puesto que muy pocos han podido realizarlas. La mayor parte de los estudiantes han reducido dichos grupos a una sola consonante, en general la primera de las que integran el grupo, desembocando las realizaciones de dichas palabras en una estructura silábica del tipo CV que es la propia del chino mandarín. (Poch e Igarreta, 2014: 155)

Las autoras apuntan que este problema está directamente relacionado con la estructura silábica de la lengua de los estudiantes, aspecto señalado ya por distintos autores.

En 2015 Igarreta Fernández publica dos trabajos. En el primero, la investigadora compara los sistemas fonéticos del español, el chino mandarín y el cantonés. Para ello, analiza el sistema consonántico y el vocálico de estas lenguas, además del sistema tonal y la estructura silábica de las mismas. En relación con este aspecto, la autora señala que la sílaba del mandarín está constituida por la parte inicial y la parte final o rima. El inicio lo forma un sonido consonántico, mientras que el final está constituido por el prenúcleo, el núcleo y la coda, los cuales son una semiconsonante, una vocal y una semivocal o nasal respectivamente. (Igarreta Fernández, 2015a: 264).

El segundo estudio de la investigadora es un trabajo de fonética aplicada cuyo objetivo es mostrar cómo se puede acercar la pronunciación de los estudiantes sinohablantes a la propia del español. Con el fin de lograr dicho objetivo, la autora realiza, en primer lugar y mediante un análisis acústico, el diagnóstico de los problemas que presentan dichas hablantes (Igarreta Fernández, 2015b). La autora detecta que los errores de tipo segmental más comunes entre el grupo de informantes sinohablantes son: pronunciación de [l] como [r]; pronunciación de [r] como [l] o [n]; elisión de [r]; pronunciación de [b] como [p]; pronunciación de [d] como [t]; y pronunciación de [g] como [k]. (Igarreta Fernández, 2015b: 191). Tras esto, se diseña una secuencia didáctica basada en los principios del método verbo-tonal. Los resultados de este trabajo muestran que la aplicación de dicha secuencia contribuye a mejorar la pronunciación de las informantes.

El trabajo de Planas Morales (2016) trata la enseñanza y evaluación de la pronunciación de los estudiantes sinohablantes en el aula de español como lengua extranjera. La investigadora indica que “los estudiantes chinos tienden a pronunciar los sonidos españoles según el sistema fonético del chino mandarín” (Planas Morales, 2016: 14), aunque, según la autora, los errores relacionados con la articulación no suponen un problema de comunicación grave, mientras que aquellos relacionados con el ritmo y la entonación sí lo son.

Pose (2017) también emplea el análisis contrastivo para determinar las posibles áreas de interferencia entre el chino y el español. El investigador señala algunos de los fenómenos que afectan al nivel fonético-fonológico de la L2, como la confusión entre las líquidas del español, o entre las oclusivas sordas y sonoras. Además, el investigador apunta que “(e)s frecuente que los sinohablantes articulen débilmente algunas consonantes en posición implosiva (...) o que no las pronuncien” (Pose, 2017: 9), aspecto que puede verse reflejado en la escritura, por ejemplo, en la de la palabra *bachiller* como *bachille*. El autor señala que “(e)s posible que este debilitamiento de consonantes -y otros fenómenos como la reducción de grupos consonánticos- que presentan los sinohablantes tenga entre sus causas las diferencias de estructura silábica que pueden observarse entre el chino y el español” (Pose, 2017: 9). En cuanto a la sílaba del chino, Pose sigue los trabajos de Duanmu (2007) y Cortés Moreno (2009a) e indica que está compuesta por dos partes: el inicio, el cual puede estar constituido por una consonante, y el final, el cual tiene una estructura más compleja. De acuerdo con el investigador, la sílaba china tiene una estructura más restringida que la del español, ya que en esta lengua no se dan grupos consonánticos y “las consonantes en final de sílaba se limitan a las nasales (alveolar y velar) y a la retrofleja” (Pose, 2017: 9). Por este motivo, señala que “(l)a transferencia de rasgos de la sílaba china puede explicar las reducciones que se observan en la interlengua de los sinohablantes” (Pose, 2017: 9). Sin embargo, el autor indica que

(l)os límites de la sílaba no están delineados solamente en el nivel fonológico, también se ve claramente afectada la pronunciación de los segmentos. Por ejemplo, algunos procesos de asimilación de rasgos, como es la nasalización de las vocales ante consonantes nasales, se producen dentro de la sílaba y no traspasan sus límites. Otro fenómeno relacionado con la delimitación de la sílaba es la presencia de ataque vocálico duro en las sílabas que comienzan con vocales medias o bajas (Duanmu, 2007: 72), este tipo de pronunciación introduce un corte entre estas sílabas y la precedente. (Pose, 2017: 10)

En cambio, en español se da el caso contrario, ya que los procesos que traspasan el límite de la sílaba, como las asimilaciones, son numerosos, y las vocales iniciales no presentan el cierre propio del ataque duro (Pose, 2017: 10). El autor relaciona este aspecto con el carácter *staccato* de la interlengua de los sinohablantes mencionado anteriormente por Cortés moreno (2009a), ya que, tal y como indica, en el español de los sinohablantes suele notarse una demarcación especial (Pose, 2017: 11). Ejemplo de ello es que aparecen cortes entre vocales que, generalmente, no están presentes en español, como en el caso de *la obra*, en el que el ataque duro afecta a la transición entre vocales y hay un corte entre la vocal final y la inicial. El autor indica que esta tendencia a reforzar la frontera silábica también afecta al resilabeo característico del español. Por ejemplo, en español la la

estructura silábica más frecuente es CV, por lo que en secuencias como *el amor* se da un proceso de resilabificación para alcanzar dicha estructura, lo que no sucede en el habla de los sinohablantes. Pose indica que los señalados errores de resilabeo y los cortes glóticos del ataque duro, además de otros fenómenos particulares relacionados con el ritmo y la entonación, dan como resultado una pronunciación entrecortada, con variaciones tonales dentro de una misma sílaba, y diferentes de una a otra, lo que puede dificultar la comunicación y afectar a segmentos menores, como consonantes y vocales (Pose, 2017: 11). El investigador señala que el tratamiento de estos problemas puede ayudar a mejorar la capacidad comunicativa de los sinohablantes, por lo que propone distintas herramientas para que el profesor pueda actuar frente a este tipo de problemas.

La tesis doctoral de Pérez García (2018) supone una importante aportación en el análisis de las vocales tónicas del español en sílaba libre realizadas por mujeres sinohablantes cuyo nivel de competencia en español es B1-B2. En primer lugar, el investigador analiza la duración de los sonidos vocálicos, y los valores medios obtenidos (en milisegundos) pueden verse en la tabla 3.

[a]	[e]	[i]	[o]	[u]
133	122	119	122	122

Tabla 3. Duración media de las vocales del español por informantes sinohablantes. Adaptado de Pérez García (2018: 163)

Tal y como se puede observar, la [i] tiene la duración más breve, [e], [o] y [u] tienen la misma duración media, y [a] es la más larga de las vocales. Pérez García (2018: 163) compara estos resultados con los obtenidos en el análisis de la producción de un grupo de informantes cuya lengua materna es el español, e indica que todas las vocales producidas por las estudiantes chinas tienen una duración media mayor que las producidas por las hispanohablantes nativas.

En segundo lugar, el investigador analiza la estructura del primer y segundo formante de estas vocales en tres puntos: inicial (punto 1), medio (punto 2) y final (punto 3). Los resultados del análisis aparecen en la tabla 4:

		<b>Pto. 1</b>	<b>Pto. 2</b>	<b>Pto. 3</b>
[i]	F1	431.2	444.7	441.3
	F2	2511.5	2644.0	2541.4
[e]	F1	620.2	671.4	634.7
	F2	2077.9	2204.2	2150.1
[a]	F1	812.2	954.6	796.2
	F2	1712.5	1695.4	1777.6
[o]	F1	579.5	613.8	558.9
	F2	1243.6	1126.0	1198.7
[u]	F1	429.4	445.0	440.7
	F2	1134.8	1013.2	1262.7

*Tabla 4.* Frecuencias medias de F1 y F2 en los tres puntos analizados en cada vocal (Adaptado de Pérez García, 2018: 131-156)

En relación con el primer formante, el investigador señala que las oscilaciones internas del mismo se deben a la interferencia del tono de la lengua materna (Pérez García, 2018: 144). Las vocales más abiertas son las que presentan una mayor variación, de manera que [a] es la vocal que mayor oscilación presenta, [e] y [o] presentan una tendencia hacia la apertura inicial y posterior cierre, y [u] e [i] muestran una mayor tendencia a la estabilidad formántica (Pérez García, 2018: 144). Sin embargo, lo que diferencia a estas vocales de las articuladas por los hablantes de lengua materna española son dichas oscilaciones, ya que las vocales realizadas por los hablantes chinos aparecen en el mismo rango de frecuencias que las españolas (Pérez García, 2018: 145). En relación con el segundo formante, el investigador señala que también en este hay oscilaciones internas, especialmente cuando son sonidos adelantados o atrasados. Las vocales anteriores [i] y [e] presentan un mayor grado de variabilidad, y en el inicio los valores ascienden, y posteriormente descienden; las vocales posteriores [o] y [u] presentan menor variabilidad que las anteriores, y el comportamiento de estas es diferente, ya que en el inicio hay un descenso, y posteriormente un ascenso. La vocal central [a] es la más estable (Pérez García, 2018: 158-159).

El investigador concluye que “las dificultades en la adquisición y producción de las vocales del español para estas informantes están relacionadas con la clasificación equivalente” (Pérez García, 2018: 249). Sin embargo, de acuerdo con los resultados de su investigación, el tono es la principal fuente de interferencia en la realización de los sonidos vocálicos, y las oscilaciones tonales alejan la articulación de la de los hablantes

nativas, aspecto que también está relacionado con la mayor duración de las vocales, la cual provoca oscilaciones en el timbre (Pérez García, 2018: 249).

Jiménez y Tang (2018) también analizan la producción del sistema vocálico del español por hablantes de chino. Para ello, en primer lugar realizan un análisis contrastivo de los sistemas vocálicos del español y del chino. Los autores aluden a la falta de consenso en lo que respecta a las unidades que componen el sistema vocálico chino y siguen la propuesta de Duanmu (2007: 47) para el chino estándar, según la cual las vocales de esta lengua se clasifican como aparece en la tabla 5:

	Anteriores	Centrales	Posterior
Altas	i / y		u
Media		ə	
Baja		a	
	No redondeadas		Redondeadas

Tabla 5. Clasificación de los sonidos vocálicos del chino propuesta por Duanmu (2007: 47) (Extraído de Jiménez y Tang, 2018: 16)

De acuerdo con esta clasificación, el chino, al igual que el español, tiene cinco vocales fonológicas, aunque las características de las mismas son distintas de las del español. Como señalan los autores, “(a)mbas lenguas coinciden en presentar las tres vocales periféricas extremas /i, u, a/; el chino, en cambio, cuenta con una vocal central media /ə/ y una vocal redondeada /y/, alta y anterior como /i/” (Jiménez y Tang, 2018: 15). Además, en este trabajo los investigadores apuntan que “el grado de variación en las vocales del español no es equiparable al que presentan algunas vocales del chino como /ə/ y, en menor medida, /a/” (Jiménez y Tang, 2018: 16).

En segundo lugar, en este trabajo se estudia la producción de un grupo de estudiantes chinos cuyo nivel de conocimiento del español es, tal y como indican, intermedio. Para ello, analizan las características de las vocales del chino y sus alófonos pronunciadas por estas informantes, y también las del español. En todos los casos, tanto el sonido consonántico anterior, como el posterior, es la oclusiva sorda [t]. El resultado del análisis acústico de estas últimas se muestra en la tabla 6:

Vocal	Casos	F1		F2		Duración	
		Media	DE	Media	DE	Media	DE
A	42	861	89,38	1795	117,81	129	26,51
E	42	551	68,91	2355	149,32	122	24,22
O	42	530	78,69	1181	197,20	129	29,90
I	42	352	56,36	2815	128,22	102	24,76
U	42	366	54,33	922	136,82	117	29,34

*Tabla 6.* Media y desviación estándar de F1 y F2, y de la duración de las vocales españolas pronunciadas por hablantes chinas. (Extraído de Jiménez y Tang, 2918: 29).

Estos datos, junto con el análisis de las áreas de dispersión de las vocales han permitido a los investigadores concluir que “el sistema vocálico español de los estudiantes chinos es estructuralmente similar al sistema nativo, pero con algunas diferencias fonéticas” (Jiménez y Tang, 2018: 13). La primera de ellas es que las vocales de estos hablantes son más largas que las de los nativos, aunque más breves que las propias del chino. La segunda es que, como consecuencia de la transferencia del chino al español, el rango de frecuencias en el que aparecen las vocales de estos informantes es mayor que las propias del español. Relacionado con estos dos aspectos, los autores indican que “la mayor abertura de las vocales medias y de la vocal /a/ en español como L2” no depende de la duración, si no que es una propiedad que transfieren de su lengua materna (Jiménez y Tang, 2018: 31). Este trabajo sirve, además, para mostrar que las vocales /a/, /e/ y /o/ son las que muestran mayor variabilidad en sus realizaciones.

### **3. RECAPITULACIÓN**

**En este segundo capítulo, dedicado al marco teórico en el que sitúa la investigación realizada en esta tesis, se han presentado, en primer lugar desde una perspectiva diacrónica, los modelos que ofrecen una explicación sobre el proceso de adquisición general de segundas lenguas y, en segundo lugar, se ha ofrecido una síntesis sobre los estudios dedicados al análisis de la adquisición del español por sinohablantes.**

**Los modelos examinados en el primer apartado del capítulo han puesto de relieve la evolución teórica que se ha producido desde los estudios pioneros de Polivanov y Troubetzkoy, pasando por el Análisis Contrastivo tradicional y acabando en los modelos más actuales surgidos de los trabajos de E. Flege. Como se ha visto dichos**



modelos han tomado como base datos obtenidos del análisis del proceso de adquisición de lenguas diversas entre las cuales, rara vez se ha tomado en consideración el caso del aprendizaje del español por sinohablantes. Es lógico que así haya sido pues la enseñanza de español al público chino comienza a tomar cuerpo a finales de la década de 1990 y es lógico, por tanto, la ausencia de datos sobre este proceso hasta esa fecha.

Como se señala en la segunda parte, dedicada al caso particular en el que las lenguas en presencia son el chino y el español, en muchos casos, las hipótesis sobre la adquisición del español por chinos no parecen reflejar los distintos fenómenos de transferencia entre lenguas que mencionan muchos de los autores que han elaborado modelos de adquisición. Ello ha conducido a que, como se ha visto, casi todos los autores que han investigado sobre la cuestión hayan adoptado un punto de vista muy tradicional tomando como base el análisis contrastivo para establecer una predicción de las posibles dificultades de los estudiantes.

Se han rastreado los estudios sobre la adquisición de la pronunciación del español por sinohablantes desde los primeros trabajos publicados hasta los más recientes, aparecidos en 2018, y el punto de vista de sus autores no ha variado con respecto a la cuestión del análisis contrastivo. La mayoría de ellos coinciden en señalar los siguientes problemas en el proceso de aprendizaje: a) confusión entre las vibrantes simple y múltiple y la lateral; b) confusión entre las consonantes oclusivas sordas y sonoras; c) confusión de la interdental [θ] y de [s]; e) algunos problemas de acentuación del español; y f) grandes problemas de adquisición de los contornos entonativos ya que el chino es una lengua tonal.

Merecen mención aparte, por su carácter sorprendente, las pocas observaciones que los distintos autores realizan sobre los problemas que presentan la realización de las vocales. La mayoría de ellos indican que el sistema vocálico del español está “comprendido” en el del chino mandarín, que consta de un mayor número de unidades, cinco de las cuales “coinciden” con las vocales del español. Como es sabido, no existen sonidos idénticos entre las lenguas y, por tanto, cabe suponer que las observaciones de los distintos autores se sitúan en la perspectiva de la “inteligibilidad”. Ello significa que los timbres de las vocales del español realizadas por sinohablantes no plantean problemas de inteligibilidad en la comunicación. Obviamente, frente a los casos de ensordecimiento de todas las oclusivas sonoras (la palabra *duda* es realizada como [‘tuta]) que dificultan la comprensión el profesor de

**ELE puede estimar que las diferencias de timbre vocálico que no inciden en la comprensión de los enunciados constituyen un problema “menor”.**

**No obstante, la elaboración de modelos que expliquen los fenómenos de transferencia e interferencia en el proceso de aprendizaje de una lengua extranjera pasa por disponer de datos precisos, como los que puede proporcionar la fonética experimental, sobre qué perciben y qué pronuncian los estudiantes, sinohablantes en este caso, cuando deben adquirir la pronunciación de una lengua extranjera (español en el presente caso).**

**Tal como se ha indicado en la Introducción, esta investigación se plantea como objetivo proporcionar datos precisos sobre la adquisición de las vocales del español y sobre los fenómenos que dicho proceso provoca en el conjunto de la cadena fónica con el fin de constituir una contribución a la comprensión del complejo proceso fonético-fonológico que consiste en adquirir la pronunciación de una lengua extranjera.**

**En el capítulo siguiente, se presentará el diseño experimental de la tesis que ha permitido recoger los datos sobre la adquisición de las vocales del español por sinohablantes.**



**CAPÍTULO III**  
**METODOLOGÍA Y DISEÑO EXPERIMENTAL**



### **III. METODOLOGÍA Y DISEÑO EXPERIMENTAL**

Atendiendo a los objetivos expuestos en la introducción en el presente capítulo se expone la metodología de trabajo seguida y el diseño experimental que ha permitido obtener los datos referentes a la adquisición de las vocales del español por alumnos sinohablantes. Como ya se ha indicado, se ha trabajado con un grupo de informantes cuya lengua materna es el chino mandarín y con un grupo de control integrado hispanohablantes cuyas realizaciones vocálicas han sido consideradas como referencia de la pronunciación del español por nativos. Posteriormente se han realizado dos tipos de análisis de las grabaciones, el primero de ellos auditivo, y el segundo acústico. Finalmente, se ha llevado a cabo un tratamiento estadístico de los datos obtenidos en el análisis acústico. En este capítulo se describen los criterios seguidos para la selección de las informantes y para el diseño del corpus, se explica en qué ha consistido el análisis auditivo, y se detallan los parámetros analizados en el análisis acústico. Finalmente, se explica el procedimiento seguido para realizar el análisis estadístico de los datos.

#### **1. SELECCIÓN DE LAS INFORMANTES**

Una de las particularidades de la presente tesis doctoral es que se centra en la producción oral en español de informantes femeninas. La mayor parte de estudiantes de español procedentes de China son mujeres de forma que puede casi afirmarse que la composición sociolingüística de la población estudiantil ha “impuesto” que el perfil estudiado sea el correspondiente a alumnas. Por ello, el grupo de control está también integrado por hablantes femeninas. Como es sabido, la mayoría de estudios que describen el sistema vocálico del español lo hacen partiendo de los datos obtenidos a partir de voces masculinas por lo que ha aparecido como indispensable que las hablantes nativas de español que constituyen la referencia fueran también mujeres habida cuenta de las diferencias variaciones existentes en los valores de frecuencias entre las voces masculinas y las voces femeninas. A continuación, se exponen las características de los dos grupos analizados en la presente tesis doctoral, el de control y el de informantes sinohablantes.

## 1.1 Grupo de control

El grupo de control está compuesto por seis mujeres de edades comprendidas entre los 19 y los 25 años. Su lengua materna es el español y la utilizan tanto en su entorno personal, como en el laboral o en el académico. La variedad que hablan es la del País Vasco. Todas ellas tienen formación universitaria, poseen conocimientos de euskera como segunda lengua, adquiridos en la escuela y poseen también conocimientos de inglés, ya que lo han estudiado como lengua extranjera desde el primer ciclo de la Educación Primaria, hasta la finalización de sus estudios de Bachillerato. La tabla 1 muestra la información relativa a cada una de las hablantes que conforman el grupo de control.

Informante	Lugar de nacimiento	L1	L2	Edad de adquisición de la L2	Edad a la que comenzó a estudiar inglés	Lengua(s) que emplea habitualmente
1	Guipúzcoa, España	Español	Euskera	4	6	Español
2	Guipúzcoa, España	Español	Euskera	4	6	Español
3	Guipúzcoa, España	Español	Euskera	4	6	Español
4	Guipúzcoa, España	Español	Euskera	4	8	Español
5	Guipúzcoa, España	Español	Euskera	4	8	Español
6	Guipúzcoa, España	Español	Euskera	3	6	Español

Tabla 1. Perfil lingüístico de las hablantes del grupo de control analizadas en esta investigación.

## 1.2 Informantes sinohablantes

El grupo de informantes sinohablantes está compuesto por seis mujeres de edades comprendidas entre los 18 y los 20 años. Todas ellas poseen el chino mandarín como lengua materna y lo emplean tanto en su vida cotidiana como en el entorno académico

y/o laboral. Cabe señalar que algunas conocen otros dialectos del chino (ver tabla 2), aunque, de acuerdo con la información proporcionada por ellas mismas, únicamente los utilizan en el ámbito doméstico, alternando el uso del mandarín con el de dicho dialecto. Las informantes sinohablantes han estudiado inglés como lengua extranjera desde el primer ciclo de la Educación Primaria y comenzaron a estudiar español como lengua extranjera cuando accedieron a la universidad. A pesar de proceder de diferentes provincias de la República Popular de China, todas ellas son estudiantes de la Licenciatura en Español de la Universidad de Ciencia y Tecnología de Macao. En el momento de realizar las grabaciones para la presente tesis, las estudiantes habían finalizado el primer año de dicha carrera y estaban realizando un curso de español como lengua extranjera en la Universidad Autónoma de Barcelona (julio de 2015). En lo que respecta a su nivel de competencia en español, las informantes presentaban, en el momento de la recogida de datos, un nivel A2<sup>1</sup>.

Informante	Lugar de nacimiento	L1	L2	Edad de adquisición de la L2	Edad a la que comenzó a estudiar inglés	Edad a la que comenzó a estudiar español	Lengua(s) que emplea habitualmente
1	Fujian, China	Mandarín	Dialecto Min	Niña (entorno familiar)	8	18	Mandarín y dialecto Min (con la familia)
2	Hunan, China	Mandarín	Inglés	12	12	18	Mandarín
3	Hunan, China	Mandarín	Dialecto de Changsha	Niña (entorno familiar)	6	17	Mandarín
4	Tianjin, China	Mandarín	Inglés	6	6	19	Mandarín e inglés
5	Shanghai, China	Mandarín	Dialecto Hu	Niña (entorno familiar)	5	18	Mandarín y dialecto Hu (con la familia)
6	Anhui, China	Mandarín	Inglés	8	8	18	Mandarín

Tabla 2. Perfil lingüístico de las informantes sinohablantes analizadas en esta investigación.

<sup>1</sup> El nivel de lengua española de las informantes se ha establecido teniendo en cuenta los niveles de propuestos en el *Marco Común Europeo de Referencia para las Lenguas* (Consejo de Europa, 2002).



## **2. DISEÑO DEL CORPUS**

Tal y como se ha visto en el capítulo II, los estudiantes sinohablantes de español como lengua extranjera hacen frente a numerosas dificultades en el proceso de adquisición del sistema fonético de esta última. Dado que el objetivo principal de la presente investigación es realizar una descripción de todos los problemas que afectan a las vocales del español producidas por las estudiantes sinohablantes, se ha diseñado un corpus que ha permitido observar, además, una serie de problemas derivados de las dificultades de la pronunciación de las vocales a los que se enfrentan los estudiantes de español y que, como se verá, otros autores no han mencionado.

### **2.1 Criterios seguidos para el diseño del corpus**

Como se ha indicado, en el momento de realizar la grabación el conocimiento del español de las informantes sinohablantes era el adquirido durante el primer año de estudios universitarios. El nivel A2 no permite, todavía, que los alumnos mantengan una conversación espontánea natural y fluida en la lengua estudian, por lo que se decidió que el corpus se obtendría mediante la lectura de una lista de palabras aisladas. Dichas palabras se extrajeron de los manuales *El Español Moderno* (Dong y Liu, 2006) y *Fonética del Español* (Wen, 2008), que eran los empleados por las alumnas en su universidad de origen durante su primer año de estudio del español. Además, se obtuvieron también algunas palabras de los materiales utilizados en el curso estival en la UAB mencionado anteriormente. Este criterio es fundamental en la constitución del corpus puesto que, para obtener una lectura lo más neutra y natural posible, era necesario que las estudiantes estuvieran familiarizadas con la acentuación y el significado de las palabras.

Las palabras que componen el corpus de investigación reúnen otra serie de requisitos relacionados con la acentuación, la estructura silábica y el entorno de las vocales que se detallan a continuación.

#### **2.1.1 Tonicidad**

El corpus contiene el mismo número de realizaciones tónicas y de realizaciones átonas de cada una de las vocales del español. Dichas cantidades se especificarán en la tabla global que figura más adelante.

### 2.1.2 Estructura silábica

Puesto que el reducido número de combinaciones en la estructura silábica del chino frente a las numerosas posibilidades del español puede crear dificultades en la producción de las estudiantes sinohablantes (ver capítulo II), especialmente cuando se trata de grupos consonánticos, en la presente tesis doctoral se han omitido ese tipo de estructuras que afectan también a la producción de las consonantes y que deben, por tanto, ser analizadas en una etapa posterior de la investigación. Así, las palabras que componen el corpus de estudio son bisílabas, y los dos tipos de estructura silábica que contienen son CV y CVC que se combinan en palabras cuya estructura es CV-CV, CV-CVC o CVC-CV. Igual que sucede con la tonicidad de las vocales, el corpus presenta el mismo número de realizaciones de cada una de las vocales del español en las distintas estructuras de palabra indicadas.

### 2.1.3 Entorno consonántico

El corpus diseñado evita, en la medida de lo posible, la presencia de un número elevado de sonidos consonánticos en el entorno anterior y en el posterior que, según la bibliografía, resultan problemáticos para los estudiantes sinohablantes de español.

## 2.2 Composición del corpus

### 2.2.1 Tonicidad

Después de extraer las palabras de los materiales anteriormente indicados, el corpus de estudio, en función de la tonicidad de la vocal analizada, está compuesto por los siguientes elementos:

Vocal	Nº total de casos	Acentuación	Nº de casos según su tonicidad
/i/	912	Tónica	480
		Átona	432
/e/	960	Tónica	480
		Átona	480
/a/	960	Tónica	480
		Átona	480
/o/	960	Tónica	480
		Átona	480
/u/	960	Tónica	480
		Átona	480
<b>Total:</b>	4752		4752

Tabla 3. Número de vocales que componen el corpus en función de la tonicidad de las mismas.

Como se puede observar en la tabla 3, mientras que en todas las vocales hay el mismo número de casos de realizaciones tónicas y átonas (480), en la vocal /i/ el número de casos en los que esta vocal es átona se reduce a 432. Esto se debe a que la composición del corpus debe seguir los criterios anteriormente descritos, y no se han podido extraer más palabras en las que la /i/ es átona de los materiales señalados.

### 2.2.2 Estructura silábica

Siguiendo los criterios indicados en el apartado 2.1.2, y después de extraer las palabras de los materiales anteriormente indicados, el corpus queda compuesto de la siguiente manera en función de las estructuras silábicas:

Estructura silábica	Nº total de casos
CV-CV	1764
CV-CVC	1692
CVC-CV	1692

Tabla 4. Número de vocales que componen el corpus según la estructura de la palabra.

Como se ha indicado, al diseñar el corpus se ha tenido en cuenta, además de la estructura de la palabra, la estructura de la sílaba en la que se encuentra la vocal analizada. De esta manera, el corpus queda compuesto de la siguiente manera atendiendo a este criterio:

Vocal analizada	Estructura de la sílaba de la vocal	Nº total de casos
/i/	CV	612
	CVC	360
/e/	CV	720
	CVC	360
/a/	CV	720
	CVC	360
/o/	CV	720
	CVC	360
/u/	CV	648
	CVC	360

Tabla 5. Número de vocales que componen el corpus según la estructura de la sílaba en la que se encuentra dicha vocal.

### 2.2.3 Entorno anterior

Los sonidos consonánticos que aparecen en el entorno anterior a las vocales analizadas se han clasificado en función de si se trata de oclusiva sorda, oclusiva sonora, aproximante, fricativa, nasal, lateral, vibrante simple o vibrante múltiple. Debe tenerse en cuenta que, como se ha indicado, el corpus se ha extraído de los manuales con los que las

informantes sinohablantes han trabajado y, además, las palabras son bisílabas y su puede ser CV-CV, CV-CVC o CVC-CV. Por esta razón, el número de casos para cada uno de los diferentes tipos de entornos difiere de unas vocales a otras, y el corpus queda compuesto de la siguiente manera atendiendo a este criterio:

<b>Entorno consonántico anterior a la vocal analizada.</b>									
<b>Total de casos en el corpus</b>									
	<b>Oclusiva sorda</b>	<b>Oclusiva sonora</b>	<b>Aproxi- mante</b>	<b>Fricativa</b>	<b>Africada</b>	<b>Nasal</b>	<b>Lateral</b>	<b>Vib. Simple</b>	<b>Vib. Múltiple</b>
/i/	240	120	96	216	24	156	60	12	24
/e/	240	144	146	228	0	144	60	36	72
/a/	396	132	72	192	12	180	36	36	48
/o/	384	84	118	122	36	108	72	48	60
/u/	288	180	24	204	12	192	36	12	84

Tabla 6. Número de vocales que componen el corpus según el sonido consonántico anterior.

#### 2.2.4 Entorno posterior

Para constituir el corpus según el entorno posterior a la vocal se han seguido los mismos criterios que los mencionados en relación con el entorno anterior. Por este motivo, el número de sonidos vocálicos que componen el corpus en función del entorno posterior también difiere de unas vocales a otras. Tras extraer las palabras de los materiales y clasificarlas, el corpus según el entorno posterior presenta la siguiente estructura:

<b>Entorno consonántico posterior a la vocal analizada.</b>									
<b>Total de casos en el corpus.</b>									
	<b>Oclusiva sorda</b>	<b>Oclusiva sonora</b>	<b>Aproxi- mante</b>	<b>Fricativa</b>	<b>Africada</b>	<b>Nasal</b>	<b>Lateral</b>	<b>Vib. Simple</b>	<b>Vib. Múltiple</b>
/i/	84	12	84	204	12	144	36	228	0
/e/	24	24	70	108	0	156	144	132	12
/a/	36	12	84	144	0	84	84	240	24
/o/	60	0	24	96	0	228	36	132	24
/u/	96	12	192	190	36	204	84	192	12

Tabla 7. Número de vocales que componen el corpus según el sonido consonántico posterior.

#### 2.2.5 Vocal en final de palabra

Para constituir el grupo de vocales que se encuentran en posición final de palabra aislada, se han aplicado los mismos criterios que los mencionados en los dos apartados anteriores. Por ello, la composición del corpus, en dicha posición, queda de la siguiente forma:

<b>Vocal</b>	<b>Vocal en posición final de palabra aislada. Total de casos en el corpus</b>
/i/	144
/e/	396
/a/	384
/o/	480
/u/	24

*Tabla 8.* Número de vocales que componen el corpus que se encuentran en posición final de palabra aislada.

### **2.2.6 Estructura global del corpus**

A modo de resumen, la figura 1 muestra la composición del corpus según los distintos puntos de vista desde los cuales se estudiarán las vocales del español.

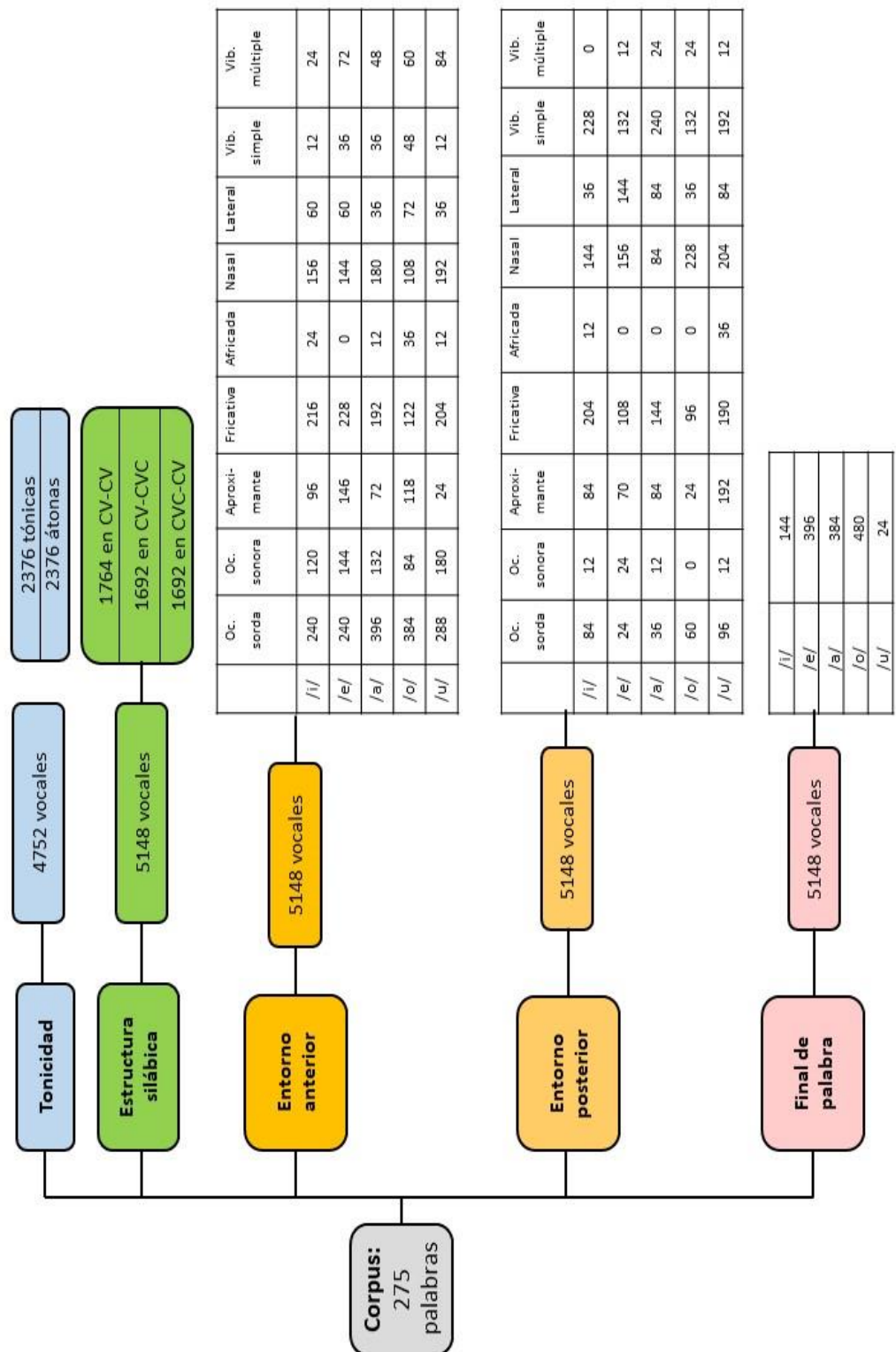


Figura 1. Composición del corpus según las distintas perspectivas de análisis.

## **2.3 Grabación**

Para la grabación del corpus, que se llevó a cabo en la cámara insonorizada del Servei de Tractament de la Parla (STP) de la Universidad Autónoma de Barcelona, se ha empleado el programa específico Praat (versión 5.3), y un ordenador MacBook Pro, con sistema Mac OS X, versión 10.7.5, con un procesador de 2.4 Ghz y memoria de 4GB 1333 Mhz DDR3.

Las hablantes han leído las palabras a partir de una presentación en *Power Point* en la que figuraba una única palabra en la pantalla para evitar el “efecto lista” en la lectura de las mismas.

## **3. ANÁLISIS AUDITIVO**

Como ya se ha indicado, el análisis del corpus obtenido mediante las grabaciones se ha dividido en 3 fases: análisis auditivo, análisis acústico y tratamiento estadístico de los datos. En primer lugar, se ha realizado un análisis auditivo de las grabaciones, tanto del grupo de control, como del grupo de informantes sinohablantes. En el caso del primero, simplemente se ha comprobado que las realizaciones respondían a las expectativas de producción de los hablantes nativos. En el caso del segundo, este análisis ha servido para comprobar si los aspectos que, según la bibliografía, son problemáticos para los sinohablantes también lo son para este grupo de informantes, y para detectar si aparecían en las grabaciones, como así ha ocurrido, fenómenos no mencionados por otros autores.

### **3.1 Problemas preliminares**

Mientras que el análisis auditivo de las grabaciones correspondientes al grupo de control ha confirmado que las realizaciones en ellas contenidas son las propias de los hablantes nativos de español, en el caso de las alumnas sinohablantes se han detectado algunos fenómenos dignos de mención que han orientado el tratamiento posterior de los datos y que, por esta razón, se resumen en este capítulo dedicado a la Metodología de trabajo.

Se exponen pues, a continuación, las observaciones realizadas sobre las realizaciones de las alumnas sinohablantes durante esta primera fase del análisis. Dichas

observaciones están clasificadas en aspectos relacionadas con la acentuación, con la estructura silábica, con los sonidos consonánticos del entorno, y con los propios sonidos vocálicos y su duración.

<b>Acentuación</b>	<ul style="list-style-type: none"> <li>- Realización de sílabas tónicas como átonas</li> <li>- Realización de sílabas átonas como tónicas</li> </ul>
<b>Duración</b>	<ul style="list-style-type: none"> <li>- Realizaciones más largas que las de los nativos en sílaba tónica</li> <li>- Alargamiento de la vocal cuando hay elisión de un sonido consonántico, tanto anterior como posterior</li> </ul>
<b>Estructura silábica</b>	<ul style="list-style-type: none"> <li>- Silabeo, especialmente cuando la sílaba tónica es la final</li> <li>- Realización de estructuras CVC como CV, tanto en posición final como en posición final</li> <li>- Realización de estructuras CVC como CCV, especialmente cuando el primer sonido consonántico es [r]</li> <li>- Realización de estructuras CVC como CVCV</li> </ul>
<b>Sonidos consonánticos</b>	<ul style="list-style-type: none"> <li>- Elisión de sonidos consonánticos y, como consecuencia, alargamiento del sonido vocálico anterior o posterior</li> <li>- Realización de consonantes oclusivas sordas como sonoras, y viceversa</li> <li>- Realización de [r] y [r] como [l]</li> <li>- Realización de [r] como [dr]</li> </ul>
<b>Sonidos vocálicos</b>	<ul style="list-style-type: none"> <li>- Timbre inestable, especialmente cuando se produce el alargamiento de algún sonido vocálico</li> <li>- Realización [ə] de [i], [e], [a], [o], [u]</li> <li>- Palatalización de las vocales [e], [a], [o], [u]</li> <li>- Nasalización de las vocales [i], [e], [a], [o], [u]</li> <li>- Labialización de la vocal [u]</li> <li>- Articulación más abierta de las vocales [i], [e], [a], [o], [u]</li> <li>- Articulación más cerrada de las vocales [e], [a], [o], [u]</li> </ul>



	<ul style="list-style-type: none"> <li>- Realización de [i] como [e]</li> <li>- Realización de [e] como [i], [ie] o [ei], especialmente en posición final de palabra</li> <li>- Realización de [a] como [e]</li> <li>- Realización de [o] como [u], [ou] o [a]</li> <li>- Realización de [u] como [e] o [ue]</li> <li>- Elisión de algunos sonidos vocálicos, especialmente [i], [a], [u]</li> </ul>
--	--

Tabla 9. Fenómenos detectados en el análisis auditivo de las realizaciones de las hablantes chinas.

## 4. ANÁLISIS ACÚSTICO

La segunda fase del estudio ha consistido en realizar un análisis acústico de las grabaciones, para lo que se ha empleado el programa de análisis de voz *Praat*, (versión 5.3). Los parámetros analizados han sido determinados tomando como punto de partida el análisis auditivo que, como se ha indicado, ha permitido poner de relieve la presencia de numerosos fenómenos relacionados con la pronunciación de las vocales que no habían sido detectados por otros autores y no figuran, por tanto, en los estudios hasta ahora publicados.

Como variables independientes no numéricas se han considerado la acentuación de la vocal, la estructura silábica en la que está insertada y el entorno consonántico porque los datos auditivos han conducido a estimar que estas variables dependientes ejercen una importante influencia sobre la realización de los timbres vocálicos. Como variables dependientes y numéricas se han analizado acústicamente, la duración y las frecuencias de los formantes. Para el almacenamiento de los datos, se ha empleado el programa *Microsoft Excel 2016*.

### 4.1 Parámetros acústicos analizados

#### 4.1.1 La duración de las vocales

Como se ha indicado, durante el análisis auditivo se ha podido apreciar que la duración de los sonidos vocálicos varía considerablemente de unos casos a otros. Por este motivo, se ha obtenido, en milisegundos (ms.) la duración de todos los sonidos vocálicos

del corpus tanto de los producidos por el grupo de control como de los producidos por las estudiantes sinohablantes. La comparación de dichos valores ha permitido establecer el diferente comportamiento que presentan los nativos y los sinohablantes con respecto a esta parámetro.

#### **4.1.2 Las frecuencias de los formantes**

En la primera fase, es decir, en el análisis auditivo, se ha detectado que el timbre de las vocales producidas por el grupo de informantes sinohablantes es, en ocasiones, inestable, fenómeno que se encuentra en estrecha relación con la duración de las realizaciones. Tradicionalmente, los estudios acústicos sobre las vocales del español han tomado como referencia el centro de la vocal para indicar los valores de frecuencia de los formantes. Sin embargo, la medición única de dicho punto no permite saber si, habida cuenta de la duración de las vocales, se producen oscilaciones en sus valores de frecuencia y, en caso de haberlas, observar su comportamiento. No debe olvidarse que el mandarín es una lengua tonal y que la duración de los tonos permite justamente su oscilación. La larga duración de las vocales del español producida por este perfil de hablantes podría estar relacionada con una proceso de transferencia de un rasgo del mandarín al español y, por dicho motivo, se ha considerado importante explorar esta dimensión de las realizaciones vocálicas. Así, se han obtenido en hercios (Hz), los valores de las frecuencias del punto inicial, del punto medio y del final de los dos primeros formantes (F1 y F2) de las vocales producidas, tanto por las informantes sinohablantes, como por el grupo de control, dado que es necesario tener datos de hablantes nativas que sirvan como referencia para interpretar el comportamiento de las estudiantes chinas.

#### **4.1.3 La acentuación**

Otro de los fenómenos detectados en el análisis auditivo es que, en ocasiones, las informantes sinohablantes realizan las sílabas tónicas como átonas, y viceversa. Por este motivo, otro de los parámetros estudiados en la producción oral de las estudiantes chinas ha sido la acentuación, correcta o no, de sus realizaciones vocálicas del español.

#### **4.1.4 El entorno**

El último parámetro no numérico analizado es el entorno. Durante el análisis auditivo se ha podido comprobar que, tal y como indica la bibliografía, las informantes

sinohablantes tienen dificultades al pronunciar algunos de los sonidos consonánticos del español. Por este motivo, se ha analizado cómo realizan dichos sonidos las estudiantes chinas y cómo influye dicha realización en su pronunciación de las vocales. Dicho estudio se ha llevado a cabo desde distintas perspectivas, como se verá más adelante.

#### **4.1.5 La estructura silábica**

El último parámetro no numérico analizado es la estructura silábica. El análisis auditivo de las grabaciones ha permitido también detectar que, en ocasiones, las informantes sinohablantes eliden algunos sonidos en su producción oral en español, o añaden nuevos sonidos a las palabras. Ejemplo de ello es la articulación de *color* como [ko.'lo] y de *rosa* como ['tro.sa]. Por este motivo, otro de los parámetros analizados es la estructura silábica, la cual se ha analizado desde distintas perspectivas, como se muestra más adelante.

#### **4.2 Almacenamiento de los datos**

Para el almacenamiento de datos se han creado dos bases de datos, una para cada uno de los grupos estudiados, aunque la codificación empleada ha sido la misma en ambas. En los dos casos se han medido dos variables cuantitativas: la duración de la vocal, la cual ha sido medida en milisegundos (ms.), y las frecuencias en las que aparecen los distintos puntos de los formantes analizados, las cuales se han medido en hercios (Hz). En cuanto a las variables no numéricas, en el caso de las informantes sinohablantes además de introducir los datos correspondientes a cada vocal en español, se ha añadido cuál ha sido la realización de cada una de ellas por parte de las estudiantes chinas. En la tabla 9 se muestra la codificación empleada para cada una de estas variables.

<b>Variable cualitativa</b>	<b>Codificación</b>
Informante	Del 1 al 6
Vocal	1 para /i/; 2 para /e/ 3 para /a/; 4 para /o/ 5 para /u/
Acentuación en español	Tónica: 1 Átona: 2
Acentuación de la vocal por informante sinohablante	
Estructura de la palabra en español	Vocal: 1 Consonante: 2
Estructura de la palabra por informante sinohablante	
Estructura de la sílaba en español	Vocal: 1 Consonante: 2
Estructura de la sílaba por informante sinohablante	
Estructura de la sílaba anterior en español	Vocal: 1 Consonante: 2
Estructura de la sílaba anterior por informante sinohablante	
Estructura de la sílaba posterior en español	Vocal: 1 Consonante: 2
Estructura de la sílaba posterior por informante sinohablante	
Entorno anterior en español	Del 1 al 9 para cada tipo de consonante
Entorno anterior por informante sinohablante	
Entorno posterior en español	Del 1 al 9 para cada tipo de consonante y 0 para elisión
Entorno posterior por informante sinohablante	
Vocal en final de palabra en español	11 para vocal y del 1 al 9 para cada tipo de consonante
Vocal en final de palabra por informante sinohablante	

*Tabla 10.* Codificación empleada en las bases de datos.

## **5. TRATAMIENTO ESTADÍSTICO DE LOS DATOS**

En el presente apartado se explica el procedimiento seguido para realizar el tratamiento estadístico de los mismos. Para realizar el tratamiento estadístico de los datos, se ha contado con la colaboración del Servicio de Estadística de la Universidad Autónoma de Barcelona. La lectura, gestión y validación de la base de datos, y la creación de áreas de dispersión de los formantes se han realizado con el software *RStudio versión 3.5.0* para Windows. Para la creación de gráficos se ha empleado el programa *Microsoft Excel 2016*. A continuación, se detallan cada una de las pruebas llevadas a cabo en la presente tesis doctoral.

### **5.1 Validación estadística del análisis auditivo**

Como se ha indicado, en ocasiones las realizaciones de las informantes sinohablantes no corresponden con las propias del español en acentuación, estructura silábica o en lo que a los sonidos consonánticos se refiere. En el capítulo IV se establece una hipótesis nula y una hipótesis alternativa para cada uno de los parámetros analizados. La primera de estas afirma que no existe diferencia significativa entre las dos variables analizadas, mientras que la segunda indica lo contrario, que las diferencias son significativas. En todos los casos la hipótesis que se prueba es la nula, y dependiendo de si el resultado es significativo o no, esta se mantiene o se rechaza.

Con el fin de determinar si se rechaza la hipótesis nula, se ha calculado el Coeficiente kappa de Cohen, el cual permite estimar la concordancia entre variables cualitativas. Si el índice kappa es igual a 0, se rechaza la hipótesis nula, por lo que se asume que la lengua materna interfiere en la realización correcta del parámetro analizado. En cambio, si el valor es cercano a 1, no se rechaza la hipótesis nula, lo que se interpreta como que la realización correcta de dicho parámetro no depende de la lengua materna de las informantes sinohablantes. Cabe señalar que el Coeficiente kappa de Cohen es una medida estadística que permite estimar la concordancia entre variables cualitativas, y es una medida más robusta que un simple cálculo del porcentaje de concordancias, ya que tienen en cuenta el acuerdo que ocurre por azar. Dado que es difícil que valores pequeños del Coeficiente kappa resulten estadísticamente significativos (diferentes de 0), en esta tesis doctoral se emplea la siguiente escala de valoración del índice kappa, ampliamente utilizada, y propuesta por Landis y Koch (1977):

<b>Coefficiente kappa</b>	<b>Grado de acuerdo</b>
< 0	Pobre (sin acuerdo)
0 – 0.2	Leve
0.21 – 0.4	Aceptable
0.41 – 0.6	Moderado
0.61 – 0.8	Sustancial
0.81 – 1	Casi perfecto

*Tabla 11.* Escala de valoración del Coeficiente kappa (Landis y Koch, 1977).

## **5.2 Análisis de la duración vocálica en función de distintas variables**

Se ha calculado la duración media y la desviación estándar de las vocales del español producidas por las informantes en función de diferentes aspectos. Las variables explicativas tenidas en cuenta son, en primer lugar, las relacionadas con la acentuación de la vocal; en segundo lugar, las relacionadas con la estructura silábica; y, en tercer lugar, las relacionadas con el entorno de la vocal. Tras realizar dichos cálculos, se ha procedido a comparar los resultados obtenidos en el análisis del grupo de control con los de las informantes sinohablantes vocal por vocal. En aquellos casos en los que los datos siguen una distribución normal, se ha aplicado la Prueba T para conocer el nivel de significación; en caso contrario, se ha realizado la prueba U de Mann-Whitney-Wilcoxon. Las decisiones estadísticas se han realizado tomando como nivel de significación el valor 0.05. Es decir, si el valor es mayor a 0.05, no existen diferencias significativas entre ambos grupos, por lo que se acepta la hipótesis nula; si dicho valor es menor a 0.05, existen diferencias significativas, por lo que se rechaza la hipótesis nula y se asume que la lengua materna influye en la producción oral de las informantes sinohablantes.

## **5.3 El análisis del timbre**

En la tercera parte del capítulo IV se han medido las frecuencias en las que aparecen los formantes de las vocales del español producidas por el grupo de control y por las informantes sinohablantes respectivamente. Como se ha indicado, se han medido tres puntos en cada uno de los formantes: el inicio, el centro y el final. En este caso, la hipótesis nula es que no existen diferencias significativas entre los distintos puntos analizados de cada una de las vocales; la hipótesis alternativa indica que sí existen

diferencias significativas entre dichos puntos. Para calcular el nivel de significación, se ha aplicado la prueba de Friedman, la cual se trata de una prueba paramétrica utilizada para comparar más de dos variables relacionadas. Dicha prueba se ha aplicado, por un lado, a los datos relativos al grupo de control y, por otro lado, a los de las informantes sinohablantes. Las decisiones estadísticas se han realizado tomando como nivel de significación el valor 0.05. Es decir, si el valor obtenido es mayor que 0.05, no hay una diferencia significativa entre las variables analizadas. Si dicho valor es menor que 0.05, la diferencia sí es significativa.

#### **5.4 El análisis del timbre: comparación entre el grupo de control y el grupo de informantes sinohablantes**

En la tercera parte del capítulo IV se realiza una comparación entre el grupo de control y el de estudiantes chinas relativa a las frecuencias en las que aparecen las distintas vocales del español. Aquí, la hipótesis nula afirma que no existen diferencias significativas entre el grupo de control y las informantes sinohablantes; la hipótesis alternativa indica que sí existen diferencias entre ambos grupos y que la lengua materna influye en la articulación de las vocales. Para este análisis, se ha aplicado la prueba de Mann Whitney, la cual se trata de una prueba paramétrica aplicada a dos muestras independientes. Las decisiones estadísticas se han realizado tomando como nivel de significación el valor 0.05. Como se ha señalado anteriormente, si el valor obtenido es mayor que 0.05, no existen diferencias significativas entre el grupo de control y las informantes sinohablantes; si dicho valor es menor que 0.05, las diferencias sí son significativas, por lo que se rechaza la hipótesis nula y se asume que la lengua materna influye en la realización de las vocales del español.

**CAPÍTULO IV**  
**RESULTADOS**





## **IV. RESULTADOS**

### **PRIMERA PARTE: INFLUENCIA DE LAS VARIABLES INDEPENDIENTES SOBRE LA REALIZACIÓN DE LAS VOCALES. PROCESOS DE RESILABACIÓN.**

En esta primera parte del capítulo se presentan los resultados del análisis de la influencia de las tres variables independientes consideradas (acentuación, entorno anterior y posterior y estructura silábica) sobre la realización de los sonidos vocálicos del español por alumnos sinohablantes.

#### **1. LA ACENTUACIÓN**

En el caso de la acentuación, la hipótesis nula que se ha puesto a prueba es:

H0: La lengua materna de las informantes sinohablantes no influye en la correcta acentuación de las vocales del español.

La hipótesis alternativa que se quiere probar es la siguiente:

H1: La lengua materna de las informantes sinohablantes sí influye en la correcta acentuación de las vocales del español.

Con el fin de determinar el grado de acuerdo entre la realización de dicho parámetro por parte de las informantes hispanohablantes del grupo de control y por las informantes sinohablantes, se ha calculado el Coeficiente kappa de Cohen, que permite estimar el grado de concordancia entre variables cualitativas. Si el valor obtenido es 1, se considera que la realización de las informantes sinohablantes es adecuada en todos los casos, por lo que se valida la hipótesis nula (el chino no tendría ninguna influencia sobre la tonicidad de las vocales del español pronunciadas por las informantes estudiadas). Si dicho valor es inferior a 1, se rechaza la hipótesis nula lo cual significa que estas hablantes tienen problemas para acentuar correctamente las vocales del español. Con el fin de determinar

el grado de acuerdo entre las dos variables cuando el valor obtenido ha sido inferior a 1, se ha empleado la escala de valoración propuesta por Landis y Koch (1977)<sup>2</sup>.

Debe recordarse que los autores que han analizado esta cuestión señalan que, generalmente, los estudiantes sinohablantes no tienen dificultades en marcar adecuadamente la sílaba tónica del español. Los resultados del cálculo del Coeficiente kappa de Cohen pueden verse en la tabla 1 y, como se aprecia, en todas las vocales el valor del mismo es superior a 0.81, lo que, según la escala de Landis y Koch (1977) indica que la acentuación tónica o átona de las vocales del español por parte de las informantes chinas es adecuada en la mayor parte de las realizaciones.

Vocal	Coeficiente kappa
/i/	0.92
/e/	0.88
/a/	0.94
/o/	0.95
/u/	0.93

Tabla 1. Coeficiente kappa para cada una de las vocales según la tonicidad.

La figura 1 muestra también que en todos los casos el porcentaje de realizaciones correctas desde el punto de vista de la acentuación es muy elevado, cercano en muchos casos al 100%.

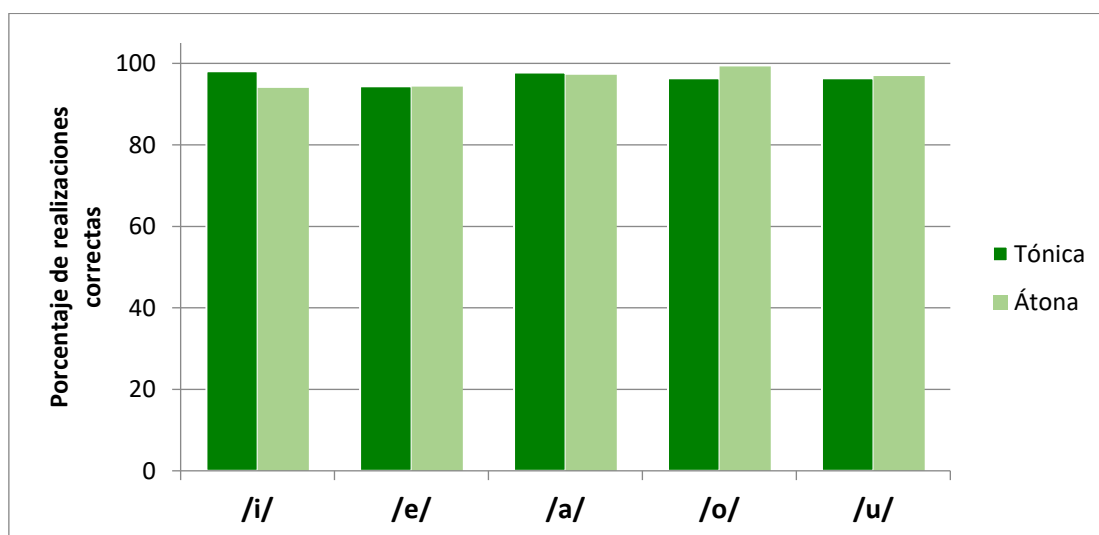


Figura 1. Porcentaje de realizaciones correctas de las vocales realizadas por las informantes sinohablantes en función de su acentuación en español.

<sup>2</sup> En el capítulo III se han descrito con mayor detalle el Coeficiente kappa de Cohen y la escala de Landis y Koch (1977).

La tabla 2 muestra en detalle los datos de la realización de las vocales por parte del grupo de informantes sinohablantes contrastadas con la realización propia del español. Las tres primeras columnas ofrecen los datos propios del español y del corpus de estudio de este trabajo. La primera de ellas, ‘vocal’, indica cuál es el sonido vocálico analizado; ‘acentuación español’ indica si dicha vocal es tónica o átona en español; y ‘total casos’ hace referencia al número de casos de dicha vocal presentes en el corpus y que se han analizado en este apartado. Las columnas restantes muestran las realizaciones tónicas o átonas correctas e incorrectas de las informantes sinohablantes<sup>3</sup>. En la fila correspondiente a las realizaciones de /i/ tónica se observa que, de los 240 casos analizados, 235 realizaciones del sonido vocálico han sido tónicas, un 98% del total, y 5 han sido átonas, lo que supone un 1.9% de total. En la siguiente fila correspondiente a las realizaciones de /i/ átona, los resultados expuestos muestran que de 216 casos que tiene el corpus y que han sido analizados, 203 han sido realizados de manera correcta, es decir, como vocales átonas, y 13 de estas 216 vocales han sido realizadas como tónicas, es decir, que el porcentaje de realizaciones correctas es de 93.9% frente al 6% de incorrectas.

Vocal	Acentuación español	Total casos	Realizaciones sinohablantes			
			Tónica		Átona	
			Nº casos	Porcentaje	Nº casos	Porcentaje
/i/	Tónica	240	235	98.0%	5	1.9%
	Átona	216	13	6.0%	203	93.9%
/e/	Tónica	240	227	94.4%	13	5.5%
	Átona	240	14	5.7%	226	94.2%
/a/	Tónica	240	235	97.7%	5	2.2%
	Átona	240	7	2.7%	233	97.2%
/o/	Tónica	240	231	96.3%	9	3.6%
	Átona	240	2	0.7%	238	99.2%
/u/	Tónica	240	231	96.3%	9	3.6%
	Átona	240	7	3.0%	233	96.9%

Tabla 2. Porcentaje de realizaciones de la acentuación silábica de las estudiantes sinohablantes contrastada con la realización propia del español.

<sup>3</sup> Los datos que aparecen en negro en la tabla 2 corresponden a las realizaciones correctas. Los marcados en rojo corresponden a las realizaciones incorrectas; es decir, los casos en los que la vocal tónica se ha realizado como átona, y los casos en los que la vocal átona se ha realizado como tónica.

El análisis de las realizaciones de las vocales del español tónicas por parte de las informantes sinohablantes muestra que el porcentaje de realizaciones correctas es muy elevado, ya que oscila entre el 94.4% de aciertos de /e/ y el 98% de /i/.

En cuanto a las realizaciones de las vocales del español átonas, el porcentaje de vocales acentuadas correctamente oscila entre el 93.9% de /i/ y el 99.2% de /o/. Al igual que con las realizaciones tónicas, en el caso de las átonas el número de realizaciones correctas también es muy elevado.

Los datos expuestos en este apartado corroboran lo indicado por los autores citados en el capítulo II, es decir, que las estudiantes sinohablantes no presentan dificultades remarcables en su realización de las distintas vocales como tónicas o átonas en palabras bisílabas. En cuanto al desplazamiento de patrones acentuales en palabras trisílabas señalado por Blanco y Nogueroles (2013: 217-218), según quienes en algunas ocasiones los estudiantes sinohablantes realizan las palabras llanas como esdrújulas, este aspecto no se ha estudiado en el presente trabajo dado que el corpus está constituido por palabras bisílabas, lo que podría ser una línea de investigación interesante de cara a futuros trabajos.

## **2. ENTORNO CONSONÁNTICO**

En el capítulo II se ha señalado cuáles son, según la bibliografía, los sonidos consonánticos que causan problemas a los estudiantes sinohablantes en su producción oral en español. En este apartado se ha analizado la realización de los entornos de la vocal, tanto el anterior, como el posterior. Igual que para la variable relacionada con la acentuación, en primer lugar, se ha calculado el Coeficiente kappa de Cohen y, a continuación, se han analizado las realizaciones correctas e incorrectas de dicho sonido con el fin de determinar qué consonantes resultan problemáticas para los estudiantes de español cuya lengua materna es el chino.

### **2.1 Entorno anterior**

La tabla 3 muestra que, en todos los casos, el valor del índice del coeficiente de kappa de Cohen es mayor que 0.61 y menor que 0.8, lo que indica que la realización de las vocales españolas por parte de las informantes sinohablantes presenta problemas (por su distancia al valor 1 del coeficiente) que están relacionadas con el entorno consonántico.

No ocurre así en los valores del mismo coeficiente cuando, en el apartado anterior, se analiza la tonicidad de las vocales pues, como se recordará, los resultados muestran valores del coeficiente situados muy cerca de 1, lo que indica que las hablantes sinohablantes no tienen problemas en marcar adecuadamente las vocales en función de dicha variable independiente.

Vocal	Coefficiente Kappa
/i/	0.73
/e/	0.67
/a/	0.72
/o/	0.66
/u/	0.65

Tabla 3. Coeficiente Kappa para cada una de las vocales según el entorno anterior.

Con el fin de determinar cuáles son los sonidos consonánticos más problemáticos para las estudiantes sinohablantes y su influencia sobre la pronunciación de las vocales, a continuación se analiza la realización del entorno anterior según el tipo de consonante. Como se ha indicado anteriormente, dichos entornos se han clasificado en: oclusiva sorda, oclusiva sonora, aproximante, fricativa, africada, nasal, lateral, vibrante simple y vibrante múltiple.

Como se puede observar globalmente en la figura 2, cuando el sonido anterior es una oclusiva sorda, fricativa, africada, nasal o lateral, las estudiantes chinas no presentan grandes dificultades, ya que el porcentaje de realizaciones correctas se sitúa entre el 90% y el 100%. Los sonidos que mayor dificultad suponen son la oclusiva sonora, la aproximante, la vibrante simple y la vibrante múltiple.

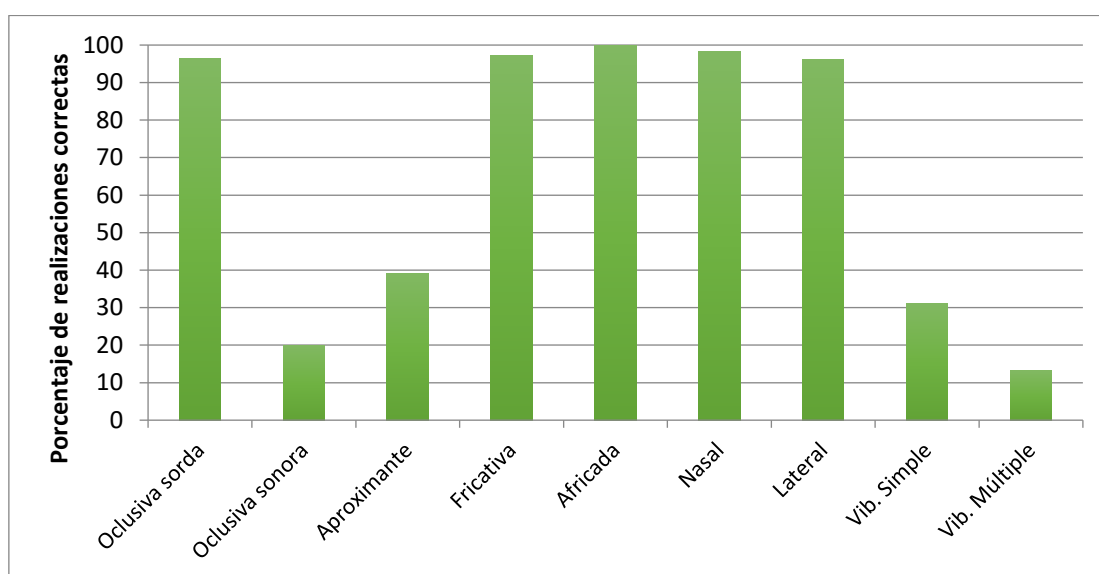


Figura 2. Porcentaje de realizaciones correctas del entorno anterior a la vocal analizada.

La tabla 4 muestra los datos detallados del porcentaje de realizaciones correctas del entorno anterior por parte del grupo de informantes cuya lengua materna es el chino. Según estos datos, cuando el sonido anterior a la vocal es una oclusiva sorda, las informantes sinohablantes lo articulan como tal en el 96.5% de los casos. La oclusiva sonora, sin embargo, representa un problema, ya que únicamente lo han realizado adecuadamente en 66 ocasiones, es decir, el 20.1% de las apariciones. La realización de las aproximantes también conlleva dificultades para las estudiantes chinas, ya que únicamente la han realizado como tal en un 39% de las ocasiones. En cuanto a la fricativa, como ya se ha dicho antes, no supone un problema importante, ya que la han articulado correctamente en un 97.2% de las veces. La africada, por su parte, ha sido realizada como tal en todas las ocasiones según los datos recogidos en esta investigación. La nasal tampoco supone una dificultad importante para estas informantes, ya que el porcentaje de realizaciones correctas es del 98.2%, y lo mismo sucede con la lateral, aunque el porcentaje es ligeramente inferior, el 96.2%. Las vibrantes, en cambio, suponen un escollo casi insalvable para las informantes sinohablantes. Por un lado, la vibrante simple únicamente la han realizado como tal en 30 ocasiones, es decir, en el 31.2% de las ocasiones, y la múltiple en menos, en el 13.2%.

Sonido anterior	Total casos analizados	Realizaciones correctas del sonido anterior	
		Nº de realizaciones	Porcentaje
<b>Oclusiva sorda</b>	773	746	96.5%
<b>Oclusiva sonora</b>	330	66	20%
<b>Aproximante</b>	228	89	39.0%
<b>Fricativa</b>	481	468	97.2%
<b>Africada</b>	42	42	100%
<b>Nasal</b>	390	383	98.2%
<b>Lateral</b>	132	127	96.2%
<b>Vib.simple</b>	96	30	31.2%
<b>Vib.múltiple</b>	144	19	13.2%

*Tabla 4.* Realizaciones correctas del sonido anterior por las informantes sinohablantes.

Estos datos permiten saber cuáles son los sonidos consonánticos más problemáticos para las informantes sinohablantes. A continuación se analiza, para cada una de las categorías de sonidos consonánticos, cómo han realizado las estudiantes chinas aquellos que no han articulado como los del español. Por ejemplo, en aquellos casos en los que el sonido anterior a la vocal es una oclusiva sonora y las informantes no la han realizado como tal, se analiza qué tipo de sonido han pronunciado.

### 2.1.1 Realización de las oclusivas sordas

Anteriormente se ha señalado que el porcentaje de realizaciones correctas de la oclusiva sorda en posición anterior al sonido vocálico es muy elevado (96.5%). En cuanto a las realizaciones incorrectas (27 en total), la figura 3 muestra una clara tendencia a pronunciar las oclusivas sordas como sonoras.

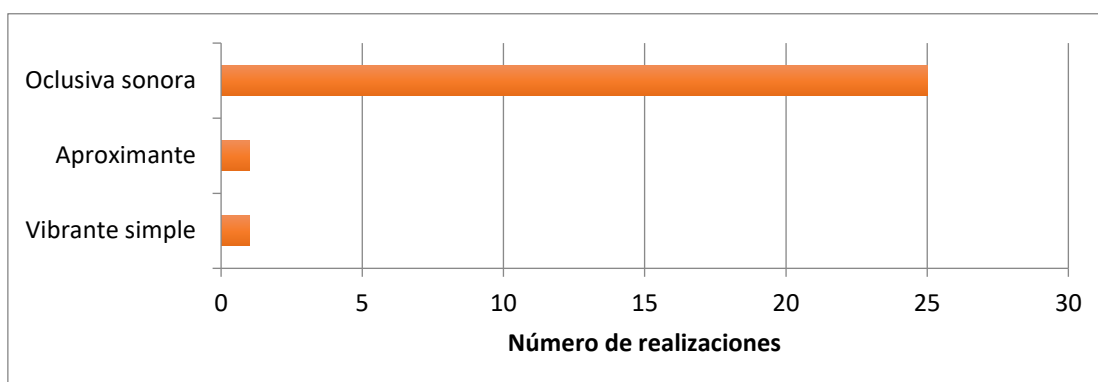
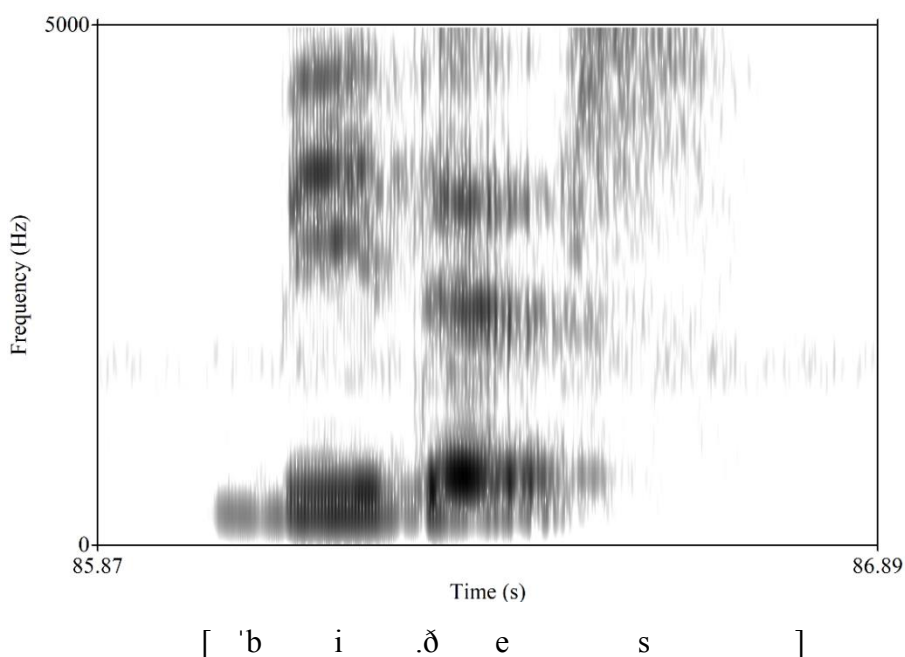


Figura 3. Número y tipo de realizaciones incorrectas de la oclusiva sorda anterior a la vocal por las informantes sinohablantes.

De las 27 realizaciones incorrectas, 25 lo son como oclusiva sonora, 1 como fricativa y 1 como vibrante, por lo que se puede decir que, aunque las informantes sinohablantes no presentan grandes dificultades para articular las oclusivas sordas, cuando no las hacen como tal la tendencia es a realizar una oclusiva sonora. Ejemplo de ello es la pronunciación de *lento* como ['len.do], *banco* como ['ban.go] o *pides* como ['bi.ðes], cuyo espectrograma aparece a continuación y en el que se aprecia que la palabra comienza con la oclusión y el inicio de sonoridad propios de una oclusiva sonora.



Espectrograma 1. Realización de la palabra *pides* por la informante sinohablante 2.



### 2.1.2 Realización de las oclusivas sonoras

Tal y como se ha indicado al comienzo de este apartado, las informantes sinohablantes tienen grandes dificultades para realizar la oclusiva sonora cuando esta se encuentra en posición anterior a la vocal (únicamente la realizan correctamente en el 20.1% de las ocasiones). La figura 4 muestra que, en la mayoría de los casos, dicho sonido es articulado como oclusiva sorda, y en algunas ocasiones como nasal, fricativa, o es elidido.

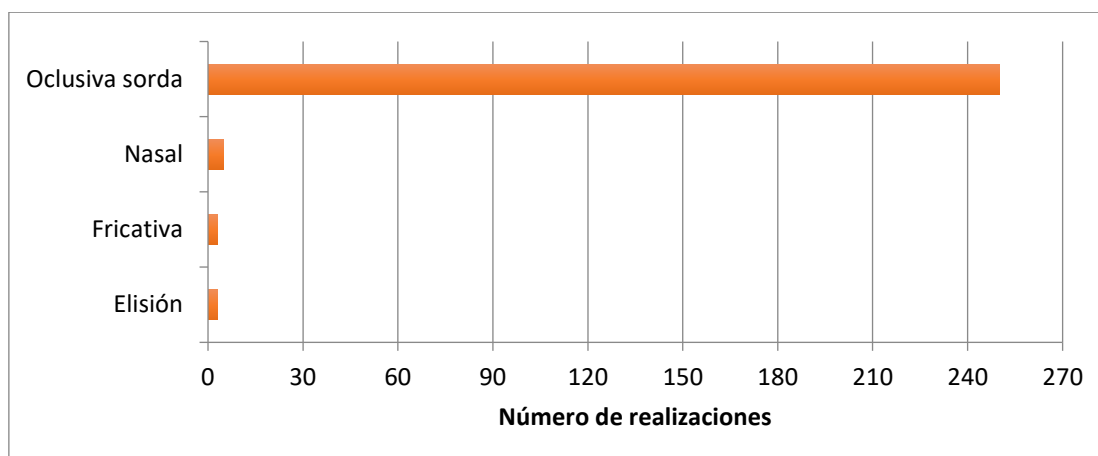
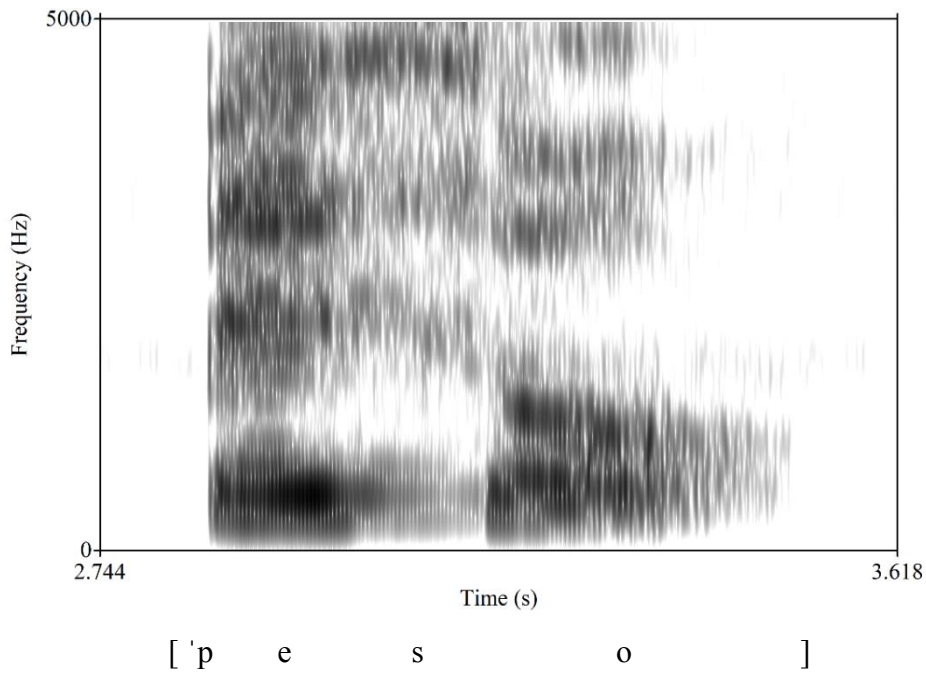


Figura 4. Número y tipo de realizaciones incorrectas de la oclusiva sonora anterior a la vocal por las informantes sinohablantes.

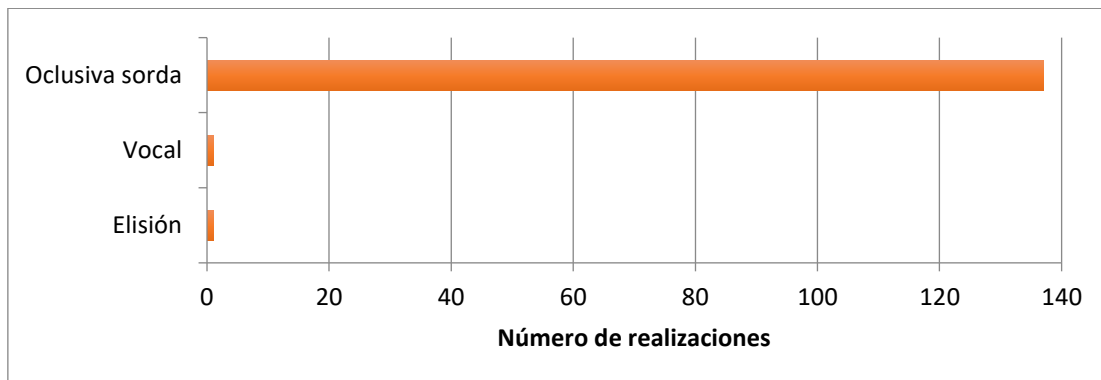
Las informantes sinohablantes han realizado la oclusiva sonora como sorda hasta en 250 casos, es decir, en un 76,4% de las ocasiones, y en el resto de realizaciones incorrectas han pronunciado una consonante nasal, una fricativa o han elidido el sonido, en 5, 3 y 3 casos respectivamente. Estos datos muestran que las estudiantes cuya lengua materna es el chino presentan dificultades para pronunciar la oclusiva sonora, y la tendencia general es a realizarla como sorda. De esta manera, las informantes pronuncian palabras como *beso* como [‘pe.so], *danza* como [‘tan.θa] o *gomas* como [‘ko.mas]. En el espectrograma 2 se muestra la realización de la palabra *beso*, la cual comienza con la barra de explosión propia de una oclusiva sorda.



*Espectrograma 2.* Realización de la palabra *beso* por la informante sinohablante 1.

### 2.1.3 Realización de las aproximantes

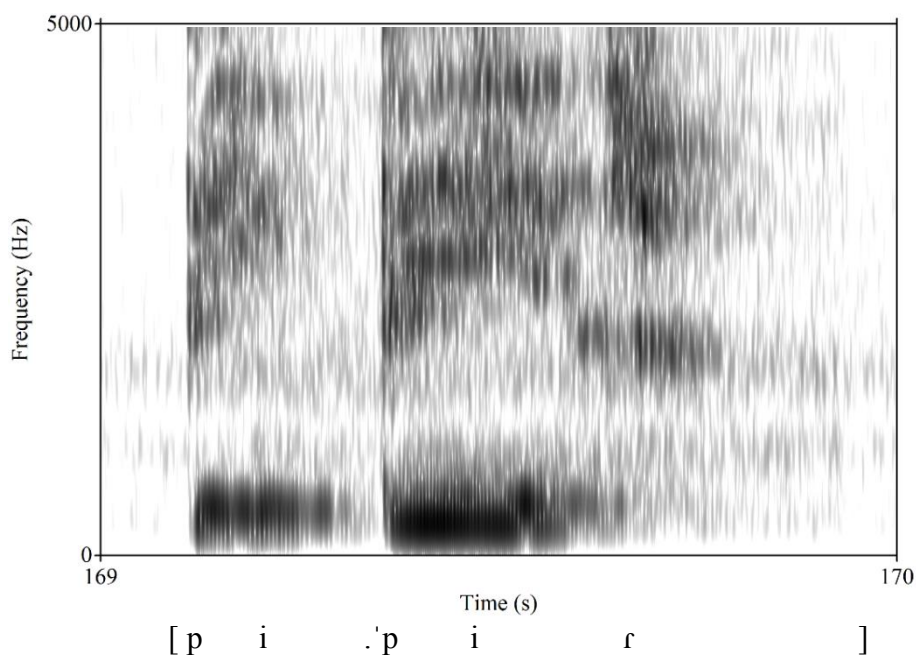
Como se ha indicado, las informantes sinohablantes presentan problemas para articular la aproximante cuando esta se encuentra en posición anterior a la vocal, es decir, en posición inicial de sílaba, ya que únicamente la han realizado correctamente en un 39% de las ocasiones. La figura 5 muestra una clara tendencia a articular dicho sonido consonántico como oclusiva sorda, y en ocasiones aisladas la realizan como vocal o la eliden.



*Figura 5.* Número y tipo de realizaciones incorrectas de la aproximante anterior a la vocal por las informantes sinohablantes.

Las informantes sinohablantes han realizado la aproximante como oclusiva sorda hasta en 139 ocasiones, en una como vocal, y en otra la han elidido. Estos datos muestran que las estudiantes cuya lengua materna es el chino tienden a realizar las aproximantes como

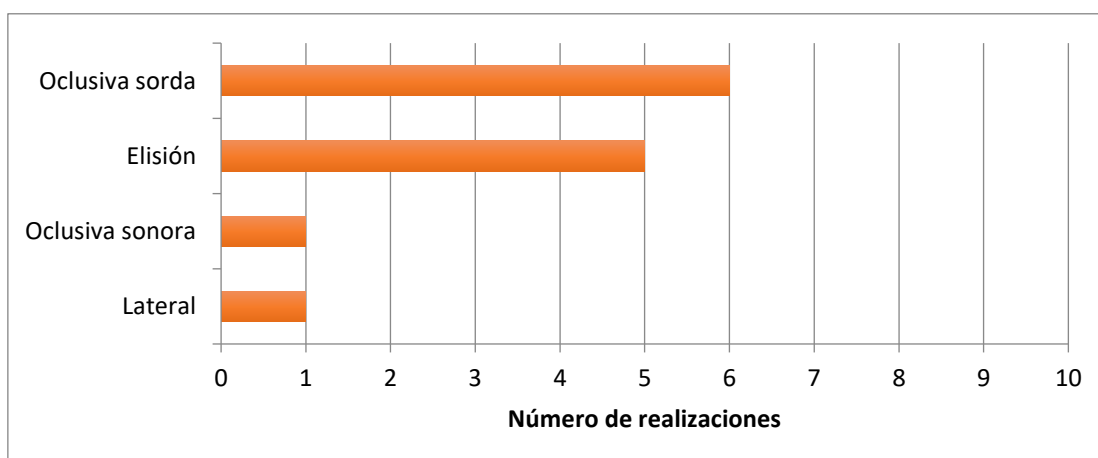
oclusivas sorda. De esta manera, las hablantes pronuncian, por ejemplo, la palabra *vivir* como [pi.'pir].



*Espectrograma 3.* Realización de la palabra *vivir* por la informante sinohablante 3.

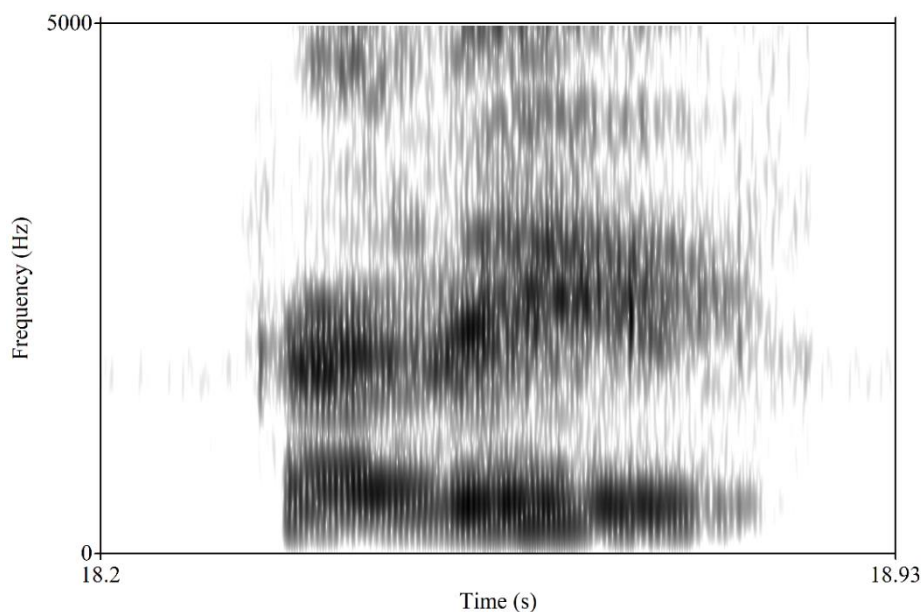
### 2.1.4 Realización de las fricativas

Anteriormente se ha señalado que cuando el sonido anterior a la vocal es una consonante fricativa el porcentaje de realizaciones adecuadas es del 97.2%, por lo que no es un sonido que represente un problema grave para las informantes sinohablantes. La figura 6 muestra que, de las realizaciones incorrectas de dicho sonido (un total de 13 en el conjunto del corpus), en la mayoría de ocasiones la fricativa es articulada como oclusiva sorda o se elide, aunque también hay casos en los que se realiza como oclusiva sonora o como lateral.



*Figura 6.* Número y tipo de realizaciones incorrectas de la fricativa anterior a la vocal por las informantes sinohablantes.

Las estudiantes chinas han realizado la fricativa anterior a la vocal como oclusiva sorda en 6 ocasiones, y la han elidido en 5. En cuanto a la realización de este sonido como oclusiva sonora, únicamente hay un caso, y como lateral, otro. Así, palabras como *feliz* las pronuncian como [ke.'li] o *jugó* como [u.'ko].



[ k e . ' l i ]

*Espectrograma 4.* Realización de la palabra *feliz* por parte de la informante sinohablante 6

Nótese que, en estos casos, la palabra presenta una estructura silábica diferente de la que corresponde a la expectativa del español con lo que la vocal se ve sometida a procesos importantes de transformación.

### 2.1.5 Realización de las africadas

Tal y como se ha indicado anteriormente, las informantes sinohablantes han realizado todos los sonidos consonánticos africados de manera correcta, por lo que se puede afirmar que esta consonante no supone un problema para las hablantes cuya lengua materna es el chino.

### 2.1.6 Realización de las nasales

Los sonidos nasales no presentan dificultades importantes para las informantes sinohablantes, ya que el porcentaje de realizaciones correctas, como se ha indicado anteriormente, es de 98.2%. La figura 7 muestra que cuando la nasal es realizada de manera errónea, las informantes la eliden, la sustituyen por una vocal, la realizan como oclusiva sorda o como lateral.

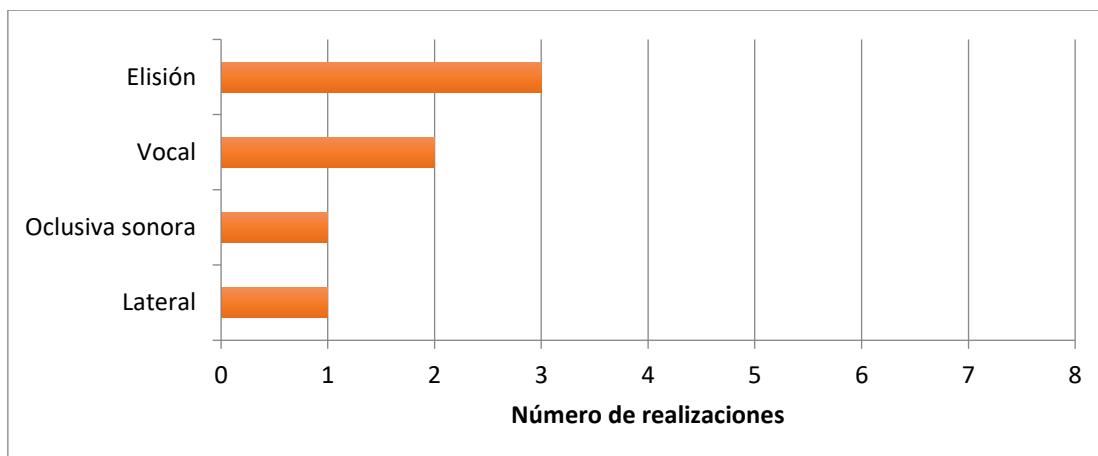
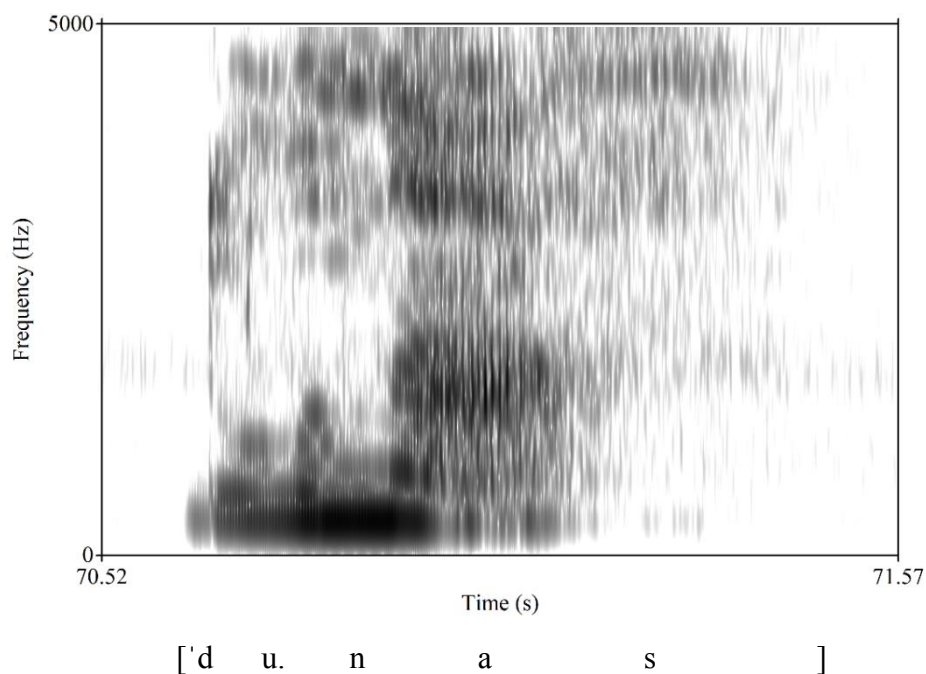


Figura 7. Número y tipo de realizaciones incorrectas de la nasal anterior a la vocal por las informantes sinohablantes.

Analizando la figura 7 en detalle se observa que hay 3 casos de elisión, 2 en los que la nasal es sustituida por una vocal, 1 en el que es realizado como oclusiva sonora, y otro en el que es realizado como lateral. Estos datos muestran que cuando una nasal se encuentra en posición anterior a la vocal las informantes sinohablantes, en la mayoría de ocasiones, la realizan como tal. En aquellas en las que la realizan de otra manera, no se detecta un patrón claro. Ejemplo de una realización de nasal como oclusiva sonora es la articulación de la palabra *nulas* como [ˈdu.nas], la cual aparece en el espectrograma 5.



Espectrograma 5. Realización de la palabra *nulas* por la informante 1.

### 2.1.7 Realización de las laterales

De acuerdo con lo explicado anteriormente, las informantes sinohablantes no presentan grandes dificultades para articular la consonante lateral como tal, ya que en un 96.2% de las ocasiones lo han hecho correctamente. Tal y como se puede apreciar en el gráfico de la figura 8, las informantes presentan cuatro comportamientos diferentes cuando no realizan de manera correcta la consonante lateral anterior a la vocal.



Figura 8. Número y tipo de realizaciones incorrectas de la lateral anterior a la vocal por las informantes sinohablantes.

En 2 ocasiones han realizado la lateral como oclusiva sonora, en una ocasión han elidido el sonido, en otra lo han realizado como vocal, y también en una ocasión lo han realizado como nasal. El alto porcentaje de realizaciones correctas, y la falta de homogeneidad en las realizaciones incorrectas de la lateral anterior a la vocal indican que las informantes sinohablantes no tienen problemas con este sonido, y que no se detecta un patrón claro de realización en los pocos casos en los que no realizan dicha consonante de manera correcta.

### 2.1.8 Realización de las vibrantes simples

Anteriormente se ha indicado que la vibrante simple en posición anterior a vocal supone un problema para las informantes sinohablantes, ya que el porcentaje de realizaciones correctas es solamente del 31.2%. La figura 9 muestra que, si bien las informantes han realizado la vibrante simple de tres maneras diferentes cuando no la han articulado correctamente, prácticamente en todos los casos la han realizado como lateral.

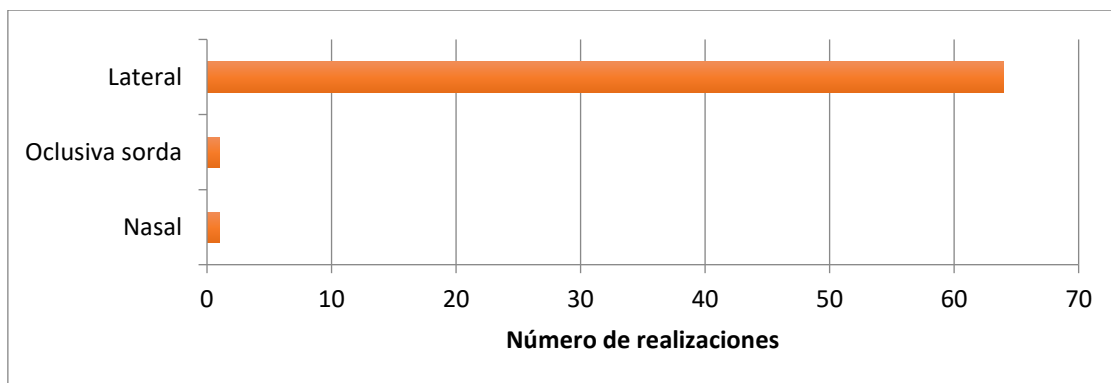
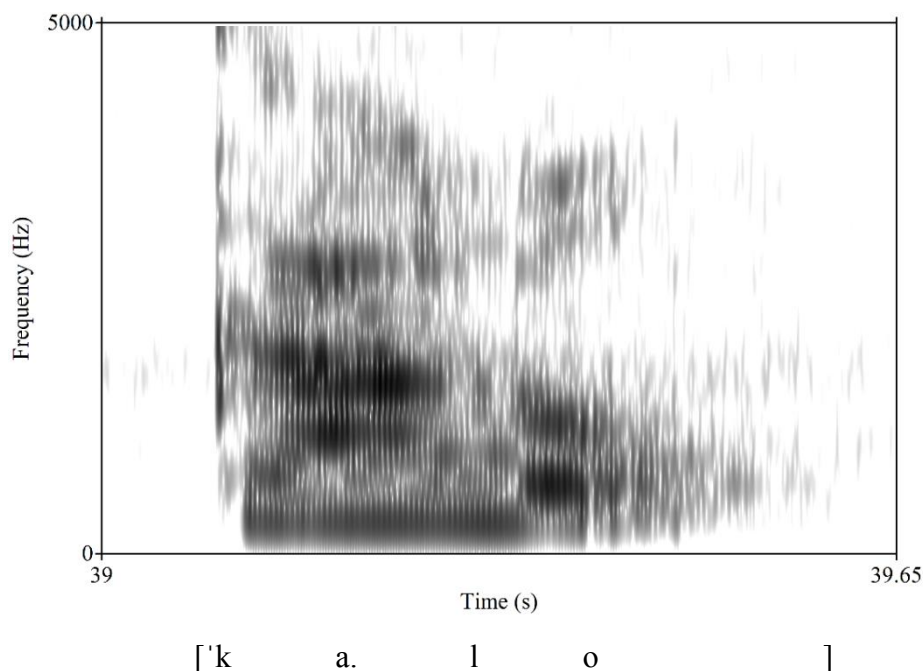


Figura 9. Número y tipo de realizaciones incorrectas de la vibrante simple anterior a la vocal por las informantes sinohablantes.

De las 96 vibrantes simples en posición anterior a vocal analizadas, las informantes han realizado dicho sonido como lateral hasta en 64 casos, es decir, en un 66.7% de las ocasiones. Además, en una ocasión la han realizado como oclusiva sorda, y en otra como nasal. Esto indica que las hablantes cuya lengua materna es el chino presentan dificultades para articular la vibrante simple como tal, y que la tendencia es a realizarla como lateral. Por eso, es común que las informantes articulen *caro* como [‘ka.lo], *cero* como [‘θe.lo] o *muro* como [‘mu.lo].



Espectrograma 6. Realización de la palabra *caro* por la informante sinohablante 2.

El espectrograma 6 muestra la realización de la palabra *caro*, la cual no presenta la fase de cierre característica de la vibrante simple.

### 2.1.9 Realización de las vibrantes múltiples

Las informantes sinohablantes han realizado la vibrante múltiple anterior a la vocal de manera correcta únicamente en el 13.2% de los casos. La figura 10 muestra que estas hablantes realizan dicho sonido de cuatro maneras diferentes: vibrante simple, lateral, oclusiva sorda y africada, siendo la primera la que más casos presenta.

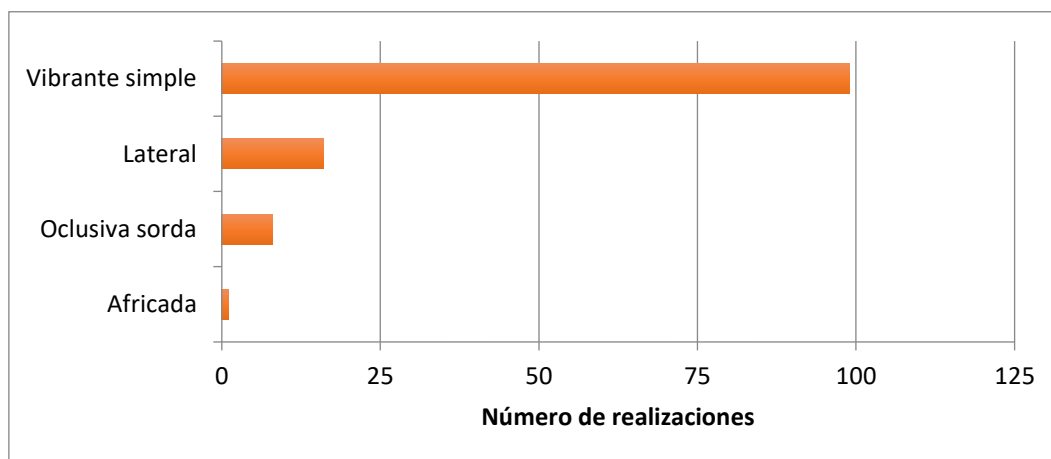
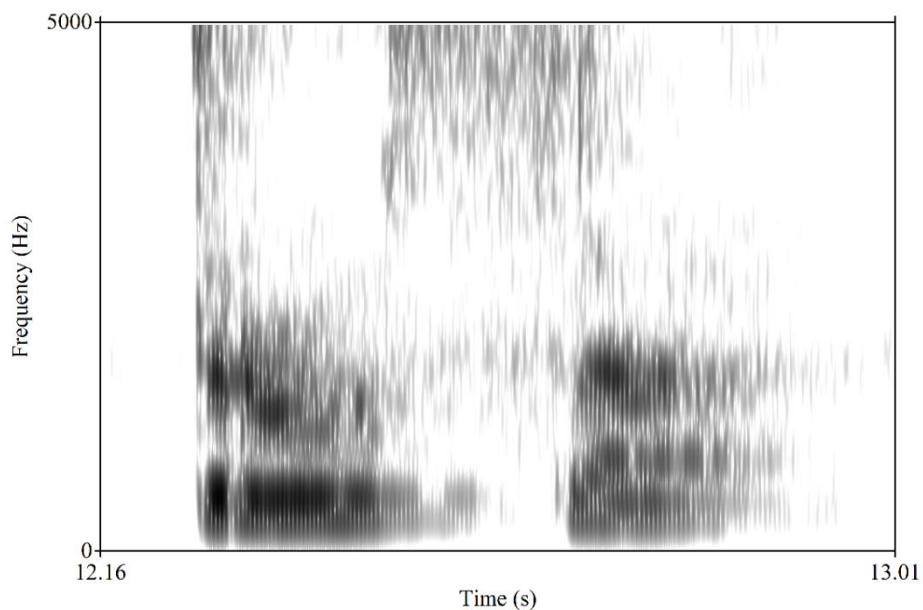


Figura 10. Número y tipo de realizaciones incorrectas de la vibrante múltiple anterior a la vocal por las informantes sinohablantes.

Las informantes sinohablantes realizan la vibrante múltiple en posición anterior a vocal como vibrante simple en 99 ocasiones, como lateral en 16, como oclusiva sorda en 8, y como africada en 1. Es decir, la tendencia que se observa es a realizar la vibrante múltiple como vibrante simple, y si no, como lateral. Estos datos muestran que las hablantes de lengua materna chino mandarín tienen dificultades para pronunciar este sonido, y que, generalmente, la realizan como vibrante simple. Ejemplo de ello es que palabras como *rosa*, *resto* o *rezan* las articulan como [ˈtro.sa], [ˈtres.to] y [ˈtre.θan] respectivamente. Esto muestra que, además, insertan otro sonido antes de la vibrante simple, generalmente una oclusiva sorda con lo que queda alterada la estructura silábica de la palabra.





[ ' t r o . s a ]

*Espectrograma 7.* Realización de la palabra *rosa* por la informante sinohablante 4.

En el espectrograma 7 ilustra dichas realizaciones, ya que en él se observa que en lugar de la vibrante múltiple, la sinohablante articula una oclusiva sorda seguida de una vibrante simple.

## 2.2 Entorno posterior

Al igual que se ha estudiado el entorno anterior a la vocal realizada por las informantes sinohablantes, en este apartado se muestran los resultados del análisis de las realizaciones del entorno posterior al sonido vocálico. Aquí, el Coeficiente kappa va desde 0.60 hasta 0.69, más bajo que en el caso del entorno anterior lo que indica que las realizaciones de estas estudiantes presentan problemas importantes en lo que respecta a la pronunciación de los sonidos consonánticos del español en posición de coda silábica.

Vocal	Coeficiente kappa
/i/	0.64
/e/	0.64
/a/	0.60
/o/	0.69
/u/	0.61

*Tabla 5.* Coeficiente Kappa para cada una de las vocales según el entorno posterior.

Igual que en el apartado anterior, se analizan a continuación las características de la realización de las consonantes que siguen a la vocal.

En una primera apreciación global, la figura 11 muestra que, cuando la vocal es seguida por una oclusiva sorda, una fricativa, una africada o una nasal, el porcentaje de realizaciones adecuadas se sitúa entre el 90% y el 100%. En cambio, la realización apropiada de las laterales es menor, se sitúa en el 60%. En cuanto a los sonidos más problemáticos para las informantes, se ha detectado que la realización de la vibrante simple se ha realizado correctamente en, aproximadamente, el 40% de los casos, y el porcentaje es menor cuando el sonido posterior a la vocal es una aproximante, ya que el porcentaje se sitúa solamente entre el 30 y 40%. La consonante oclusiva sonora y la vibrante múltiple son los sonidos que menor porcentaje de realizaciones correctas presenten, alcanzando el 30% la primera, y no llegando al 10% la segunda.

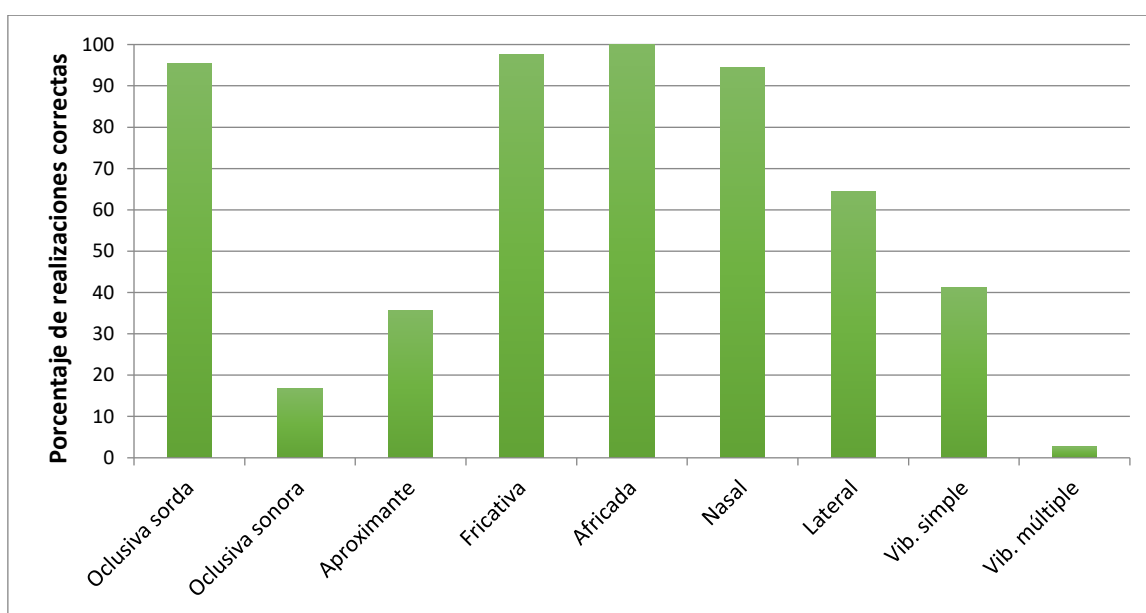


Figura 11. Porcentaje de realizaciones correctas del entorno posterior a la vocal analizada.

La tabla 6 muestra los datos detallados del porcentaje de realizaciones correctas del entorno posterior por parte del grupo de informantes cuya lengua materna es el chino. De la misma manera que sucede cuando la oclusiva sorda está en posición anterior, cuando está en posición posterior a la vocal el porcentaje de realizaciones correctas es muy elevado, del 95.3%, lo que indica que este sonido no parece suponer un problema para estas hablantes. En cuanto a las oclusivas sonoras, al igual que cuando preceden a la vocal, en posición posterior también parecen suponer un escollo para las informantes sinohablantes, ya que únicamente la han realizado como tal en el 16.6% de las ocasiones. Sin embargo, cabe señalar que el corpus únicamente presenta 24 casos de oclusiva sonora posterior a vocal, como corresponde a la estructura propia del español, por lo que este resultado muestra una tendencia, pero no un comportamiento generalizado. En cuanto a

la realización de las aproximantes, estas también suponen un escollo para las estudiantes chinas, ya que el porcentaje de realización adecuada de las mismas es del 35.6%. Como se ha indicado, las informantes sinohablantes han realizado la consonante fricativa posterior a vocal en un porcentaje de ocasiones muy elevado, concretamente del 97.5%, lo que indica que este sonido no supone un escollo para aquellos estudiantes de español cuya lengua materna es el chino. Cuando la consonante que sigue a la vocal es una africada, las informantes no tienen problemas para articular la misma, ya que lo han realizado bien en el 100% de las ocasiones, aunque cabe señalar que aquí también el número de casos en el corpus es reducido, 24 concretamente. Si el sonido que sigue a la vocal es una consonante nasal, el porcentaje de realizaciones correctas también es muy elevado, del 94.3%, lo que indica que este sonido no resulta problemático para las informantes sinohablantes. Sin embargo, mientras que en el apartado anterior los resultados muestran que cuando el sonido anterior a la vocal es una lateral las estudiantes chinas no parecen tener problemas para articularla correctamente, cuando la lateral está en posición posterior al sonido vocálico el porcentaje de realizaciones correctas desciende al 64.5%, lo que indica que estas hablantes tienen más problemas para articular la lateral cuando sigue a una vocal. El porcentaje de realizaciones adecuadas de vibrante simple es mayor cuando esta se encuentra en posición posterior es superior que cuando es anterior a la vocal, pero sigue siendo bajo, ya que las informantes sinohablantes la han realizado como corresponde en el 41.1%. Estos datos muestran que, si bien la vibrante simple es un sonido problemático para las informantes sinohablantes, causa más dificultades cuando está en posición anterior que cuando está en posición posterior. En cuanto a la vibrante múltiple, las informantes la han articulado como tal en una única ocasión, lo que muestra que cuando la vocal es seguida por una vibrante múltiple, esta supone un escollo totalmente insalvable para las hablantes cuya lengua materna es el chino. Sin embargo, ha de tenerse en cuenta que el corpus únicamente presenta 36 casos de este sonido en posición posterior a la vocal, por lo que este dato muestra una tendencia, pero no un comportamiento generalizado.

Sonido posterior	Total casos analizados	Realizaciones correctas	
		Nº de realizaciones	Porcentaje
Oclusiva sorda	149	142	95.3%
Oclusiva sonora	24	4	16.6%
Aproximantes	227	81	35.6%
Fricativa	372	363	97.5%
Africada	24	24	100%
Nasal	414	391	94.4%
Lateral	192	124	64.5%
Vib.simple	462	190	41.1%
Vib.múltiple	36	1	2.7%

Tabla 6. Realizaciones correctas del sonido posterior por las informantes sinohablantes.

A continuación se realiza un análisis de los sonidos que las informantes sinohablantes no han realizado como los del español. Los gráficos que se muestran a continuación reflejan cómo han articulado los distintos sonidos consonánticos y cuántos casos hay de cada realización.

### 2.2.1 Realización de las oclusivas sordas

Tal y como se ha indicado anteriormente, el número de realizaciones adecuadas de la oclusiva sorda en posición posterior a la vocal es muy elevado (95.3%), por lo que son pocas las otras realizaciones de dicho sonido. La figura 12 muestra que son tres las maneras en las que han realizado dicho sonido cuando no lo han hecho adecuadamente: como oclusiva sonora, como nasal, o han elidido el sonido.

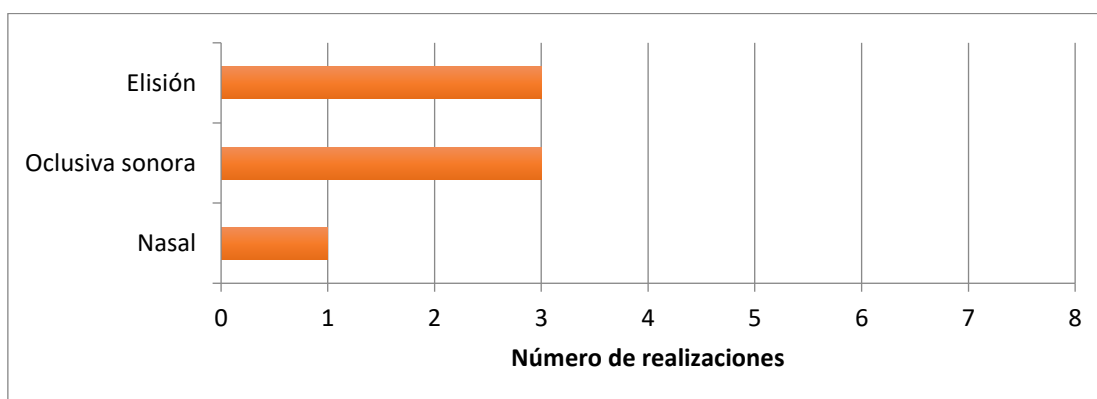


Figura 12. Número y tipo de realizaciones incorrectas de la oclusiva sorda posterior a la vocal por las informantes sinohablantes.

Las informantes han elidido la oclusiva sorda en 3 ocasiones, como en *mamut*, pronunciado como [ma.'mu]; la han realizado como oclusiva sonora en 3 ocasiones

también, como en *súper*, la cual pronuncian como [‘su.ber]; y la han realizado como nasal en una única ocasión, *botón* como [pon.‘ton]. Estos datos indican que las hablantes cuya lengua materna es el chino no presentan problemas a la hora de articular el sonido oclusivo sordo, y que, en caso de no hacerlo correctamente, no hay una tendencia clara de realización, aunque podría ser a una elisión o a realizarlo como oclusiva sonora.

### 2.2.2 Realización de las oclusivas sonoras

Anteriormente se ha indicado que la realización de la oclusiva sonora es una de las que más dificultades parece presentar para las informantes sinohablantes, ya que únicamente la han realizado de manera correcta en el 16.6% de las ocasiones, aunque cabe señalar que el número de casos en los que dicho sonido aparece tras vocal en el corpus es de 24, por lo que, como se ha indicado, este dato muestra una tendencia, no un comportamiento generalizado. Cuando la realización no ha sido la adecuada, las informantes han elidido la oclusiva sonora, o la han realizado como oclusiva sorda o como vibrante múltiple, tal y como muestra la figura 13.

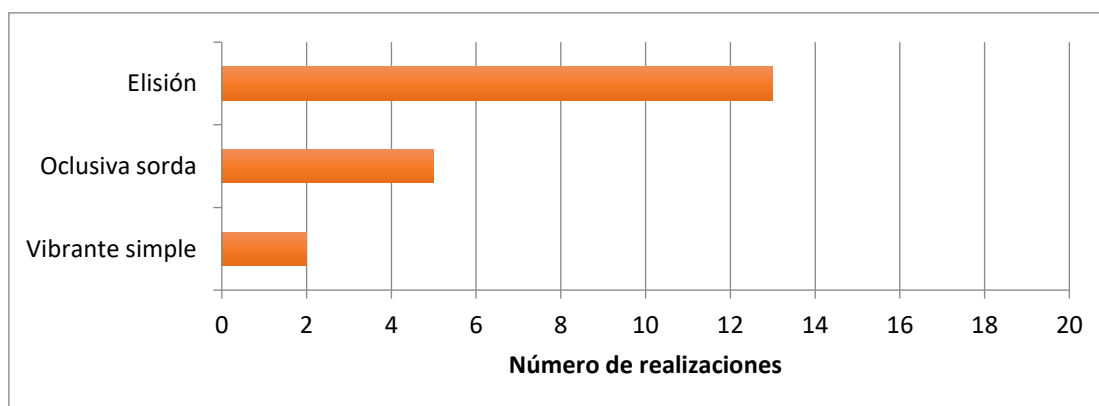
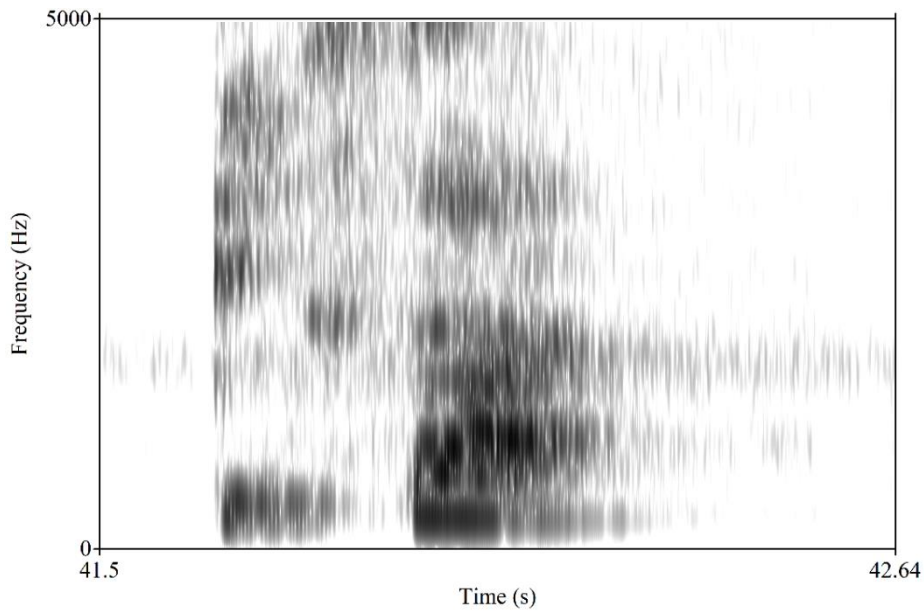


Figura 13. Número y tipo de realizaciones incorrectas de la oclusiva sonora posterior a la vocal por las informantes sinohablantes.

Las estudiantes chinas han elidido la oclusiva sonora en 13 ocasiones, la han realizado como oclusiva sorda 5 veces y 2 como vibrante simple. Estos datos revelan que la tendencia general es a elidir el sonido oclusivo sonoro, pronunciando palabras como *pisad* o *comed* como [pi.‘sa] o [ko.‘me] respectivamente, o a realizarlo como sordo, *comed* como [ko.‘met] o *pisad* como [pi.‘sat]. El espectrograma 6 muestra un ejemplo en el que la informante elide la oclusiva sonora de la palabra *pisad*.



[ p i . ' s a ]

Espectrograma 8. Realización de la palabra *pisad* por la informante sinohablante 3.

### 2.2.3 Realización de las aproximantes

Las informantes sinohablantes han realizado la consonante aproximante como tal en el 35.6% de los casos. Como se puede observar en la figura 14, hay una clara tendencia a realizar dicho sonido como oclusiva sorda, y en un número más reducido de ocasiones la han articulado como nasal.

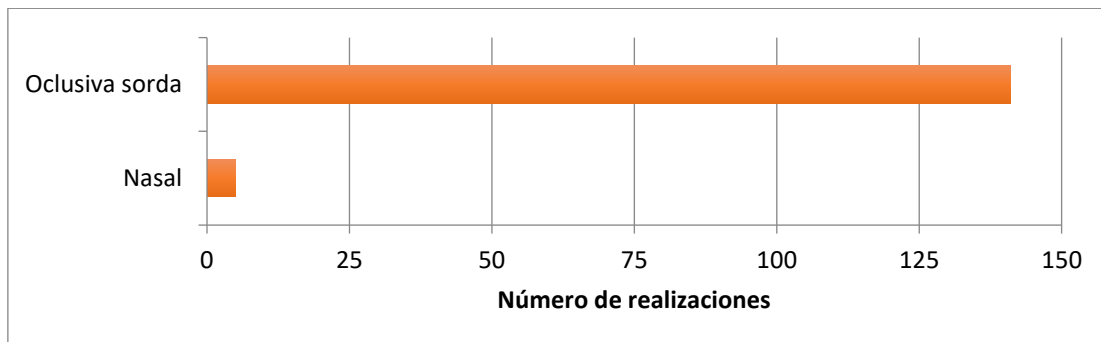


Figura 14. Número y tipo de realizaciones incorrectas de la aproximante posterior a la vocal por las informantes sinohablantes.

De los 227 casos de aproximante posterior a vocal que figuran en el corpus, las informantes sinohablantes han realizado dicho sonido consonántico como oclusiva sorda en 141 ocasiones, y como nasal en 5. Esto muestra que este sonido supone un escollo para estas hablantes y que, además, hay una clara tendencia a realizarlo como oclusiva sorda. Es decir, las estudiantes chinas han articulado palabras como *dudó* o *viví* como [tun.'to] o [pi.'pi] respectivamente.

#### 2.2.4 Realización de las fricativas

Cuando el sonido que sigue a la vocal es una fricativa el porcentaje de realizaciones correctas es del 97.5%, por lo que este sonido no supone un problema importante para las informantes sinohablantes. La figura 15 muestra que, en las ocasiones en las que el sonido no es realizado adecuadamente, se ha elidido, o se ha articulado como oclusiva sorda o como vocal.

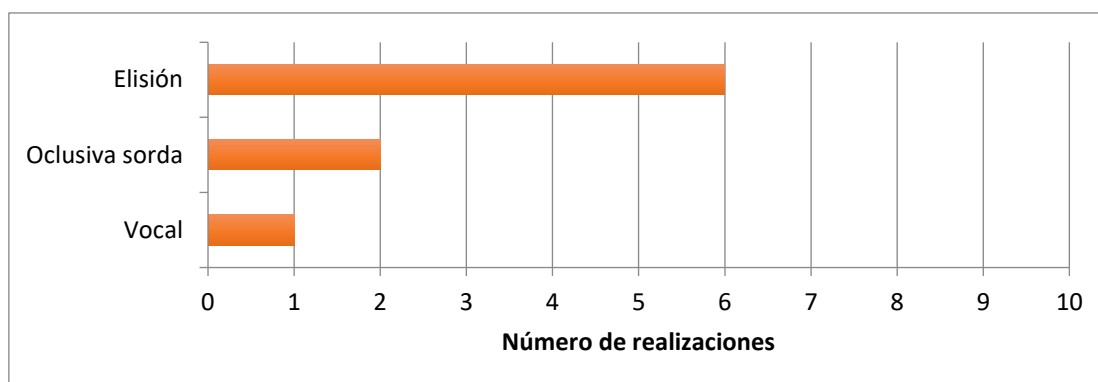


Figura 15. Número y tipo de realizaciones incorrectas de la fricativa posterior a la vocal por las informantes sinohablantes.

Las informantes sinohablantes han elidido la fricativa posterior a vocal en 6 ocasiones, en 2 la han realizado como oclusiva sorda, y en una como vocal. Ejemplo de ello es la articulación de palabras como *reloj*, *lápiz* o *ruge* como [re.'lo], ['la.pi] y ['ru.ke] respectivamente. El alto porcentaje de realizaciones adecuadas, y la falta de homogeneidad en las demás indican que las estudiantes chinas no tienen problemas con este sonido, y que no se detecta un patrón de realización cuando no realizan la fricativa como tal.

#### 2.2.5 Realización de las africadas

Tal y como se ha indicado anteriormente, las informantes sinohablantes han realizado todos los sonidos consonánticos africados de manera correcta, por lo que, teniendo en cuenta que el número de casos de aparición de africadas es reducido, se puede considerar que esta consonante no supone un problema para las hablantes que constituyen el grupo de estudio.

#### 2.2.6 Realización de las nasales

El porcentaje de realizaciones correctas de la nasal cuando sigue a una vocal es muy elevado (94.3%). Cuando las informantes no han realizado dicho sonido como tal,

lo han elidido o lo han realizado como oclusiva sorda o como lateral, tal y como muestra la figura 16.

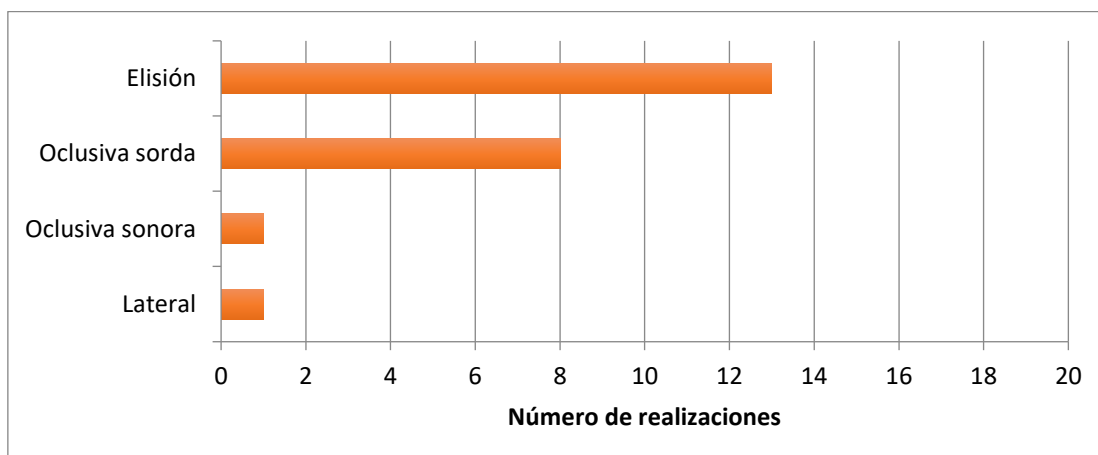
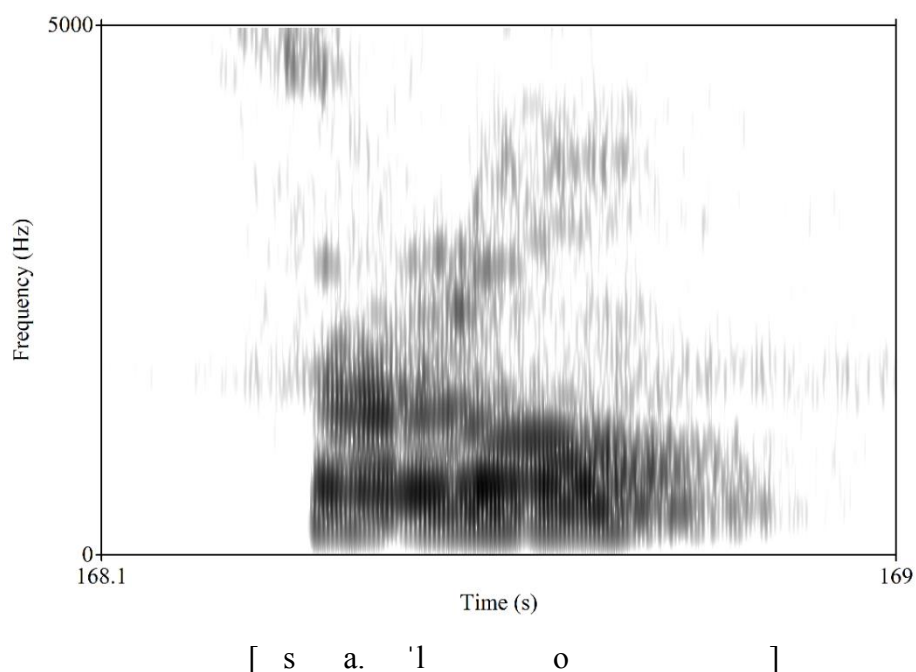


Figura 16. Número y tipo de realizaciones incorrectas de la nasal posterior a la vocal por las informantes sinohablantes.

Las informantes han elidido la nasal en 13 ocasiones, la han realizado como oclusiva sorda 8 veces, una vez como oclusiva sonora, y otra como lateral. Estos datos indican que las hablantes de lengua materna china no tienen problemas para realizar la nasal, y que cuando no la realizan como tal, tienden a elidirla o a realizarla como oclusiva sorda. Cuando la eliden, articulan palabras como *limón* como [li.'mo] o *salón* como [sa.'lo], la cual aparece reflejada en el espectrograma 7, en el cual se aprecia que la palabra acaba en vocal, en lugar de nasal.



Espectrograma 9. Realización de la palabra *salón* por la informante sinohablante 6.



### 2.2.7 Realización de las laterales

Anteriormente se ha indicado que el porcentaje de realizaciones correctas de la lateral cuando sigue a un sonido consonántico es del 64.5%, lo que indica que las informantes sinohablantes tienen dificultades para realizar dicho sonido del español. La figura 17 muestra que las hablantes han realizado la lateral de maneras muy diversas: como vibrante simple, la han elidido, como fricativa, como oclusiva sorda, como nasal o incluso como vocal.

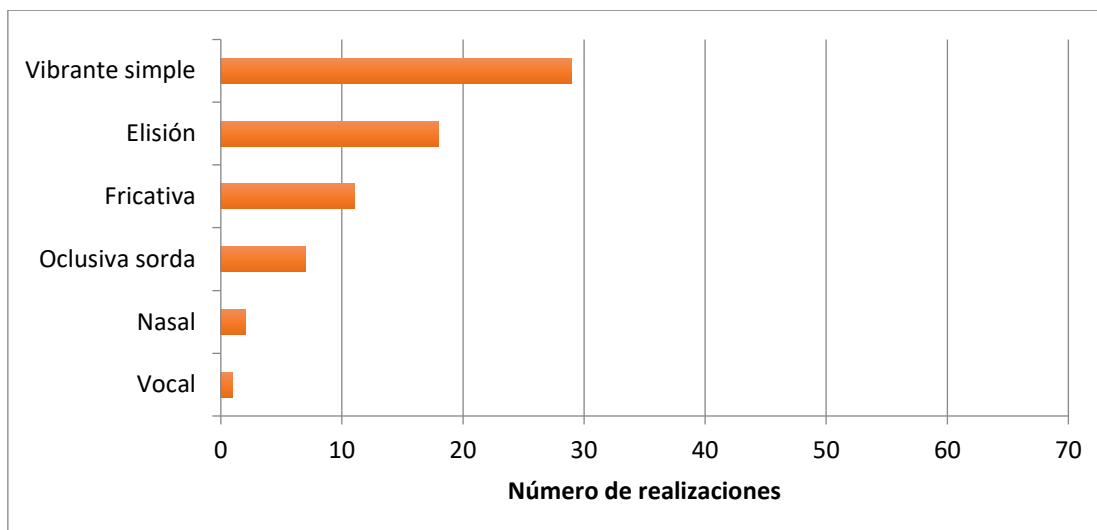
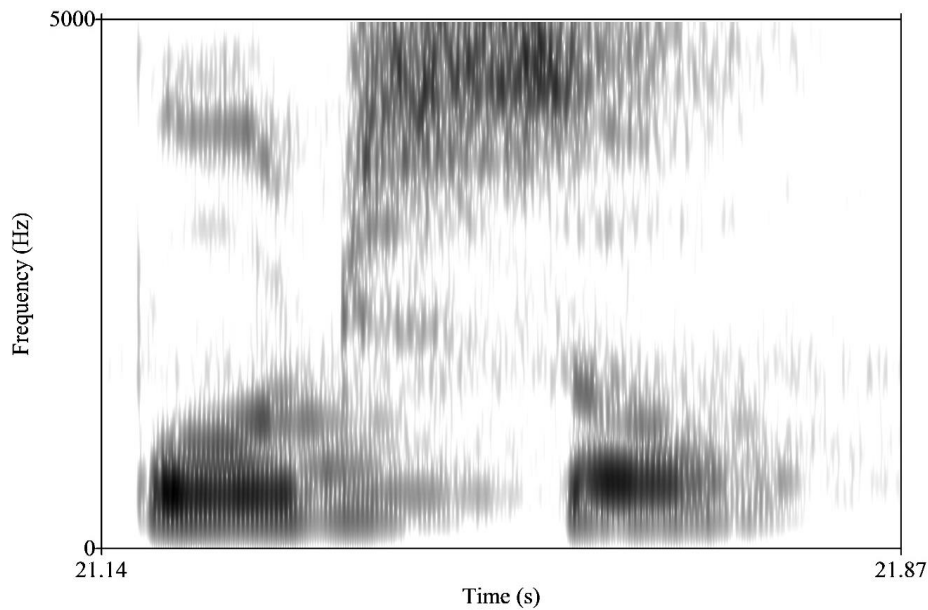


Figura 17. Número y tipo de realizaciones incorrectas de la lateral posterior a la vocal por las informantes sinohablantes.

Las informantes sinohablantes han realizado la lateral de manera incorrecta hasta en 68 ocasiones, de las cuales 29 veces la han pronunciado como vibrante simple, en 18 ocasiones la han elidido, 11 veces la han realizado como fricativa, 7 como oclusiva sorda, 2 como nasal y una como vocal. Esta amplia variedad de realizaciones indica que, si bien las hablantes cuya lengua materna es el chino mandarín tienen dificultades para pronunciar la lateral correctamente, no hay homogeneidad a la hora de realizar dicho sonido, aunque sí se observa cierta tendencia a articularlo como vibrante simple o a elidirlo. Así, se pueden encontrar realizaciones de *bolso* como [ˈpor.so], *culto* como [ˈkur.to] o *papel* como [pa.ˈpe].



[ˈp o r s o]

*Espectrograma 10.* Realización de la palabra *bolsa* por una informante sinohablante 1.

En el espectrograma 10 se puede apreciar que, en lugar de la lateral, la informante articula una vibrante simple al producir la palabra *bolsa*.

### 2.2.8 Realización de las vibrantes simples

Como se ha indicado, el porcentaje de realizaciones adecuadas de la vibrante simple en posición posterior a la vocal es del 41.1%. La figura 18 muestra que cuando las informantes sinohablantes no realizan este sonido como tal, la variedad de realizaciones es amplia, aunque destaca que en más de 175 ocasiones la han articulado como lateral. También se observa un número considerable de casos en los que las estudiantes eliden la vibrante y no articulan otro sonido, y hay ocasiones en las que la realizan como oclusiva sorda, fricativa, o nasal. Además, hay casos aislados en los que las informantes articulan una oclusiva sonora o una vibrante múltiple.

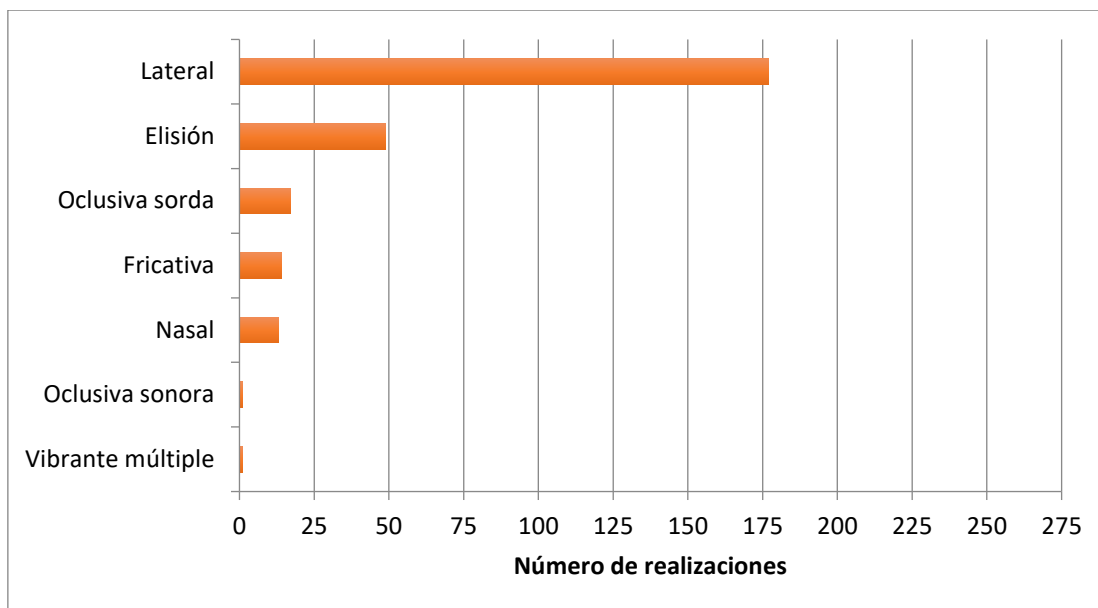
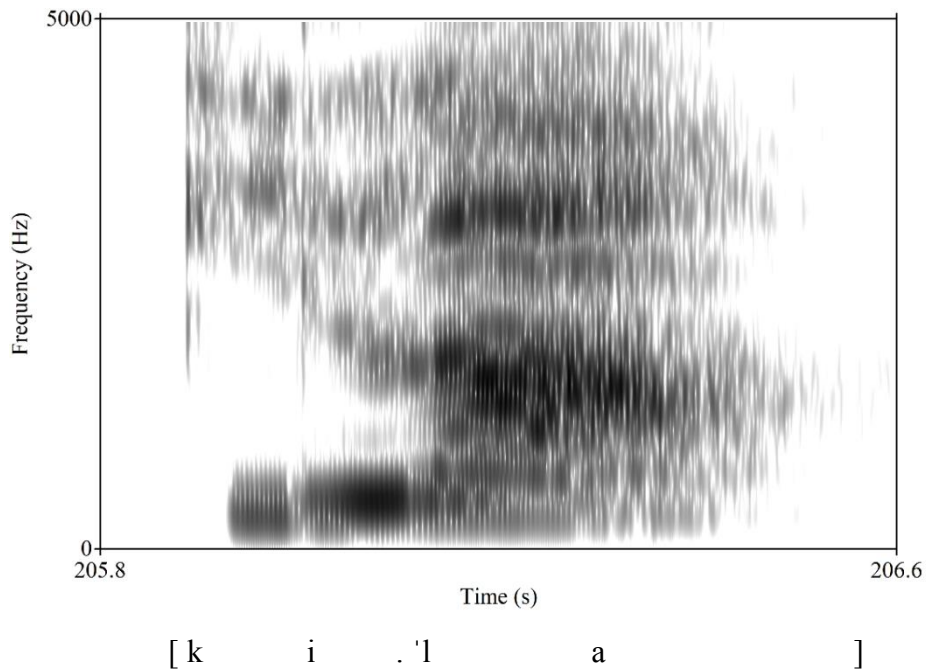


Figura 18. Número y tipo de realizaciones incorrectas de la vibrante simple posterior a la vocal por las informantes sinohablantes.

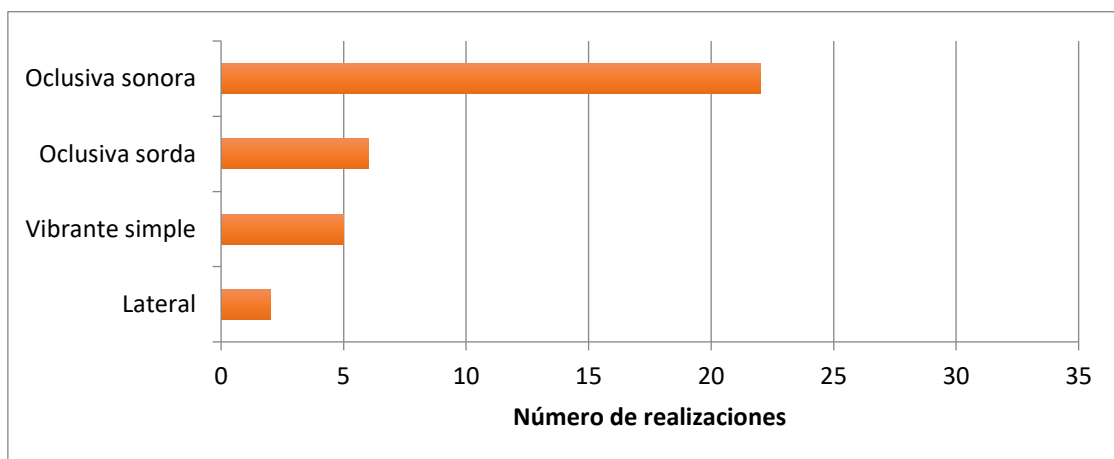
El número de ocasiones en las que la realización de la vibrante simple se ha pronunciado como lateral es claramente superior a las otras realizaciones posibles, y las informantes la han realizado como tal en 177 ocasiones, es decir, en un 38.3% de las ocasiones. Como ya se ha indicado, el número de elisiones también es considerable, 49 en total. En cuanto a otros tipos de realizaciones de este sonido por parte de las informantes sinohablantes, éstas han realizado la vibrante simple hasta 17 veces como oclusiva sorda, 14 como fricativa, 13 como nasal, 1 como oclusiva sonora, y otra como vibrante múltiple. Estos datos muestran que las hablantes cuya lengua materna es el chino tienen dificultades para pronunciar la vibrante simple como tal, y que hay una clara tendencia a realizar la vibrante simple como lateral o a elidirla, como en las palabras *coger*, *forma*, *comer* o *salir*, las cuales han sido realizadas como [ko.'xel], ['fol.ma], [ko.'me] y [sa.'li] respectivamente.



*Espectrograma 11.* Realización de la palabra *girar* por la informante sinohablante 1.

### 2.2.9 Realización de las vibrantes múltiples

Las informantes sinohablantes objeto de estudio de esta investigación han realizado la vibrante múltiple del español en posición posterior a vocal como tal en una única ocasión (2.7%). El resto de realizaciones se dividen entre oclusiva sonora, oclusiva sorda, vibrante simple y lateral. Sin embargo, cabe señalar que, como ya se ha indicado, el número de casos en los que este sonido aparece en esta posición en el corpus es más reducido (36).



*Figura 19.* Número y tipo de realizaciones incorrectas de la vibrante múltiple posterior a la vocal por las informantes sinohablantes.

El análisis ha permitido detectar que las informantes han realizado la vibrante múltiple como oclusiva sonora 22 veces, 6 como oclusiva sorda, 5 como vibrante simple y 2 como

lateral. Estos datos muestran que las hablantes cuya lengua materna es el chino tienen dificultades para pronunciar la vibrante múltiple del español y que, generalmente, la realizan como oclusiva sonora. Ejemplo de ello es que pronuncian *currar* como [ku.ˈdra], o *torre* como [ˈto.dre]. Esto muestra que las informantes tienden a articular la vibrante múltiple como un conjunto de oclusiva sonora y vibrante simple.

### 2.3 Realización de la vocal en posición final de palabra aislada

Además de estudiar la articulación del entorno anterior y posterior a la vocal, se ha analizado la realización de los sonidos vocálicos del español producidos por las informantes sinohablantes cuando estos se encuentran en posición final de palabra aislada. La tabla 7 muestra que el Coeficiente kappa va desde 0.9 hasta 1, lo que indica que el grado de acuerdo entre el español y lo realizado por las informantes sinohablantes es muy elevado.

Vocal	Coeficiente Kappa
/i/	0.9
/e/	1
/a/	0.9
/o/	0.9
/u/	1

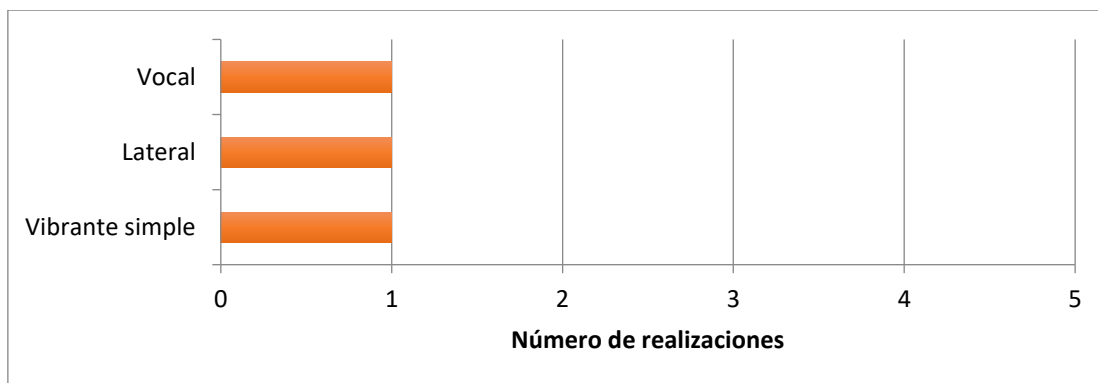
Tabla 7. Coeficiente Kappa para cada una de las vocales cuando se encuentran en posición final de palabra.

La tabla 8 muestra que en el 99.5% de los casos las realizaciones de vocal en posición final de palabra se han realizado de manera correcta, es decir, no se ha articulado otro sonido posterior a dicha vocal.

Posición	Total casos analizados	Realizaciones correctas	
		Nº de realizaciones	Porcentaje
Final de palabra	716	713	99.5%

Tabla 8. Realizaciones correctas de la vocal en posición final de palabra por las informantes sinohablantes.

La figura 20 muestra que en las tres únicas ocasiones en las que las informantes sinohablantes han realizado un sonido posterior a la vocal analizada, han realizado una vocal en una ocasión, una lateral en otra ocasión, y una vez una vibrante simple.



*Figura 20.* Número y tipo de realizaciones incorrectas por las informantes sinohablantes cuando la vocal se encuentra en posición final de palabra.

Estos datos muestran que la articulación de vocales del español en posición final de palabra aislada no supone un problema para las informantes sinohablantes.

## 2.4 Recapitulación

El cálculo del Coeficiente kappa de Cohen ha permitido observar que la articulación de los sonidos consonánticos que preceden o siguen al sonido vocálico presenta numerosos problemas que, como se verá, influyen en las características del timbre de las realizaciones. El resultado del análisis de dichos problemas puede resumirse como se indica a continuación.

Las consonantes oclusivas sordas se encuentran entre los sonidos que son menos problemáticos para las informantes sinohablantes, ya que el porcentaje de realizaciones correctas es superior al 95%, tanto cuando se encuentran en posición anterior a la vocal, como cuando siguen al sonido vocálico. En aquellos casos en los que no han articulado la oclusiva sorda como tal, las estudiantes chinas la han realizado, en la mayoría de los casos, como oclusiva sonora. Sin embargo, también se han detectado elisiones de este sonido, o articulaciones aisladas de este como fricativa, vibrante simple, o nasal.

En cambio, las consonantes oclusivas sonoras resultan problemáticas para las estudiantes chinas, ya que el porcentaje de realizaciones correctas es inferior al 20%. Se ha podido observar que estas hablantes, en lugar de articular este sonido, en la mayoría de los casos han producido una oclusiva sorda en su lugar. También hay casos en los que el sonido consonántico se ha elidido, y un número menor de ocasiones en las que lo han realizado como nasal, fricativa o vibrante simple.

Las consonantes aproximantes también suponen un gran escollo para estas hablantes, dado que el porcentaje de realizaciones correctas no supera el 39%. En aquellos casos en los que no han realizado la aproximante como tal, las hablantes han mostrado una importante tendencia a articularla como oclusiva sorda. En cuanto a otro tipo de realizaciones, se han detectado casos aislados en los que la aproximante se ha articulado como nasal, como vocal, o se ha elidido dicho sonido.

En cambio, las fricativas son menos problemáticas. Como se ha mostrado, las informantes sinohablantes las han realizado correctamente en más del 97% de los casos. En las ocasiones en las que no las han articulado como tal, las estudiantes han elidido la consonante, o la han realizado como oclusiva sorda, oclusiva sonora, lateral o como vocal.

El análisis del entorno de la vocal, tanto del anterior, como del posterior, muestra que las informantes sinohablantes no tienen dificultades para articular las consonantes africadas, ya que las han realizado correctamente en el 100% de los casos.

En cuanto a la correcta articulación de las consonantes nasales, estas han sido realizadas como tal en más del 94% de las ocasiones, lo que indica que las informantes sinohablantes tienen menos problemas para realizar estos sonidos que para articular otros. En cuanto a las realizaciones incorrectas de las nasales, los datos muestran que las estudiantes chinas la han elidido o la han realizado como oclusiva sorda, y en ocasiones aisladas la han articulado como lateral.

Mientras que en el porcentaje de realizaciones correctas de los sonidos vocálicos anteriormente analizados es similar cuando se encuentra tanto en entorno anterior como en posterior, en el caso de las consonantes laterales la diferencia es considerable. Cuando el sonido lateral se encuentra en posición anterior a la vocal, el porcentaje de realizaciones correctas es del 96%, mientras que, si este sonido se encuentra en posición posterior a la misma, dicho porcentaje desciende considerablemente y es del 64%. Estos datos muestran que la lateral es un sonido considerablemente menos problemático para las estudiantes sinohablantes cuando se encuentra en posición anterior a la vocal, que cuando sigue a la misma. En los casos en los que la realización no ha sido correcta, las informantes han mostrado cierta tendencia a realizarla como vibrante simple, o a elidirla. Además, hay casos en los que las estudiantes han articulado la lateral como fricativa, oclusiva sorda, oclusiva sonora, nasal, e incluso como vocal.

En cuanto a la articulación de la vibrante simple, esta ha sido realizada correctamente en un 31% de los casos cuando se encuentra en posición anterior, y en el 41% cuando se encuentra en posición posterior. Por lo tanto, si bien las informantes sinohablantes tienen problemas para articular correctamente este sonido, es menos problemático cuando sigue a la vocal que cuando la precede. En cuanto a los casos en los que la articulación no ha sido la correcta, las estudiantes chinas, generalmente, la han articulado como lateral, aunque también en muchas ocasiones la han elidido. Menos frecuentemente, han realizado la vibrante simple como oclusiva sorda, fricativa, nasal, y en casos aislados como oclusiva sonora o vibrante múltiple.

En lo que concierne a la realización de la vibrante múltiple, las informantes sinohablantes la han realizado como tal únicamente en el 13% de las ocasiones cuando precede a la vocal, y en el 2% de las ocasiones cuando sigue al sonido vocálico. Estos datos muestran que la articulación de este sonido por parte de las estudiantes chinas es un escollo prácticamente insalvable, y es el que más problemas supone para estas hablantes. En cuanto a los sonidos articulados en lugar de la vibrante múltiple, las informantes sinohablantes han mostrado una clara tendencia a realizarla como vibrante simple. Se ha comprobado que en un número considerable de ocasiones el sonido producido ha sido una oclusiva sonora, y en menor medida una lateral. También se encuentran realizaciones de vibrante múltiple como oclusiva sorda, lateral o africada.

Se ha analizado también la realización de las vocales en posición final de palabra aislada, y se ha comprobado que las informantes sinohablantes han respetado esta estructura en el 99% de las ocasiones, por lo que se puede considerar que esta realización no supone un problema para las informantes sinohablantes. En las tres ocasiones en las que no la han realizado como tal, las informantes han añadido una vocal, una lateral o una vibrante simple.

Finalmente, y de manera global, se aprecia que la realización de las consonantes del entorno anterior y posterior a la vocal considerada pueden cambiar sustancialmente la estructura de la palabra pues fenómenos como la pronunciación de un tipo de consonante diferentes a la que es propia del español o su elisión influye, mediante la coarticulación, en las características de las vocales que forman parte de la sílaba. Ello significa que las estudiantes realizan vocales en los lugares de la cadena fónica en los que así lo hacen los hispanohablantes nativos pero, como se verá



en el apartado dedicado a estudiar el timbre de las realizaciones vocálicas, las características de dichos sonidos con frecuencia no son equiparables a las de las vocales de los hablantes nativos.

### 3. ESTRUCTURA SILÁBICA

En este apartado se analiza la realización de las distintas estructuras silábicas del español por parte de las informantes sinohablantes. Igual que en los casos anteriores, se ha calculado el Coeficiente kappa, esta vez en función de la estructura de la sílaba en la que se encuentra la vocal. Como se puede apreciar en la tabla 9, los valores del coeficiente oscilan desde 0.17 hasta 0.21, de lo que cabe inferir que las alumnas sinohablantes tienen grandes dificultades para realizar la estructura silábica del español, tan diferente de la propia del chino mandarín.

Vocal	Coeficiente Kappa
/i/	0.21
/e/	0.19
/a/	0.18
/o/	0.17
/u/	0.21

Tabla 9. Coeficiente kappa para cada una de las vocales según la estructura de la sílaba.

De acuerdo con la escala de valoración de Landis y Koch (1977), estos Coeficientes indican que el grado de acuerdo entre la estructura del español y la producida por las informantes sinohablantes es bajo. Ello se debe a que las dos únicas estructuras silábicas posibles en mandarín son CV y CVC, contrariamente al español, cuya estructura silábica admite más combinaciones de sonidos. Así, se han analizado las estructuras silábicas producidas por las informantes sinohablantes y cómo las han integrado en la articulación de la palabra. A continuación se presenta, en primer lugar, el porcentaje de realizaciones correctas de cada una de las estructuras consideradas en la presente tesis doctoral (CV-CV, CV-CVC y CVC-CV) con el fin de mostrar cuáles son más o menos problemáticas para los alumnos en el proceso de aprendizaje del español. En segundo lugar, se analizan los casos en los que las estudiantes chinas no han realizado las estructuras de la palabra correctamente, y se muestran ejemplos de las distintas realizaciones.

La figura 21 muestra el porcentaje de realizaciones correctas de cada una de las estructuras analizadas. Tal y como se puede observar en la figura 21, las informantes sinohablantes respetan la estructura de las palabras CV-CV en más del 90% de ocasiones; la estructura CVC-CV es realizada como tal en más del 80% de ocasiones, y la estructura que más dificultades parece plantear a las informantes es la formada por CV-CVC, ya que el porcentaje de realizaciones correctas no alcanza el 70%.

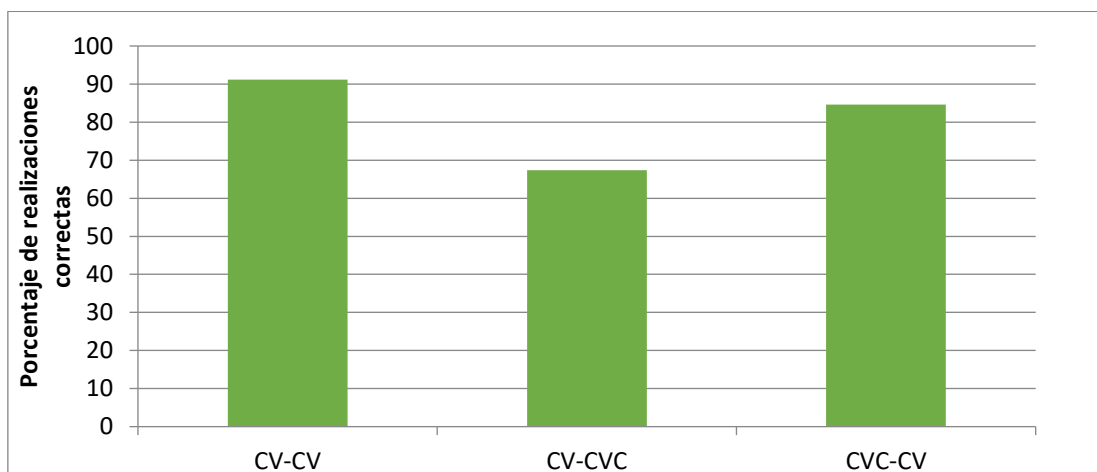


Figura 21. Porcentaje de realizaciones correctas de las distintas estructuras silábicas del español por las informantes sinohablantes.

La tabla 10 muestra los datos detallados correspondientes a la figura 21. Como se ha observado, la estructura CV-CV es la menos problemática para las informantes, ya que, de 882 casos, han realizado dicha estructura correctamente en 804 ocasiones, es decir, en el 91.1% de las veces han mantenido la estructura CV-CV. Como ya se ha indicado, esta es también la estructura más habitual en chino mandarín. La estructura CV-CVC es la que más problemas causa a las informantes, ya que únicamente la han realizado como tal en el 67.3% de las ocasiones. En relación con la estructura CVC-CV, las estudiantes chinas tienen menos problemas para realizarla correctamente que la estructura CV-CVC, ya que han mantenido la estructura propia del español en el 84.6% de las ocasiones.

Estructura palabra español	Total casos analizados	Realizaciones correctas	
		Nº casos	Porcentaje
CV-CV	882	804	91.1%
CV-CVC	846	570	67.3%
CVC-CV	846	716	84.6%

Tabla 10. Porcentaje de realizaciones de la estructura de la palabra de las estudiantes sinohablantes contrastada con la realización propia del español.

A continuación, se analizan de manera detallada los casos en los que las informantes sinohablantes no mantienen las estructuras silábicas propias del español.

La Real Academia Española (2011: 326) indica que el término silabación hace referencia a “la división en sílabas tanto en la pronunciación como en la escritura, división basada en el establecimiento de los límites o fronteras entre las agrupaciones de segmentos en sílabas distintas”. Tal y como se ha señalado en el capítulo II (estado de la cuestión) y se ha podido observar en el análisis auditivo del corpus de sinohablantes, los estudiantes de español cuya lengua materna es el chino mandarín no realizan la silabación propia del español, ya que, dado que en chino no hay grupos consonánticos, estos suponen un problema para dichos alumnos (Cortés Moreno, 2009a, 2009b). Por ello, estos hablantes recurren a una simplificación de los mismos (Blanco Nogueroles 2013), y generalmente dicha reducción afecta a la primera consonante que integra el grupo: CV, de forma que palabras como *frontal*, *flexión* y *bravo* se realizan como [fon.'tal], [fe.'ksion] y ['ba.βo] respectivamente (Poch e Igarreta, 2014:155). Este fenómeno se asemeja al denominado resilabación, el cual, según la Real Academia Española (2011: 326) “se emplea para designar los procesos de modificación que una secuencia de segmentos puede experimentar en sus límites silábicos por diversas razones”. Por ejemplo, cuando el artículo *el* [e] y el sustantivo *arma* ['ar.ma] aparecen juntos en la secuencia *el arma* se resilabea como [e.'lar.ma] (RAE, 2011: 331). Si bien este término no describe exactamente la modificación en la estructura de la sílaba que realizan los estudiantes sinohablantes, es el que más se asemeja al fenómeno que se analiza a continuación. Por esta razón, en adelante se empleará el término *resilabación*.

### 3.1 Estructura CV-CV

Como ya se ha indicado, el porcentaje de realizaciones correctas de la estructura CV-CV es elevado. Sin embargo, resulta de gran interés analizar qué tipo de estructuras realizan las informantes sinohablantes cuando no respetan la propia de las palabras en español. La figura 22 muestra que las hablantes han realizado hasta 8 tipos de estructura diferentes, y, como se observa, aquellas que aparecen con mayor frecuencia son CCV-CV, CV-CCV y CVC-CV.

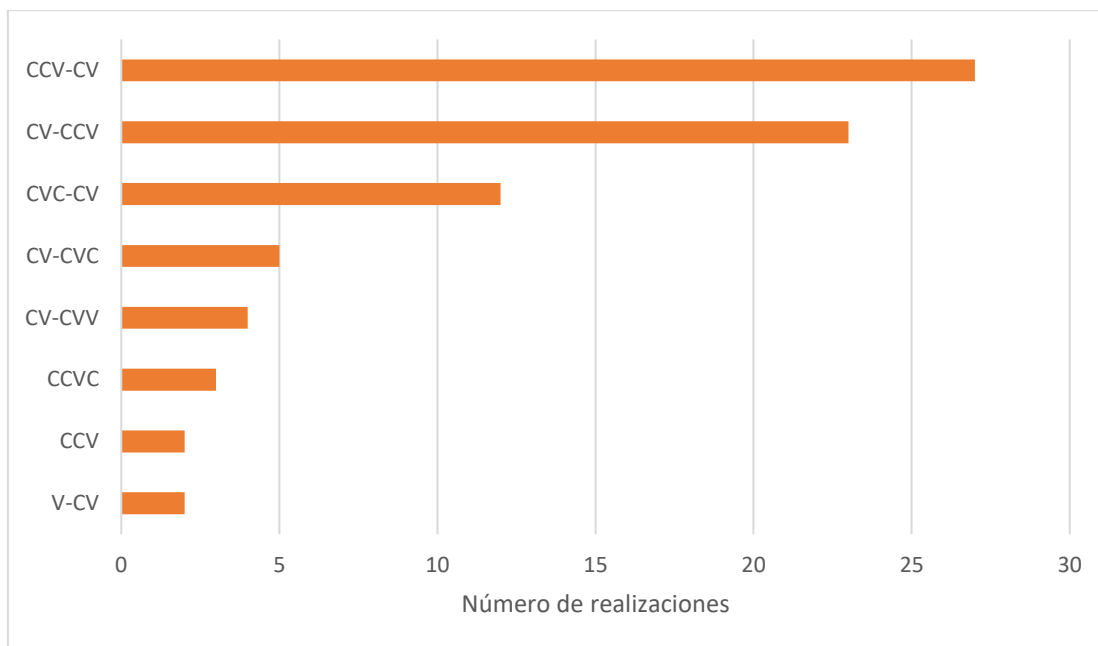


Figura 22. Número y tipo de realizaciones incorrectas de la estructura CV-CV por las informantes sinohablantes.

Las informantes sinohablantes han realizado la estructura CV-CV como CCV-CV en 27 ocasiones, como CV-CCV en 23 ocasiones y CVC-CV y 12 ocasiones. Ejemplo de ello es la articulación de palabras como *rosa*, *torre* o *tabú* como ['tro.sa], ['to.dre] y [tam.'pu] respectivamente. Por tanto, se observa que la tendencia general es a insertar un sonido entre consonante y vocal, tanto en la primera como en la segunda sílaba, especialmente cuando el sonido anterior o posterior es uno de los más problemáticos para las sinohablantes, como se ha mostrado en el apartado 2.

Sin embargo, tal y como se ha indicado, las informantes han realizado también otros tipos de estructura, aunque en un número de ocasiones menor: CV-CVC aparece en 5 casos, CV-CVV en 4 y V-CV en 2. Ejemplo de estas realizaciones es la articulación de palabras como *leña* o *jugó* como ['le.nia] y [u'go]. Por tanto, se observa que la tendencia general es a insertar un sonido entre consonante y vocal, tanto en la primera como en la segunda sílaba. Como se ha señalado anteriormente, todas las palabras del corpus son bisílabas. Sin embargo, las informantes sinohablantes han articulado palabras cuya estructura es CV-CV como monosílabas, y las estructuras realizadas son CCVC y CCV, las cuales han producido en 3 y 2 ocasiones respectivamente, articulando así la palabra *mutó* como ['mto], por ejemplo.

### 3.2 Estructura CV-CVC

Los datos mostrados en la tabla 10 muestran que la estructura CV-CVC es la que mayores dificultades causa a las informantes sinohablantes, ya que únicamente la han realizado correctamente en el 67.3% de las ocasiones. La figura 23 muestra las estructuras producidas por estas hablantes en lugar de CV-CVC. Se observa que han producido hasta 13 tipos de estructura diferentes, aunque destaca la realización de CV-CVC como CV-CV.

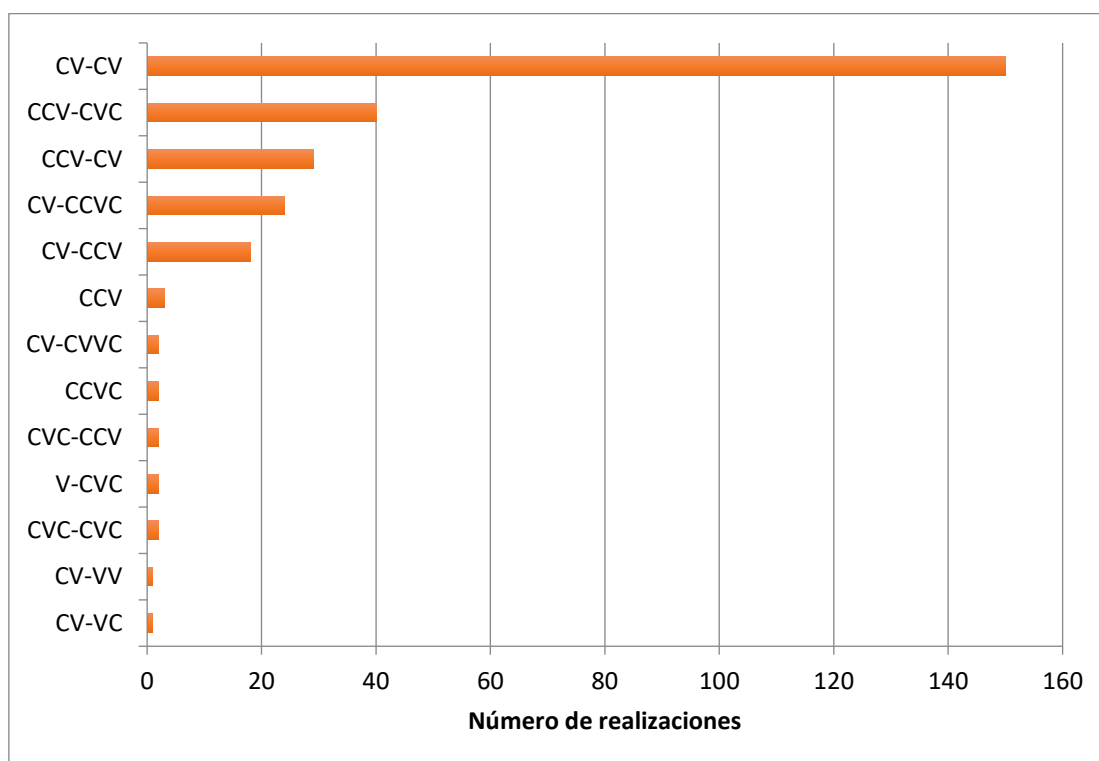


Figura 23. Número y tipo de realizaciones incorrectas de la estructura CV-CVC por las informantes sinohablantes.

Estos datos indican que las informantes sinohablantes muestran una gran tendencia a elidir la consonante en posición final de palabra, ya que han realizado la estructura CV-CVC como CV-CV hasta en 150 ocasiones, articulando *total* como [to.'ta] o *papel* como [pa.'pe], por ejemplo. Otro cambio que han realizado en la estructura de la palabra es insertar un sonido consonántico entre la consonante y la vocal de la primera sílaba de la palabra, resultando así la estructura CCV-CVC, la cual han reproducido 40 veces, y de la cual son ejemplo las realizaciones de *ratón* como [tra.'ton] o *risas* como ['tri.sas]. En 29 ocasiones han realizado ese mismo cambio, y además han elidido la consonante en posición final de palabra, dando como resultado la estructura CCV-CV, como *reloj*, pronunciada como [tre.'lo]. En 24 ocasiones han respetado la estructura de la primera sílaba, pero en la segunda han insertado un sonido consonántico entre la primera

consonante y la vocal, realizando así la estructura CV-CCVC en palabras como *correr* [ko.ˈdɾer]. 18 veces han realizado este mismo cambio y además han elidido la consonante en posición final, dando como resultado una estructura CV-CCV también en *correr*, pronunciada como [ko.ˈdre]. El resto de realizaciones presentan un número de realizaciones menor, así, realizan la estructura CV-CVC como CV-CVVC, CVC-CCV, V-CVC y CVC-CVC en 2 ocasiones respectivamente, y se ha contabilizado un caso en el que la estructura realizada es CV-VV, y otro de CV-VC. Cabe señalar que las informantes sinohablantes también han articulado palabras cuya estructura es CV-CVC como monosílabas. Ejemplo de ello es que la han articulado como CCV en 3 ocasiones, y como CCVC en 2, pronunciando palabras como *Miguel* y *muchos* como ['mke] y ['mtʃos] respectivamente.

### 3.3 Estructura CVC-CV

La estructura CVC-CV es realizada de manera correcta en el 84.6% de las ocasiones, lo que indica que las informantes sinohablantes tienen menos problemas con este tipo de estructura que con CV-CVC, pero más dificultades que con la estructura CV-CV. En aquellos casos en los que las estudiantes chinas no han realizado correctamente la estructura propia del español, la figura 24 muestra que han articulado hasta 8 tipos diferentes, aunque la tendencia general es a elidir la segunda consonante de la primera sílaba y articular la palabra como CV-CV.

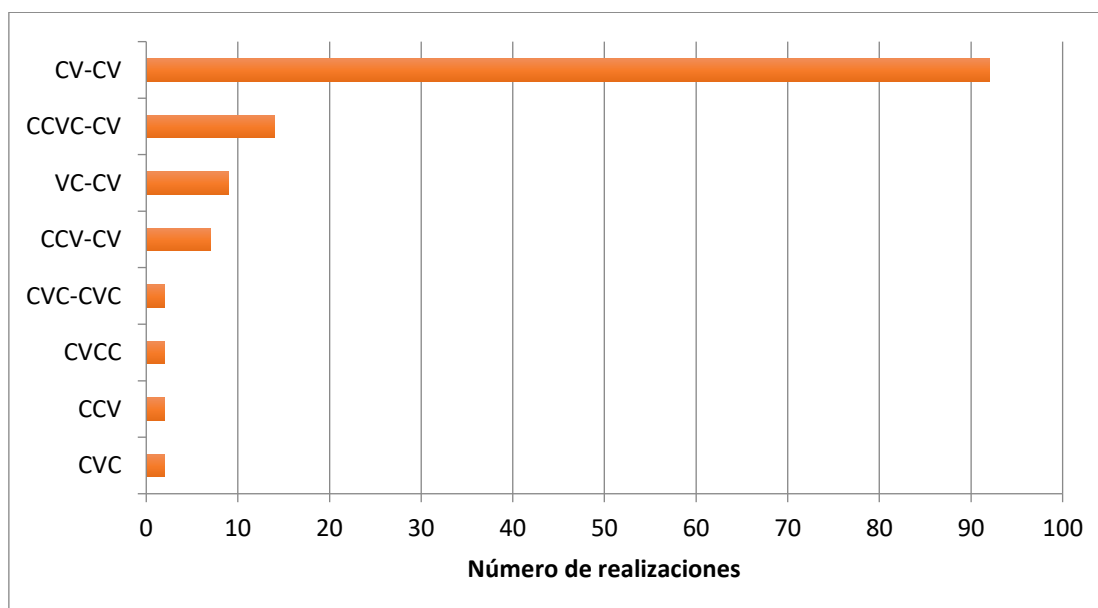


Figura 24 Número y tipo de realizaciones incorrectas de la estructura CVC-CV por las informantes sinohablantes.

Las informantes han elidido la segunda vocal de la primera sílaba en 92 ocasiones, dando lugar a la estructura CV-CV, así, palabras como *barba*, *pinta* o *culto* son pronunciadas como [‘pa.pa], [‘pi.ta] y [‘ku.to] respectivamente. Como se ha indicado, hay otro tipo de estructuras que las estudiantes chinas han realizado en un menor número de ocasiones. Las hablantes han añadido una consonante entre la primera consonante y la vocal de la primera sílaba en 14 ocasiones, produciendo así palabras con la estructura CCVC-CV, como *resto* y *ritmo*, articuladas como [‘tres.to] y [‘trit.mo] respectivamente. La elisión también ha tenido lugar en la primera consonante de algunas palabras, y la estructura resultante es VC-CV, la cual han realizado como tal 9 veces, como en *busca*, pronunciada como [‘us.ka]. En 7 ocasiones han añadido una consonante entre la primera consonante y la vocal de la primera sílaba, y además han elidido la segunda consonante de dicha sílaba, dando como resultado la estructura CCV-CV. Ejemplo de ello son las palabras *turnó* o *recto*, pronunciadas como [tru.‘no] y [‘tre.to] respectivamente. En 2 ocasiones las estudiantes sinohablantes han añadido una segunda consonante en el final de la segunda sílaba, dando como resultado la estructura CVC-CVC. Cabe señalar que aquí también se han detectado casos en los que la palabra bisílaba pasa a ser articulada como monosílaba, y la han realizado como CVCC, CCV o CVC en 2 ocasiones cada una de ellas, articulando así palabras como *nunca* o *cursé* como [‘nka] o [‘curs] respectivamente.

### 3.4 Recapitulación

**El análisis de la realización de la estructura de las palabras del español por parte de las informantes sinohablantes ha permitido observar que estas hablantes tienen problemas con los tres tipos de estructura presentes en el corpus de esta investigación.**

**Las palabras cuya estructura es CV-CV son las que menos dificultades presentan para estas estudiantes. En este caso, las estructuras que han articulado con mayor frecuencia cuando la realización no ha sido la correcta han sido CCV-CV, CV-CCV y CVC-CV, aunque se han detectado un total de 8 tipos de estructura diferentes. Es necesario añadir que, generalmente, cuando las informantes han insertado otro sonido entre la consonante y la vocal, lo han hecho cuando la consonante, anterior o posterior, es una de aquellas que las estudiantes chinas tienen más dificultades para articular correctamente. Además, se ha señalado que hay casos en los que las palabras han sido articuladas como monosílabas.**

Los datos muestran que la estructura CV-CVC es la más problemática para estas informantes, y se ha comprobado que hay una gran tendencia a elidir el sonido consonántico de final de palabra y articularla como CV-CV. Además, se han detectado hasta 12 tipos más de realizaciones diferentes, siendo los más numerosos CCV-CVC, CCV-CV o CV-CCVC, es decir, aquellos en los que, nuevamente, las estudiantes han insertado un sonido consonántico que no corresponde a la palabra en español. Cabe señalar que en este tipo de palabra también se han detectado casos en los que se ha articulado como monosílaba, en lugar de bisílaba.

En relación con la estructura CVC-CV, se ha comprobado que, si bien las informantes sinohablantes tienen problemas para realizarla correctamente, es menos problemática que la estructura CV-CVC. Aquí, nuevamente, la tendencia general ha sido a elidir la segunda consonante de la primera sílaba, articulando la palabra como CV-CV. Además, se han detectado 7 tipos más de realizaciones, siendo el segundo más realizado CCVC-CV. Igual que en las anteriores estructuras, también aquí se ha detectado que las informantes sinohablantes han articulado algunas de estas palabras como monosílabas.

A modo de conclusión, se han podido observar dos tendencias generales entre las informantes sinohablantes. La primera de ellas es la elisión de la segunda consonante en aquellas sílabas cuya estructura silábica es CVC, dando como resultado la articulación CV-CV para palabras cuya estructura es CV-CVC o CVC-CV. La otra tendencia detectada es la integración de un sonido consonántico que no corresponde a dicha palabra, fenómeno que tiene lugar especialmente en casos en los que el sonido consonántico anterior o posterior es problemático para las hablantes cuya lengua materna es el chino. Además, cabe señalar que en algunas ocasiones las informantes sinohablantes han modificado hasta tal punto la estructura silábica que la palabra bisílaba ha pasado a ser monosílaba.

Como se ha indicado al comienzo de este capítulo, en aquellos casos en los que una secuencia de segmentos experimenta procesos de modificación en sus límites silábicos se emplea el término resilabación para hacer referencia a estos cambios. Si bien este fenómeno no describe la modificación en la estructura de la sílaba y de la palabra que realizan las informantes sinohablantes, sí se considera que se produce un proceso de resilabación dentro de la palabra en la producción oral de las estudiantes de español cuya lengua materna es el chino.



**Como conclusión general de este apartado dedicado a la influencia de las variables cualitativas independientes en la pronunciación de las vocales del español por sinohablantes puede afirmarse que se han puesto de manifiesto los siguientes fenómenos:**

**Las informantes sinohablantes no presentan graves problemas en el momento de marcar las sílabas tónicas del español pues acentúan las mismas sílabas que acentúan los hispanohablantes.**

**El entorno consonántico anterior y posterior a la vocal presenta, en ocasiones, serios problemas de pronunciación para las hablantes de mandarín, especialmente en determinadas clases de consonantes (oclusivas sonoras, aproximantes, vibrantes, etc.). Los fenómenos de coarticulación que, como es sabido, provocan modificaciones en la pronunciación de los sonidos vecinos, actúan aquí sobre las vocales objeto de análisis de esta tesis por lo que los timbres de las mismas producidas por sinohablantes, como se verá en el apartado correspondiente, presentan importantes diferencias con respecto a los realizados por las hablantes nativas.**

**Finalmente, las dificultades en la realización de las estructuras silábicas del español, tan distintas de las del mandarín, conduce a la aparición de fenómenos de resilabación importantes como se ha visto a través de los ejemplos presentados. Las vocales, pues, verán también sus timbres alterados en función de los cambios sufridos por la estructura silábica de la palabra.**

**Como conclusión de carácter general de este apartado parece claro que el estudio de la adquisición de una clase de sonidos no depende exclusivamente de las características intrínsecas de los sonidos estudiados sino que, para comprender las razones que conducen a los sinohablantes a pronunciar las vocales del español de forma diferente a la de los nativos aunque, fonológicamente, las cinco vocales de esta lengua estén también presentes en el sistema del mandarín, deben considerarse todos los factores que influyen en la realización de la clase de sonidos estudiada. De lo contrario, se adoptaría un enfoque reduccionista cercano a planteamientos tradicionales que no tienen en cuenta los condicionamientos que actúan globalmente sobre el conjunto de los elementos segmentales integrantes de la cadena fónica.**

## **SEGUNDA PARTE: LA DURACIÓN DE LAS VOCALES**

En esta segunda parte del capítulo de resultados se analiza la duración de las vocales del español en función de los parámetros indicados en el capítulo III: duración general, duración de las vocales según la tonicidad, según el entorno y según la estructura silábica. En el primer apartado, se muestran los datos correspondientes a las realizaciones de las hablantes que conforman el grupo de control; en el segundo, se estudian los datos obtenidos en el análisis de las vocales producidas por las informantes sinohablantes; y en el tercer apartado se comparan los resultados de ambos grupos con el fin de establecer cuáles son las diferencias entre la realización de las hablantes cuya lengua materna es el español y las de las estudiantes chinas, y comprobar si dichas diferencias son significativas.

### **1. GRUPO DE CONTROL**

#### **1.1 Duración general**

En primer lugar, se ha analizado la duración media de las distintas vocales del español producidas por las informantes que constituyen el grupo de control. Para ello, se han estudiado 2376 vocales, y, en este caso, no se ha tenido en consideración la estructura silábica de la palabra o el entorno anterior y posterior, ya que dichos aspectos se estudian más adelante. La figura 25 muestra los resultados del análisis general de la duración de las vocales del español producidas por el grupo de control, además del número de casos que se han analizado de dicha vocal. Estas vocales siguen, de menor a mayor duración, el orden /u/, /o/, /i/, /e/, /a/, y exceptuando la /u/, la cual no alcanza los 100 ms. de duración media, las demás oscilan entre 100 y 120 ms. de duración media. La cifra que figura junto a las barras indica, en todos los casos, el número de vocales analizado.

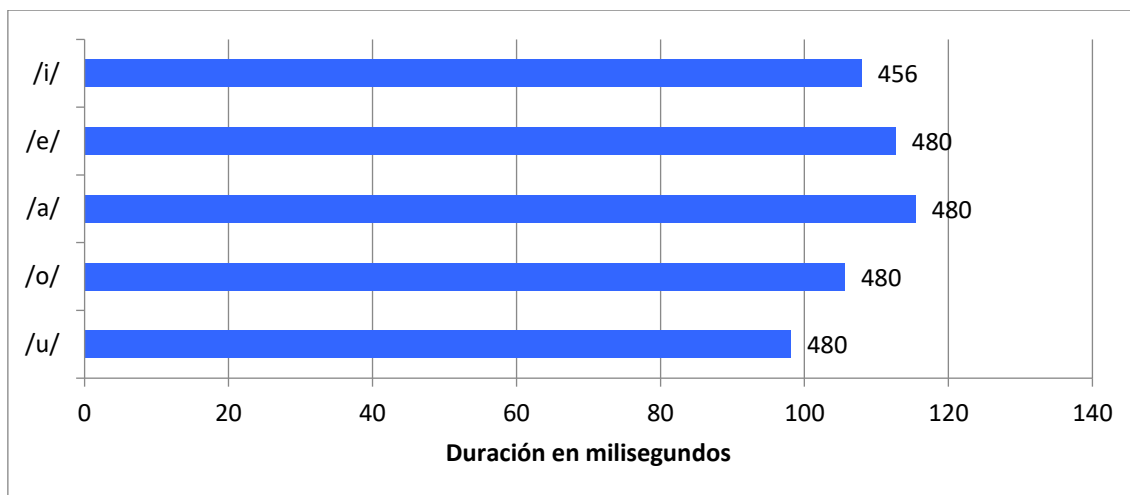


Figura 25. Duración media de las vocales producidas por el grupo de control.

La tabla 11 muestra los valores medios de duración de cada una de las vocales y la desviación estándar correspondiente a las mismas.

Vocal	Nº de vocales analizadas	Duración media (ms)	Desviación estándar
/i/	456	107.9	20.4
/e/	480	112.6	23.5
/a/	480	115.4	21.3
/o/	480	105.6	22.2
/u/	480	98.1	19.8

Tabla 11. Duración media y desviación estándar de las vocales producidas por el grupo de control.

La duración media oscila entre 98 y 115 ms., siendo las vocales posteriores /u/ y /o/ las más breves, /i/ y /e/ algo más largas, y la central /a/ es la vocal de mayor duración media. Como puede apreciarse, la desviación estándar de la duración de las vocales producidas por el grupo de control oscila entre 19 y 23, lo que indica que en todos los sonidos vocálicos la duración posee valores homogéneos.

## 1.2 Acentuación de la vocal

En segundo lugar, se ha analizado la duración de las vocales producidas por el grupo de control en función de la tonicidad de estas. Tal y como se observa en la figura 26, la duración media de las vocales tónicas es, en todos los casos, mayor que la de las

átonas correspondientes. En el caso de las primeras, la media oscila entre 100 y 120 ms., mientras que en el de las segundas, la duración de /i/, /o/ y /u/ oscila entre 80 y 100 ms., y la de /e/ y /a/ entre 100 y 120 ms.

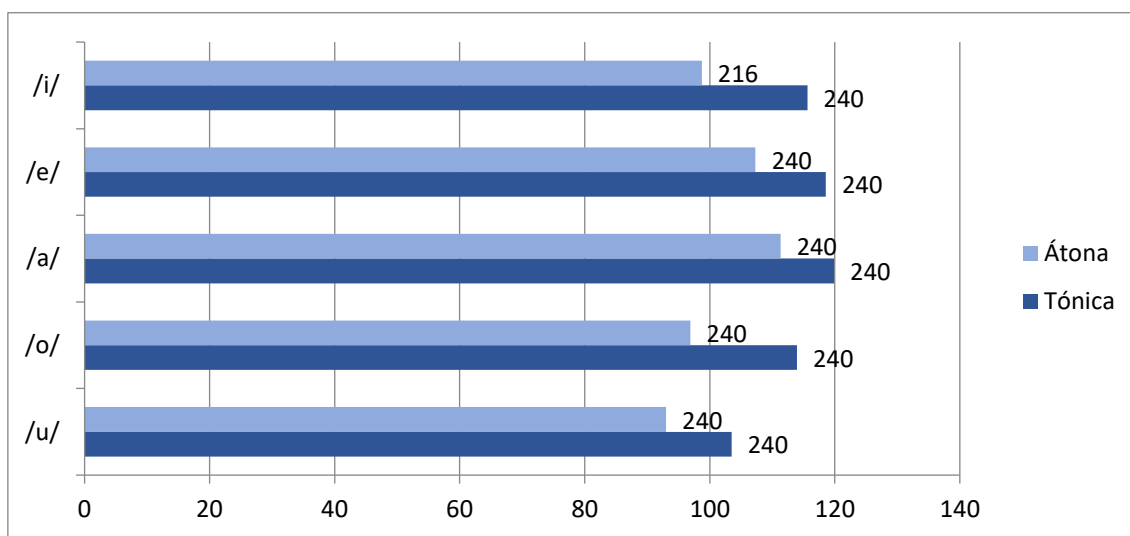


Figura 26. Duración media de las vocales producidas por el grupo de control según la tonicidad.

La tabla 12 permite analizar detalladamente los datos extraídos del análisis de la duración de las vocales en función de la acentuación. La primera columna indica cuál es la vocal analizada; la segunda, si dicha vocal es tónica o átona; la tercera señala el número de vocales analizadas para cada uno de los casos; la cuarta muestra la duración media de dicha vocal en milisegundos; y la quinta columna indica la desviación estándar.

Los datos muestran que cuando /i/ es tónica, la duración es de 115 ms., mientras que la vocal es más breve si es átona (98 ms.). La /e/ presenta una duración media de 118 ms. cuando es tónica, y de 107 ms. cuando es átona, valores similares a los de la /a/, los cuales son de 119 y 111 ms. cuando es tónica y átona respectivamente. En cuanto a /o/, la media también es mayor si es tónica (113 ms.) que si es átona (96 ms.). La vocal /u/, por su parte, tiene una duración media de 103 ms. cuando es tónica, y de 93 ms. cuando es átona.

Vocal	Acentuación	Nº de vocales analizadas	Duración media (ms)	Desviación estándar
/i/	Tónica	240	115.6	24.1
	Átona	216	98.7	18.6
/e/	Tónica	240	118.5	24.0
	Átona	240	107.3	20.8
/a/	Tónica	240	119.8	20.6
	Átona	240	111.3	20.2
/o/	Tónica	240	113.9	23.1
	Átona	240	96.9	16.9
/u/	Tónica	240	103.4	20.0
	Átona	240	93.0	18.0

Tabla 12. Duración media y desviación estándar de las vocales producidas por el grupo de control según la tonicidad de estas.

Cuando el sonido vocálico analizado se encuentra en sílaba tónica, la vocal de menor duración media es la /u/, la cual es de 103 ms. La /o/ presenta una duración media mayor, de 113 ms., y es similar a la de /i/ (115 ms.), la de /e/ (118 ms.) y la de /a/ (119 ms.), la cual presenta las realizaciones más largas. Si la vocal analizada se encuentra en posición átona, la /u/ también es la más breve, con una duración media de 93 ms., seguida por /o/ (96 ms.), /i/ (98 ms.) y /e/ (107 ms.). La vocal más larga es la /a/ , con una media de 111 ms. Esto muestra que, tanto si la vocal es tónica, como si es átona, siguen, de más breves a más largas, el orden /u/, /o/, /i/, /e/, /a/.

En cuanto a la desviación estándar, oscila entre 20 y 24 en el caso de las realizaciones tónicas, y entre 16 y 20 en el de las átonas. Esto muestra que, si bien en todos los casos las realizaciones presentan gran homogeneidad en la duración, es incluso mayor cuando las vocales son átonas, ya que la desviación es menor.

### 1.3 Entorno consonántico

En el presente apartado se analiza, en primer lugar, la duración de las distintas vocales producidas por el grupo de control en función del entorno consonántico anterior; en segundo lugar, la duración de estas según el entorno posterior; y, en tercer lugar, se estudia la duración de las mismas cuando se encuentran en posición final de palabra. Igual que en la primera parte de este capítulo de resultados, los sonidos consonánticos han sido

clasificados según las siguientes categorías: oclusiva sorda, oclusiva sonora, aproximante, fricativa, africada, nasal, lateral, vibrante simple y vibrante múltiple.

### 1.3.1 Duración de la vocal según el entorno anterior

En este apartado se analiza la duración de las vocales del español producidas por las hablantes que conforman el grupo de control según el entorno anterior. Es decir, si se analiza la vocal /i/ en palabras como *pico* o *venir*, el entorno anterior de dicha vocal es la oclusiva sorda [t] en el caso de la primera, y la nasal [n] en el caso de la segunda.

#### 1.3.1.1 Vocal /i/

La figura 27 muestra la duración media de la vocal /i/ según el entorno anterior y el número de casos en los que ha aparecido dicha realización. Se observa, por ejemplo, que la oclusiva sorda precede a /i/ en 120 ocasiones, mientras que la oclusiva sonora lo hace en 60 casos. Cuando el sonido anterior a la vocal es una oclusiva sorda, una africada, o una vibrante múltiple, la duración media es menor de 100 ms. Si dicho sonido consonántico es una oclusiva sonora, una fricativa, una nasal o una lateral, la duración oscila entre 100 y 120 ms., mientras que si los sonidos que preceden a la vocal son una aproximante o una vibrante simple, la duración media es superior a 120 ms.

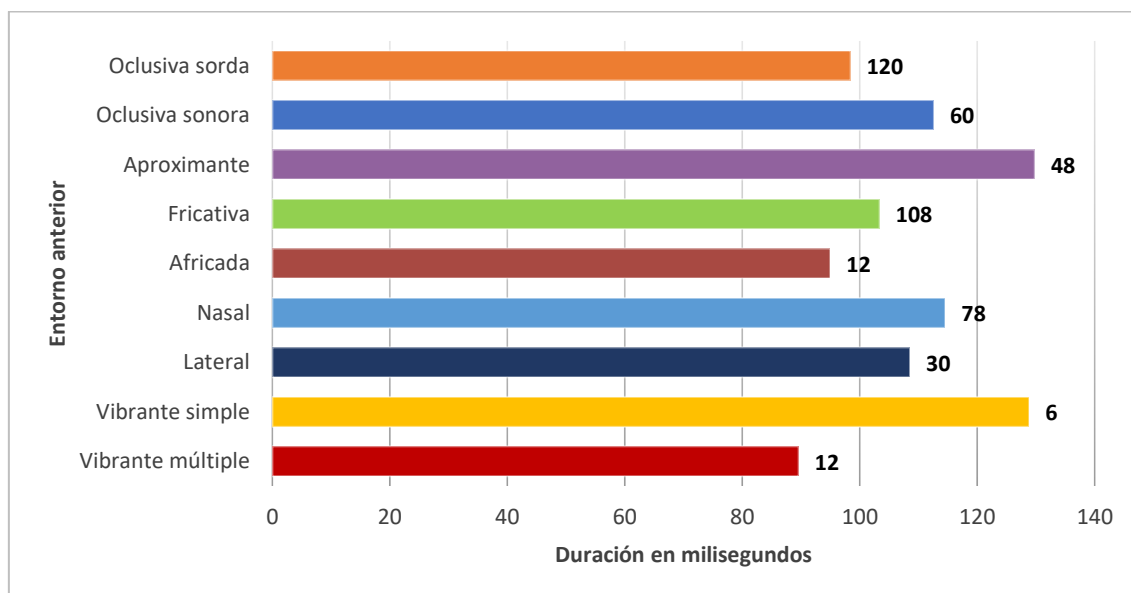


Figura 27. Duración media de la vocal /i/ producida por el grupo de control según el entorno anterior.

La tabla 13 muestra que, tal y como se ha indicado, la vocal /i/ es más breve cuando el sonido anterior a la vocal es una vibrante simple (89 ms.); y la duración media es

ligeramente superior si la consonante que precede a la vocal es una africada (95 ms.), o una oclusiva sorda (98 ms.). Esta duración aumenta si se trata de una fricativa (103 ms.), una lateral (108 ms.), una oclusiva sonora (112 ms.) o una nasal (114 ms.). Las duraciones más largas corresponden a /i/ precedida por vibrante simple (128 ms.) y por aproximante (129 ms.).

<i>/i/</i>			
<b>Entorno anterior realizado por las informantes sinohablantes</b>	<b>Casos</b>	<b>Duración media</b>	<b>Desviación estándar</b>
Oclusiva sorda	120	98.4	24.2
Oclusiva sonora	60	112.6	26.3
Aproximante	48	129.8	28.5
Fricativa	108	103.4	25.1
Africada	12	95.0	8.7
Nasal	78	114.5	21.1
Lateral	30	108.5	25.6
Vibrante simple	6	128.8	33.7
Vibrante múltiple	12	89.6	10.5

*Tabla 13.* Duración media y desviación estándar de la vocal /i/ producida por el grupo de control según el entorno anterior.

En cuanto a la desviación estándar en la duración de /i/, esta es menor cuando el sonido anterior es una consonante africada (8) y mayor cuando antes de esta vocal se encuentra una vibrante simple (33). Sin embargo, estos datos reflejan solamente tendencias puesto que el corpus contiene únicamente 12 casos de la primera consonante, y 6 de la segunda. En cuanto al resto de sonidos, se observa que la desviación estándar oscila entre 20 y 28. Esto muestra que, generalmente, la homogeneidad en la duración de /i/ en función del sonido anterior es similar.

### 1.3.1.2 Vocal /e/

Como se puede observar en la figura 28 que muestra la duración media de la vocal /e/ según el sonido que la precede, igual que sucede en /i/, las realizaciones más breves de /e/ corresponden a la articulación de esta vocal precedida por vibrante múltiple. En el resto de casos, la duración media oscila entre 100 y 120 ms., exceptuando aquellos en los que el sonido anterior a /e/ es una vibrante simple, en los cuales la media supera los 120 ms. Cabe señalar que no hay casos en los que la vocal del español /e/ esté precedida por una consonante africada.

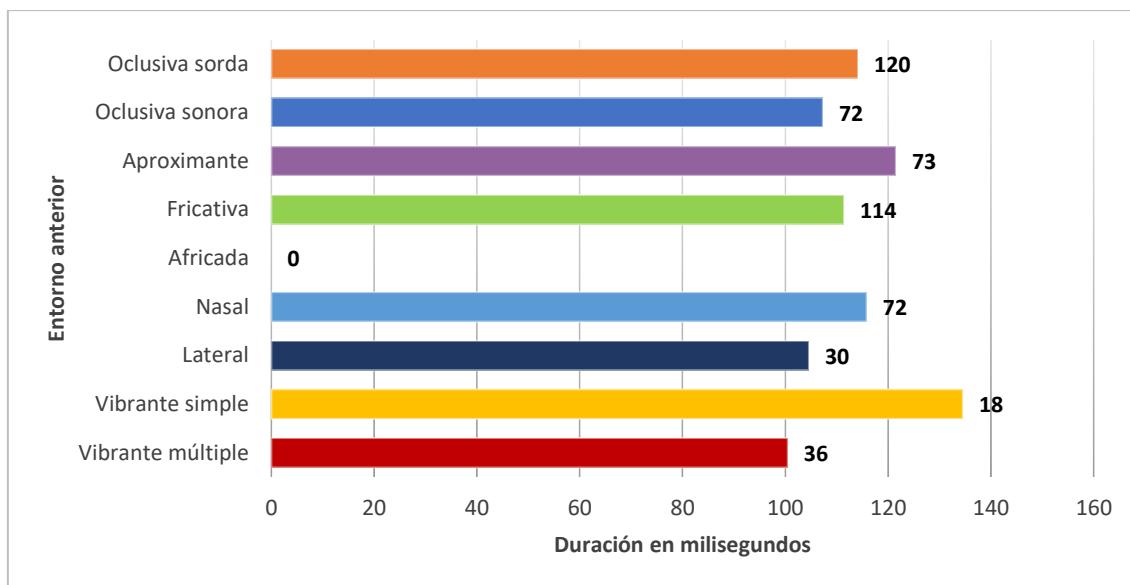


Figura 28. Duración media de la vocal /e/ producida por el grupo de control según el entorno anterior.

La tabla 14 muestra que, como se ha indicado, la duración de /e/ más breve corresponde a la de esta vocal precedida por una vibrante múltiple (100 ms.). Si dicho sonido consonántico es una lateral o una oclusiva sonora, la duración media es similar, de 104 y 107 ms. respectivamente. Esta vocal presenta una duración levemente superior cuando el sonido anterior es una fricativa (111 ms.), una oclusiva sorda (114 ms.) o una nasal (115 ms.). Las realizaciones más largas corresponden a las de /e/ precedida por aproximante (121 ms.) o por vibrante simple (134 ms.).

/e/			
Entorno anterior realizado por las informantes sinohablantes	Casos	Duración media	Desviación estándar
Oclusiva sorda	120	114.1	27.3
Oclusiva sonora	72	107.3	27.0
Aproximante	73	121.5	28.6
Fricativa	114	111.4	28.5
Africada	0	--	--
Nasal	72	115.8	22.8
Lateral	30	104.6	13.2
Vibrante simple	18	134.5	22.9
Vibrante múltiple	36	100.4	17.6

Tabla 14. Duración media y desviación estándar de la vocal /e/ producida por el grupo de control según el entorno anterior.

En cuanto a la desviación estándar, esta oscila entre 13 y 28, lo que indica que la homogeneidad en la duración de /e/ según el sonido anterior es similar en todos los casos.



### 1.3.1.3 Vocal /a/

La figura 29 muestra que, mientras que en los casos de /i/ y /e/ las diferencias de duración según el entorno anterior son mayores, en el caso de /a/ la vocal presenta una duración media similar independientemente del sonido anterior, y en todos los casos oscila entre 100 y 120 ms. Cabe señalar que, igual que sucede con /i/ y /e/, la vocal es ligeramente más larga si el sonido consonántico que precede a la vocal es una vibrante simple.

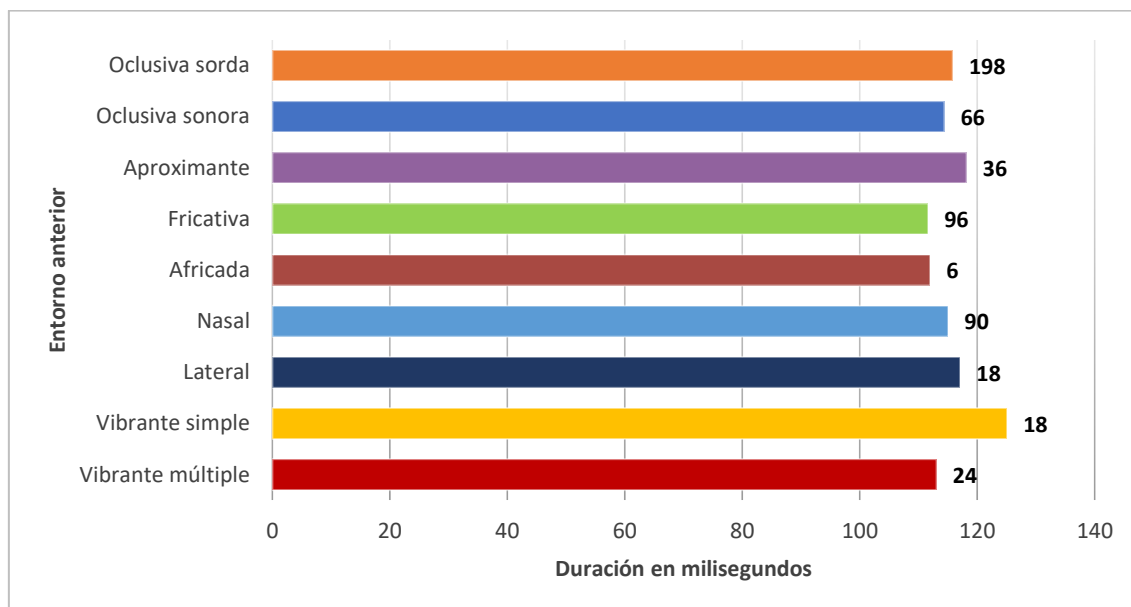


Figura 29. Duración media de la vocal /a/ producida por el grupo de control según el entorno anterior.

La duración media de /a/ es menor cuando el sonido anterior es una fricativa (111 ms.), aunque es muy similar cuando dicho sonido es una africada (112 ms.), una vibrante múltiple (113 ms.), una oclusiva sonora (114 ms.), una nasal (115 ms.), una oclusiva sorda (115 ms.), una lateral (117 ms.) o una aproximante (118 ms.). Si el sonido que precede a la vocal es una vibrante simple, tal y como se ha indicado, la /a/ es más larga, con una duración media de 125 ms. 7

<i>/a/</i>			
<b>Entorno anterior realizado por las informantes sinohablantes</b>	<b>Casos</b>	<b>Duración media</b>	<b>Desviación estándar</b>
Oclusiva sorda	198	115.8	26.2
Oclusiva sonora	66	114.4	23.8
Aproximante	36	118.2	26.2
Fricativa	96	111.6	25.4
Africada	6	112.0	21.3
Nasal	90	115.0	21.7
Lateral	18	117.1	11.5
Vibrante simple	18	125.1	23.0
Vibrante múltiple	24	113.1	21.8

*Tabla 15.* Duración media y desviación estándar de la vocal /a/ producida por el grupo de control según el entorno anterior.

En cuanto a la desviación estándar, exceptuando los casos en los que el sonido anterior es una lateral, en los que su desviación es de 11, en el resto de casos la desviación oscila entre 20 y 26, por lo que la homogeneidad en la duración de /a/ apenas varía según el sonido consonántico anterior a esta vocal.

#### **1.3.1.4 Vocal /o/**

De acuerdo con los datos mostrados en la figura 30, cuando el sonido consonántico que precede a /o/ es una vibrante múltiple, la duración es inferior que cuando se trata de otra consonante, y no alcanza los 100 ms. de media, como tampoco lo hace cuando dicho sonido es una oclusiva sorda. También se observa que la duración media de esta vocal oscila entre 100 y 120 ms. cuando el sonido anterior es una oclusiva sonora, fricativa, africada, nasal, lateral, vibrante simple o aproximante, siendo esta última la que tiene una mayor duración media.

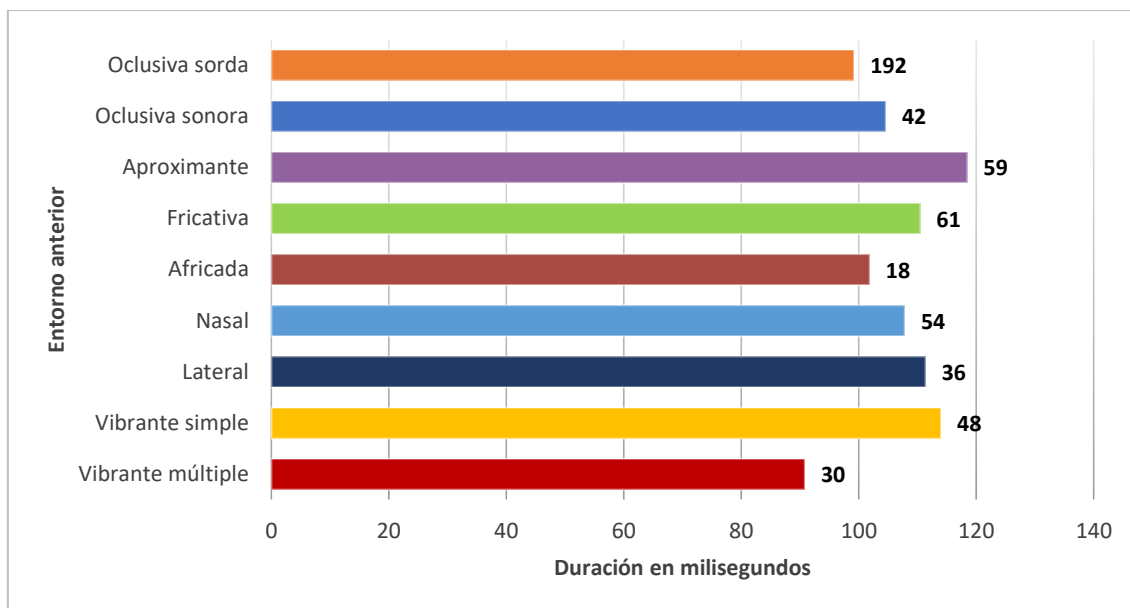


Figura 30. Duración media de la vocal /o/ producida por el grupo de control según el entorno anterior.

Como se ha indicado, la duración media de /o/ según el entorno anterior es bastante similar en todos los casos. Por un lado, las duraciones más breves aparecen cuando la consonante anterior es una vibrante múltiple (90 ms.). La media es similar si el sonido que precede a la vocal en cuestión es una oclusiva sorda (99 ms.), una africada (101 ms.), una oclusiva sonora (104 ms.) o una nasal (107 ms.). Las mayores duraciones medias, las cuales no distan mucho de las mencionadas hasta el momento, corresponden a las realizaciones de /o/ precedida por fricativa (110 ms.), lateral (111 ms.), vibrante simple (114 ms.) y aproximante (118 ms.).

/o/			
Entorno anterior realizado por las informantes sinohablantes	Casos	Duración media	Desviación estándar
Oclusiva sorda	192	99.2	24.9
Oclusiva sonora	42	104.6	19.4
Aproximante	59	118.5	30.5
Fricativa	61	110.5	26.6
Africada	18	101.8	12.4
Nasal	54	107.8	19.4
Lateral	36	111.3	16.3
Vibrante simple	48	114.0	28.7
Vibrante múltiple	30	90.8	17.9

Tabla 16. Duración media y desviación estándar de la vocal /o/ producida por el grupo de control según el entorno anterior.

La desviación estándar es bastante similar en todos los casos. Las realizaciones más homogéneas corresponden a la realización de /o/ precedida por una africada (12) o una lateral (16), mientras que las menos homogéneas son las de aproximante seguida por esta vocal (30). Estos datos muestran que la duración media de /o/ es similar en todos los casos, y que, generalmente, la duración de /o/ en función del sonido anterior presenta bastante homogeneidad.

### 1.3.1.5 Vocal /u/

La figura 31 muestra que, mientras que cuando se analiza la duración de /a/ y /o/ en función del sonido anterior se ha detectado que la media es bastante similar en todos los casos, cuando dicha vocal es /u/ hay mayor variación en la duración. Esta vocal es más breve cuando el sonido consonántico anterior es una africada, y la duración media no alcanza los 80 ms. Esta oscila entre 80 y 100 ms. cuando el sonido que precede a /u/ es una oclusiva sorda, una fricativa o una vibrante simple, mientras que si es una oclusiva sonora, una nasal o una lateral, la duración media supera los 100 ms. Las realizaciones de mayor duración corresponden a las /u/ precedida por aproximante o vibrante simple, las cuales superan los 120 ms. Sin embargo, cabe señalar que son pocos los casos en los que estos sonidos consonánticos preceden a la vocal (6 y 12 respectivamente).

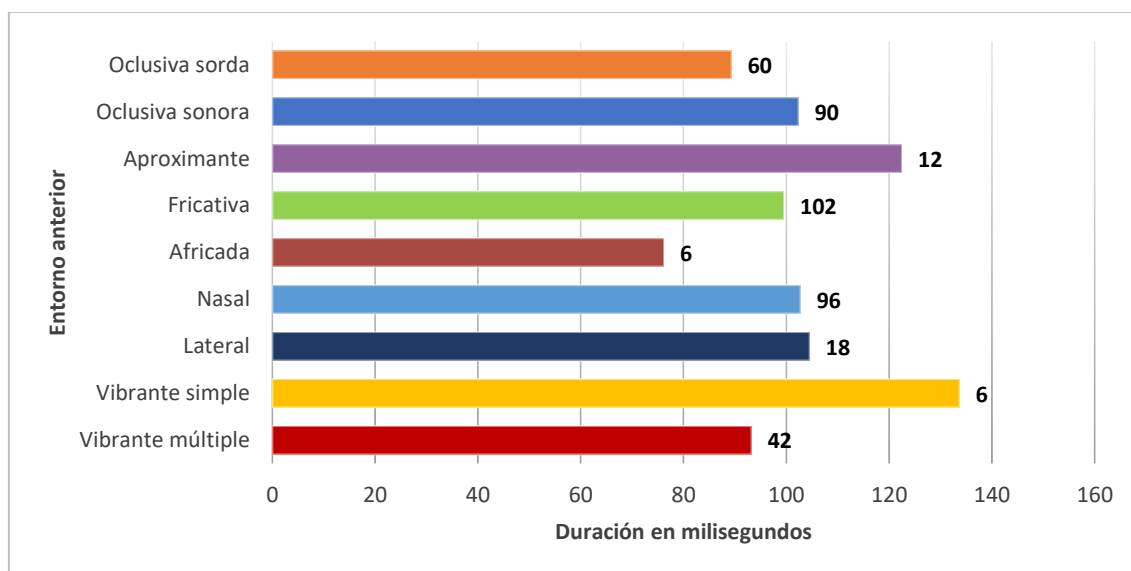


Figura 31. Duración media de la vocal /u/ producida por el grupo de control según el entorno anterior.

Como se ha indicado, la duración más breve corresponde a la de /u/ tras africada (76 ms.). La media es ligeramente superior cuando la consonante que precede a la vocal es una oclusiva sorda (89 ms.) o una vibrante múltiple (93 ms.). Si la consonante que precede a

la vocal es una fricativa (99 ms.), una oclusiva sonora (102 ms.), una nasal (102 ms.) o una lateral (104 ms.) la duración media es similar. Sin embargo, si dicho sonido es una aproximante (122 ms.) o una vibrante simple (133 ms.), la duración media es considerablemente superior.

<b>/u/</b>			
<b>Entorno anterior realizado por las informantes sinohablantes</b>	<b>Casos</b>	<b>Duración media</b>	<b>Desviación estándar</b>
Oclusiva sorda	60	89.3	21.7
Oclusiva sonora	90	102.4	22.3
Aproximante	12	122.5	31.6
Fricativa	102	99.5	25.0
Africada	6	76.2	21.4
Nasal	96	102.7	18.9
Lateral	18	104.5	18.7
Vibrante simple	6	133.7	26.8
Vibrante múltiple	42	93.2	15.8

Tabla 17. Duración media y desviación estándar de la vocal /u/ producida por el grupo de control según el entorno anterior.

En cuanto a la desviación estándar, la tabla 17 muestra que, igual que sucede en las vocales anteriormente analizadas, esta es similar y la realización es bastante homogénea, ya que la menor desviación corresponde a la realización de vibrante múltiple seguida por /u/ (15), y la mayor, se produce cuando dicho sonido es una aproximante (31).

### 1.3.2 Duración de la vocal según el entorno posterior

En este apartado se analiza la duración de las vocales del español producidas por las hablantes que constituyen el grupo de control según el entorno posterior. Es decir, si se analiza la vocal /e/ en palabras como *reloj* o *verbo*, el sonido posterior a dicha vocal es la lateral [l] en el caso de la primera, y la vibrante simple [r] en el caso de la segunda.

#### 1.3.2.1 Vocal /i/

La figura 32 muestra la duración media de la vocal /i/ según el entorno posterior y el número de casos en los que ha aparecido dicha realización. Se observa, por ejemplo, que la oclusiva sorda sigue a /i/ en 42 ocasiones, mientras que la oclusiva sonora lo hace en 6. Tal y como se puede observar, cuando el sonido que sigue a la vocal es una nasal o

una lateral, la duración media oscila entre 80 u 100 ms., mientras que si la consonante posterior a /i/ es una oclusiva sorda, oclusiva sonora, aproximante, fricativa, africada o una vibrante simple, la duración media oscila entre 100 y 120 ms. Cabe señalar que no hay casos en los que una vibrante múltiple se encuentra en posición posterior a /i/.

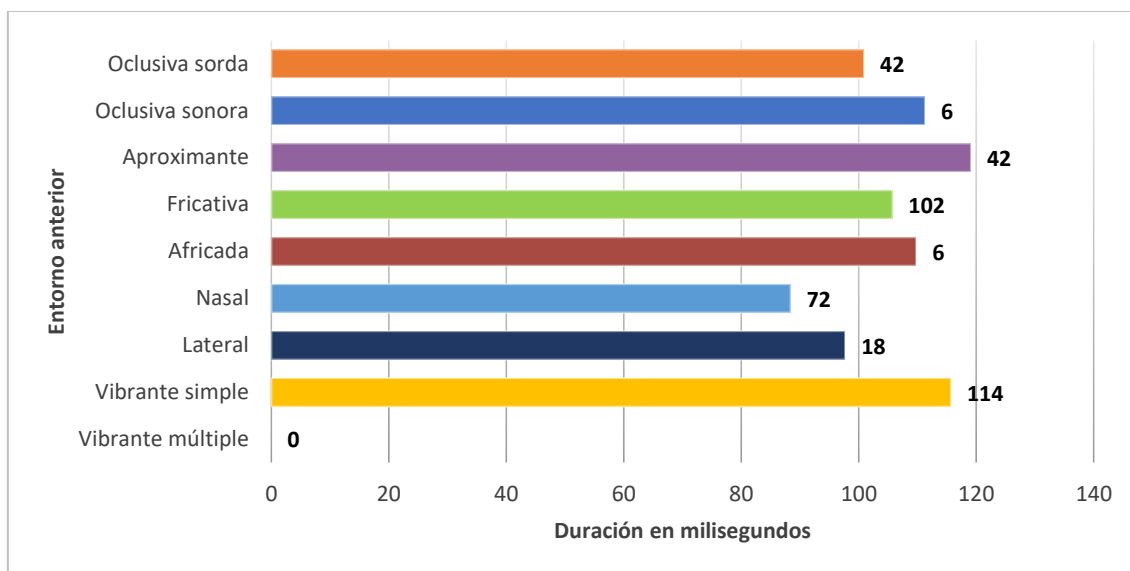


Figura 32. Duración media de la vocal /i/ producida por el grupo de control según el entorno posterior.

La tabla 18 muestra que las realizaciones más breves de /i/ se dan cuando el sonido que sigue a esta vocal es una consonante nasal (88 ms.), y la duración es algo superior si la vocal es seguida por una lateral (97 ms.). Cuando el sonido posterior es una oclusiva sorda (100 ms.), una fricativa (105 ms.), una africada (109 ms.), una oclusiva sonora (111 ms.) o una vibrante simple (115 ms.), la duración media es similar, y la realización más larga corresponde a la de /i/ seguida por aproximante (119 ms.), también similar a las otras.

/i/			
Entorno anterior realizado por las informantes sinohablantes	Casos	Duración media	Desviación estándar
Oclusiva sorda	42	100.9	18.9
Oclusiva sonora	6	111.3	21.3
Aproximante	42	119.1	27.2
Fricativa	102	105.7	25.3
Africada	6	109.8	33.3
Nasal	72	88.4	14.1
Lateral	18	97.7	16.4
Vibrante simple	114	115.7	23.2
Vibrante múltiple	0	--	--

Tabla 18. Duración media y desviación estándar de la vocal /i/ producida por el grupo de control según el entorno posterior.

En cuanto a la desviación estándar en la duración de /i/, esta es menor cuando después de la vocal se encuentra una consonante nasal (14), y mayor cuando el sonido posterior es una africada (33), aunque de esta únicamente hay 6 casos. En cuanto al resto de sonidos, se observa que la desviación estándar oscila entre 16 y 27. Esto muestra que, generalmente, la homogeneidad en la duración de /i/ en función del sonido posterior es similar.

### 1.3.2.2 Vocal /e/

Como se puede observar en la figura 33, cuando el sonido posterior a /e/ es una vibrante múltiple o una oclusiva sorda, la vocal es más breve y la duración media no alcanza los 100 ms. Si este es una fricativa, una nasal, una lateral o una vibrante simple, la duración media es similar y supera los 100 ms. Si /e/ es seguida por una aproximante, la vocal es más larga que en los otros casos, aunque la realización más larga de todas corresponde a la de la vocal seguida por oclusiva sonora, la cual supera los 120 ms. de media. Cabe señalar que, como se puede observar, en este caso no hay realizaciones de /e/ seguida por africada.

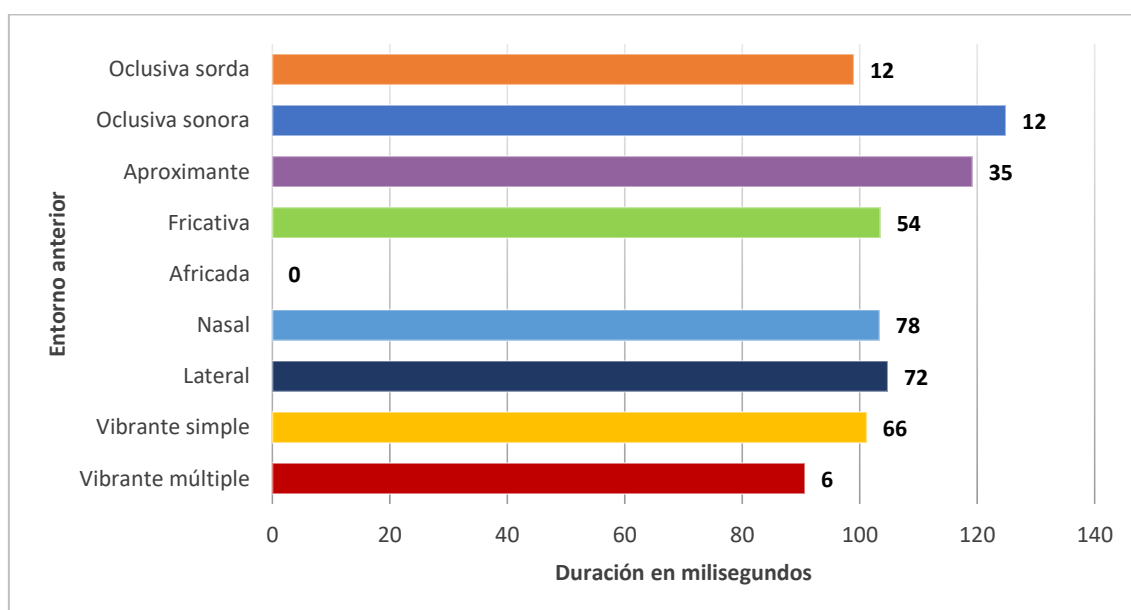


Figura 33. Duración media de la vocal /e/ producida por el grupo de control según el entorno posterior.

Como se ha indicado, la /e/ articulada por las hablantes del grupo de control es más breve cuando el sonido consonántico posterior es una vibrante múltiple, con una duración media de 90 ms. La duración de esta vocal es similar cuando la consonante que sigue a la vocal es una oclusiva sorda (99 ms.), una vibrante simple (101 ms.), una fricativa (103 ms.),

una nasal (103 ms.) o una lateral (104 ms.). En cambio, la vocal es más larga cuando el sonido consonántico posterior a la vocal es una aproximante, con una media de 119 ms., o una oclusiva sonora, con una media de 124 ms.

<i>/e/</i>			
<b>Entorno anterior realizado por las informantes sinohablantes</b>	<b>Casos</b>	<b>Duración media</b>	<b>Desviación estándar</b>
Oclusiva sorda	12	99.0	26.1
Oclusiva sonora	12	124.9	30.4
Aproximante	35	119.2	31.4
Fricativa	54	103.5	26.9
Africada	0	--	--
Nasal	78	103.4	16.9
Lateral	72	104.8	16.9
Vibrante simple	66	101.2	16.8
Vibrante múltiple	6	90.7	27.9

*Tabla 19.* Duración media y desviación estándar de la vocal /e/ producida por el grupo de control según el entorno posterior.

En cuanto a la desviación estándar, la tabla 19 muestra que es menor cuando el sonido posterior a /e/ es una vibrante simple, una nasal o una lateral (16), y mayor si la consonante es una aproximante (31). Los datos muestran que, en general, la homogeneidad en la duración de /e/ según el sonido posterior es similar.

### 1.3.2.3 Vocal /a/

La figura 34 muestra que cuando el sonido consonántico posterior a la vocal /a/ del español es una oclusiva sorda, una aproximante, una fricativa, una nasal, una lateral o una vibrante múltiple, la duración media oscila entre 100 y 120 ms. Si la consonante que sigue a la vocal es una vibrante simple, la media supera los 120 ms., mientras que, si dicho sonido es una oclusiva sonora, la duración supera los 140 ms. Sin embargo, es necesario tener en cuenta que en el corpus únicamente aparecen 6 casos de /a/ seguida por oclusiva sonora.



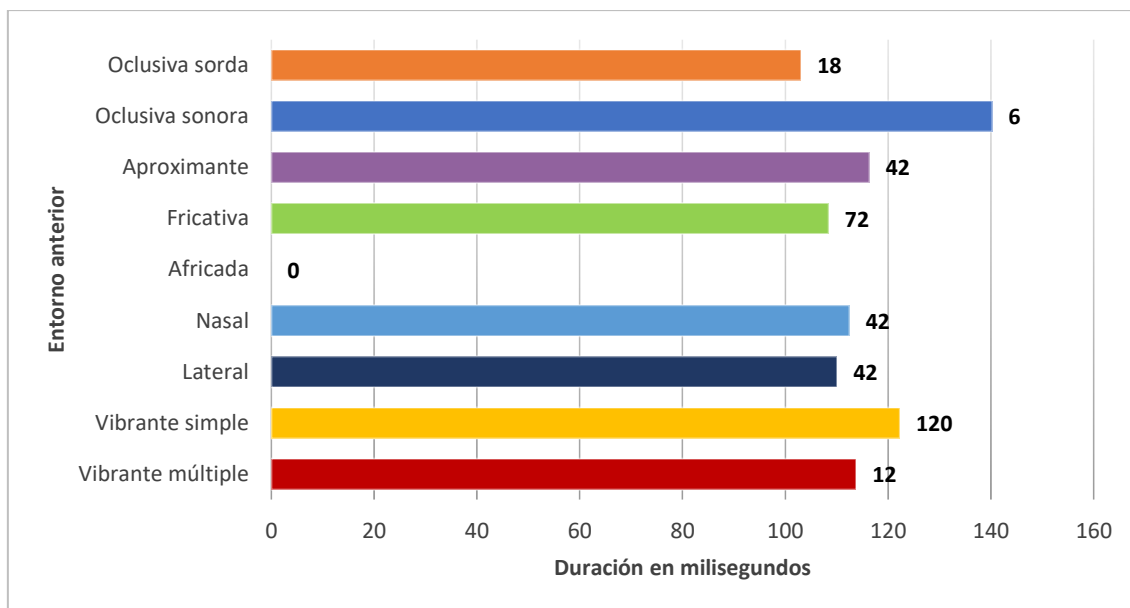


Figura 34. Duración media de la vocal /a/ producida por el grupo de control según el entorno posterior.

La /a/ es más breve cuando el sonido posterior es una oclusiva sorda, con una duración media de 103 ms. Si dicho sonido es una fricativa (108 ms.), una lateral (110 ms.), una nasal (112 ms.), una vibrante múltiple (113 ms.) o una aproximante (116 ms.) la duración media es similar. Si el sonido posterior a /a/ es una vibrante simple, la es más larga, con una media de 122 ms., mientras que si es una oclusiva sonora, la vocal tiene una duración media considerablemente superior, de 140 ms.

/a/			
Entorno anterior realizado por las informantes sinohablantes	Casos	Duración media	Desviación estándar
Oclusiva sorda	18	103.1	14.0
Oclusiva sonora	6	140.3	23.1
Aproximante	42	116.0	23.9
Fricativa	72	108.4	21.8
Africada	0	--	--
Nasal	42	112.5	17.0
Lateral	42	110.1	21.4
Vibrante simple	120	122.3	20.7
Vibrante múltiple	12	113.8	11.7

Tabla 20. Duración media y desviación estándar de la vocal /a/ producida por el grupo de control según el entorno posterior.

En cuanto a la desviación estándar, esta es muy similar en todos los casos, ya que la realización que mayor homogeneidad presenta es la /a/ seguida por vibrante múltiple (11), y la menos homogénea es la de aproximante posterior a la vocal (23). Esto muestra que

la duración de /a/ no presenta una gran variación en función del sonido posterior a dicha vocal.

#### 1.3.2.4 Vocal /o/

La figura 35 muestra que, cuando el sonido posterior a /o/ es una oclusiva sorda, una nasal o una lateral la duración media es muy similar, y es ligeramente inferior a 100 ms. Si la consonante que sigue a la vocal es una aproximante, una fricativa o una vibrante simple, la duración media supera los 100 ms. y también si es una vibrante múltiple, aunque en este caso la duración es ligeramente superior que en los demás.

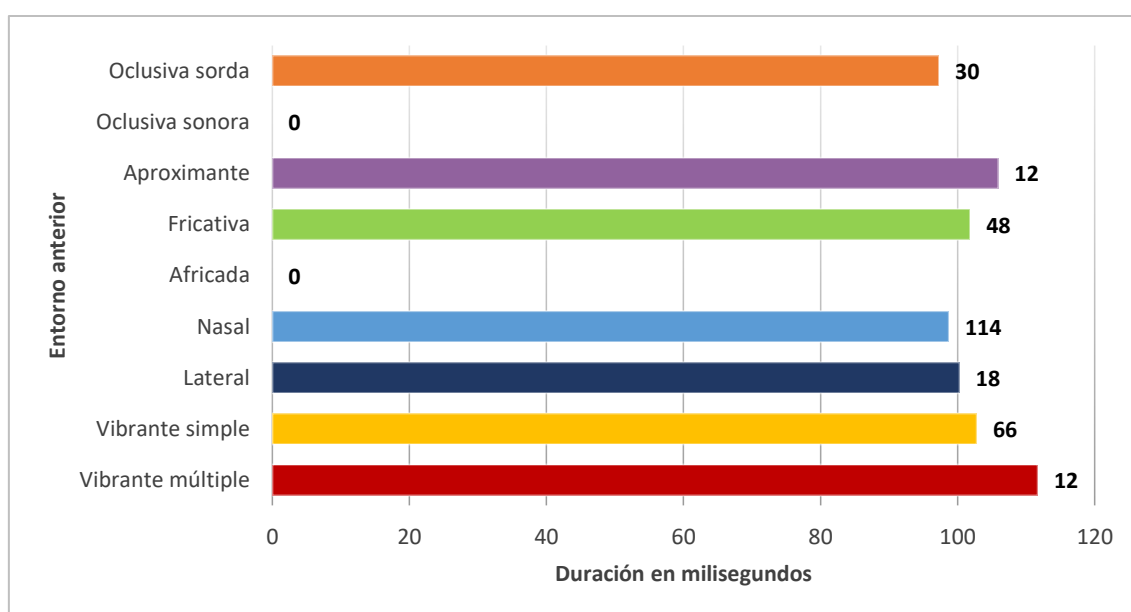


Figura 35. Duración media de la vocal /o/ producida por el grupo de control según el entorno posterior.

Como se puede observar en la tabla 21, la duración de la vocal /o/ según el entorno posterior no varía mucho en función de la consonante que sigue a la vocal, ya que si este es una oclusiva sorda, la media es de 97 ms., de 98 ms. si es una nasal, de 100 ms. si es una lateral, de 101 ms. si es una fricativa, de 102 ms. si es una vibrante simple y de 105 ms. si es una aproximante. Como se ha indicado, la realización más larga de /o/ corresponde a la articulación de esta vocal seguida de vibrante múltiple, la cual tiene una duración media de 111 ms.

<i>/o/</i>			
<b>Entorno anterior realizado por las informantes sinohablantes</b>	<b>Casos</b>	<b>Duración media</b>	<b>Desviación estándar</b>
Oclusiva sorda	30	97.2	24.4
Oclusiva sonora	0	--	--
Aproximante	12	105.9	25.6
Fricativa	48	101.8	22.1
Africada	0	--	--
Nasal	114	98.7	17.9
Lateral	18	100.3	15.3
Vibrante simple	66	102.8	17.4
Vibrante múltiple	12	111.7	24.2

*Tabla 21.* Duración media y desviación estándar de la vocal /o/ producida por el grupo de control según el entorno posterior.

En cuanto a la desviación estándar, esta es muy similar en todos los casos, ya que la realización que mayor homogeneidad presenta es la de /o/ seguida por lateral (15), y la menos homogénea es la de aproximante posterior a la vocal (25). Esto muestra que la duración de /o/ no presenta una gran variación en función del sonido posterior a dicha vocal.

### **1.3.2.5 Vocal /u/**

La figura 36 muestra que la duración de la vocal /u/ producida por las hablantes del grupo de control es muy similar en todos los casos, ya que oscila entre 80 y 100 ms., aunque cuando el sonido que sigue a la vocal es una aproximante, la duración media es superior a 120 ms.

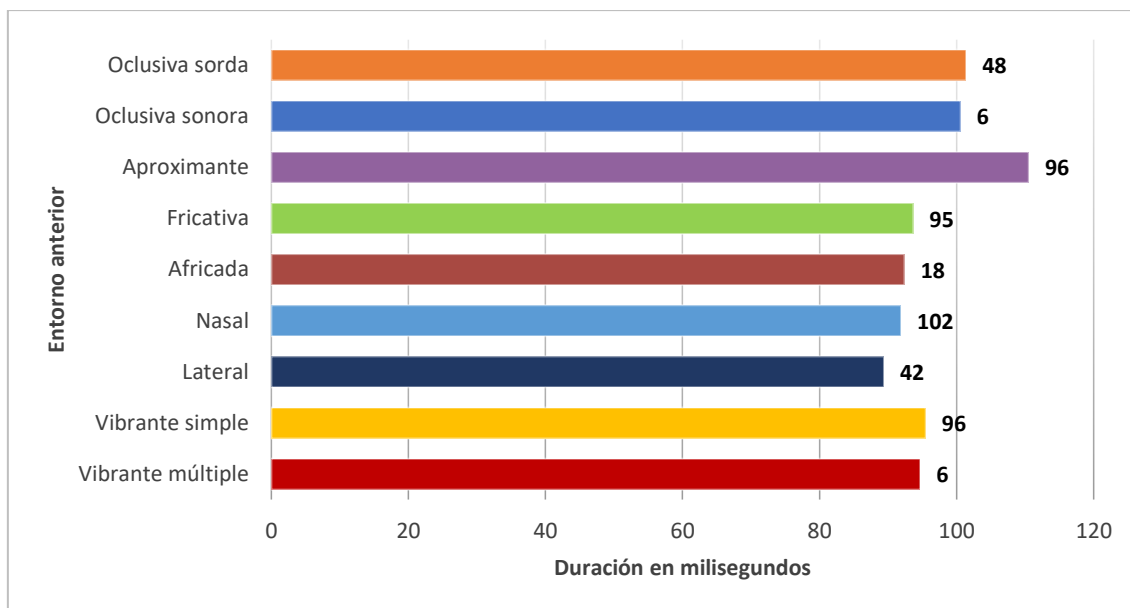


Figura 36. Duración media de la vocal /u/ producida por el grupo de control según el entorno posterior.

La duración media de /u/ cuando es seguida por lateral es de 89 ms., es decir, es la más breve. Sin embargo, esta media es similar a las demás, ya que si el sonido posterior es una nasal, la media es de 91 ms., si es una africada, de 92 ms., si es una fricativa, de 93 ms., si es vibrante múltiple, de 94 ms., si es una vibrante simple, de 95 ms., y si es una oclusiva sonora, de 100 ms. En cambio, cuando /u/ precede a una aproximante, la vocal es más larga y tiene una duración media de 110 ms.

/u/			
Entorno anterior realizado por las informantes sinohablantes	Casos	Duración media	Desviación estándar
Oclusiva sorda	48	101.4	24.0
Oclusiva sonora	6	100.6	24.3
Aproximante	96	110.5	27.5
Fricativa	95	93.6	21.3
Africada	18	92.4	10.0
Nasal	102	91.9	19.0
Lateral	42	89.4	12.5
Vibrante simple	96	95.5	18.3
Vibrante múltiple	6	94.7	23.0

Tabla 22. Duración media y desviación estándar de la vocal /u/ producida por el grupo de control según el entorno posterior.

En cuanto a la desviación estándar, la /u/ tiene una duración más homogénea cuando el sonido posterior es una africada (10), mientras que, si dicho sonido es una aproximante (27), las duraciones de la vocal son las menos homogéneas. Los datos muestran que, dado

que la diferencia entre la menor y mayor desviación estándar no es elevada, la homogeneidad en la duración de /u/ según el sonido posterior es similar.

### 1.3.3 Duración de la vocal en posición final de palabra

De la misma manera que se ha analizado la duración de las vocales del español producidas por el grupo de control según el entorno anterior y el entorno posterior, en este apartado se analizan aquellas realizaciones de vocal que se encuentran en posición final de palabra aislada. Por ejemplo, la vocal /o/ de la palabra *lago* o la vocal /e/ de la palabra *cine*.

La figura 37 destaca que cuando los sonidos vocálicos /i/, /e/, /a/ y /o/ se encuentran en posición final de palabra aislada, la duración media es similar y oscila entre 110 y 120 ms. En cambio, si /u/ se encuentra en dicha posición, la vocal es más larga y la duración media supera los 120 ms. Sin embargo, ha de tenerse en cuenta que el corpus únicamente contiene 12 casos de /u/ en esta posición (la vocal /u/ es la menos frecuente en español).

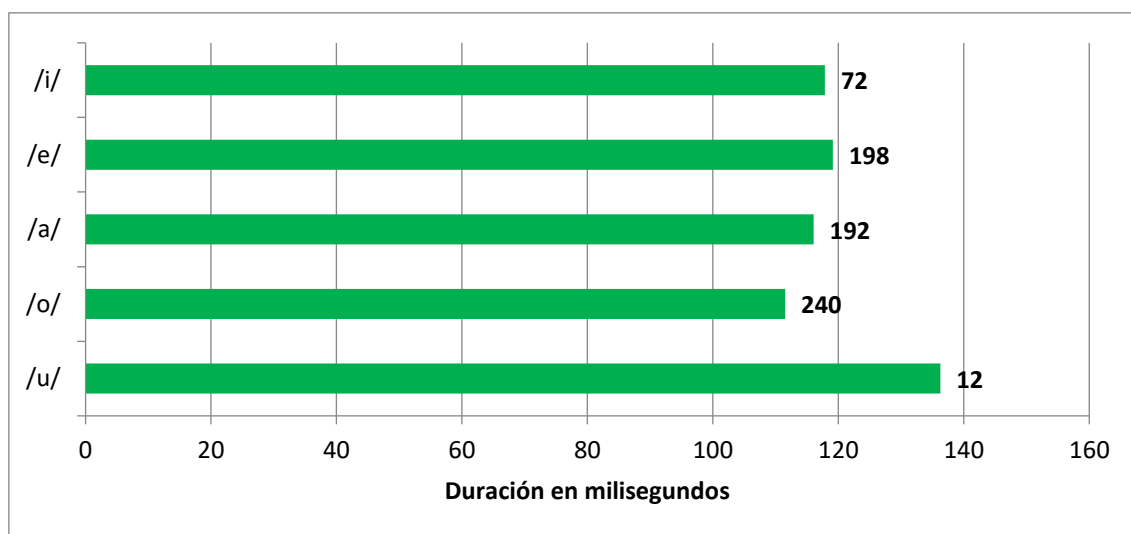


Figura 37. Duración media de las vocales producidas por el grupo de control cuando estas se encuentran en sílaba final de palabra.

La tabla 23 muestra que la duración media de /o/ en posición final de palabra es de 111 ms., la de /a/ es de 116 ms., la de /i/ es de 117 ms. y la de /e/ es de 119 ms., es decir, como se ha indicado, estas vocales tienen una duración media muy similar. En cambio, la /u/ en posición final es más larga y tiene una duración media de 136 ms.

<b>Duración de la vocal cuando se encuentra en final de palabra</b>			
<b>Vocal</b>	<b>Nº de vocales analizadas</b>	<b>Duración media (ms)</b>	<b>Desv. Estándar</b>
/i/	72	117.8	26.9
/e/	198	119.1	24.1
/a/	192	116.1	21.3
/o/	240	111.5	24.2
/u/	12	136.2	16.9

*Tabla 23.* Duración media y desviación estándar de las vocales producidas por el grupo de control cuando estas se encuentran en final de palabra aislada.

En cuanto a la desviación estándar, esta no es elevada, ya que oscila entre 16 y 26, siendo la vocal que mayor homogeneidad presenta la /u/, y la que mayor variación tiene la /i/. Esto muestra que la realización de las vocales en posición final de palabra aislada por parte de las informantes que conforman el grupo de control presenta una homogeneidad similar a los demás casos analizados.

### **1.3.4 Recapitulación**

**En este apartado se ha estudiado, en primer lugar, la duración de las vocales del español según el entorno anterior a estas. De este análisis se concluye que, generalmente, las vocales presentan una duración menor si el sonido consonántico anterior es una oclusiva sorda, una vibrante múltiple o una africada, y esta duración es, normalmente, inferior a 100 ms. En cambio, se realizan las vocales con una mayor duración cuando el sonido anterior es una vibrante simple o una aproximante, y esta supera, generalmente, los 120 ms. El análisis de las vocales según el entorno anterior ha mostrado que la desviación estándar, generalmente, es superior a 8, y menor que 30, lo que significa que la duración es homogénea y similar en la mayoría de los casos.**

**En segundo lugar, se ha estudiado la duración de las vocales según el entorno posterior. En el análisis se ha podido comprobar que /u/ es, generalmente, más breve que el resto de vocales, ya que la duración únicamente supera los 100 ms. cuando la consonante posterior es una aproximante o una oclusiva sorda. En cambio, la duración de la mayoría de realizaciones de /i/, /e/, /a/ y /o/ oscila entre 100 y 120 ms.,**

si bien hay excepciones. Por ejemplo, cuando el sonido posterior a /i/ y /o/ es nasal, la media no alcanza los 100 ms., y lo mismo sucede cuando la consonante que sigue a /e/ y /o/ es una oclusiva sorda. También hay casos en los que la duración es superior a 120 ms., como cuando /e/ es seguida por oclusiva sorda, o /a/ por vibrante simple o por oclusiva sonora. En cuanto a la desviación estándar, esta oscila, generalmente, entre 11 y 33, por lo que la homogeneidad en la duración de las vocales según el entorno posterior es similar a cuando estas se analizan en función de la consonante que las precede.

En tercer lugar, se ha estudiado la duración de las vocales cuando se encuentran en posición final de palabra. De este análisis se concluye que, a excepción de la /u/, la cual presenta una duración media mayor, pero el número de casos analizados es más reducido, las demás vocales tienen una duración media que oscila entre 110 y 120 ms., y la desviación estándar oscila entre 16 y 26, lo que indica que la duración de las vocales del español que se encuentran en final de palabra aislada tienen una duración similar y que esta, a su vez, es homogénea.

Estos datos, referidos a la duración de las vocales producidas por hispanohablantes nativas constituyen la referencia, para las voces femeninas, que permitirá establecer las diferencias de comportamiento entre hablantes nativas y estudiantes de español sinohablantes.

#### **1.4 Estructura silábica**

En tercer lugar, se ha analizado la duración media de las vocales articuladas por las informantes del grupo de control según la estructura silábica. Dicho análisis se ha realizado en función de diversos parámetros, como la estructura de la sílaba en la que se encuentra la vocal analizada, si la vocal se encuentra en sílaba inicial o final, la estructura de la sílaba anterior a la de la vocal analizada, la estructura de la sílaba posterior a la del sonido vocálico estudiado, o si la vocal se encuentra en posición final.

##### **1.4.1 Duración de la vocal según la sílaba en la que se encuentra**

Los resultados del análisis de la duración de la vocal en función de la estructura de la sílaba en la que se encuentra dicho sonido se muestran en la figura 38. Como se puede observar, en general, la duración media de las vocales es muy similar tanto si se encuentran en sílaba CV, como en sílaba CVC. En todos los casos la duración media

oscila entre 100 y 120 ms., excepto cuando /u/ se encuentra en sílaba CVC, en cuyo caso la duración es menor que 100 ms.

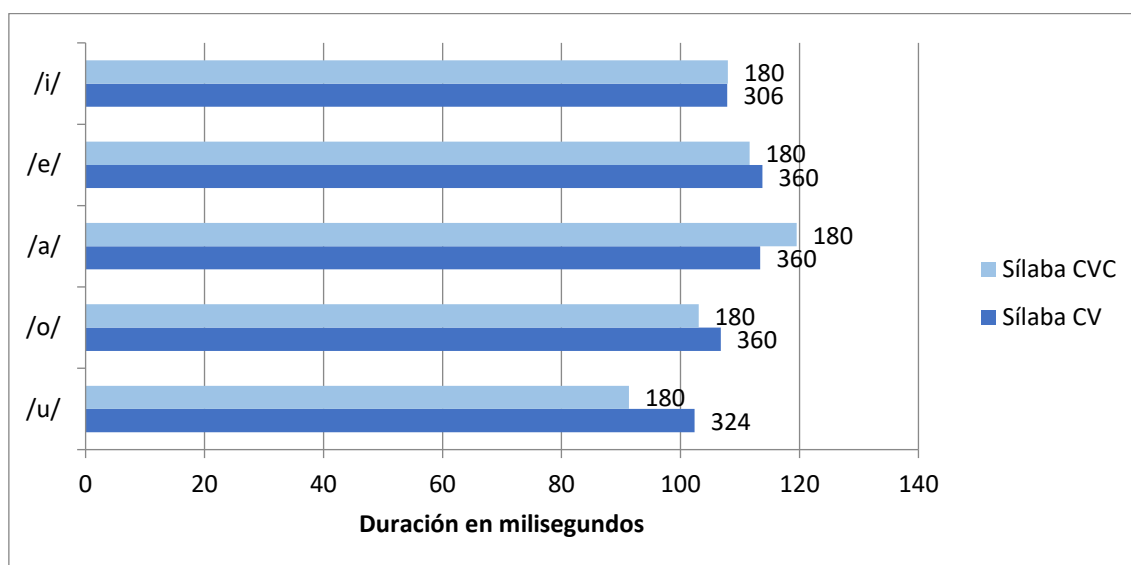


Figura 38. Duración media de las vocales producidas por el grupo de control según la sílaba en la que se encuentran dichos sonidos vocálicos.

En la primera columna de la tabla 24 aparece la vocal estudiada; en la segunda, se indica si la vocal analizada se encuentra en sílaba CV o CVC; en la tercera aparece el número de casos analizados; en la cuarta se indica la duración media de la vocal en función de la estructura de la sílaba en la que se encuentra, y en la quinta columna se muestra la desviación estándar correspondiente.

La tabla 24 muestra que la duración media de /i/ es la misma, tanto si se encuentra en una sílaba cuya estructura es CV, como si esta es CVC (107 ms.). La /e/ es más larga que /i/, y en este caso la duración también es similar tanto si se encuentra en estructura CV (113 ms.) como en CVC (111 ms.). En cuanto a la /a/, es la vocal más larga. En este caso, si este sonido se encuentra en una sílaba cuya estructura es CV, la duración media es de 113 ms., y es ligeramente más larga si se encuentra en CVC (119 ms.). En cuanto a la /o/, es más breve que las vocales analizadas, y la duración media es de 106 ms. si la estructura de la sílaba en la que se encuentra es CV, y de 103 ms. si es CVC. La /u/ es la vocal más breve de todas, y, a su vez, es la que mayor diferencia presenta según la estructura en la que se encuentra, ya que, si la estructura de la sílaba es CV, la media es de 102 ms., y si se encuentra en CVC, de 91 ms.



<b>Duración de la vocal según la estructura de la sílaba de la vocal</b>				
<b>Vocal</b>	<b>Estructura de la sílaba de la vocal</b>	<b>Nº de vocales analizadas</b>	<b>Duración media (ms)</b>	<b>Desviación estándar</b>
<b>/i/</b>	CV	306	107.8	22.3
	CVC	180	107.9	25.1
<b>/e/</b>	CV	360	113.7	24.2
	CVC	180	111.6	21.2
<b>/a/</b>	CV	360	113.4	21.1
	CVC	180	119.5	21.1
<b>/o/</b>	CV	360	106.7	23.6
	CVC	180	103.0	17.9
<b>/u/</b>	CV	324	102.3	20.0
	CVC	180	91.3	17.3

*Tabla 24.* Duración media y desviación estándar de la vocal según la estructura silábica en la que se encuentra dicha vocal.

Tal y como se ha podido observar, cuando la estructura de la sílaba es CV, la duración oscila entre 102 y 113 ms., siendo el orden de duración de las vocales, de más breves a más largas, el siguiente: /u/, /o/, /i/, /e/, /a/. En el caso de la estructura CVC, la duración media oscila entre 91 y 119 ms., y el orden de duración, de más breve a más larga, también es /u/, /o/, /i/, /e/, /a/.

En cuanto a la desviación estándar, esta oscila entre 20 y 24 cuando la estructura de la sílaba es CV, siendo estos datos los correspondientes a las vocales /u/ y /e/ respectivamente. Si las vocales se encuentran en una sílaba cuya estructura es CVC, la desviación estándar oscila entre 17 y 25, siendo /u/ la vocal que menor variabilidad presenta, e /i/ la de mayor variabilidad. Esto indica que la homogeneidad en la duración de las vocales en función de la estructura de la sílaba en la que se encuentran es similar en todos los casos, y la variabilidad es baja.

#### **1.4.2 Vocal en sílaba inicial**

Los resultados del análisis de la duración de las vocales del español cuando estas se encuentran en sílaba inicial de palabra aislada se muestran en la figura 39. Tal y como se puede observar, cuando la vocal en cuestión se encuentra en sílaba inicial de palabra aislada, la duración media es similar en todos los casos, aunque destaca /a/ por ser la vocal más larga.

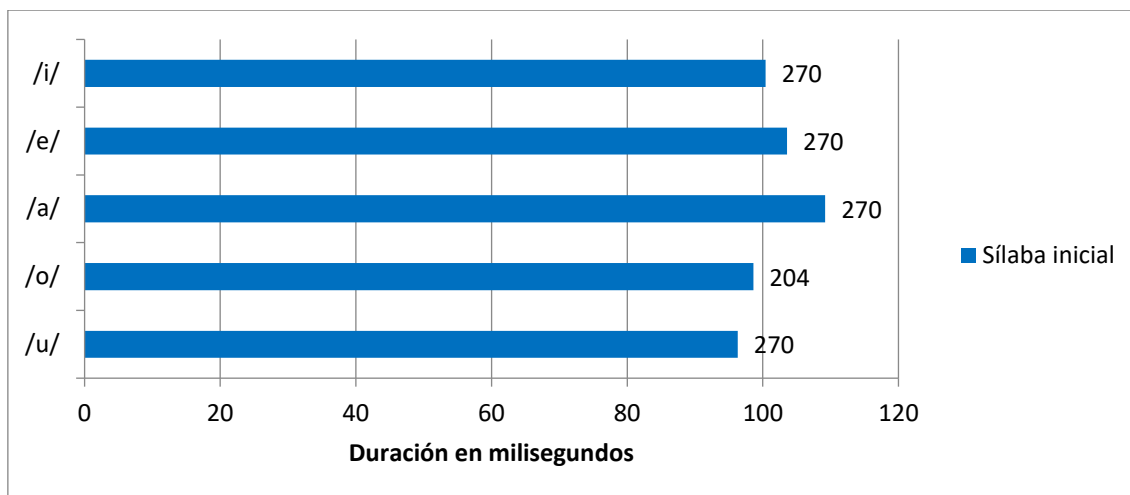


Figura 39. Duración media de las vocales producidas por el grupo de control cuando la vocal analizada se encuentra en sílaba inicial.

La vocal más breve es la posterior /u/, la cual tiene una duración media de 96 ms. La duración de /o/ es muy similar, y la media es de 98 ms. La diferencia entre esta y la /i/ también es pequeña, ya que la segunda tiene una duración media de 100 ms. La /e/ es ligeramente más larga, y su duración media es de 103 ms., y la /a/ es la vocal más larga, con una duración media de 109 ms. Estos datos muestran que hay poca diferencia entre la vocal más breve y la más larga.

Vocal	Vocal en sílaba inicial		
	Nº de vocales analizadas	Duración media (ms)	Desviación estándar
/i/	270	100.4	18.1
/e/	270	103.5	18.6
/a/	270	109.2	15.9
/o/	204	98.6	17.3
/u/	270	96.3	18.2

Tabla 25. Duración media y desviación estándar de las vocales producidas por el grupo de control cuando la vocal analizada se encuentra en sílaba inicial.

En cuanto a la desviación estándar, /a/ es la vocal que presenta mayor homogeneidad, ya que la desviación es de 15, y /e/ presenta la desviación estándar más elevada, de 18, aunque, como se puede apreciar, la diferencia es muy pequeña. Estos valores muestran que cuando las vocales se encuentran en sílaba inicial de palabra, la duración de estas es muy homogénea, y además la diferencia en la duración es muy baja, ya que la media oscila entre 96 y 109 ms.

### 1.4.3 Sílabas anteriores

La figura 40 muestra los resultados del análisis de la duración de la vocal en función de la estructura de la sílaba anterior a la sílaba en la que se encuentra el sonido vocálico estudiado. Por ejemplo, si en la palabra *total* la vocal analizada es /a/, la duración de dicha vocal se analiza en función de la sílaba anterior, en este caso 'to-', la cual tiene estructura CV. Se observa que en todos los casos la vocal es más larga cuando la estructura de la sílaba anterior es CV. Cabe señalar que, en este caso, no hay realizaciones de /u/ cuando la sílaba anterior es CVC.

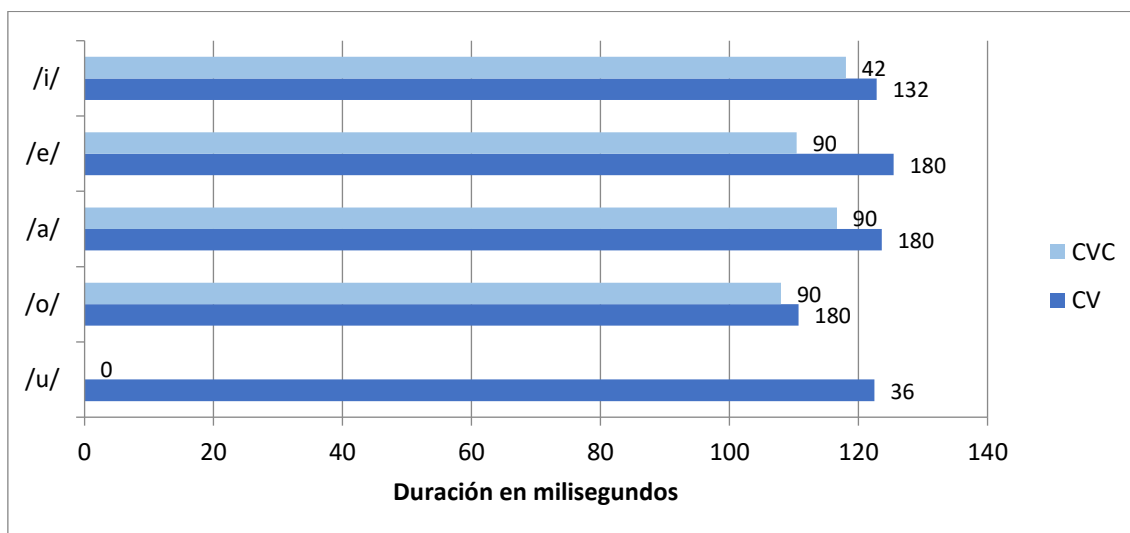


Figura 40. Duración media de las vocales producidas por el grupo de control según la sílaba anterior a dichos sonidos vocálicos.

La tabla 26 muestra que, en el caso de /i/, la duración media de esta vocal es de 122 ms. cuando la sílaba anterior tiene estructura CV, mientras que, si esta es CVC, la media es ligeramente inferior, de 118 ms. En cuanto a /e/, si la sílaba anterior es CV, la media es de 125 ms., mientras que la vocal es más breve, con una media de 110 ms., si la estructura de la sílaba anterior es CVC. Los datos relativos a /a/ muestran que la duración media es prácticamente igual que en /i/, es decir, de 123 ms. si la sílaba anterior es CV, y de 116 ms. si es CVC. La vocal /o/ es la más breve, ya que la duración media es de 110 ms. cuando la sílaba anterior es CV, y de 107 ms. cuando es CVC. Como se ha indicado, no hay casos en los que la estructura de la sílaba anterior a /u/ es CVC, y cuando esta es CV, la duración media es de 122 ms.

Duración de la vocal según la estructura de la sílaba anterior				
Vocal	Estructura de la sílaba anterior a la vocal	Nº de vocales analizadas	Duración media (ms)	Desv. Estándar
/i/	CV	132	122.8	23.8
	CVC	42	118.0	22.7
/e/	CV	180	125.4	22.0
	CVC	90	110.4	21.0
/a/	CV	180	123.5	22.3
	CVC	90	116.6	19.2
/o/	CV	180	110.6	24.4
	CVC	90	107.9	20.5
/u/	CV	36	122.4	21.8
	CVC	0	--	--

Tabla 26. Duración media y desviación estándar de la vocal según la estructura silábica de la sílaba anterior a la sílaba en la que se encuentra dicha vocal.

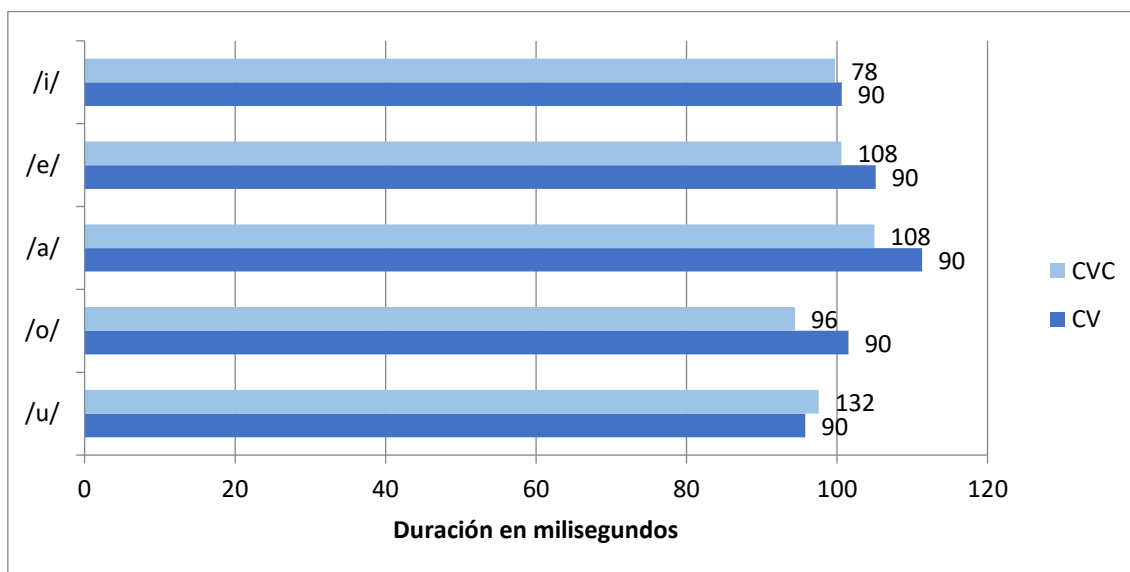
Si la estructura de la sílaba anterior a la de la vocal analizada es CV, la vocal más breve es la /o/, y la más larga la /e/, siendo la duración media de cada una de estas de 110 ms. y 125 ms. respectivamente. En este caso se observa que, si bien la /e/ es la de mayor duración media, las duraciones medias de /u/, /i/, /a/ son muy similares a la de esta (122, 122 y 123 ms. respectivamente). En los casos en los que la estructura de la sílaba anterior es CVC, la vocal más breve es la /o/, la cual tiene una duración media de 107 ms., la /e/ es ligeramente más larga, con una media de 110 ms., y /a/ e /i/ son las vocales más largas, con una duración media de 116 y 118 ms. respectivamente.

En cuanto a la desviación estándar, cuando la estructura de la sílaba anterior es CV, esta oscila entre 21 y 24, lo que indica que, en este caso, la homogeneidad es similar en todas las vocales. Lo mismo sucede cuando la estructura de la sílaba anterior es CVC, ya que la desviación estándar oscila entre 19 y 22.

#### 1.4.4 Sílaba posterior

En este apartado se analizan los resultados del análisis de la duración de las distintas vocales según la sílaba posterior a la sílaba del sonido vocálico estudiado. Por ejemplo, si en la palabra *lista* la vocal analizada es /i/, la duración de dicha vocal se estudia en función de la sílaba posterior, en este caso ‘-ta’. La figura 41 permite observar que la duración de las vocales es muy similar cuando la sílaba posterior tiene estructura CV o CVC. En los casos de las vocales /e/, /a/ y /o/, la vocal es levemente más larga cuando la

sílaba posterior tiene estructura CV. Sin embargo, en el caso de /u/ la vocal es levemente más larga cuando la estructura de la sílaba posterior es CVC.



*Figura 41.* Gráfico que muestra la duración media de las vocales producidas por el grupo de control según la sílaba posterior a dichos sonidos vocálicos.

De acuerdo con los datos mostrados en la tabla 27, la vocal /i/ presenta una duración media prácticamente idéntica, tanto si la sílaba posterior tiene estructura CV (100 ms.), como si es CVC (99 ms.). En /e/ la diferencia entre ambos casos no es elevada, ya que cuando la sílaba posterior es CV la duración media es de 105 ms., y cuando es CVC, de 100 ms. En cuanto a la /a/, la diferencia entre ambos casos es similar a la de /e/, aunque esta vocal presenta unas medias ligeramente superiores, de 111 ms. cuando la sílaba siguiente tiene estructura CV, y de 104 ms. si tiene estructura CVC. Cuando la sílaba posterior a la de /o/ es CV, la duración media de esta vocal es de 101 ms., y si dicha sílaba es CVC, la media es de 94 ms. En estos casos la media es siempre ligeramente superior cuando la estructura de la sílaba posterior es CV. En cambio, en el caso de la vocal /u/, la duración media es muy similar en ambos casos, siendo esta de 95 ms. ante estructura CV, y de 97 ms. ante una sílaba cuya estructura es CVC.

Duración de la vocal según la estructura de la sílaba posterior				
Vocal	Estructura de la sílaba posterior a la vocal	Nº de vocales analizadas	Duración media (ms)	Desviación estándar
/i/	CV	90	100.6	17.8
	CVC	78	99.7	18.0
/e/	CV	90	105.1	19.1
	CVC	108	100.5	16.5
/a/	CV	90	111.3	15.9
	CVC	108	104.9	14.6
/o/	CV	90	101.5	17.0
	CVC	96	94.4	15.5
/u/	CV	90	95.7	18.6
	CVC	132	97.5	17.2

Tabla 27. Duración media y desviación estándar de la vocal según la estructura silábica de la sílaba posterior a la sílaba en la que se encuentra dicha vocal.

Cuando la sílaba que sigue a la vocal analizada tiene la estructura CV, la vocal más breve es la /u/, con una duración media de 95 ms., y la más larga es la /a/, con una duración media de 111 ms. No obstante, cabe señalar que la duración de las otras vocales es muy similar, de 100 ms. en el caso de /i/, de 101 ms. en el de /o/ y de 105 ms. en el de /e/. Cuando la estructura de la sílaba posterior a la vocal analizada es CVC, la vocal más breve es la /o/, con una duración media de 94 ms., y la más larga es la /a/, con una media de 104 ms. Las vocales /u/, /i/ y /e/ tienen una duración media de 97, 99 y 100 ms. respectivamente.

En cuanto a la desviación estándar, esta oscila entre 15 y 19 cuando la sílaba posterior tiene estructura CV, y entre 14 y 18 cuando dicha sílaba es CVC. Esto indica que la realización de las vocales del español producidas por las hablantes que constituyen el grupo de control es homogénea y la variación en la duración es baja.

#### 1.4.5 Vocal en sílaba final

En cuanto a las vocales del español cuando dicha vocal se encuentra en sílaba final de palabra, tal y como se muestra en la figura 42, la duración media de las mismas es muy similar en todos los casos, excepto en /o/, siendo esta más breve que los otros sonidos vocálicos.

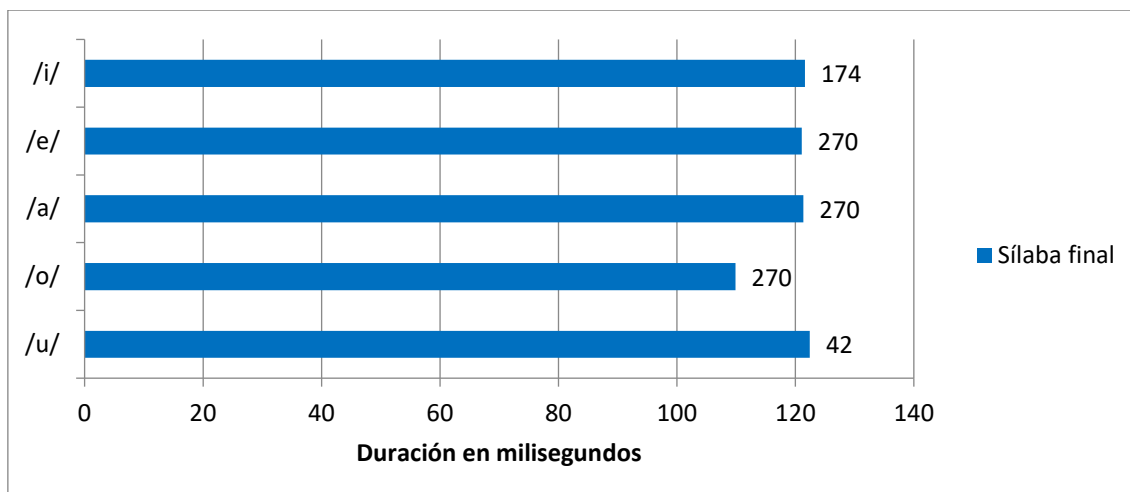


Figura 42. Duración media de las vocales producidas por el grupo de control cuando la vocal analizada se encuentra en sílaba final.

Como se ha indicado, la duración media de todas las vocales producidas por las hispanohablantes nativas es muy similar cuando la vocal se encuentra en sílaba final de palabra. En el caso de /o/, la duración media es de 109 ms., en los casos de /i/, /e/ y /a/ la duración media es de 121 ms., y en el de /u/ es de 122 ms.

Duración de la vocal cuando se encuentra en sílaba final			
Vocal	Nº de vocales analizadas	Duración media (ms)	Desv. Estándar
/i/	174	121.6	24.1
/e/	270	121.0	22.9
/a/	270	121.3	21.7
/o/	270	109.8	23.6
/u/	42	122.4	21.8

Tabla 28. Duración media y desviación estándar de las vocales producidas por el grupo de control cuando la vocal analizada se encuentra en sílaba final.

En cuanto a la desviación estándar, esta oscila entre 21 y 24, lo que indica que la realización de las vocales en sílaba final de palabra por las hablantes que conforman el grupo de control es homogénea y, además, la duración de todas ellas es muy similar.

#### 1.4.6 Recapitulación

A modo de resumen, los datos arrojan que tanto si la vocal analizada se encuentra en sílaba cuya estructura es CV, o si es CVC, la duración media es similar, siendo las vocales posteriores las que presentan una duración levemente más breve.

Además, las realizaciones son homogéneas, dado que la desviación oscila entre 17 y 25. Cuando las vocales analizadas se encuentran en posición inicial, la duración media oscila entre 96 y 109 ms., y la desviación estándar entre 15 y 19., lo que indica que la duración de las distintas vocales es muy similar y las realizaciones homogéneas. Sin embargo, la duración de las vocales es ligeramente superior cuando se analizan según la sílaba anterior que cuando se encuentran en sílaba inicial. En este caso la duración de las mismas es similar entre ellas, aunque cuando la sílaba anterior tiene estructura CV las vocales tienden a ser levemente más largas. También en este caso las realizaciones son homogéneas, puesto que la desviación estándar oscila entre 15 y 24. Si se analiza la duración de la vocal en función de la sílaba posterior, los datos muestran que la duración media oscila entre 95 y 111 ms. si la estructura de la sílaba posterior es CV, y entre 94 y 105 ms. si la estructura es CVC. Igual que en los casos anteriores, también en este caso las realizaciones son homogéneas, puesto que la desviación estándar es baja y oscila entre 14 y 19. Finalmente, se ha analizado la duración de las vocales cuando estas se encuentran en sílaba final de palabra, y se ha podido comprobar que, en este caso, los sonidos vocálicos son ligeramente más largos que en los anteriormente analizados. En cuanto a la desviación estándar, esta indica que las realizaciones de las informantes del grupo de control son homogéneas.

Cabe añadir que, del análisis de la duración de las vocales en función de la estructura silábica en la que se encuentran, de la sílaba anterior, o de la posterior, se concluye que generalmente las vocales más breves son las posteriores, es decir, /u/ y /o/, /i/ presenta una mayor duración que estas, y /e/ y /a/ suelen ser las más largas.

## **2. INFORMANTES SINOABLANTES**

En este apartado se muestran los resultados obtenidos tras realizar el análisis acústico de las grabaciones y el posterior tratamiento estadístico de los datos correspondientes al grupo de informantes sinohablantes. Igual que en el caso de las informantes que conforman el grupo de control, en este apartado se analiza la duración de las vocales producidas por las estudiantes chinas en función de distintos parámetros, como la acentuación, el entorno y la estructura silábica.



## 2.1 Duración general

En primer lugar, se ha analizado la duración media de las distintas vocales del español producidas por las informantes sinohablantes de manera global. La figura 43 muestra que las vocales siguen, de más breves a más largas, el orden /u/, /i/, /o/, /a/, /e/. La duración media de /u/ supera los 120 ms., la de /i/ y /o/ oscila entre 140 y 160 ms., y la de /e/ y /a/ supera los 160 ms.

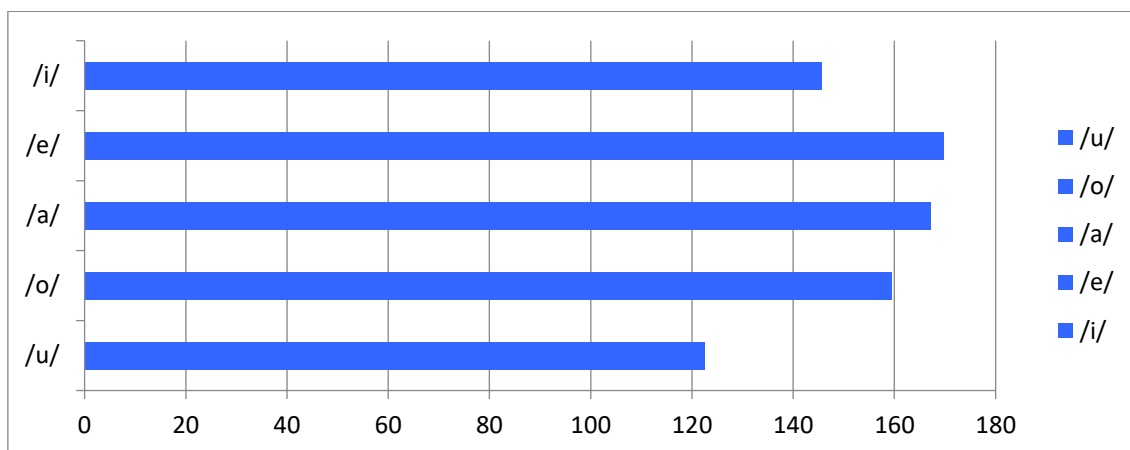


Figura 43. Duración media de las vocales producidas por las informantes sinohablantes.

La tabla 29 muestra la duración media de cada una de las vocales y la desviación estándar correspondiente a las mismas. Como se puede observar, la duración media de las vocales articuladas por las informantes sinohablantes oscila entre 122.5 y 169.7 ms.

Vocal	Nº de vocales analizadas	Duración media (ms)	Desviación estándar
/i/	456	145.7	53.9
/e/	480	169.7	56.7
/a/	480	167.2	46.2
/o/	480	159.5	50.0
/u/	429	122.5	40.2

Tabla 29. Duración media y desviación estándar de las vocales producidas por las informantes sinohablantes.

Tal y como se ha indicado, la vocal posterior /u/ es la más breve de todas las vocales producidas por el grupo de informantes sinohablantes, con una duración media de 122.5 ms., seguida por la anterior /i/, con una media de 145.7 ms. La vocal posterior /o/ es más larga y tiene una media de 159.5 ms.; las vocales /a/ y /e/ son las más largas, con una duración media de 167.2 y 169.7 ms. respectivamente. En cuanto a la desviación estándar,

esta es considerablemente elevada, ya que oscila entre 40.2 en el caso de /u/ y 56.7 en el de /e/, que además son la vocal más breve y más larga respectivamente.

## 2.2 Acentuación

En segundo lugar, se ha analizado la duración de las vocales producidas por las informantes sinohablantes en función de la tonicidad de las mismas. La figura 44 permite tener una visión general de las realizaciones del grupo de informantes sinohablantes.

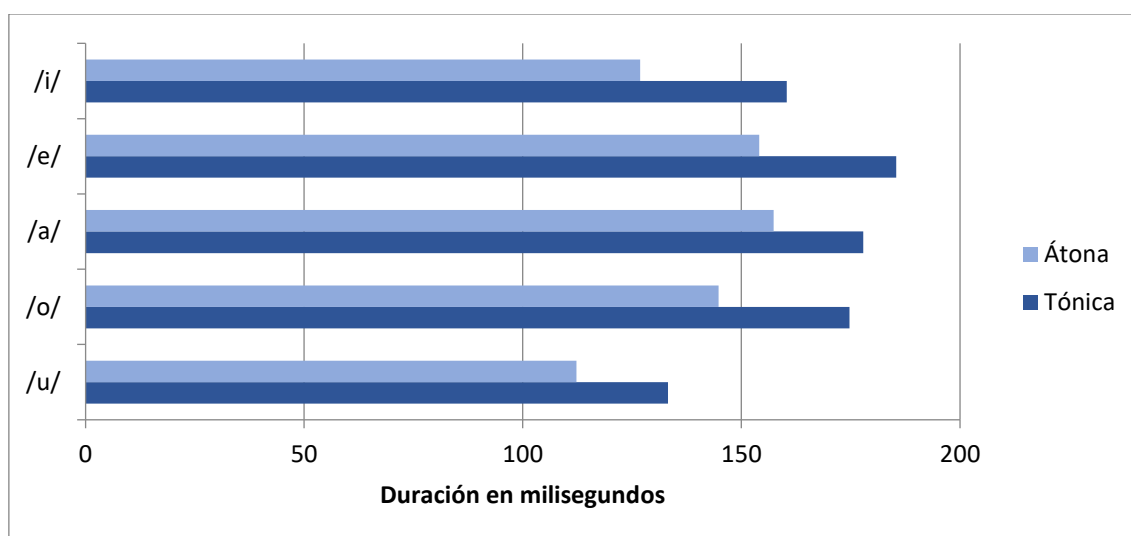


Figura 44. Duración media de las vocales producidas por las informantes sinohablantes.

Se observa que en todos los casos la duración de las vocales realizadas como tónicas es mayor que las realizadas como átonas. En las primeras, la duración en orden de más breve a más larga es /u/, /i/, /o/, /a/, /e/, mientras que en las segundas, el orden es /u/, /i/, /o/, /e/, /a/.

La tabla 30 permite realizar un análisis exhaustivo de los datos. La primera columna indica cuál es la vocal analizada, la segunda señala si dicha vocal es tónica o átona, la tercera columna muestra el número de vocales analizadas para cada uno de los casos, la cuarta columna muestra la duración media de dicha vocal en milisegundos y la quinta columna indica la desviación estándar, también en milisegundos, de cada una de las medias calculadas.

Vocal	Acentuación	Nº de vocales analizadas	Duración media (ms)	Desviación estándar
/i/	Tónica	240	160.3	57.1
	Átona	216	126.8	42.0
/e/	Tónica	240	185.3	60.8
	Átona	240	154.0	42.3
/a/	Tónica	240	177.8	48.6
	Átona	240	157.4	36.6
/o/	Tónica	240	174.7	56.1
	Átona	240	144.7	35.2
/u/	Tónica	210	133.1	42.9
	Átona	219	112.2	33.9

Tabla 30. Duración media y desviación estándar de las vocales producidas por las informantes sinohablantes en función de la tonicidad de las mismas.

En cuanto a la duración media de las vocales que las informantes sinohablantes han realizado como tónicas, ya se ha indicado que la /u/ es la más breve, con una duración media de 133.1 ms., mientras que la /e/ es la más larga, con una media de 185.3 ms. Las vocales átonas son, como se ha señalado, más breves, y en este caso la duración media va desde los 112.2 ms. de /u/, hasta los 157.4 ms. de /a/. Además, se observa que en el caso de las vocales tónicas hay menor homogeneidad que en el de las átonas, ya que la desviación estándar de las primeras es superior y va desde 42.9 hasta 60.8, mientras que en el de las segundas va desde 33.9 hasta 42.3.

### 2.3 Entorno

En este apartado se analiza la duración de las distintas vocales producidas por las informantes sinohablantes en función del entorno, tanto anterior, como posterior, y de cuando el sonido vocálico se encuentra en posición final de palabra aislada. Igual que en el análisis correspondiente a las hablantes del grupo de control, los sonidos han sido clasificados en las siguientes categorías: oclusiva sorda, oclusiva sonora, aproximante, fricativa, africada, nasal, lateral, vibrante simple y vibrante múltiple. Como se ha visto en la parte 1 del presente capítulo, en algunas ocasiones las informantes sinohablantes no realizan correctamente el entorno, y en ocasiones articulan un sonido vocálico. Por este motivo, en este apartado, se ha añadido una nueva categoría a las anteriormente señaladas, la categoría *vocal*.

### 2.3.1 Duración de la vocal según el entorno anterior

En este apartado se analiza la duración de las vocales del español producidas por los hablantes de mandarín según el entorno anterior. Es decir, si se analiza la vocal /o/ en las palabras *total* o *muro*, el entorno anterior de dicha vocal es la oclusiva sorda [t] en el caso de la primera palabra y la vibrante simple [r] en el caso de la segunda.

#### 2.3.1.1 Vocal /i/

La figura 45 muestra la duración media de la vocal /i/ según el entorno anterior, y el número de casos en que ha aparecido dicha realización. Se observa, por ejemplo, que la oclusiva sorda precede a /i/ en 198 ocasiones, mientras que la oclusiva sonora únicamente lo hace en 16 casos. Tal y como se puede observar, cuando el entorno anterior a /i/ es una oclusiva sorda, oclusiva sonora, fricativa, africada, nasal, vibrante simple o vocal, la duración media oscila, aproximadamente, entre 130 y 150 ms., mientras que cuando el entorno anterior está compuesto por una aproximante, un sonido lateral o una vibrante múltiple, la duración es superior, especialmente en el caso de la aproximante. Sin embargo, es necesario tener en cuenta que el número de casos en los que aparece /i/ detrás de la aproximante, por ejemplo, es más reducido que el de oclusivas sordas o fricativas, por lo que cuando el número de casos es escaso, los datos únicamente permiten observar una tendencia en el comportamiento, y no un comportamiento generalizado.

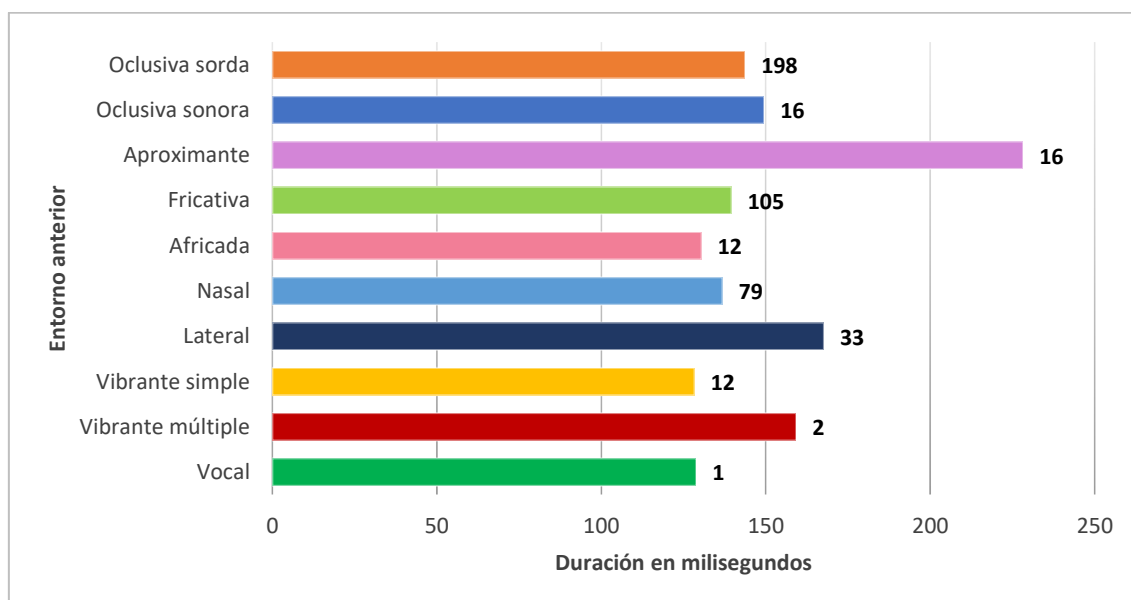


Figura 45. Duración media de la vocal /i/ producida por las informantes sinohablantes según el entorno anterior realizado.

Cuando el sonido anterior a /i/ es otra vocal o una vibrante simple, la duración media es la más breve, ya que es de 128 ms. en ambos casos. Esta media es ligeramente superior si el sonido que precede a la vocal es una africada (130.5 ms.), una nasal (136.8 ms.) o una fricativa (139.6 ms.). Si la consonante anterior es una oclusiva sorda, la media es similar a las anteriores, de 143.7 ms., y ligeramente superior si la oclusiva es sonora (149.5 ms.). Si el sonido que precede a la vocal /i/ es una vibrante múltiple, la duración media es mayor, de 159.3 ms., y es superior si se trata de una lateral (167.6 ms.). Como se ha indicado, la mayor duración media de esta vocal corresponde a la realización de aproximante seguida de /i/, en cuyo caso asciende a 228.2 ms.

/i/			
Entorno anterior realizado por las informantes sinohablantes	Casos	Duración media	Desviación estándar
Oclusiva sorda	198	143.7	55.9
Oclusiva sonora	16	149.5	49.0
Aproximante	16	228.2	73.8
Fricativa	105	139.6	52.6
Africada	12	130.5	49.4
Nasal	79	136.8	56.8
Lateral	33	167.6	53.9
Vibrante simple	12	128.3	38.8
Vibrante múltiple	2	159.3	25.0
Vocal	1	128.9	--

*Tabla 31.* Duración media y desviación estándar de la vocal /i/ producida por las informantes sinohablantes según el entorno anterior.

En el caso de la duración de la vocal /i/, la desviación estándar va desde 25 cuando el sonido anterior es una vibrante múltiple, hasta 73.8 cuando se trata de una aproximante. Sin embargo, cabe señalar que ambos sonidos han sido realizados en 2 y 16 ocasiones respectivamente, por lo que únicamente muestran una tendencia. En cuanto a aquellos sonidos cuya presencia es mayor en el corpus, se observa que la desviación oscila entre 52 y 60, por lo que puede considerarse que la falta de homogeneidad en la duración de la vocal /i/ producida por las informantes sinohablantes es considerable.

### 2.3.1.2 Vocal /e/

La figura 46 muestra que cuando el sonido anterior a /e/ es una oclusiva sonora o una vibrante múltiple, la duración media de esta vocal es más breve que en los otros casos. Si dicho sonido consonántico es una nasal, una lateral, una vibrante simple o una fricativa,

la duración es ligeramente superior y similar en estos casos. Tal y como se puede observar, si el sonido que precede a la vocal es una aproximante u otra vocal, la duración de /e/ es la más larga de todos los casos. Sin embargo, cabe señalar que únicamente hay un caso en el que el sonido anterior ha sido realizado como vocal, por lo que este dato es indicador de una tendencia, y no de un comportamiento generalizado.

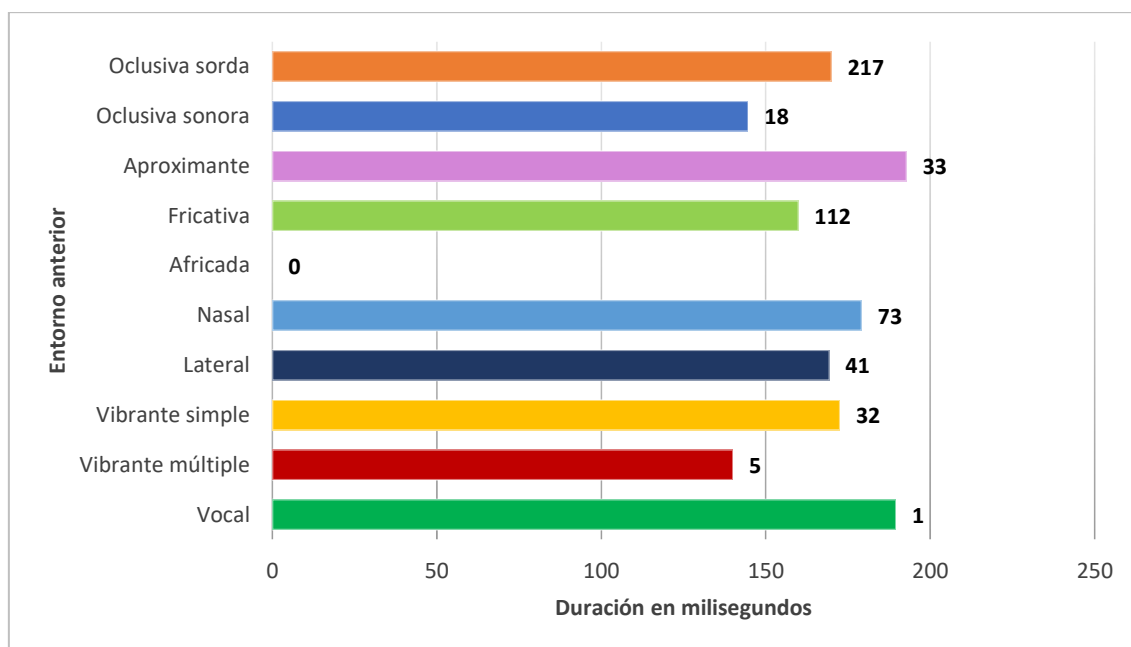


Figura 46. Duración media de la vocal /e/ producida por las informantes sinohablantes según el entorno anterior realizado.

Como se ha indicado, /e/ es más breve cuando el sonido anterior es una vibrante múltiple (140 ms.), una oclusiva sonora (144.7 ms.) o una fricativa (159.9 ms.). Este sonido vocálico tiene una duración media similar cuando las informantes sinohablantes antes han realizado una lateral (169.3 ms.), una oclusiva sorda (170 ms.), una vibrante simple (172.5 ms.) o una nasal (176.1 ms.). En el caso en el que el sonido anterior a /e/ es otra vocal, la duración es de 189.6 ms., y las realizaciones más largas corresponden a los casos en los que el sonido anterior es una aproximante (192.8 ms.).

<i>/e/</i>			
<b>Entorno anterior realizado por las informantes sinohablantes</b>	<b>Casos</b>	<b>Duración media</b>	<b>Desviación estándar</b>
Oclusiva sorda	217	170.0	65.3
Oclusiva sonora	18	144.7	41.1
Aproximante	33	192.8	68.1
Fricativa	112	159.9	61.3
Africada	0	--	--
Nasal	73	179.1	70.1
Lateral	41	169.3	61.7
Vibrante simple	32	172.5	53.1
Vibrante múltiple	5	140.0	36.9
Vocal	1	189.6	--

*Tabla 32.* Duración media y desviación estándar de la vocal /e/ producida por las informantes sinohablantes según el entorno anterior.

En cuanto a la desviación estándar en la duración de /e/, esta es menor cuando el sonido anterior es una vibrante múltiple (36.9) o una oclusiva sonora (41.1), aunque este sonido no ha sido realizado en un número elevado de casos. Si el sonido anterior es una oclusiva sorda, una fricativa, una nasal o una lateral, la desviación estándar se sitúa entre 61 y 70, lo que indica que en la realización de /e/ hay incluso menor homogeneidad que en la de /i/.

### 2.3.1.3 Vocal /a/

Tal y como se puede observar en la figura 47, cuando el sonido anterior es una fricativa o una africada la duración media de /a/ es más breve que cuando es precedida por los otros sonidos. Si la consonante que precede a la vocal es una oclusiva sorda, una oclusiva sonora, una aproximante, una nasal, una lateral, una vibrante simple u otra vocal, la duración media de /a/ oscila entre 150 y 200 ms. Cabe señalar que las informantes sinohablantes han realizado el sonido anterior a /a/ como aproximante, africada o vocal en un número reducido de ocasiones, por lo que estos datos muestran una tendencia, pero no comportamiento generalizado de /a/ cuando dichos sonidos preceden a esta vocal.

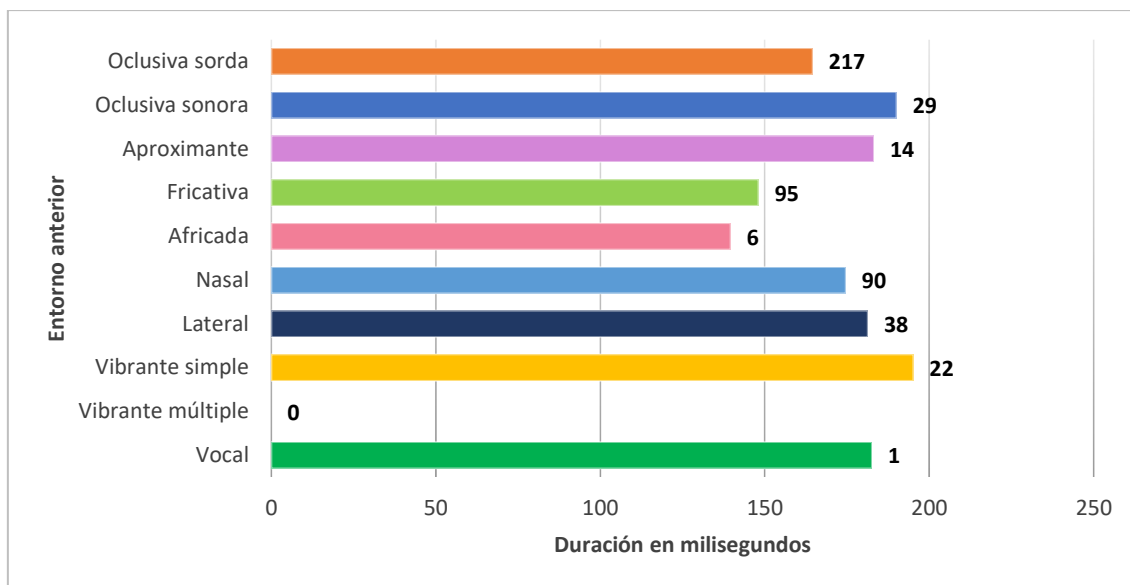


Figura 47. Duración media de la vocal /a/ producida por las informantes sinohablantes según el entorno anterior realizado.

De acuerdo con los datos mostrados en la tabla 33, la duración de la vocal oscila entre 139.6 y 195.1 ms. Se observa que /a/ es más breve cuando el sonido anterior es una africada (139.6 ms.), y la duración es algo superior si dicha consonante es una fricativa (148.1 ms.). Cuando el sonido anterior es una oclusiva sorda, la /a/ es más larga que en los casos anteriores (164.4 ms.), y ligeramente superior si es una nasal (174.6 ms.). Las duraciones más largas de esta vocal se han detectado cuando el sonido anterior es una aproximante (183.2 ms.), una vocal (182.6 ms.) o una lateral (181.3 ms.).

/a/			
Entorno anterior realizado por las informantes sinohablantes	Casos	Duración media	Desviación estándar
Oclusiva sorda	217	164.4	48.0
Oclusiva sonora	29	190.0	58.2
Aproximante	14	183.2	55.3
Fricativa	95	148.1	44.2
Africada	6	139.6	40.1
Nasal	90	174.6	59.1
Lateral	38	181.3	47.2
Vibrante simple	22	195.1	47.6
Vibrante múltiple	0	--	--
Vocal	1	182.6	--

Tabla 33. Duración media y desviación estándar de la vocal /a/ producida por las informantes sinohablantes según el entorno anterior.

Los casos en los que el sonido anterior a la vocal /a/ es una oclusiva sorda, una fricativa o una nasal, la desviación estándar oscila entre 44 y 69. Esta desviación es más elevada



que la correspondiente a otras vocales, lo que indica falta de homogeneidad en la duración de la vocal en cuestión realizada por las informantes sinohablantes.

### 2.3.1.4 Vocal /o/

En la figura 48 se aprecia que la duración media de /o/ según el entorno anterior a la vocal no presenta diferencias tan pronunciadas como en las vocales anteriormente analizadas. Esta vocal es ligeramente más breve cuando es precedida por una oclusiva sorda, una fricativa o una africada, mientras que si el sonido anterior es una oclusiva sonora, una vibrante simple o una vibrante múltiple, /o/ es más larga, y es superior cuando es precedida por una aproximante. Como sucede en apartados anteriores, hay algunos sonidos anteriores a /o/ que las informantes sinohablantes han articulado en pocas ocasiones, como la vibrante múltiple o la oclusiva sonora, por lo que estos datos obtenidos del análisis de los mismos muestran una tendencia, y no un comportamiento generalizado.

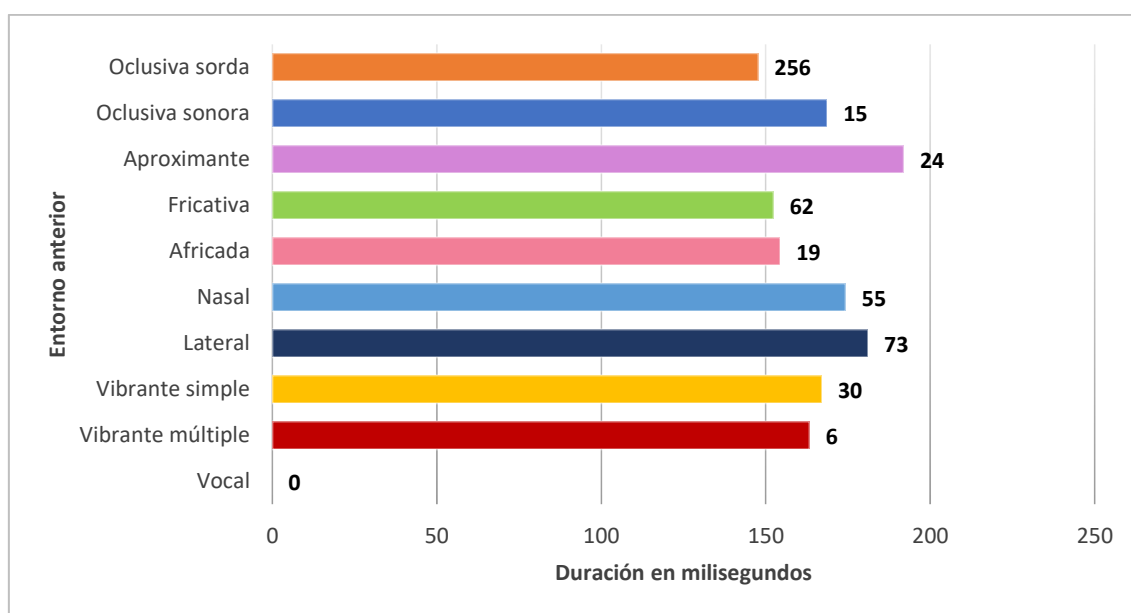


Figura 48. Duración media de la vocal /o/ producida por las informantes sinohablantes según el entorno anterior realizado.

La duración media de /o/ según el sonido anterior es de 147,8 ms. cuando el sonido anterior es una oclusiva sorda. Si es precedida por una fricativa o una africada, la media es muy similar en ambos casos, de 152.4 y 154.4 ms. respectivamente. También es similar y algo más larga cuando el sonido anterior es una vibrante múltiple (163.3 ms.), una vibrante simple (167 ms.) o una oclusiva sonora (168.7 ms.), e incluso superior si se trata de una nasal (174.2 ms.). Las duraciones más largas de esta vocal corresponden a la realización de la misma precedida por lateral (181.1 ms.) y por aproximante (191.9 ms.).

<i>/o/</i>			
<b>Entorno anterior realizado por las informantes sinohablantes</b>	<b>Casos</b>	<b>Duración media</b>	<b>Desviación estándar</b>
Oclusiva sorda	256	147.8	52.4
Oclusiva sonora	15	168.7	56.1
Aproximante	24	191.9	69.7
Fricativa	62	152.4	45.7
Africada	19	154.4	34.2
Nasal	55	174.2	51.0
Lateral	73	181.1	64.4
Vibrante simple	30	167.0	67.8
Vibrante múltiple	6	163.3	78.3
Vocal	0	--	--

*Tabla 34.* Duración media y desviación estándar de la vocal /o/ producida por las informantes sinohablantes según el entorno anterior.

En cuanto a la desviación estándar, se observa que esta oscila entre 34 y 69, lo que muestra que esta es considerablemente alta, y que la duración de la vocal /o/ producida por las informantes presenta variación de unos casos a otros.

### **2.3.1.5 Vocal /u/**

La figura 49 muestra que cuando el sonido anterior es una africada o una fricativa, la vocal /u/ es más breve. Si dicha consonante es una oclusiva sorda, oclusiva sonora, nasal, lateral o vibrante múltiple, la duración media de esta vocal es superior, y similar en estos casos. En cambio, /u/ es más larga cuando el sonido anterior es una vibrante simple, y destaca la duración de esta vocal cuando es precedida por una aproximante. Sin embargo, cabe señalar que las informantes sinohablantes han realizado esta consonante únicamente en tres ocasiones, por lo que estos datos no reflejan un comportamiento generalizado en la duración de esta vocal.

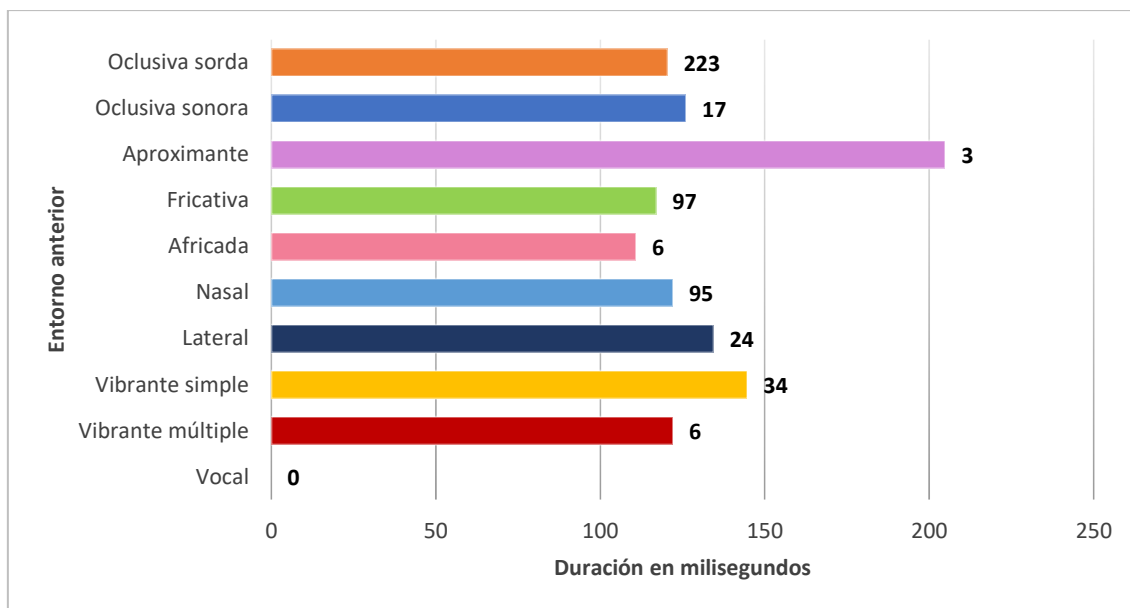


Figura 49. Duración media de la vocal /u/ producida por las informantes sinohablantes según el entorno anterior realizado.

Como se ha indicado, la duración más breve de /u/ corresponde a la realización de esta precedida por una consonante africada, y es de 110.8 ms. Esta vocal tiene una duración muy similar cuando el sonido anterior es una fricativa (117 ms.), una oclusiva sorda (120.4 ms.), una nasal (122.1 ms.), una vibrante múltiple (122.1 ms.), o una oclusiva sonora (126 ms.). /u/ es más larga cuando el sonido que la precede es una lateral (134.4 ms.) o una vibrante simple (144.7 ms.), y las realizaciones más largas corresponden a las de aproximante seguida por /u/ (204.8 ms.).

/u/			
Entorno anterior realizado por las informantes sinohablantes	Casos	Duración media	Desviación estándar
Oclusiva sorda	223	120.4	40.2
Oclusiva sonora	17	126.0	45.0
Aproximante	3	204.8	104.6
Fricativa	97	117.0	38.8
Africada	6	110.8	27.3
Nasal	95	122.1	40.6
Lateral	24	134.4	41.8
Vibrante simple	34	144.7	48.9
Vibrante múltiple	6	122.1	31.1
Vocal	0	--	--

Tabla 35. Duración media y desviación estándar de la vocal /u/ producida por las informantes sinohablantes según el entorno anterior.

Se observa que la desviación estándar en la duración de /u/ oscila entre 38 y 48. Esta desviación es menor que la de las vocales anteriormente analizadas (/i/, /e/, /a/, /o/), lo que indica que existe mayor homogeneidad en la duración de /u/ articulada por las informantes sinohablantes que en la de cualquiera de las otras vocales.

### 2.3.2 Duración de la vocal según el entorno posterior

En este apartado se analiza la duración de las vocales del español producidas por las hablantes de mandarín según el entorno posterior. Es decir, si se analiza la vocal /a/ en las palabras *total* o *tomar*, el entorno posterior de dicha vocal es la lateral [l] en el caso de la primera y la vibrante simple [r] en el caso de la segunda.

#### 2.3.2.1 Vocal /i/

La figura 50 muestra que cuando el sonido posterior a /i/ es una nasal u otra vocal, /i/ es más breve que en el resto de casos, y la duración media oscila entre 100 y 120 ms. Si el sonido que sigue a esta vocal es una oclusiva sorda, una aproximante, una fricativa, una africada o una lateral, la duración media de /i/ oscila entre 120 y 140 ms., mientras que es considerablemente más larga cuando el sonido realizado por las informantes sinohablantes es una vibrante simple. Cabe señalar que algunos de los sonidos, como la africada o la vocal, han aparecido en un reducido número de ocasiones después de /i/, por lo que los datos relativos a los mismos únicamente muestran una tendencia.

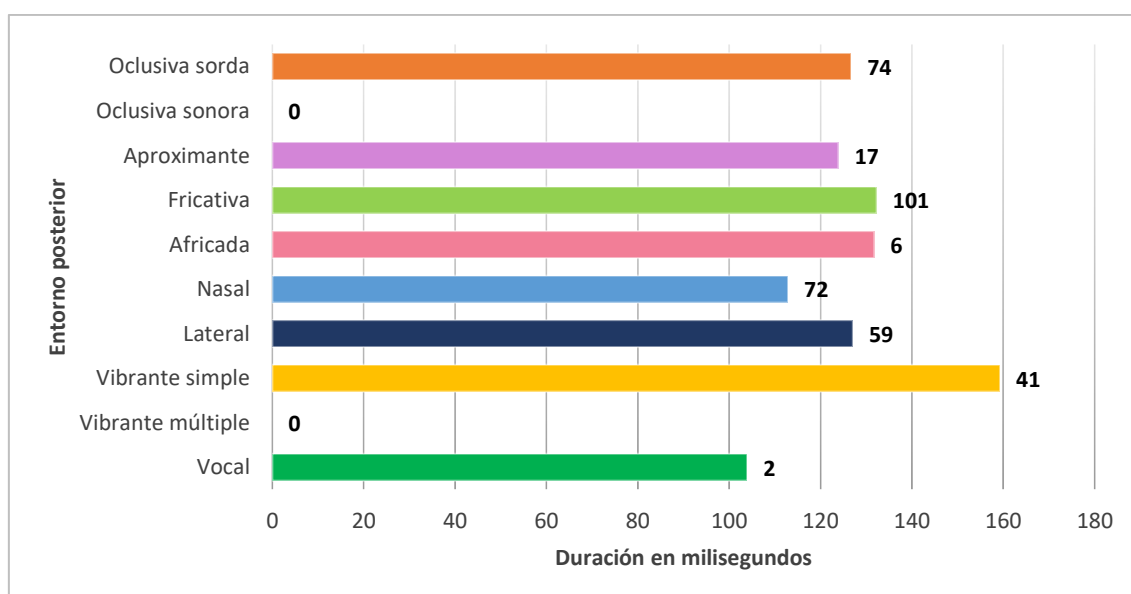


Figura 50. Duración media de la vocal /i/ producida por las informantes sinohablantes según el entorno posterior realizado.

La tabla 36 muestra que la vocal /i/ es más breve cuando el sonido posterior realizado por las estudiantes chinas es otra vocal (103.9 ms.). En los casos en los que /i/ es seguida por una consonante, esta vocal es más breve si dicho sonido es una nasal (112.8 ms.). La duración media es mayor cuando la consonante posterior es una aproximante (123.9 ms.), una oclusiva sorda (126.7 ms.), una lateral (127 ms.), una africada (131.7 ms.) o una fricativa (132.2 ms.). En cambio, la vocal /i/ realizada por las informantes sinohablantes es más larga cuando el sonido posterior es una vibrante simple, siendo la media de 159.3 ms.

<i>/i/</i>			
<b>Entorno posterior realizado por las informantes sinohablantes</b>	<b>Casos</b>	<b>Duración media</b>	<b>Desviación estándar</b>
Oclusiva sorda	74	126.7	34.9
Oclusiva sonora	0	--	--
Aproximante	17	123.9	27.2
Fricativa	101	132.2	42.8
Africada	6	131.7	44.2
Nasal	72	112.8	33.3
Lateral	59	127.0	46.1
Vibrante simple	41	159.3	38.8
Vibrante múltiple	0	--	--
Vocal	2	103.9	8.0

*Tabla 36.* Duración media y desviación estándar de la vocal /i/ producida por las informantes sinohablantes según el entorno posterior.

En cuanto a la desviación estándar, en aquellos sonidos en los que hay un número representativo de casos se ha detectado que la desviación oscila entre 33 y 46, lo que indica que hay una cierta falta de homogeneidad en la realización de /i/.

### 2.3.2.2 Vocal /e/

La figura 51 muestra que la duración media de /e/ es más breve cuando el sonido posterior a la misma es una nasal. Si dicha consonante es una fricativa o una oclusiva sorda, la vocal es más larga, y la duración media supera los 160 ms. cuando /e/ precede a oclusiva sonora, aproximante, lateral o vibrante múltiple. Cabe señalar que no se han detectado casos en los que esta vocal es seguida por africada, vibrante simple o vocal. Además, el número de casos relativos a la oclusiva sonora es escaso, por lo que estos datos únicamente permiten observar una tendencia, no un comportamiento generalizado.

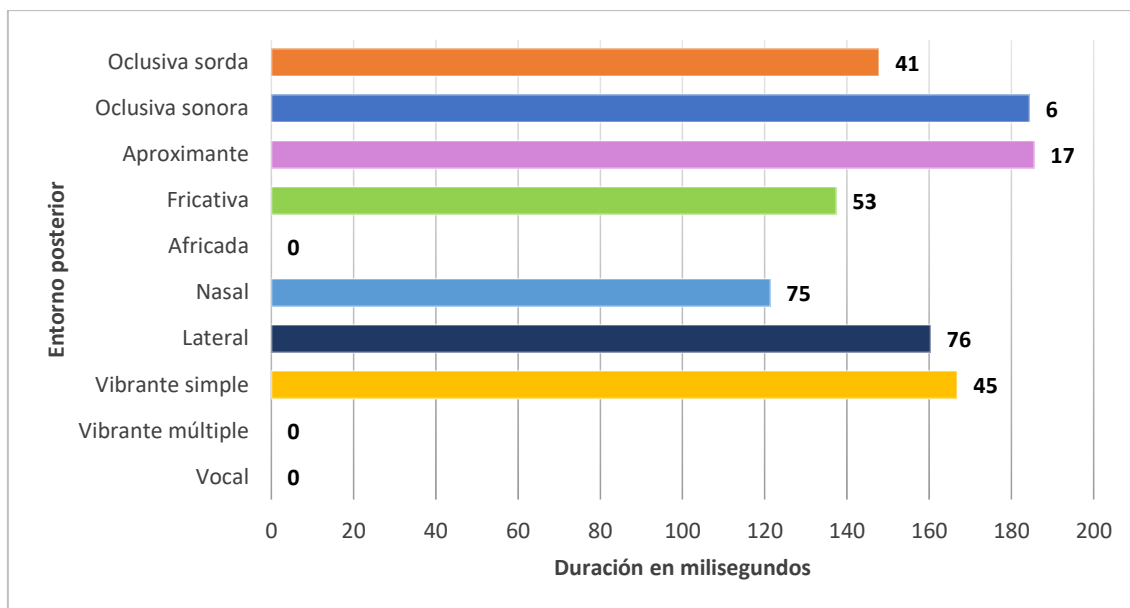


Figura 51. Duración media de la vocal /e/ producida por las informantes sinohablantes según el entorno posterior realizado.

Tal y como se ha indicado, la duración más breve de /e/ corresponde a la realización de esta seguida por una consonante nasal (121.4 ms.). Esta vocal es más larga cuando el sonido posterior es una fricativa (137.4 ms.), una oclusiva sorda (147.8 ms.). La duración media es mayor que en los casos anteriores cuando el sonido que sigue a la vocal en cuestión es una vibrante simple (166.8 ms.) o una lateral (160.3 ms.). Cuando las informantes sinohablantes articulan /e/ seguida de oclusiva sonora o aproximante, la vocal tiene la duración media mayor de todos los casos analizados, ya que es de 184.4 y 185.5 ms. respectivamente.

/e/			
Entorno posterior realizado por las informantes sinohablantes	Casos	Duración media	Desviación estándar
Oclusiva sorda	41	147.8	40.9
Oclusiva sonora	6	184.4	74.7
Aproximante	17	185.5	91.1
Fricativa	53	137.4	35.4
Africada	0	--	--
Nasal	75	121.4	32.7
Lateral	76	160.3	45.2
Vibrante simple	45	166.8	44.2
Vibrante múltiple	0	--	--
Vocal	0	--	--

Tabla 37. Duración media y desviación estándar de la vocal /e/ producida por las informantes sinohablantes según el entorno posterior.

En cuanto a la desviación estándar en la realización de /e/, en aquellos casos en los que hay un número representativo de realizaciones, la desviación oscila entre 32.7 y 45.2, muy similar a la vocal /i/, lo que indica que hay falta de homogeneidad en la realización de dicho sonido.

### 2.3.2.3 Vocal /a/

Cuando las informantes sinohablantes articulan la /a/ seguida por una oclusiva sorda, una oclusiva sonora, una fricativa o una nasal, la duración media es menor que en otros casos, pero similar entre estos. Si el sonido posterior a esta vocal es una aproximante, lateral o vibrante simple, la duración media oscila entre 150 y 200 ms., mientras que, si dicho sonido es una vibrante múltiple, /a/ supera los 200 ms. Sin embargo, cabe señalar que solo hay un caso en el que esta consonante es posterior a /a/, y lo mismo sucede cuando /a/ es seguida por otra vocal, por lo que este dato no muestra un comportamiento generalizado.

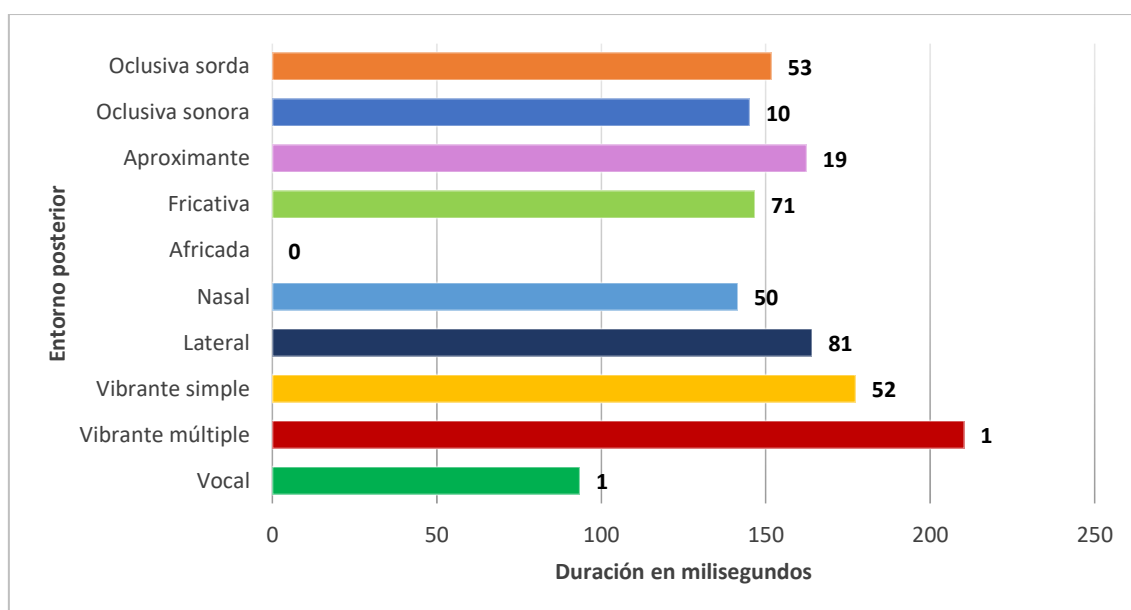


Figura 52. Duración media de la vocal /a/ producida por las informantes sinohablantes según el entorno posterior realizado.

La tabla 38 muestra que la duración de /a/ cuando el sonido posterior es otra vocal es de 93.5 ms., aunque, como ya se ha indicado, únicamente hay un caso en el que otra vocal sigue a /a/. Cuando la consonante que sigue a esta vocal es una nasal, la duración media es de 141.5 ms., si es una oclusiva sonora, de 145.1 ms., si es una fricativa, de 146.7 ms., y si es una oclusiva sorda, de 151.8 ms. Esta vocal es más larga cuando el sonido posterior es una aproximante (162.4 ms.) o una lateral (164.1 ms.), aunque las realizaciones más

largas corresponden a la realización de /a/ seguida por vibrante simple (177.3 ms.) o vibrante múltiple (210.4 ms.).

<i>/a/</i>			
<b>Entorno posterior realizado por las informantes sinohablantes</b>	<b>Casos</b>	<b>Duración media</b>	<b>Desviación estándar</b>
Oclusiva sorda	53	151.8	46.3
Oclusiva sonora	10	145.1	43.2
Aproximante	19	162.4	52.0
Fricativa	71	146.7	29.6
Africada	0	--	--
Nasal	50	141.5	35.0
Lateral	81	164.1	37.4
Vibrante simple	52	177.3	53.1
Vibrante múltiple	1	210.4	--
Vocal	1	93.5	--

*Tabla 38.* Duración media y desviación estándar de la vocal /a/ producida por las informantes sinohablantes según el entorno posterior.

En relación con la desviación estándar de la vocal /a/ realizada por las hablantes cuya lengua materna es el chino, esta oscila entre 29.6 y 53.1, lo que indica hay falta de homogeneidad en la duración de esta vocal según el entorno posterior.

#### **2.3.2.4 Vocal /o/**

En relación con la duración de /o/ según el entorno posterior a esta, la figura 53 muestra que cuando la consonante que sigue a la vocal es una vibrante múltiple la /o/ es mucho más breve que en el resto de casos. Sin embargo, cabe señalar que solo hay un caso de esta vocal seguida por este sonido consonántico, por lo que esto no muestra un comportamiento generalizado. Cuando el sonido posterior es una aproximante, una nasal, una lateral o una vibrante simple, la duración es similar y oscila entre 120 y 140 ms., mientras que la vocal es ligeramente más larga si la consonante es una oclusiva sorda o una oclusiva sonora.



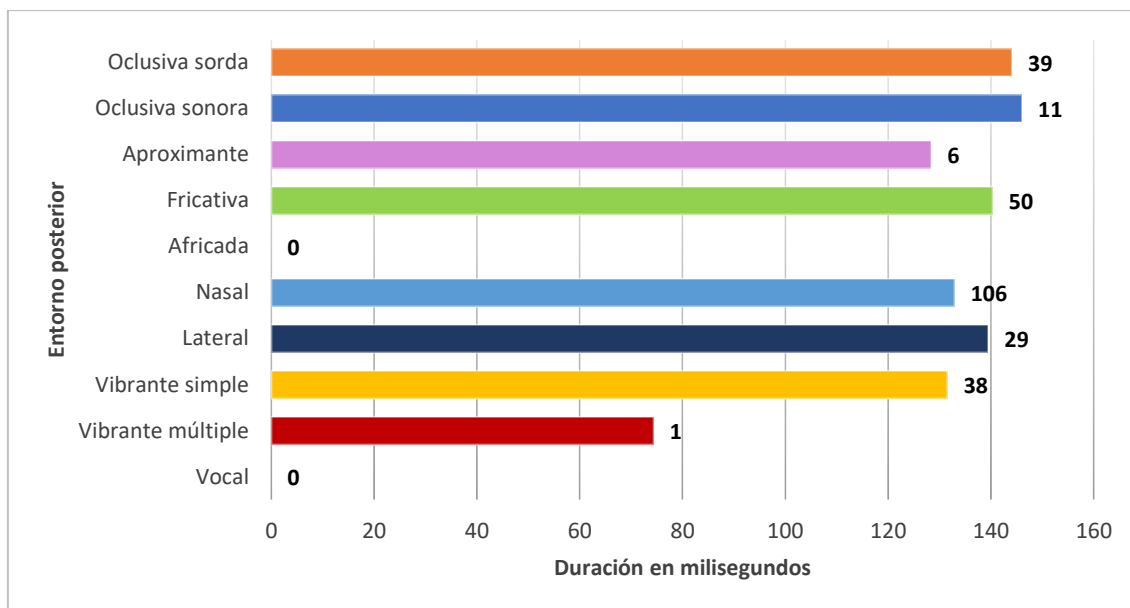


Figura 53. Duración media de la vocal /o/ producida por las informantes sinohablantes según el entorno posterior realizado.

La duración de /o/ es más breve cuando el sonido posterior a esta vocal es una aproximante (128.4 ms.), una vibrante simple (131.5 ms.), o una nasal (132.9 ms.), y ligeramente más larga si la consonante que sigue a /o/ es una lateral (139.4 ms.), una fricativa (140.3 ms.), una oclusiva sorda (144.2 ms.) o una oclusiva sonora (146 ms.). En el único caso en el que se ha articulado vibrante múltiple en posición posterior a esta vocal, la duración de la misma es de 74.4 ms.

/o/			
Entorno posterior realizado por las informantes sinohablantes	Casos	Duración media	Desviación estándar
Oclusiva sorda	39	144.1	47.9
Oclusiva sonora	11	146.0	39.5
Aproximante	6	128.4	20.3
Fricativa	50	140.3	31.0
Africada	0	--	--
Nasal	106	132.9	39.8
Lateral	29	139.5	39.1
Vibrante simple	38	131.5	31.0
Vibrante múltiple	1	74.4	--
Vocal	0	--	--

Tabla 39. Duración media y desviación estándar de la vocal /o/ producida por las informantes sinohablantes según el entorno posterior.

La desviación estándar en la realización de /o/ mostrada en la tabla 39 indica que hay falta de homogeneidad en la realización de dicha vocal, ya que la desviación oscila entre 31 y 47 en aquellos casos representativos de un comportamiento generalizado.

### 2.3.2.5 Vocal /u/

Tal y como se puede observar en la figura 54, cuando el sonido posterior a /u/ es una oclusiva sonora, una fricativa, una africada, o una nasal, la duración media de la vocal oscila entre 100 y 120 ms., mientras que la vocal es más larga si precede a una oclusiva sorda, aproximante, lateral o vibrante simple. Cabe señalar que no hay casos en los que las informantes realicen una vibrante múltiple o una vocal tras /u/, y que el número de realizaciones de oclusiva sonora o africada tras este sonido vocálico es escaso, por lo que únicamente se puede hablar de una tendencia en estos casos, y no de un comportamiento generalizado.

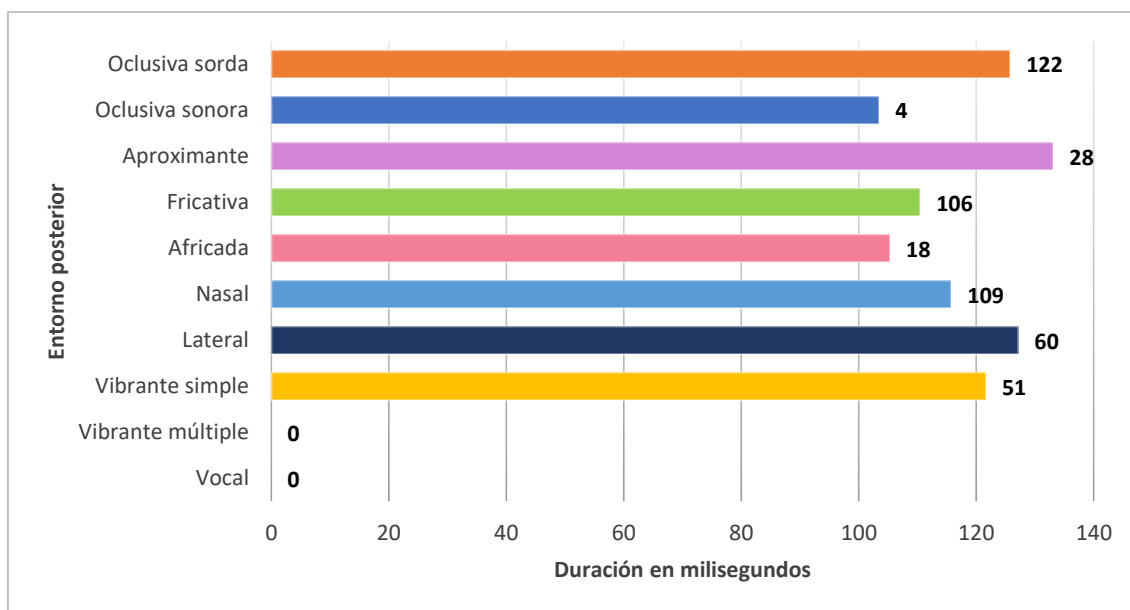


Figura 54. Duración media de la vocal /u/ producida por las informantes sinohablantes según el entorno posterior.

Cuando la vocal /u/ es seguida por una oclusiva sonora, la duración media es de 103.4 ms., la cual es muy similar a la correspondiente a la africada (105.3 ms.). Si el sonido posterior es una consonante fricativa o una nasal, la vocal es ligeramente más larga (110.4 y 115.7 ms. respectivamente). Si /u/ precede a una vibrante simple, una oclusiva sorda o una lateral, la duración media es muy similar en estos casos, ya que es de 121.6 ms. en el caso de la primera, de 125.7 ms. en el caso de la segunda, y de 127.1 ms. en el caso de la

tercera. Como se ha indicado, esta vocal es más larga cuando el sonido posterior es una aproximante, y es de 133.1 ms.

<b>/u/</b>			
<b>Entorno posterior realizado por las informantes sinohablantes</b>	<b>Casos</b>	<b>Duración media</b>	<b>Desviación estándar</b>
Oclusiva sorda	122	125.7	38.4
Oclusiva sonora	4	103.4	31.0
Aproximante	28	133.1	31.3
Fricativa	106	110.4	39.7
Africada	18	105.3	29.9
Nasal	109	115.7	34.6
Lateral	60	127.1	41.2
Vibrante simple	51	121.6	34.0
Vibrante múltiple	0	--	--
Vocal	0	--	--

*Tabla 40.* Duración media y desviación estándar de la vocal /u/ producida por las informantes sinohablantes según el entorno posterior.

En relación con la homogeneidad en la duración de /u/ articulada por las informantes sinohablantes, esta oscila entre 34 y 41, por lo que, de la misma manera que sucede con las otras vocales anteriormente analizadas, hay falta de homogeneidad en la realización de la vocal /u/ por parte de estas hablantes.

### **2.3.3 Duración de la vocal en posición final de palabra**

En este apartado se analiza la duración de las vocales del español producidas por las informantes sinohablantes cuando dichas vocales se encuentran en posición final de palabra. Por ejemplo, se analiza la duración de la vocal /e/ en la palabra *torre* o la duración de la vocal /o/ en la palabra *barco*.

Tal y como se puede observar en la figura 55, /a/ es la vocal más breve cuando se encuentra en posición final de palabra aislada, aunque la duración media de esta es similar a la de /o/. /i/ y /e/ son más largas, aunque la duración de las mismas también es similar entre ellas, y destaca que la /u/ es la vocal más larga cuando se encuentra en final de palabra. Sin embargo, cabe señalar el número de casos en los que /u/ aparece en esta posición es reducido, por lo que esta duración muestra una tendencia, no un comportamiento generalizado.

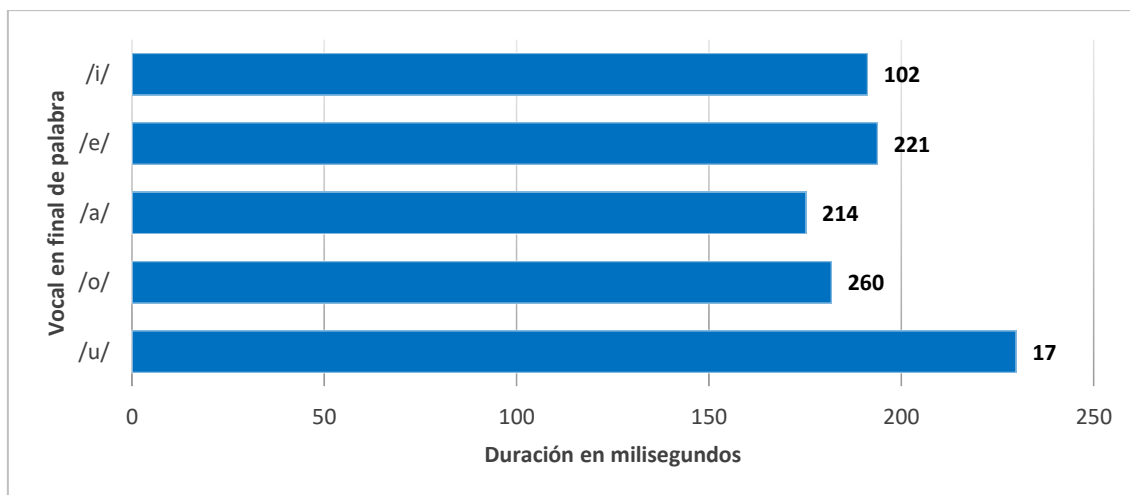


Figura 55. Duración media de las vocales producidas por las informantes sinohablantes cuando estas se encuentran en posición final de palabra.

Como se ha indicado, /a/ y /o/ son las vocales más breves en esta posición. La primera tiene una duración media de 175.2 ms., y la segunda de 181.8 ms. /i/ y /e/ son más largas que las anteriores, siendo la duración de las mismas de 191.2 y 193.8 ms. respectivamente. En cuanto a /u/, la cual es la más larga de todas, tiene una duración media de 229.8 ms.

Vocal	Vocal situada en posición final de palabra		
	Casos	Duración media	Desviación estándar
/i/	102	191.2	62.5
/e/	221	193.8	66.3
/a/	214	175.2	52.7
/o/	260	181.8	62.7
/u/	17	229.8	39.7

Tabla 41. Duración media y desviación estándar de las vocales producidas por las informantes sinohablantes cuando se encuentran en posición final de palabra.

Tal y como se ha indicado, la vocal /u/ se ha realizado en posición final únicamente en 17 casos, y la desviación estándar es de 39.7. En cambio, las otras vocales se han realizado en número de ocasiones considerablemente mayor, y la desviación oscila entre 52 y 62, lo que indica que hay falta de homogeneidad en la realización de las vocales del español en posición final por parte de las informantes sinohablantes, aunque dicha desviación es muy similar en todos los casos.

### **2.3.4 Recapitulación**

**El análisis de la duración de las vocales producidas por las informantes sinohablantes según el entorno ha mostrado que la duración media de estos sonidos vocálicos oscila entre 110 y 228 ms. Generalmente, las realizaciones son más breves cuando el sonido anterior a la vocal es una oclusiva sorda, oclusiva sonora, fricativa o africada, y más largas cuando dicho sonido es una nasal o una lateral, y las realizaciones más largas corresponden a las de vocal precedida por vibrante simple, vibrante múltiple o aproximante.**

**En cuanto a la duración media de las vocales según el entorno posterior, esta oscila entre 74 y 210 ms. Se ha observado que las vocales son más breves cuando preceden a una fricativa, africada o nasal, más largas si dicho sonido es una lateral, una oclusiva sorda o una oclusiva sonora, y las realizaciones más largas corresponden a los casos en los que el sonido posterior a la vocal es una aproximante, vibrante simple o vibrante múltiple.**

**El análisis de la duración de las vocales que se encuentran en posición final de palabra aislada ha permitido comprobar que la duración media oscila entre 175 y 229 ms. La vocal más breve es la /a/, seguida por /o/, /i/ y /e/, y la más larga es la /u/.**

### **2.4 Estructura silábica**

Como se ha indicado anteriormente, una de las variables que se han tenido en cuenta al analizar la duración de las vocales es la estructura silábica. Igual que en el caso del grupo de control, se ha analizado dicha duración en función de diferentes aspectos: la estructura de la sílaba en la que se encuentra la vocal analizada, si la vocal se encuentra en sílaba inicial, la estructura de la sílaba anterior a la de la vocal estudiada, la de la sílaba posterior a la de dicho sonido vocálico y si la vocal analizada se encuentra en sílaba final. Si bien el corpus de la presente investigación está compuesto por palabras cuya estructura silábica es CV-CV, CV-CVC o CVC-CV, tal y como se ha podido comprobar en la primera parte del presente capítulo, las informantes sinohablantes no siempre realizan dichas estructuras como tal, sino que hay una resilabación en la palabra, por lo que se ha analizado la duración de las vocales según las distintas estructuras realizadas por las informantes sinohablantes.

### 2.4.1 Duración de la vocal según la sílaba en la que se encuentra

En este apartado se muestran los resultados de la duración de las vocales producidas por el grupo de informantes sinohablantes según la estructura de la sílaba en la que se encuentra la vocal analizada. Es decir, si la vocal analizada es /a/ en la palabra *total*, se indica que la estructura de la sílaba es CVC, mientras que, si el sonido vocálico analizado es /o/, la estructura es CV.

#### 2.4.1.1 Vocal /i/

La figura 56 muestra que las informantes sinohablantes han realizado las estructuras CV o CVC en 306 y 155 ocasiones respectivamente. Cuando /i/ se encuentra en sílaba cuya estructura es CV, la duración de la misma es superior a cuando la estructura de la sílaba es CVC. En cuanto a las otras estructuras realizadas por las informantes cuya lengua materna es el chino, se han detectado 4 tipos diferentes: CCV (pronunciación de *firma* como ['fri.ma]), CCVC (*ritmo* como ['trit.mo]), VV y VC. En estos casos destaca la mayor duración cuando /i/ se encuentra en una estructura la cual la informante ha realizado como VV, y en las otras estructuras la duración de /i/ es similar a cuando se encuentra en estructura CVC.

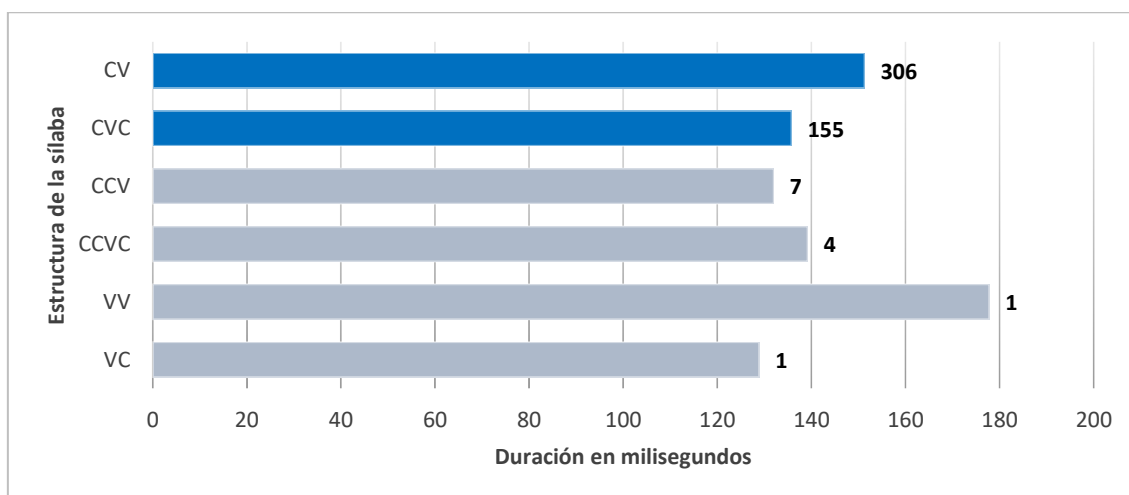


Figura 56. Duración de la vocal /i/ producida por las informantes sinohablantes según la sílaba en la que se encuentra.

La tabla 42 muestra que la duración media de /i/ cuando se encuentra en sílaba CV es de 151.3 ms., y de 135.7 ms. cuando la estructura de la sílaba es CVC. Cuando dicha vocal se encuentra en las estructuras CCV, CCVC o VC la duración media es muy similar (132, 139.2 y 128.9 ms. respectivamente), mientras que cuando la estructura producida es VV la duración asciende a 177.8 ms. Sin embargo, dado que el número de veces en que las

informantes han realizado estas estructuras es muy reducido, estos datos no son indicadores del comportamiento general, si no que muestran una tendencia en la realización de /i/ cuando las informantes sinohablantes producen estas estructuras.

/i/			
Estructura de la sílaba en la que se encuentra la vocal	Casos	Duración media	Desviación estándar
CV	306	151.3	62.9
CVC	155	135.7	43.1
CCV	7	132.0	58.4
CCVC	4	139.2	51.8
VV	1	177.8	--
VC	1	128.9	--

Tabla 42. Duración media y desviación estándar de la vocal /i/ producida por las informantes sinohablantes según la sílaba en la que se encuentra.

En cuanto a la desviación estándar, se observa que esta es elevada (entre 43.1 y 62.9), y que es superior cuando la estructura es CV que cuando es CVC, lo que indica que hay poca homogeneidad en la duración de /i/, y que es incluso cuando las informantes realizan la estructura como CV. En cuanto a CCV y CCVC, la desviación estándar es similar, y bastante elevada también. Para las estructuras VV y VC no se ha podido calcular la desviación estándar dado que únicamente se ha producido cada estructura en una ocasión.

#### 2.4.1.2 Vocal /e/

La figura 57 muestra que las informantes sinohablantes, además de realizar la vocal /e/ en las estructuras CV y CVC, en 359 y 146 ocasiones respectivamente, han realizado 3 tipos diferentes más: CCV (*recto* como ['tre.to]), CCVC (*resto* como ['tres.to]) y CVVC (*Miguel* como [mi.'yuel]). De la misma manera que sucede con /i/, cuando /e/ se encuentra en CV la duración de la misma es superior que cuando se encuentra en sílaba cuya estructura es CVC. En cuanto a las otras estructuras realizadas por las informantes sinohablantes, la duración de /e/ es considerablemente superior cuando se encuentra en sílaba CCV. Lo mismo sucede cuando la estructura es CVVC. En cambio, la duración es menor cuando la estructura realizada es CCVC. Sin embargo, igual que en el caso de /i/, estos datos únicamente permiten ver una serie de tendencias en la duración de /e/ cuando las informantes producen las estructuras CCV, CCVC o CVVC, ya que el número de casos es reducido.

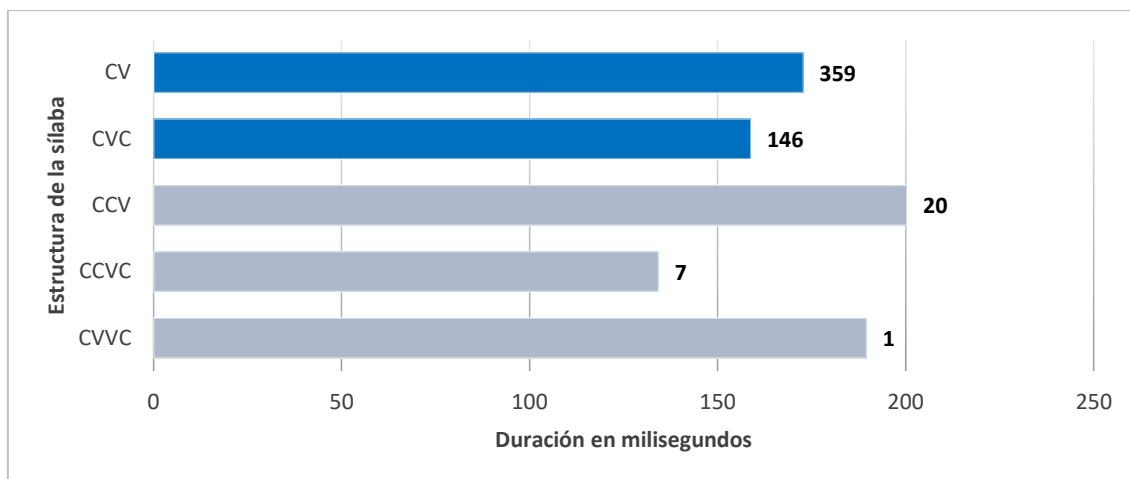


Figura 57. Duración de la vocal /e/ producida por las informantes sinohablantes según la sílaba en la que se encuentra.

La tabla 43 muestra detalladamente la duración media de la vocal /e/ y la desviación estándar de la misma. Cuando se encuentra en sílaba CV, la duración media es de 172.8 ms., y, como se ha dicho, es mayor que cuando la estructura de la sílaba es CVC, en cuyo caso la duración es de 158.8 ms. Las informantes sinohablantes han realizado la estructura CCV en 20 ocasiones, dando como resultado una duración media considerablemente superior a las anteriores, de 200.3 ms. En cuanto a las otras estructuras, en el caso de CCVC la duración media es de 134.2 ms., mientras que en el de CVVC es mayor, de 189.6 ms.

/e/			
Estructura de la sílaba en la que se encuentra la vocal	Casos	Duración media	Desviación estándar
CV	359	172.8	67.2
CVC	146	158.8	50.9
CCV	20	200.3	79.7
CCVC	7	134.2	30.0
CVVC	1	189.6	--

Tabla 43. Duración media y desviación estándar de la vocal /e/ producida por las informantes sinohablantes según la sílaba en la que se encuentra.

La desviación estándar en la duración de /e/ es similar a la de /i/. En las estructuras realizadas como CV y CVC, la desviación es de 67.2 y 50.9 respectivamente. En el caso de CCV, esta desviación aumenta a 79.7. Estos datos indican que la duración de la vocal /e/ producida por las informantes sinohablantes es poco homogénea.



### 2.4.1.3 Vocal /a/

Tal y como se observa en la figura 58, las informantes sinohablantes, además de realizar las estructuras silábicas CV y CVC en 381 y 146 casos respectivamente, han realizado las estructuras CCV (*jarra* como ['xa.dra]), CCVC (*currar* como [ku.'drar]) y CVV (*leña* como ['le.nia]) en 18, 5 y 2 ocasiones respectivamente. En este caso la duración media de /a/ cuando se encuentra en sílabas cuya estructura es CV o CVC es muy similar. Cuando la estructura de la sílaba producida por las informantes sinohablantes es CCV, la duración media es mayor que en los dos casos anteriores. En cuanto a CCVC, la duración de /a/ es similar a cuando se encuentra en CV o CVC, mientras que cuando se encuentra en sílaba CVV la duración es menor. Sin embargo, igual que en los casos de /i/ y /e/, las estructuras CCV, CCVC y CVV se han realizado en un número reducido de ocasiones, por lo que estos datos únicamente muestran una tendencia, y no un comportamiento generalizado.

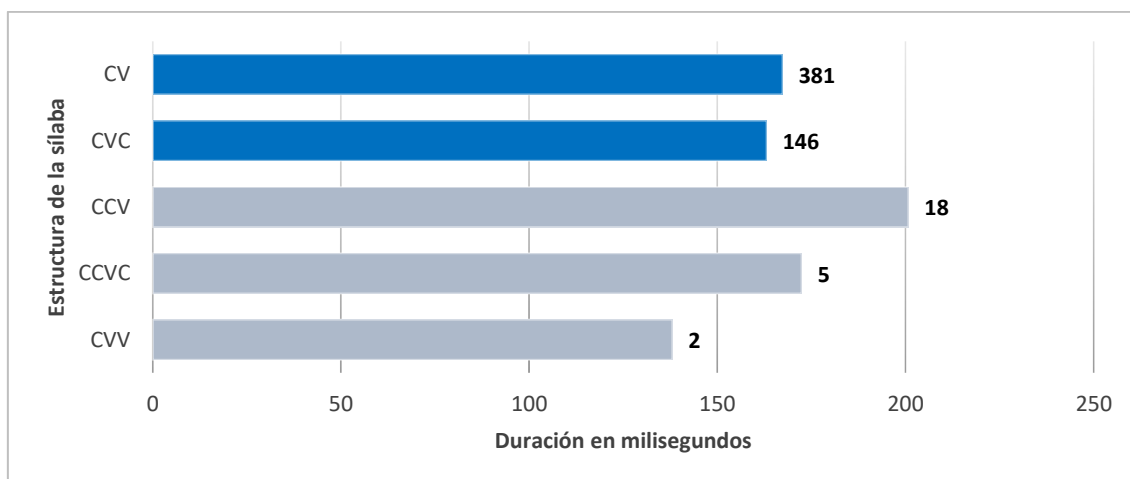


Figura 58. Duración de la vocal /a/ producida por las informantes sinohablantes según la sílaba en la que se encuentra.

La tabla 44 muestra detalladamente la duración media de la vocal /a/ y la desviación estándar de la misma. Tal y como se ha indicado, la media es muy similar cuando /a/ se encuentra en las estructuras CV (167.3 ms.) y CVC (163.1 ms.) En cuanto a la duración de esta vocal en estructura CCV, la duración media asciende a 200.7 ms., y cuando se trata de las estructuras CCVC y CVV, las medias son de 172.3 y 138 ms. respectivamente.

<b>/a/</b>			
<b>Estructura de la sílaba en la que se encuentra la vocal</b>	<b>Casos</b>	<b>Duración media</b>	<b>Desviación estándar</b>
CV	381	167.3	53.7
CVC	146	163.1	44.4
CCV	18	200.7	43.0
CCVC	5	172.3	43.2
CVV	2	138.0	63.0

Tabla 44. Duración media y desviación estándar de la vocal /a/ producida por las informantes sinohablantes según la sílaba en la que se encuentra.

En lo que a la homogeneidad en la duración de las vocales se refiere, la desviación estándar es de 53.7 y 44.4 en CV y CVC, y en CCV es muy similar a CVC, de 43, lo que indica que la duración de la vocal /a/ producida por las informantes sinohablantes es ligeramente más homogénea que en los casos de /i/ y /e/.

#### 2.4.1.4 Vocal /o/

El análisis de la duración de la vocal /o/ en función de la estructura de la sílaba en la que se encuentra muestra que las informantes sinohablantes, además de realizar las estructuras CV y CVC en 390 y 129 ocasiones respectivamente, han realizado la estructura CCV (*robar* como [tro.'pa]) en 13 casos, y CCVC (*ronca* como ['tron.ka]) en 8. La figura 59 muestra que la duración de /o/ es considerablemente superior cuando se encuentra en sílaba CV que en CVC.

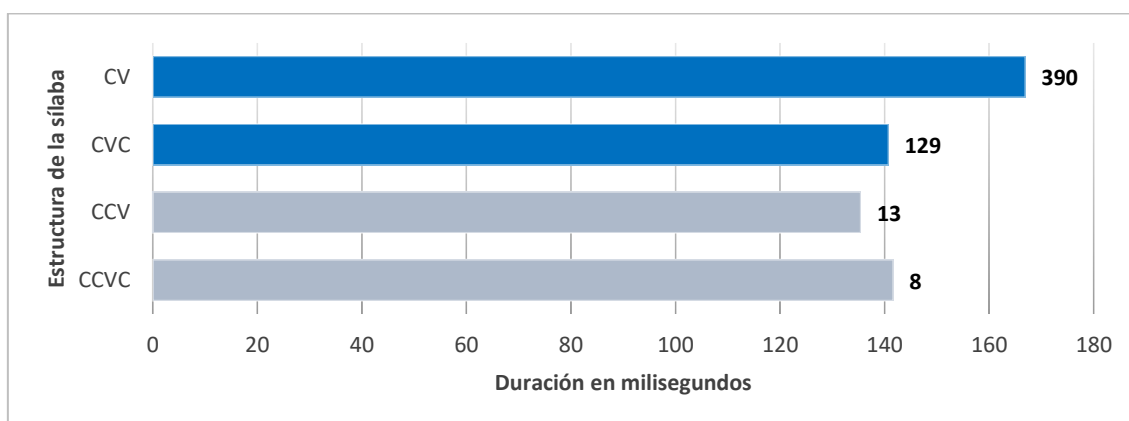


Figura 59. Duración de la vocal /o/ producida por las informantes sinohablantes según la sílaba en la que se encuentra.

La tabla 45 muestra que la duración media de /o/ cuando se encuentra en sílaba cuya estructura es CV es de 166.9 ms., mientras que cuando se encuentra en CVC la media desciende a 140.7 ms. En cuanto a CCV, la duración media es menor, de 135.4 ms., y de 141.6 ms. cuando /o/ se encuentra en CCVC. Sin embargo, el reducido número de

ocasiones en las que se realizan estas dos últimas estructuras únicamente permite ver una tendencia, pero no el comportamiento general de /o/ en dichas estructuras.

<b>/o/</b>			
<b>Estructura de la sílaba en la que se encuentra la vocal</b>	<b>Casos</b>	<b>Duración media</b>	<b>Desviación estándar</b>
CV	390	166.9	60.2
CVC	129	140.7	39.5
CCV	13	135.4	39.8
CCVC	8	141.6	40.9

*Tabla 45.* Duración media y desviación estándar de la vocal /o/ producida por las informantes sinohablantes según la sílaba en la que se encuentra.

En cuanto a la desviación estándar, se observa que cuando la estructura de la sílaba en la que se encuentra /o/ es CV, hay menor homogeneidad, ya que la desviación es de 60.2, mientras que cuando la estructura es CVC, la homogeneidad es ligeramente superior, ya que la desviación estándar es menor (39.5). En los casos de las estructuras CCV y CCVC, la desviación es prácticamente igual que la de CVC. Estos datos muestran que, igual que sucede con las vocales anteriormente analizadas, la duración de /o/ no presenta gran homogeneidad cuando es realizada por las informantes sinohablantes.

#### **2.4.1.5 Vocal /u/**

El análisis de la duración de la vocal /u/ según la estructura de la sílaba en la que se encuentra muestra que las informantes sinohablantes han realizado hasta 6 tipos de estructura silábica diferentes. En este caso, además de realizar las estructuras CV y CVC en 300 y 165 ocasiones respectivamente, las informantes sinohablantes han realizado la estructura CCV (*turnó* como [tru.'no]) en 40 casos, VC (*busca* como ['us.ka]) en 7, V (*jugó* como [u.'go]) en 2, y CVCC (*cursé* como ['kurs]) en 1 ocasión. Igual que con la vocal /o/, cuando /u/ se encuentra en sílaba CV la duración es considerablemente mayor que cuando se encuentra en sílaba CVC, y es incluso mayor cuando /u/ se encuentra en sílaba CCV. La tendencia que se observa cuando la estructura es VC o V es que la vocal sea más breve, mientras que si la estructura es CVCC, la vocal es considerablemente mayor.

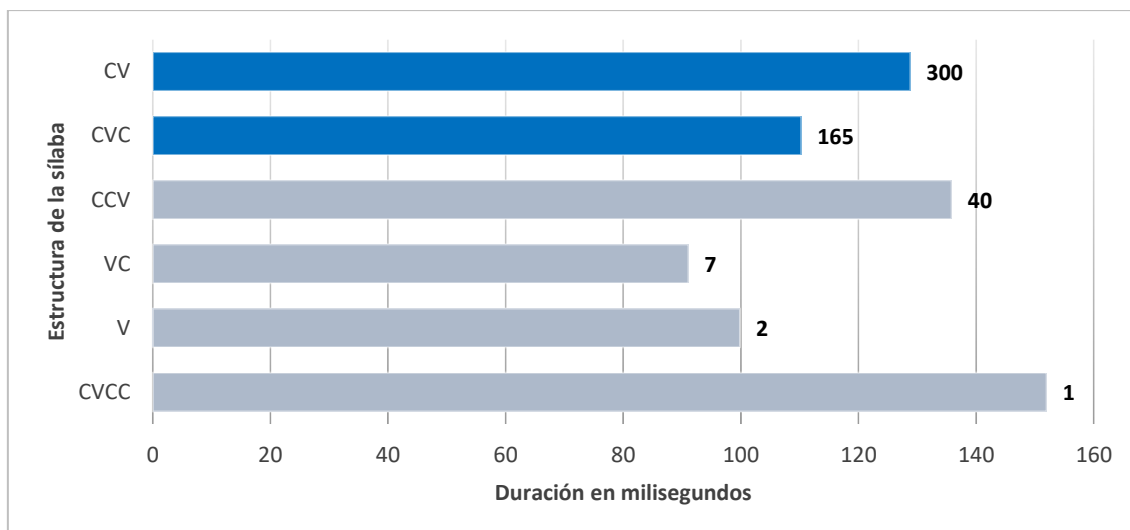


Figura 60. Duración de la vocal /u/ producida por las informantes sinohablantes según la sílaba en la que se encuentra.

La tabla 46 muestra que la duración media de /u/ cuando se encuentra en sílaba CCV es de 135.8 ms., y disminuye cuando la sílaba es CV (128.8 ms.) o CVC (110,2 ms.). En cuanto a las otras estructuras, la duración es de 91 y 99.8 ms. en las estructuras VC y V respectivamente, y asciende a 152 ms. en la estructura CVCC.

/u/			
Estructura de la sílaba en la que se encuentra la vocal	Casos	Duración media	Desviación estándar
CV	300	128.8	42.7
CVC	165	110.2	36.7
CCV	40	135.8	44.3
VC	7	91.0	12.5
V	2	99.8	17.2
CVCC	1	152.0	--

Tabla 46. Duración media y desviación estándar de la vocal /u/ según la sílaba en la que se encuentra.

Se observa que la desviación estándar de la duración de la /u/ es la menor de todas las vocales, lo que indica que la duración de la vocal /u/ producida por las informantes sinohablantes es más homogénea que el resto de vocales.

#### 2.4.2 Duración de la vocal en sílaba inicial

En este apartado se analiza la duración de las vocales producidas por el grupo de informantes sinohablantes cuando dichos sonidos vocálicos se encuentran en sílaba inicial de palabra, es decir, cuando no hay otra sílaba que las preceda. Por ejemplo, si se analiza la vocal /o/ de la palabra *total*, esta se encuentra en sílaba inicial.

Tal y como se observa en la figura 61, la duración de las vocales de más breves a más largas es /u/, /i/, /o/, /e/, /a/.

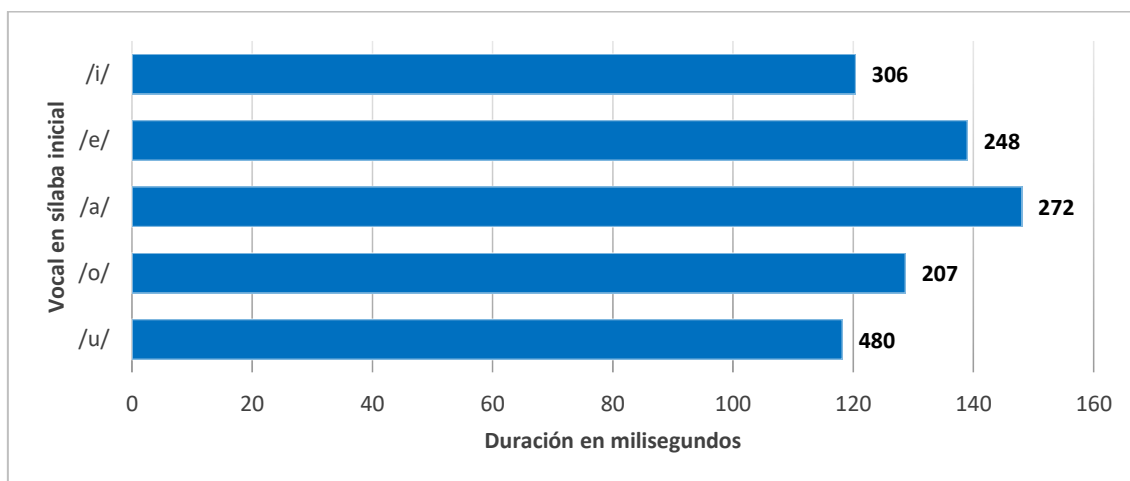


Figura 61. Duración de las vocales producidas por las informantes sinohablantes cuando se encuentran en sílaba inicial.

Tal y como se puede observar en la tabla 47, la duración media de las vocales /u/ e /i/ es muy similar, ya que en el primer caso es de 118.2 ms., y en el segundo es de 120.4 ms. La duración es ligeramente superior cuando la vocal que se encuentra en sílaba inicial es /o/ (128.7 ms.), mientras que las vocales /e/ y /a/ presentan medias superiores, de 139 y 148 ms. respectivamente.

Análisis de la vocal cuando se encuentra en sílaba inicial			
Vocal	Casos	Duración media	Desviación estándar
/i/	306	120.4	35.7
/e/	248	139.0	48.7
/a/	272	148.0	38.0
/o/	207	128.7	35.2
/u/	480	118.2	36.6

Tabla 47. Duración media y desviación estándar de las vocales producidas por las informantes sinohablantes cuando se encuentran en sílaba inicial de palabra.

En cuanto a la desviación estándar, esta oscila entre 35.2 y 48.7, siendo la /e/ la de mayor desviación, lo que indica que hay falta de homogeneidad en la realización de las distintas vocales del español por parte de las informantes sinohablantes.

#### 2.4.3 Duración de la vocal según la sílaba anterior

En el presente apartado se analiza la duración de las vocales del español producidas por las informantes sinohablantes según la estructura de la sílaba anterior a la sílaba en la que se encuentra la vocal. Por ejemplo, si se analiza el sonido vocálico /a/ de

la palabra *total*, la estructura de la sílaba anterior es CV, y si se analiza la misma vocal, pero en esta ocasión en la palabra *lista*, la estructura anterior es CVC.

### 2.4.3.1 Vocal /i/

La figura 62 muestra que las informantes sinohablantes, además de realizar las estructuras CV y CVC, las propias del corpus, han realizado la estructura CCV (*rugid* como [tru.'xið]). En el caso de las primeras, la duración media es muy similar. Cuando la estructura de la sílaba anterior es CCV la duración media es ligeramente superior a las dos estructuras anteriores. Sin embargo, el número de casos en los que las informantes sinohablantes han realizado estas dos últimas estructuras es reducido, por lo que los datos únicamente muestran una tendencia.

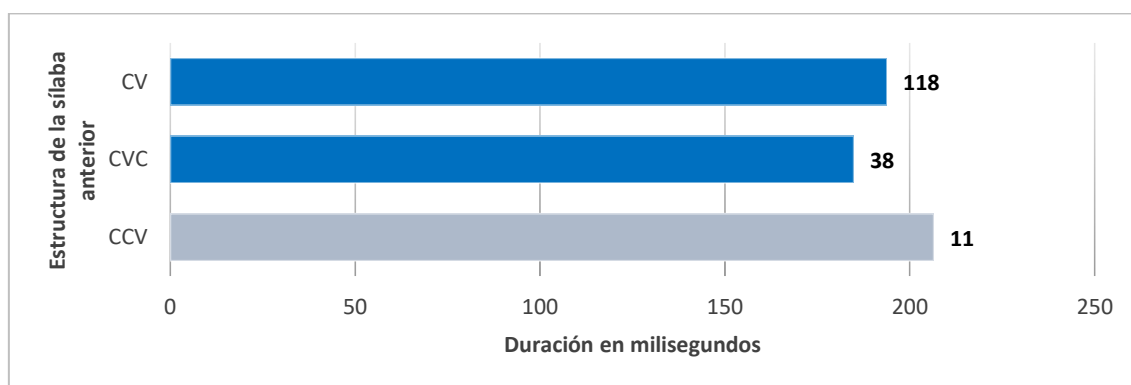


Figura 62. Duración de la vocal /i/ producida por las informantes sinohablantes según la estructura de la sílaba anterior.

Tal y como se ha indicado, cuando la sílaba anterior tiene estructura CV o CVC la duración de /a/ es similar, ya que en el primer caso la media es de 193.8 ms., y en el segundo de 184.8 ms. En cuanto a las otras estructuras realizadas, cuando la estructura de la sílaba anterior es realizada como CCV, la duración media asciende a 206.5 ms.

/i/			
Estructura de la sílaba anterior a la de la vocal	Casos	Duración media	Desviación estándar
CV	118	193.8	61.7
CVC	38	184.8	54.9
CCV	11	206.5	57.6

Tabla 48. Duración media y desviación estándar de la vocal /i/ producida por las informantes sinohablantes según la sílaba anterior.

En cuanto a la desviación estándar, esta oscila entre 54.9 y 61.7, por lo que, a pesar de que es similar en todos los casos, es elevada, lo que indica que hay falta de homogeneidad en la duración de la /i/.

### 2.4.3.2 Vocal /e/

Las informantes sinohablantes han producido únicamente las estructuras CV y CVC en sílaba anterior a la de /e/. En este caso, la duración de la vocal es mayor cuando la sílaba anterior es CV que cuando es CVC.

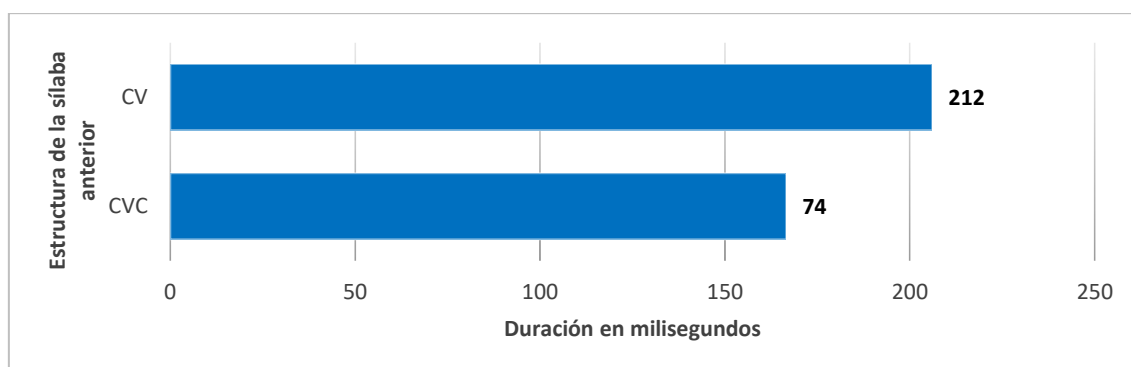


Figura 63. Duración de la vocal /e/ según la estructura de la sílaba anterior.

La tabla 49 muestra que, como se ha indicado, la duración media de /e/ según la sílaba anterior es considerablemente superior cuando la estructura anterior es CV, ya que la media es de 206 segundos, mientras que cuando es CVC, la media desciende a 166.5 ms.

/e/			
Estructura de la sílaba anterior a la de la vocal	Casos	Duración media	Desviación estándar
CV	212	206.0	63.7
CVC	74	166.5	54.4

Tabla 49. Duración media y desviación estándar de la vocal /e/ producida por las informantes sinohablantes según la sílaba anterior.

En lo que se refiere a la desviación estándar, igual que sucede con la /i/, esta es elevada, lo que indica falta de homogeneidad en la realización de la vocal /e/.

### 2.4.3.3 Vocal /a/

La figura 64 muestra que, además de realizar las estructuras CV y CVC en sílaba anterior, las informantes sinohablantes han producido las estructuras CCV (*risas* como ['tri.sas]), V (*Lucas* como ['u.kas]) y CCVC (*ronca* como ['tron.ka]). Se observa que la duración de /a/ es mayor cuando la sílaba anterior es CV que cuando es CVC. En cuanto a las estructuras CCV y V, la duración de la vocal es similar a cuando la estructura anterior es CVC, y ligeramente inferior cuando la estructura es CCVC.

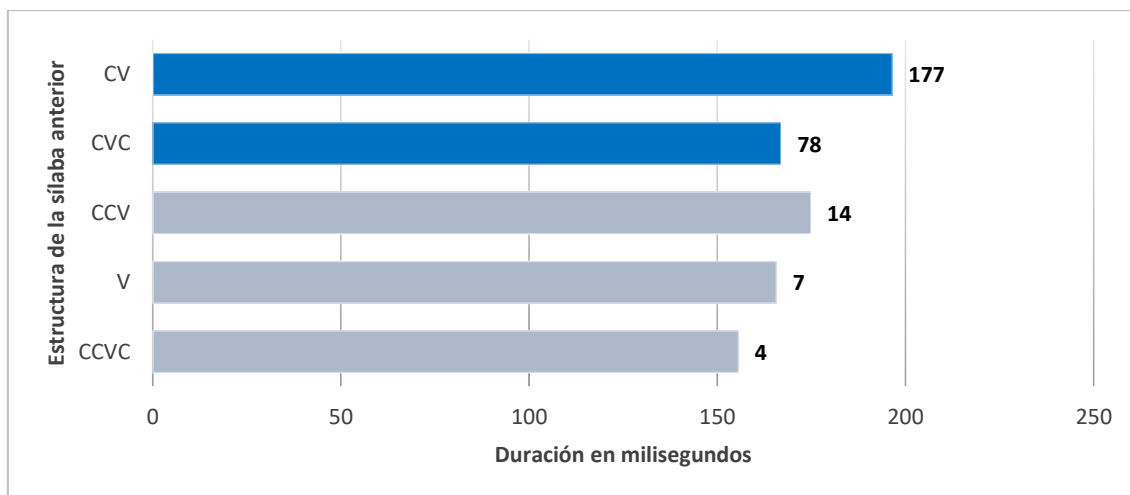


Figura 64. Duración de la vocal /a/ producida por las informantes sinohablantes según la estructura de la sílaba anterior.

Como se ha indicado, la duración de /a/ es superior cuando la estructura de la sílaba anterior es CV (196.5 ms.) que cuando es CVC (166.9 ms.). Cuando la estructura realizada por las informantes es CCV, la duración media es de 174.8 ms., y desciende cuando la estructura es V (165.7 ms.) o CCVC (155.6 ms.).

/a/			
Estructura de la sílaba anterior a la de la vocal	Casos	Duración media	Desviación estándar
CV	177	196.5	56.5
CVC	78	166.9	49.7
CCV	14	174.8	49.2
V	7	165.7	48.8
CCVC	4	155.6	35.8

Tabla 50. Duración media y desviación estándar de la vocal /a/ producida por las informantes sinohablantes según la sílaba anterior.

En cuanto a la desviación estándar, igual que en los casos anteriores, es considerablemente alta, lo que muestra que falta homogeneidad en la duración de la vocal /a/.

#### 2.4.3.4 Vocal /o/

Tal y como se puede observar en la figura 65, las informantes sinohablantes han realizado las estructuras CV y CVC en 221 y 94 ocasiones respectivamente en sílaba anterior a la de la vocal /o/. Además, han realizado la estructura CCV (*rato* como ['tra.to]) en 15 ocasiones, VC (*busto* como ['us.to]) en dos casos y V (*jugó* como [u.'yo]) en un caso. La duración de /o/ es superior cuando la estructura anterior se ha realizado como



CV que cuando se ha realizado como CVC. Si las informantes sinohablantes han realizado la sílaba anterior como CCV, la duración media es muy similar a la de CV. En cuanto a las otras dos estructuras, la duración media es superior a la de las otras. No obstante, cabe señalar que estos datos muestran una tendencia, ya que son pocos los casos en los que las informantes han realizado dicha estructura.

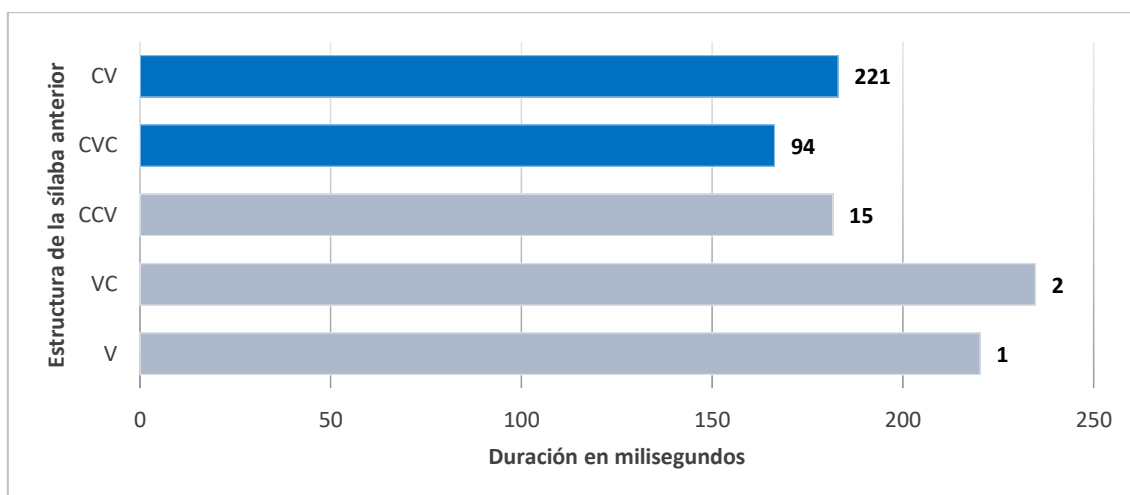


Figura 65. Duración de la vocal /o/ producida por las informantes sinohablantes según la estructura de la sílaba anterior.

De acuerdo con los datos mostrados en la tabla 51, la duración del sonido vocálico /o/ producido por las informantes sinohablantes es mayor cuando la estructura de la sílaba anterior es CV (183 ms.) que cuando es CVC (166,4 ms.). Como se ha indicado, la duración media es muy similar cuando la sílaba anterior es CCV (181.6 ms.) que cuando es CV. En cuanto a las estructuras VC y V, la duración media asciende a 234.8 y 220.3 ms. respectivamente.

<b>/o/</b>			
<b>Estructura de la sílaba anterior a la de la vocal</b>	<b>Casos</b>	<b>Duración media</b>	<b>Desviación estándar</b>
CV	221	183.0	59.2
CVC	94	166.4	54.8
CCV	15	181.6	63.0
VC	2	234.8	112.7
V	1	220.3	--

Tabla 51. Duración media y desviación estándar de la vocal /o/ producida por las informantes sinohablantes según la sílaba anterior.

En cuanto a la desviación estándar, aunque esta es muy elevada cuando la estructura realizada es VC, dado que únicamente hay dos casos de esta realización, no ha de considerarse como un comportamiento generalizado. En el resto de casos, la desviación

estándar es elevada (entre 54.8 y 63), lo que indica una falta de homogeneidad en la duración de la vocal /o/ realizada por las hablantes cuya lengua materna es el chino.

#### 2.4.3.5 Vocal /u/

Tal y como se puede observar en la figura 66, las informantes sinohablantes únicamente han realizado las estructuras que aparecen en el corpus en sílaba anterior, aunque el número de casos es muy reducido, ya que /u/ precedida por CV se ha realizado en 34 casos y por CVC únicamente en 2. El reducido número de datos al respecto únicamente proporciona una posible tendencia en la realización, pero no caracteriza el comportamiento general.

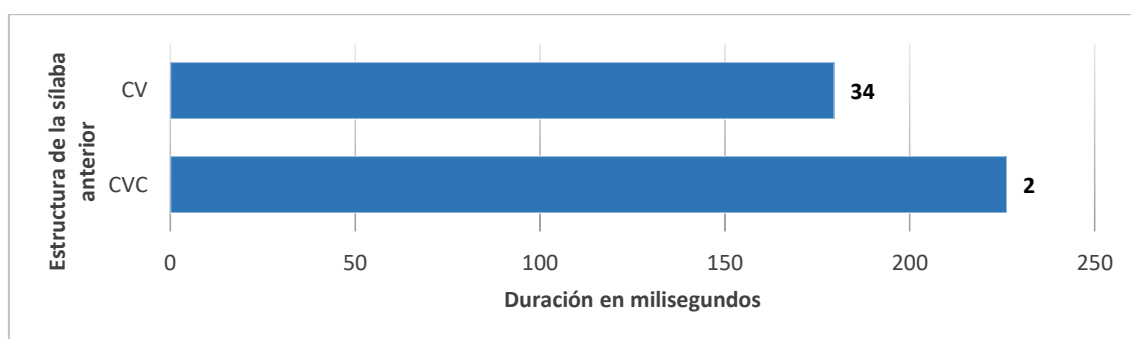


Figura 66. Duración de la vocal /u/ producida por las informantes sinohablantes según la estructura de la sílaba anterior.

Se observa que cuando la sílaba anterior tiene estructura CVC la vocal es más larga que cuando la estructura es CV, ya que en el caso de la primera la media es de 226.3 ms., mientras que en el de la segunda es de 179.6 ms.

/u/			
Estructura de la sílaba anterior a la de la vocal	Casos	Duración media	Desviación estándar
CV	34	179.6	58.0
CVC	2	226.3	11.3

Tabla 52. Duración media y desviación estándar de la vocal /u/ producida por las informantes sinohablantes según la sílaba anterior.

En cuanto a la desviación estándar, en el caso de la estructura CV es de 58, lo que muestra que no hay homogeneidad en la realización de /u/, mientras que en el caso de CVC la desviación es mucho menor, de 11.3, pero cabe señalar nuevamente que el número de casos es muy reducido.

#### 2.4.4 Duración de la vocal según la sílaba posterior

En el presente apartado se analiza la duración de las vocales del español producidas por las informantes sinohablantes según la estructura de la sílaba posterior a la sílaba en la que se encuentra la vocal. Por ejemplo, si se analiza el sonido vocálico /o/ de la palabra *total*, la estructura de la sílaba posterior es CVC, y si se analiza la vocal /i/ en la palabra *lista*, la estructura posterior es CV.

##### 2.4.4.1 Vocal /i/

La figura 67 muestra que las informantes sinohablantes han realizado las estructuras CV y CVC en sílaba posterior a la de la /i/ en 247 y 56 casos respectivamente. Además, han realizado en una ocasión la estructura CVVC (*Miguel* como [mi. 'ɣuel]) y en otra la estructura VV (*chillar* como [tʃi. 'ia]). Se observa que la duración de /i/ es similar tanto si la sílaba posterior tiene estructura CV como si tiene estructura CVC. Cuando realizan las otras dos estructuras mencionadas, la duración es menor. Sin embargo, cabe señalar que estos casos aislados no permiten conocer el comportamiento general.

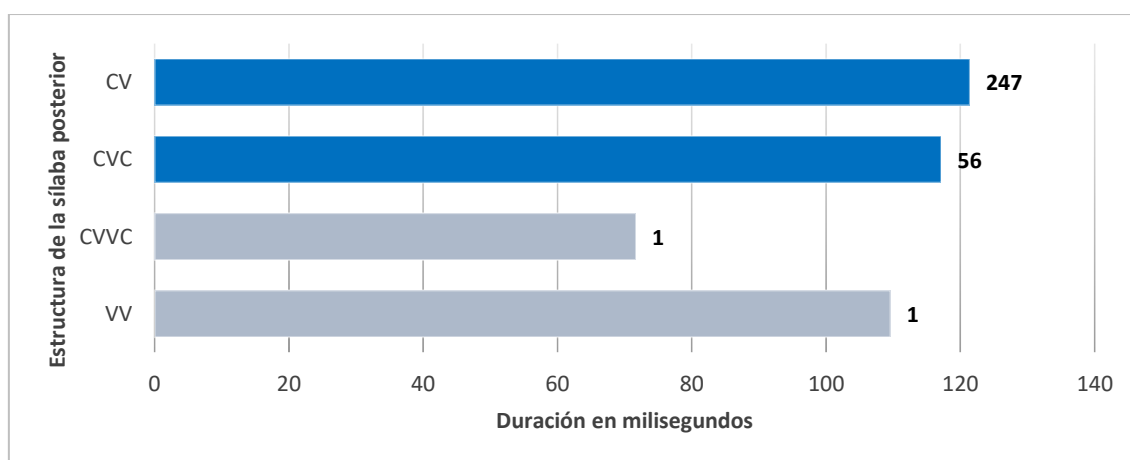


Figura 67. Duración de la vocal /i/ producida por las informantes sinohablantes según la sílaba posterior.

Tal y como se ha indicado, la duración media cuando la sílaba posterior es CV o CVC es muy similar, ya que en el caso de la primera es de 121.4 ms., y en el de la segunda de 117.1 ms.

<i>/i/</i>			
Estructura de la sílaba posterior a la de la vocal	Casos	Duración media	Desviación estándar
CV	247	121.4	35.0
CVC	56	117.1	38.6
CVVC	1	71.7	--
VV	1	109.6	--

Tabla 53. Duración media y desviación estándar de la vocal /i/ producida por las informantes sinohablantes según la sílaba posterior.

En cuanto a la desviación estándar, esta es menos elevada que cuando se analiza la duración de la vocal según el entorno anterior, lo que indica que si se analiza la duración del sonido vocálico /i/ según el entorno posterior, esta es más homogénea.

#### 2.4.4.2 Vocal /e/

En cuanto a la duración de /e/ según el entorno posterior, las informantes sinohablantes han producido las estructuras CV y CVC en 170 y 66 casos respectivamente. Además, han realizado las estructuras CCV (*cerrar* como [θe'dra]), CCVC (*cerrar* como [θe'draɪ]) y CVV (*leña* como ['le.nia]). Tal y como se puede observar, la duración de /e/ es mayor cuando la sílaba posterior tiene estructura CV que cuando dicha sílaba es CVC. Si la sílaba posterior se realiza como CCVC, la duración es muy similar a la de cuando la estructura es CV. Sin embargo, si la estructura es CVV, la duración es mayor que en cualquier de los otros casos, y si es CCV, es menor que cualquiera de los otros casos.

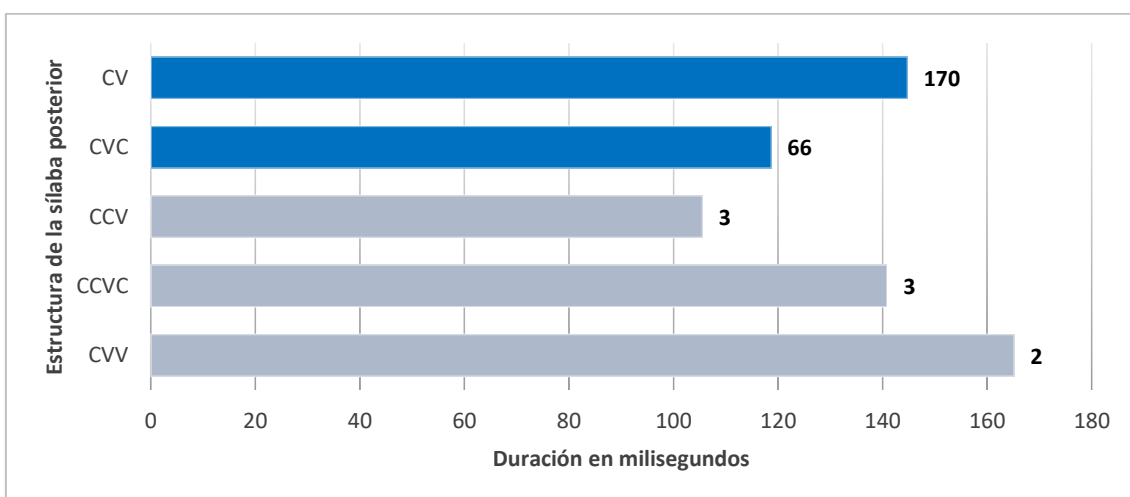


Figura 68. Duración de la vocal /e/ producida por las informantes sinohablantes según la sílaba posterior.

Como se ha indicado, la duración de /e/ es superior cuando la sílaba posterior es CV (144.7 ms.) que cuando es CVC (118.7 ms.). La duración media de la vocal /e/ cuando precede a las estructuras CCV, CCVC y CVV es de 105.5, 140.8 y 165.2 ms. respectivamente.

/e/			
Estructura de la sílaba posterior a la de la vocal	Casos	Duración media	Desviación estándar
CV	170	144.7	49.9
CVC	66	118.7	23.2
CCV	3	105.5	10.9
CCVC	3	140.8	41.8
CVV	2	165.2	70.4

Tabla 54. Duración media y desviación estándar de la vocal /e/ producida por las informantes sinohablantes según la sílaba posterior.

En cuanto a la desviación estándar, esta es más elevada cuando la sílaba posterior es CV (49.9) que cuando es CVC (23.2), lo que indica que hay mayor homogeneidad en la duración de /e/ cuando la sílaba posterior tiene estructura CVC.

#### 2.4.4.3 Vocal /a/

La figura 69 muestra que las informantes sinohablantes han realizado las estructuras CV y CVC en posición posterior a la sílaba en la que se encuentra la vocal /a/ en 190 y 64 casos respectivamente. También han realizado otras estructuras como CCV (*marrón* como [ma.'tro]), V (*saga* como ['sa.a]) o CCVC (*París* como [pa.'dris]). Igual que sucede con la vocal /i/, cuando la sílaba posterior a la de /a/ tiene estructura CV, la duración es mayor que cuando dicha estructura es CVC. En cuanto a las otras estructuras realizadas por las informantes sinohablantes, si la sílaba es CCV, la duración es similar a cuando es CV. En cambio, si la sílaba posterior es V, la duración es considerablemente superior, mientras que si es CCVC, la duración es ligeramente superior a CVC.

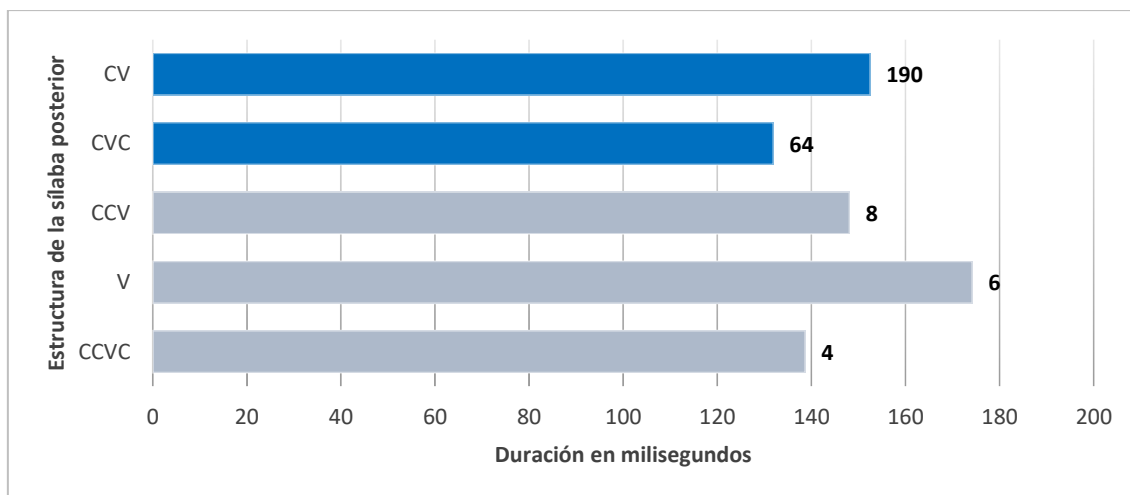


Figura 69. Duración de la vocal /a/ producida por las informantes sinohablantes según la sílaba posterior.

Tal y como muestra la tabla 55, la duración media de /a/ cuando la sílaba posterior tiene estructura CV es de 152.5 ms., y disminuye a 131.9 ms. si dicha sílaba tiene estructura CVC. Cuando /a/ precede a una sílaba cuya estructura es CCV, la duración media es de 148.1 ms., si a la sílaba que sigue a la de la vocal /a/ está formada por un único sonido vocálico (V), la duración media es considerablemente superior, de 174.2 ms., mientras que si dicha estructura es CCVC, la media es de 138.7 ms.

/a/			
Estructura de la sílaba posterior a la de la vocal	Casos	Duración media	Desviación estándar
CV	190	152.5	37.7
CVC	64	131.9	32.5
CCV	8	148.1	30.9
V	6	174.2	53.3
CCVC	4	138.7	53.9

Tabla 55. Duración media y desviación estándar de la vocal /a/ producida por las informantes sinohablantes según la sílaba posterior.

En cuanto a la desviación estándar, se observa que en las estructuras que aparecen en un mayor número de casos esta oscila entre 32.5 y 37.7, lo que indica que la homogeneidad en la realización de /a/ es similar tanto si la estructura de la sílaba posterior es CV como si es CVC. En los casos de V y CCVC la desviación es mayor, lo que indica que hay menor homogeneidad. Sin embargo, estos datos únicamente permiten observar cuál es la tendencia cuando la estructura es CCV, V o CCVC.

#### 2.4.4.4 Vocal /o/

Las informantes sinohablantes han realizado las estructuras CV y CVC en posición posterior a la sílaba en la que se encuentra /o/ en 128 y 65 ocasiones respectivamente. La duración de dicha vocal cuando precede a CV es mayor que cuando precede a CVC. Además, las informantes sinohablantes han realizado la estructura CCV (*torre* como ['to.dre]) en 8 casos, en los cuales la duración de /o/ es mayor que con las anteriores estructuras, y CCVC (*corre* como [ko.'dref]) en 3 casos, en los cuales la duración de /o/ es similar a cuando precede a la estructura CVC.

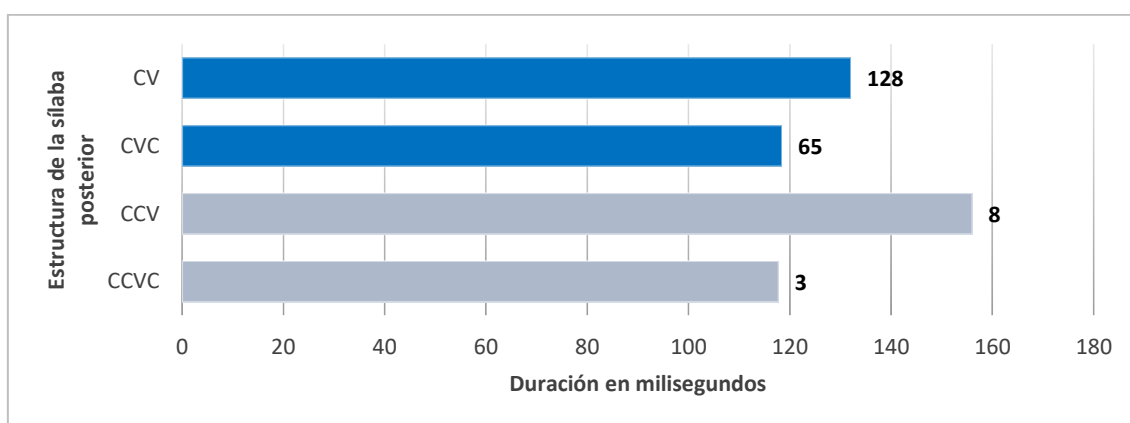


Figura 70. Duración de la vocal /o/ producida por las informantes sinohablantes según la sílaba posterior.

La duración media de /o/ cuando precede a la estructura CV es de 132 ms., y de 118.4 ms. cuando la sílaba posterior tiene estructura CVC. La media cuando la estructura es CCVC es muy similar a CVC, ya que es de 117.8 ms., y cuando la estructura posterior es CCV, la duración media de /o/ es superior y asciende a 156.9 ms.

/o/			
Estructura de la sílaba posterior a la de la vocal	Casos	Duración media	Desviación estándar
CV	128	132.0	34.3
CVC	65	118.4	35.3
CCV	8	156.0	30.5
CCVC	3	117.8	3.9

Tabla 56. Duración media y desviación estándar de la vocal /o/ producida por las informantes sinohablantes según la sílaba posterior.

En cuanto a la desviación estándar, esta es muy similar cuando la estructura posterior es CV, CVC o CCV, y oscila entre 30 y 35.3. Esto indica que, si bien no hay una gran homogeneidad en la duración de la vocal /o/, dicha homogeneidad no varía en función de la estructura de la sílaba posterior.

### 2.4.4.5 Vocal /u/

En cuanto a la duración de la vocal /u/ realizada por las informantes sinohablantes según la estructura de la sílaba posterior, las hablantes han realizado las estructuras CV y CVC en 368 y 100 ocasiones respectivamente. Además, en tres ocasiones han realizado la estructura CCV (*currar* como [ku.'dra]) y en dos ocasiones CCVC (*currar* como [ku.'drar]). La duración de la vocal /u/ es igual tanto cuando precede a la sílaba CV como a CVC. Sin embargo, la duración es mayor si la estructura posterior es CCV.

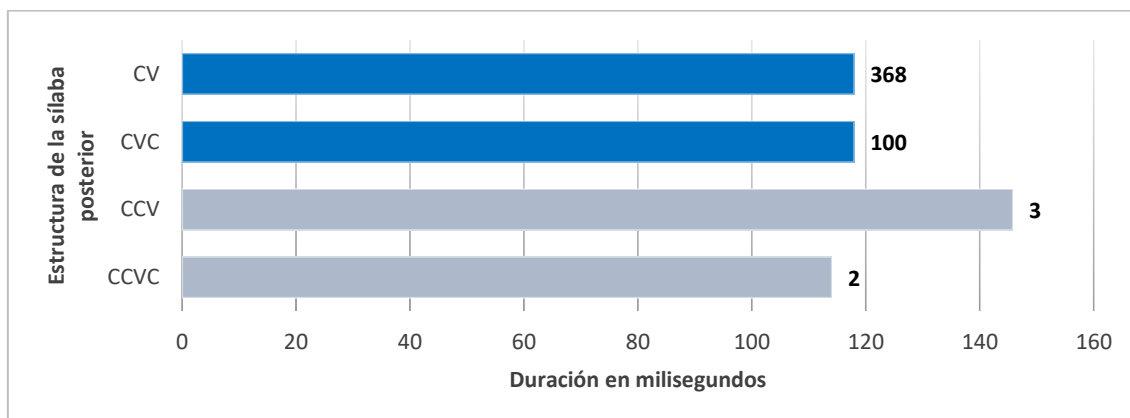


Figura 71. Duración de la vocal /u/ producida por las informantes sinohablantes según la sílaba posterior.

Tal y como se ha indicado, la duración media de /u/ seguida por CV o CVC es igual (117.9 ms.), igual que cuando la sílaba posterior es CCVC (114 ms.). Sin embargo es mayor cuando la sílaba tiene estructura CCV (145.8 ms.).

/u/			
Estructura de la sílaba posterior a la de la vocal	Casos	Duración media	Desviación estándar
CV	368	117.9	37.2
CVC	100	117.9	33.0
CCV	3	145.8	78.2
CCVC	2	114.0	48.0

Tabla 57. Duración media y desviación estándar de la vocal /u/ producida por las informantes sinohablantes según la sílaba posterior.

En cuanto a la desviación estándar, esta es muy similar cuando la sílaba es CV o CVC, mientras que es mucho mayor cuando la estructura posterior es CCV, lo que indica una menor homogeneidad en la duración de /u/ si es seguida por dicha estructura. Sin embargo, el reducido número de casos en los que aparece dicha estructura únicamente permite ver una línea de tendencia, no el comportamiento general.



### 2.4.5 Duración de la vocal en sílaba final

En este apartado se analiza la duración de las vocales producidas por las informantes cuya lengua materna es el chino cuando dichos sonidos vocálicos se encuentran en sílaba final de palabra, es decir, cuando no hay otra sílaba posterior. Por ejemplo, si se analiza la vocal /a/ de la palabra *total*, esta se encuentra en sílaba final. Tal y como se observa en la figura 72, la duración de las vocales de menor a mayor es /o/, /u/, /a/, /i/, /e/. Sin embargo, cabe señalar que, tal y como se puede observar, la duración media de las distintas vocales es muy similar.

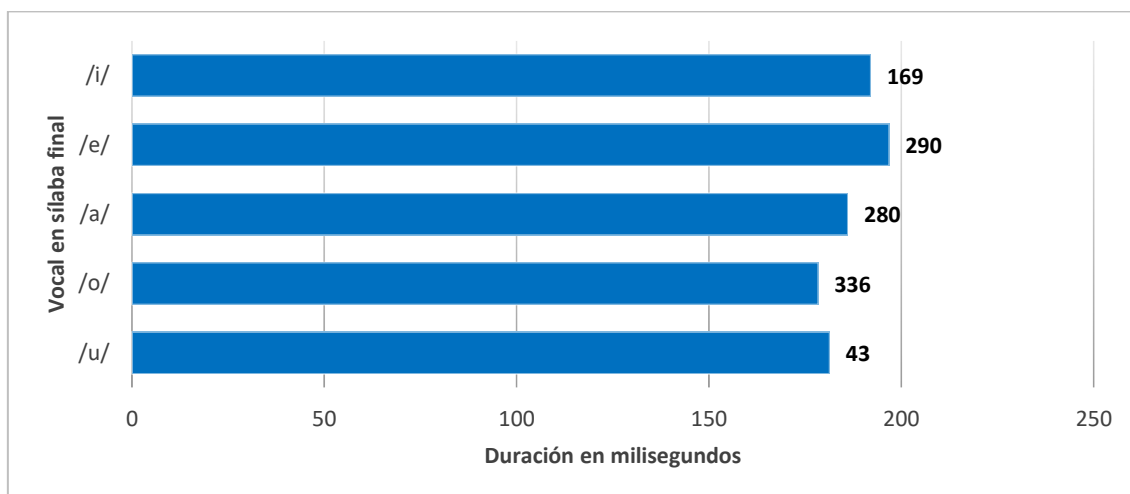


Figura 72. Duración de las vocales producidas por las informantes sinohablantes cuando se encuentran en sílaba final.

Tal y como se ha indicado, la duración media de las distintas vocales es muy similar. /o/ tiene una duración media de 178.4 ms., /u/ y /a/ de 181.3 y 186.1 ms. respectivamente, /i/ de 192 ms. y /e/ de 196.8 ms.

Análisis de la vocal cuando se encuentra en sílaba final			
Vocal	Casos	Duración media	Desviación estándar
/i/	169	192.0	60.1
/e/	290	196.8	64.6
/a/	280	186.1	55.4
/o/	336	178.4	58.6
/u/	43	181.3	56.8

Tabla 58. Duración media y desviación estándar de las vocales producidas por las informantes sinohablantes cuando se encuentran en sílaba final.

En cuanto a la desviación estándar, es muy similar entre unas vocales y otras. Sin embargo, cabe señalar que es elevada, lo que indica que hay una falta de homogeneidad en la duración de las vocales producidas por las informantes sinohablantes cuando se encuentran en posición final de palabra.

#### **2.4.6 Recapitulación**

**El análisis de la duración de la vocal según la estructura de la sílaba en la que se encuentra ha mostrado que si la estructura de dicha sílaba es CV, la duración media es ligeramente superior que cuando la estructura de la sílaba en la que se encuentra es CVC, ya que en el caso de la primera la duración media oscila entre 128 y 172 ms., y entre 110 y 163 ms. en el caso de la segunda. Además, las informantes sinohablantes han realizado otro tipo de estructuras, y abunda la inclusión de consonante entre la consonante y la vocal que conforman la sílaba, como la realización CCV, en cuyo caso la duración media va de 132 a 200 ms., o la realización CCVC, donde la media oscila entre 134 y 172 ms. Esto muestra que cuando se cambia la estructura de la sílaba en la que se encuentra la vocal, la duración de dicho sonido vocálico aumenta respecto a los casos en los que se mantiene la estructura propia de la palabra en español.**

**En cuanto a la duración de las vocales cuando se encuentran en sílaba inicial de palabra aislada, esta oscila entre 118 y 148 ms. y, aunque las duraciones no son homogéneas, la variación de unas a otras vocales es similar.**

**El análisis de la duración de la vocal en función de la sílaba anterior muestra que si la estructura de esta es CV, la media oscila entre 179 y 206 ms., mientras que si dicha sílaba tiene estructura CVC, la duración media oscila entre 166 y 226. Cuando las informantes han añadido una consonante a la sílaba anterior y la han realizado como CCV, la duración media oscila entre 174 y 206 ms.**

**De acuerdo con los datos extraídos del análisis de la duración de las vocales según la estructura de la sílaba posterior, cuando esta es CV, la duración media oscila entre 117 y 152 ms., mientras que si la estructura es CVC, la media oscila entre 117 y 131. Cuando las informantes han realizado la estructura posterior como CCV, la duración media oscila entre 105 y 156 ms., y cuando la han hecho como CCVC, la duración media oscila entre 114 y 140 ms.**

**Cuando la vocal analizada se encuentra en sílaba final, la duración media de las mismas oscila entre 178 y 196 ms. Esto muestra que cuando las informantes sinohablantes realizan vocales de mayor duración si se encuentran en sílaba final de palabra que en sílaba inicial, ya que, en este caso, como se ha indicado, la duración media oscila entre 118 y 148 ms.**

### 3. COMPARACIÓN: GRUPO DE CONTROL E INFORMANTES SINOABLANTES

En los apartados 1 y 2 se ha estudiado la duración de las vocales producidas por las hablantes que conforman el grupo de control y por las informantes sinohablantes respectivamente. En el presente apartado se comparan los resultados obtenidos por ambos grupos y se indica si las diferencias entre las hablantes nativas y las estudiantes chinas son significativas.

#### 3.1 Duración general

Tras exponer los datos extraídos del análisis acústico de las vocales producidas por las hablantes que conforman el grupo de control y por las informantes sinohablantes, en el presente apartado se comparan estos resultados. La figura 73 muestra un gráfico en el que se comparan las duraciones medias de las vocales producidas por el grupo de control (en verde) y por el grupo de las informantes sinohablantes (en azul). Tal y como se puede observar, las vocales producidas por las hablantes cuya lengua materna es el chino son considerablemente más largas que las realizadas por las hispanohablantes nativas.

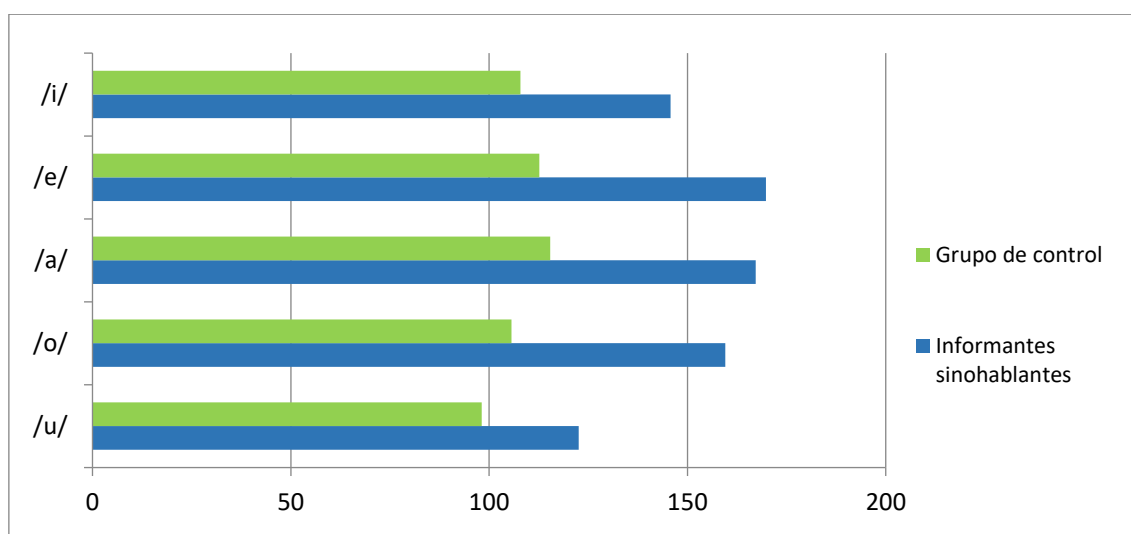


Figura 73. Duración media de las vocales articuladas por el grupo de control y por las informantes sinohablantes respectivamente.

Mientras que en el grupo de control el orden de duración de más breves a más largas es /u/, /o/, /i/, /e/, /a/, en el caso de las informantes sinohablantes el orden es /u/, /i/, /o/, /a/, /e/. Como se puede observar en la tabla 59, en el caso del grupo de control, los valores oscilan entre 98 y 115 ms., mientras que en el caso de las informantes sinohablantes dichos valores son superiores y oscilan entre 122.5 y 169.7 ms. Esto muestra que, como

se ha indicado, las vocales producidas por las informantes sinohablantes presentan una duración superior a las articuladas por las hablantes nativas de español.

Vocal	Grupo de control		Informantes sinohablantes		Nivel de significación
	DM	DS	DM	DS	
/i/	107.9	20.4	145.7	53.9	< 0.001
/e/	112.6	23.5	169.7	56.7	< 0.001
/a/	115.4	21.3	167.2	46.2	< 0.001
/o/	105.6	22.2	159.5	50.0	< 0.001
/u/	98.1	19.8	122.5	40.2	< 0.001

Tabla 59. Duración media y desviación de las vocales articuladas por los dos grupos, y nivel de significación<sup>4</sup>.

En cuanto a la desviación estándar, se observa que en todas las vocales producidas por el grupo de informantes sinohablantes es el doble que las producidas por el grupo de control. En el caso de las hablantes cuya lengua materna es el chino, la desviación estándar se sitúa entre 40.2 y 56.7, mientras que en el del grupo de control la desviación estándar va desde 19.8 hasta 23.5, por lo que estos datos indican que las realizaciones de las hablantes cuya lengua materna es el chino son menos homogéneas y presentan mayor variabilidad que las hablantes cuya lengua materna es el español.

En lo que concierne al nivel de significación, este es, en todos los casos, menor que 0.05, por lo que se considera que las diferencias entre ambos grupos son significativas y que, por lo tanto, la lengua materna de las informantes influye en la duración de las vocales.

### 3.2 Acentuación

El presente apartado muestra las diferencias en la duración de las vocales articuladas por las informantes sinohablantes y por el grupo de control. En primer lugar, la figura 73 muestra dichas diferencias en lo que a las vocales tónicas se refiere.

<sup>4</sup> En adelante, en las tablas de datos se emplearán las siglas *DM* para hacer referencia a la duración media, y *DS* para la desviación estándar.

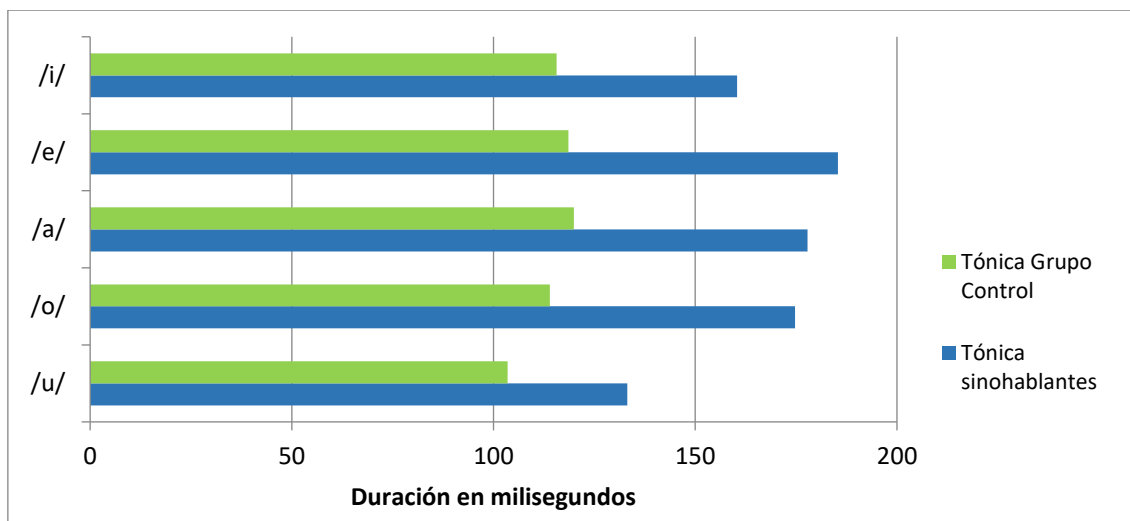


Figura 73. Duración media de las vocales tónicas producidas por el grupo de control y por las informantes sinohablantes.

Tal y como se puede observar, la duración de las vocales tónicas producidas por las informantes sinohablantes es considerablemente mayor que las realizadas por el grupo de control en todos los casos. Cabe señalar que, si bien en el caso de las informantes que conforman el grupo de control la duración de las vocales de menor a mayor es /u/, /o/, /i/, /e/, /a/, en el caso de las informantes sinohablantes el orden, como ya se ha indicado, es /u/, /i/, /o/, /a/, /e/.

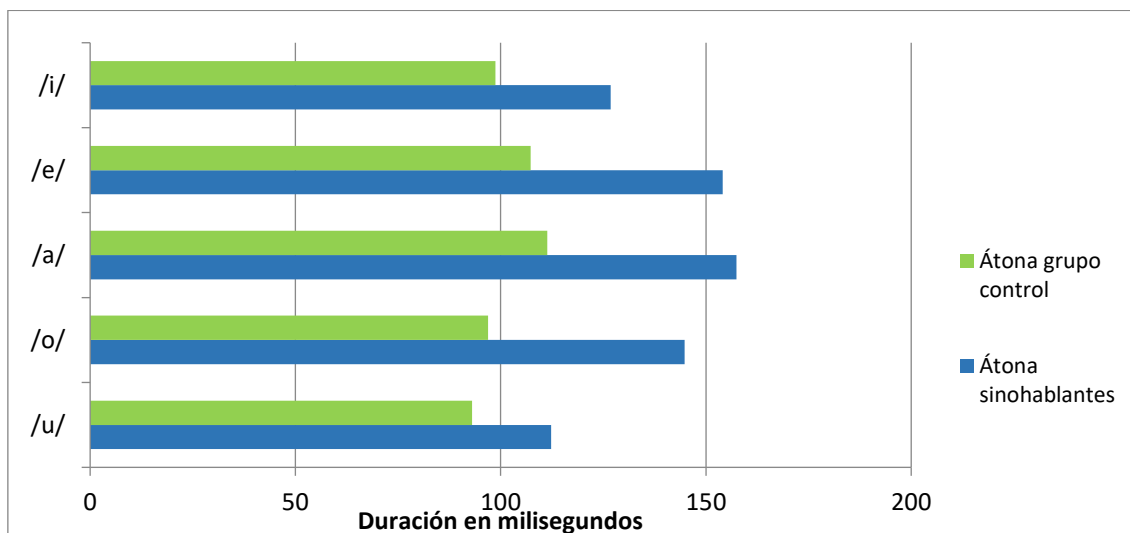


Figura 74. Duración media de las vocales átonas producidas por el grupo de control y por las informantes sinohablantes.

En el caso de las vocales átonas, las cuales aparecen reflejadas en la figura 74, las producidas por las hablantes cuya lengua materna es el chino también presentan una duración mayor que las realizadas por las informantes del grupo de control. En este caso, la duración por orden de menor a mayor es /u/, /i/, /o/, /e/, /a/ cuando las articulan las

informantes sinohablantes, mientras que el orden cambia cuando las produce el grupo de control: /u/, /o/, /i/, /e/, /a/. También se puede observar que, tanto cuando la vocal es realizada como tónica, como cuando es realizada como átona, los sonidos vocálicos cuya duración es más breve son /u/ e /i/, mientras que en el caso del grupo de control, las más breves son /u/ y /o/.

La tabla 60 muestra que, en el caso del grupo de control, la duración de las vocales tónicas oscila entre 103 y 119 ms., mientras que en el de las sinohablantes oscila entre 133 y 185 ms. En cuanto a las vocales átonas, estas tienen una duración media que oscila entre 93 y 111 ms. cuando son realizadas por las hablantes nativas de español, y entre 112 y 157 ms. cuando las pronuncian las estudiantes chinas.

Vocal	Acentuación	Grupo de control		Informantes sinohablantes		Nivel de significación
		DM	DS	DM	DS	
/i/	Tónica	115.6	24.1	160.3	57.1	< 0.001
	Átona	98.7	18,6	126.8	42.0	< 0.001
/e/	Tónica	118.5	24.0	185.3	60.8	< 0.001
	Átona	107.3	20.8	154.0	42.3	< 0.001
/a/	Tónica	119.8	20.6	177.8	48.6	< 0.001
	Átona	111.3	20.2	157.4	36.6	< 0.001
/o/	Tónica	113.9	23.1	174.7	56.1	< 0.001
	Átona	96.9	16.6	144.7	35.2	< 0.001
/u/	Tónica	103.4	20.0	133.1	42.9	< 0.001
	Átona	93.0	18.0	112.2	33.9	< 0.001

Tabla 60. Duración media y desviación estándar de las vocales tónicas y átonas articuladas por los dos grupos, y nivel de significación.

La tabla 60 muestra que la desviación estándar es considerablemente mayor en las vocales producidas por las informantes sinohablantes que en las producidas por el grupo de control, lo que indica que las realizaciones de las hablantes cuya lengua materna es el chino son menos homogéneas que las de las hispanohablantes nativas. Cabe señalar que ambos grupos coinciden en que la desviación es mayor cuando la vocal es realizada como tónica que cuando es realizada como átona, lo que muestra una mayor homogeneidad en la duración de las vocales átonas que en las tónicas.

El nivel de significación de estos datos es, en todos los casos, < 0.001, lo que muestra que existen diferencias significativas en la realización de las vocales del español, tanto tónicas como átonas, entre el grupo de control y las informantes sinohablantes.

### **3.3 Entorno**

En este apartado se comparan los resultados extraídos del análisis del grupo de control con los de las informantes sinohablantes en función del entorno de la vocal. Como se ha mostrado en la parte 1 del presente capítulo, las informantes sinohablantes realizan cambios, tanto en el entorno anterior como en el posterior, por lo que el número de realizaciones de cada una de las vocales varía del grupo de control al de informantes sinohablantes.

#### **3.3.1 Comparación de la duración de las vocales según el entorno consonántico anterior**

A continuación, se muestra la comparación de la duración de las vocales del español producidas por los dos grupos de estudio de la presente tesis doctoral, según el entorno consonántico anterior a las mismas.

##### **3.3.1.1 Vocal /i/**

En la figura 75 se muestra la duración media de las vocales según el entorno posterior y el número de casos de cada una de las vocales. Por ejemplo, el grupo de control (en verde) ha producido /i/ después de oclusiva sorda en 120 ocasiones, mientras que las informantes sinohablantes (en azul) la han realizado como tal 198 veces. Esto se debe a que, como se ha indicado, las hablantes cuya lengua materna es el chino realizan cambios en la estructura de la palabra y en los distintos sonidos del español.

Cuando el sonido anterior a la vocal es una vibrante simple, la duración media de la vocal es la misma para ambos grupos. Destaca la mayor duración de /i/ producida por las hablantes chinas cuando el sonido anterior es una aproximante, aunque ha de tenerse en cuenta que, en este caso, el número de realizaciones de las hablantes chinas (16) es inferior al del grupo de control (48).

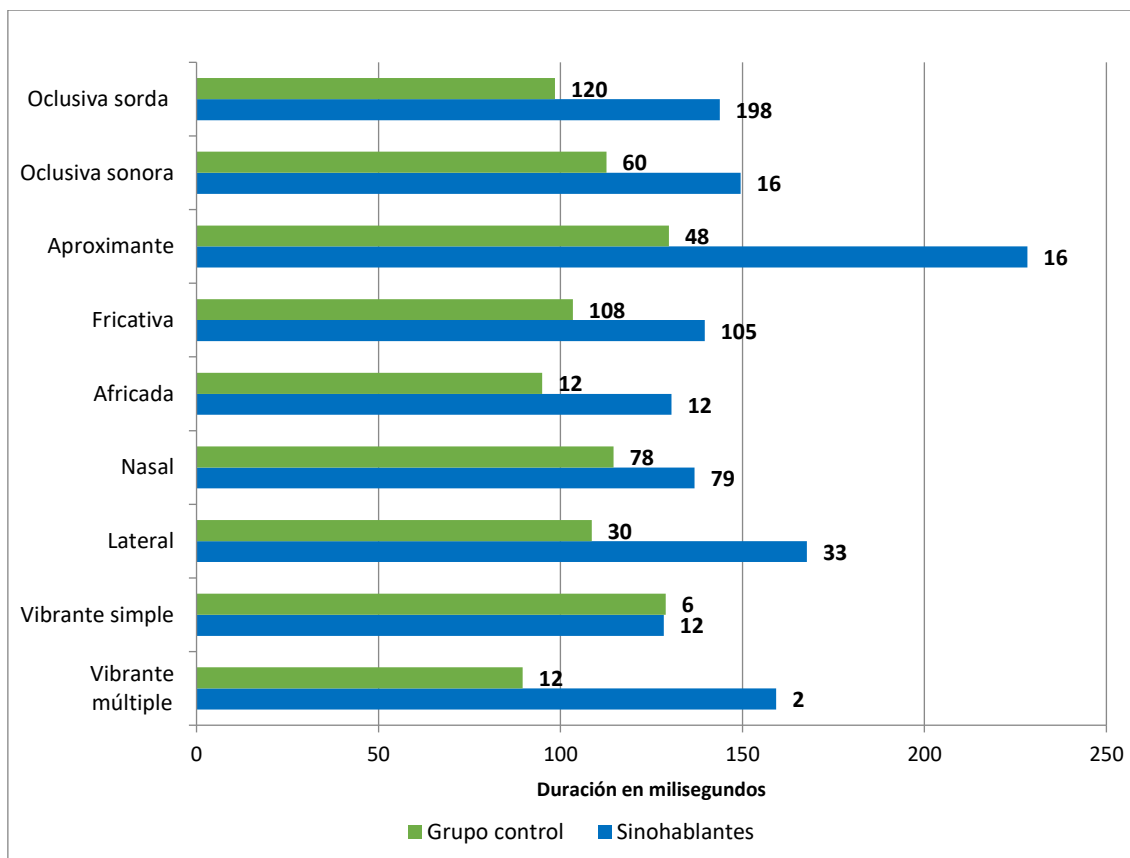


Figura 75. Comparación entre el grupo de control y el de informantes sinohablantes de la duración de la vocal /i/ según el entorno anterior.

Como se ha indicado, la duración media de las vocales del español producidas por las hablantes que conforman el grupo de control es inferior a la de las informantes sinohablantes. En el caso de las primeras, la media oscila entre 89 y 128, mientras que en el de las segundas, entre 128 y 228, es decir, las vocales producidas por las hablantes chinas tienen una duración de casi el doble que las de las hablantes nativas de español. La duración media de la vocal /i/ producida por el grupo de control de más breves a más largas según el entorno anterior es la siguiente: vibrante múltiple, africada, oclusiva sorda, fricativa, lateral, oclusiva sonora, nasal, vibrante simple y aproximante. En cambio, este orden cambia cuando /i/ es producida por las informantes sinohablantes: vibrante simple, africada, nasal, fricativa, oclusiva sonora, oclusiva sorda, vibrante múltiple, lateral y aproximante.



<b>Comparación de la duración de la vocal /i/ según el entorno anterior a la vocal</b>							
	<b>Grupo de control</b>			<b>Informantes sinohablantes</b>			<b>Nivel de significación</b>
	<b>Casos</b>	<b>DM</b>	<b>DS</b>	<b>Casos</b>	<b>DM</b>	<b>DS</b>	
<b>Oclusiva sorda</b>	120	98.4	24.2	198	143.7	55.9	< 0.001
<b>Oclusiva sonora</b>	60	112.6	26.3	16	149.5	49.0	0.004
<b>Aproximante</b>	48	129.8	28.5	16	228.2	73.8	< 0.001
<b>Fricativa</b>	108	103.4	25.1	105	139.6	52.6	< 0.001
<b>Africada</b>	12	95.0	8.7	12	130.5	49.4	0.021
<b>Nasal</b>	78	114.5	21.1	79	136.8	56.8	0.016
<b>Lateral</b>	30	108.5	25.6	33	167.6	53.9	< 0.001
<b>Vibrante simple</b>	6	128.8	33.7	12	128.3	38.8	0.262
<b>Vibrante múltiple</b>	12	89.6	10.5	2	159.3	25.0	0.006

*Tabla 61.* Duración media y desviación estándar de la vocal /i/ según el entorno consonántico anterior articulada por los dos grupos, y nivel de significación.

La desviación estándar oscila entre 8 y 28 en las vocales producidas por el grupo de control, y entre 25 y 73 cuando /i/ es realizada por las informantes sinohablantes. Esto muestra que la vocal /i/ tiene una duración considerablemente más homogénea cuando es articulada por las informantes hispanohablantes nativas que por las hablantes cuya lengua materna es el chino.

En cuanto al nivel de significación de los resultados obtenidos del análisis de la duración de /i/ según el entorno anterior, se observa que la diferencia entre el grupo de control y las informantes sinohablantes es significativa en todos los casos, excepto cuando el sonido anterior es una vibrante simple. Por este motivo, se puede asumir que existen diferencias significativas entre ambos grupos y que, por lo tanto, la lengua materna de las informantes sinohablantes interfiere en la duración de /i/.

### 3.3.1.2 Vocal /e/

La figura 76 muestra la duración media de /e/ según el entorno anterior producida por el grupo de control y por las informantes sinohablantes. Tal y como se puede observar, la duración es mayor cuando esta vocal es producida por las estudiantes chinas que por

las hablantes nativas de español. Cabe señalar que no hay casos en los que el sonido anterior a /e/ es una consonante africada.

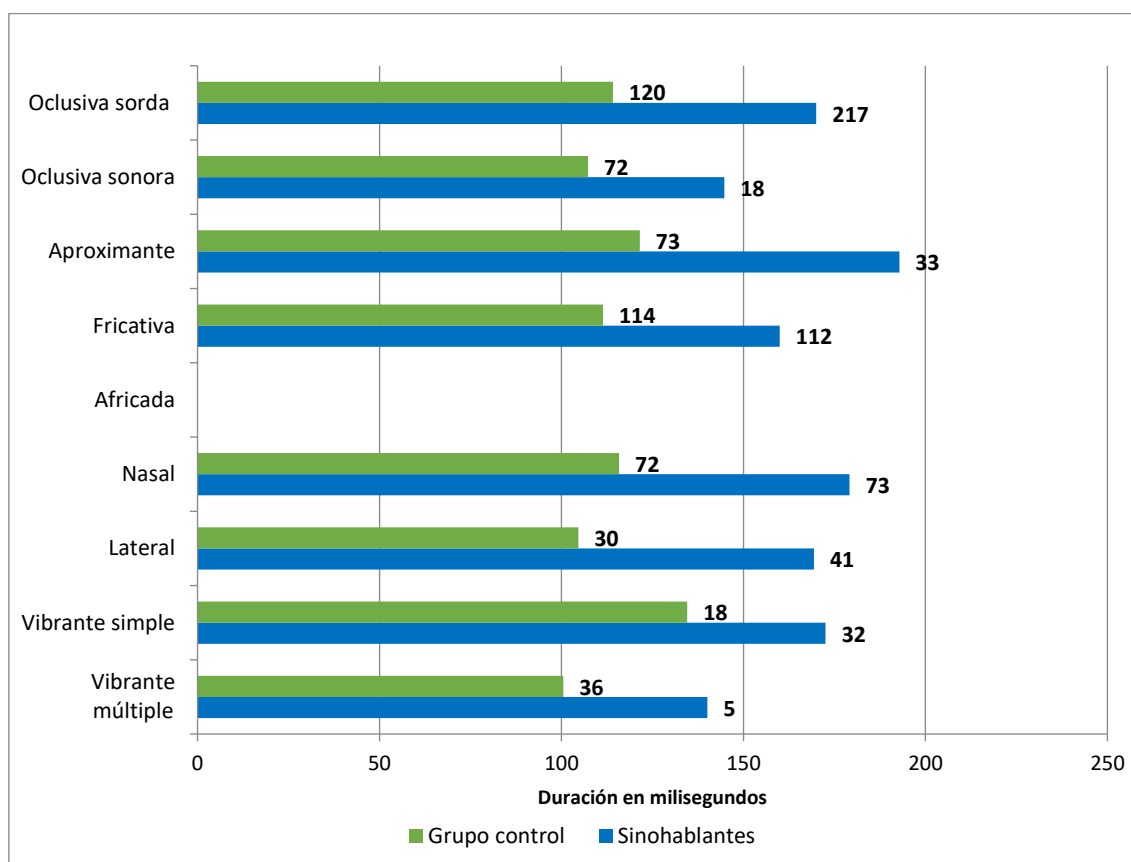


Figura 76. Comparación entre el grupo de control y el de informantes sinohablantes de la duración de la vocal /e/ según el entorno anterior.

Como se ha indicado, la duración media de las vocales del español producidas por las hablantes cuya lengua materna es el español es inferior a la de las hablantes chinas. En el caso de las primeras, la media oscila entre 100 y 134 ms., y la duración de /e/ según el entorno anterior es, de más breves a más largas, la siguiente: vibrante múltiple, lateral, oclusiva sonora, fricativa, oclusiva sorda, nasal, aproximante y vibrante simple. En el caso de las estudiantes sinohablantes, la duración media de /e/ oscila entre 140 y 192 ms. El orden de más breves a más largas cambia respecto al grupo de control, ya que en estas hablantes es: vibrante múltiple, oclusiva sonora, fricativa, lateral, oclusiva sorda, vibrante simple, nasal y aproximante.

<b>Comparación de la duración de la vocal /e/ según el entorno anterior a la vocal</b>							
	<b>Grupo de control</b>			<b>Informantes sinohablantes</b>			<b>Nivel de significación</b>
	<b>Casos</b>	<b>DM</b>	<b>DS</b>	<b>Casos</b>	<b>DM</b>	<b>DS</b>	
<b>Oclusiva sorda</b>	120	114.1	27.3	217	170.0	65.3	< 0.001
<b>Oclusiva sonora</b>	72	107.3	27.0	18	144.7	41.1	< 0.001
<b>Apro- ximante</b>	73	121.5	28.6	33	192.8	68.1	< 0.001
<b>Fricativa</b>	114	111.4	28.5	112	159.9	61.3	< 0.001
<b>Africada</b>	0	--	--	0	--	--	--
<b>Nasal</b>	72	115.8	22.8	73	179.1	70.1	< 0.001
<b>Lateral</b>	30	104.6	12.2	41	169.3	61.7	< 0.001
<b>Vibrante simple</b>	18	134.5	22.9	32	172.5	53.1	< 0.001
<b>Vibrante múltiple</b>	36	100.4	17.7	5	140.0	36.9	< 0.001

*Tabla 62.* Duración media y desviación estándar de la vocal /e/ según el entorno consonántico anterior articulada por los dos grupos, y nivel de significación.

En cuanto a la desviación estándar, en el caso del grupo de control oscila entre 12 y 28, mientras que en el de las informantes sinohablantes, entre 36 y 70, lo que indica que las realizaciones de las estudiantes chinas presentan una homogeneidad considerablemente menor que las de las hablantes del grupo de control.

En lo que al nivel de significación se refiere, la diferencia entre ambos grupos es significativa, ya que en todos los casos es menor que 0.05, por lo que se asume que la lengua materna de las informantes sinohablantes interfiere en su articulación de la vocal /e/ según el entorno anterior.

### 3.3.1.3 Vocal /a/

En cuanto a la comparación de la duración de la vocal /a/ según el entorno anterior articulada por los dos grupos de estudio, la figura 77 muestra que en todos los casos la duración de dicho sonido vocálico es superior cuando es articulado por las hablantes cuya lengua materna es el chino, que cuando lo realizan las hablantes del grupo de control. Cabe señalar que, en este caso, si bien el grupo de control presenta 24 casos de vibrante múltiple anterior a /a/, las informantes sinohablantes no han realizado dicha estructura en ninguna ocasión, por lo que no aparece en el gráfico.

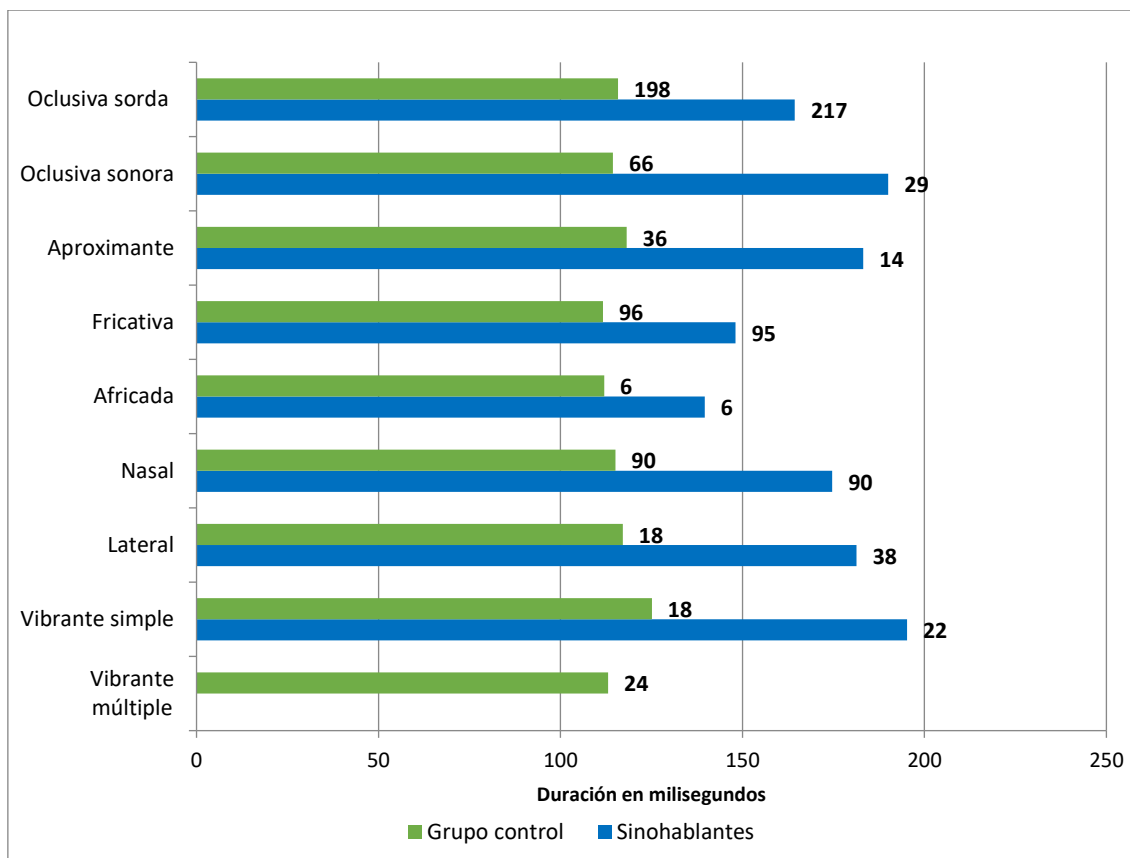


Figura 77. Comparación entre el grupo de control y el de informantes sinohablantes de la duración de la vocal /a/ según el entorno anterior.

Como se ha indicado, la duración media de /a/ producida por las hablantes que conforman el grupo de control es inferior a la de las informantes sinohablantes. Cuando esta es articulada por las primeras, la duración media es muy similar en todos los casos, y oscila entre 111 y 125 ms. En cambio, cuando es producida por las hablantes chinas, la duración media va desde los 139 hasta los 195 ms. Igual que sucede con las vocales /i/ y /e/, la duración de /a/ de más breve a más larga tampoco es la misma en un grupo y en el otro. Cuando esta vocal es producida por el grupo de control, el orden es el siguiente: fricativa, africada, vibrante múltiple, oclusiva sonora, nasal, oclusiva sorda, lateral, aproximante y vibrante simple. Cabe señalar que, en este caso, la duración de todas es muy similar (entre 111 y 118 ms), y es ligeramente superior cuando el sonido anterior es una vibrante simple (125 ms). En cuanto al orden de más breves a más largas cuando /a/ es articulada por las informantes sinohablantes, es el siguiente: africada, fricativa, oclusiva sorda, nasal, lateral, aproximante, oclusiva sonora, y vibrante simple.

<b>Comparación de la duración de la vocal /a/ según el entorno anterior a la vocal</b>							
	<b>Grupo de control</b>			<b>Informantes sinohablantes</b>			<b>Nivel de significación</b>
	<b>Casos</b>	<b>DM</b>	<b>DS</b>	<b>Casos</b>	<b>DM</b>	<b>DS</b>	
<b>Oclusiva sorda</b>	198	115.8	26.2	217	164.4	48.0	< 0.001
<b>Oclusiva sonora</b>	66	114.4	23.8	29	190.0	58.2	< 0.001
<b>Aproxi mante</b>	36	118.2	26.2	14	183.2	55.3	< 0.001
<b>Fricativa</b>	96	111.6	25.4	95	148.1	44.2	< 0.001
<b>Africada</b>	6	112.0	21.3	6	139.6	40.1	0.200
<b>Nasal</b>	90	115.0	21.7	90	174.6	59.1	< 0.001
<b>Lateral</b>	18	117.1	11.5	38	181.3	47.2	< 0.001
<b>Vibrante simple</b>	18	125.1	23.0	22	195.1	47.6	< 0.001
<b>Vibrante múltiple</b>	24	113.1	21.8	0	--	--	< 0.001

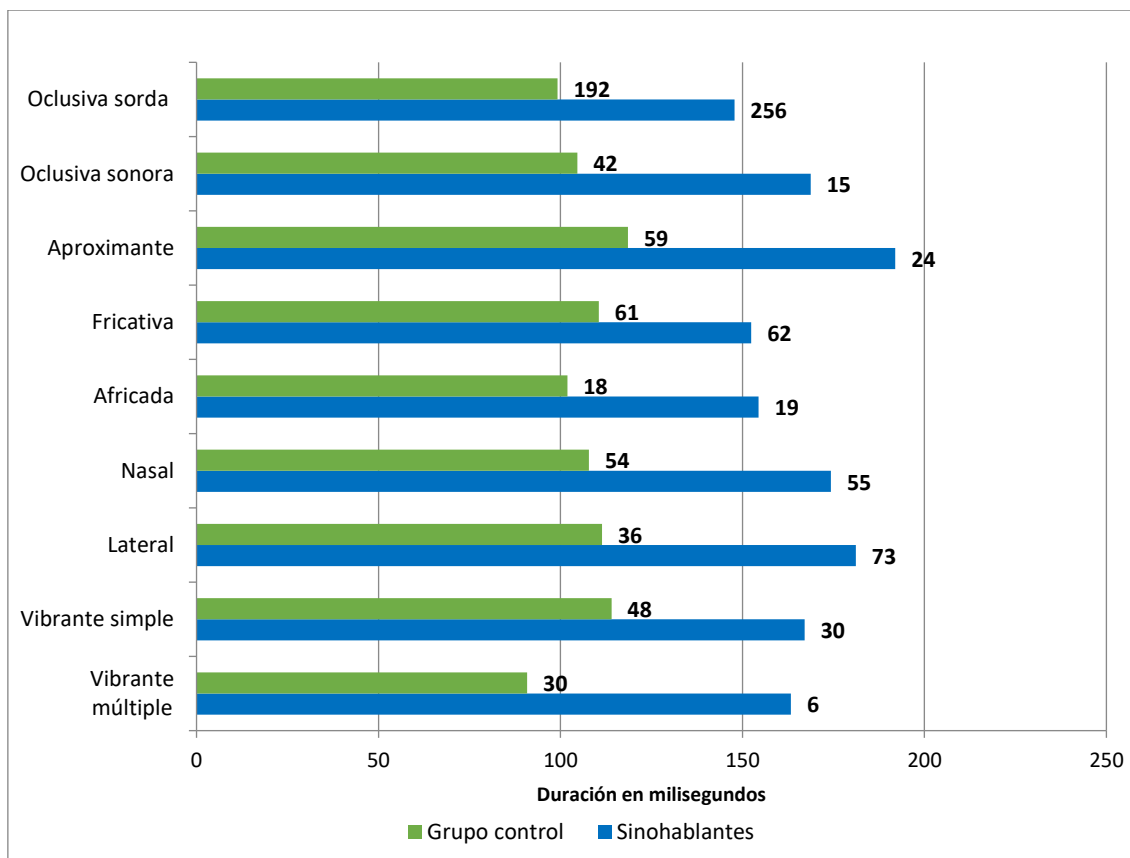
*Tabla 63.* Duración media y desviación estándar de la vocal /a/ según el entorno consonántico anterior articulada por los dos grupos, y nivel de significación.

En cuanto a la desviación estándar, en el caso del grupo de control oscila entre 11 y 26, mientras que en el de las informantes sinohablantes, entre 40 y 59, lo que indica que las realizaciones de las estudiantes chinas presentan menor homogeneidad que las de las hablantes del grupo de control.

En lo que al nivel de significación se refiere, exceptuando cuando /a/ es precedida por una consonante africada, la diferencia entre ambos grupos es significativa, ya que es menor que 0.05, por lo que se asume que la lengua materna de las informantes, en este caso el chino, interfiere en su realización de /a/.

#### **3.3.1.4 Vocal /o/**

La figura 78 muestra la duración media de /o/ según el entorno anterior producida por el grupo de control y por las informantes sinohablantes. Tal y como se puede observar, la duración media es mayor cuando este sonido vocálico es realizado por las estudiantes chinas que por las hablantes nativas de español.



*Figura 78.* Comparación entre el grupo de control y el de informantes sinohablantes de la duración de la vocal /o/ según el entorno anterior.

Cuando /o/ es articulada por las hablantes que conforman el grupo de control, la duración media oscila entre 90 y 118 ms., mientras que si esta vocal es producida por las informantes sinohablantes, la duración es considerablemente superior, y va desde los 147 hasta los 191 ms. En el caso del primer grupo, la duración media de /o/ de menor a mayor según el entorno anterior es la siguiente: vibrante múltiple, oclusiva sorda, africada, oclusiva sonora, nasal, fricativa, lateral, vibrante simple y aproximante. Cabe señalar que en estos casos la duración media es muy similar, ya que la diferencia entre las más largas y las más breves es de menos de 20 ms. En cambio, este orden cambia cuando la vocal es producida por las hablantes chinas: oclusiva sorda, fricativa, africada, vibrante múltiple, vibrante simple, oclusiva sonora, nasal, lateral y aproximante.

Comparación de la duración de la vocal /o/ según el entorno anterior a la vocal							
	Grupo de control			Informantes sinohablantes			Nivel de significación
	Casos	DM	DS	Casos	DM	DS	
<b>Oclusiva sorda</b>	192	99.2	24.9	256	147.8	52.4	< 0.001
<b>Oclusiva sonora</b>	42	104.6	19.4	15	168.7	56.1	< 0.001
<b>Aproxi- mante</b>	59	118.5	30.5	24	191.9	69.7	< 0.001
<b>Fricativa</b>	61	110.5	26.6	62	152.4	45.7	< 0.001
<b>Africada</b>	18	101.8	12.4	19	154.4	34.2	< 0.001
<b>Nasal</b>	54	107.8	10.4	55	174.2	51.0	< 0.001
<b>Lateral</b>	36	111.3	16.3	73	181.1	64.4	< 0.001
<b>Vibrante simple</b>	48	114.0	28.7	30	167.0	67.8	< 0.001
<b>Vibrante múltiple</b>	30	90.8	17.9	6	163.3	78.3	< 0.001

Tabla 64. Duración media y desviación estándar de la vocal /o/ según el entorno consonántico anterior articulada por los dos grupos, y nivel de significación.

En cuanto a la desviación estándar, en el caso del grupo de control oscila entre 10 y 30, mientras que en el de las informantes sinohablantes, entre 34 y 78, lo que indica que, igual que sucede con /i/, /e/ y /a/, las realizaciones de las estudiantes chinas presentan menor homogeneidad que las de las hablantes del grupo de control.

En relación con el nivel de significación, en todos los casos la diferencia entre ambos grupos es significativa, ya que es menor que 0.05, por lo que se asume que la lengua materna de las informantes sinohablantes interfiere en su realización de /o/.

### 3.3.1.5 Vocal /u/

En cuanto a la comparación de la duración de la vocal /u/ según el entorno anterior articulada por los dos grupos de estudio, la figura 79 muestra que en todos los casos la duración de dicho sonido vocálico es superior cuando es articulado por las hablantes cuya lengua materna es el chino, que cuando lo realizan las hablantes del grupo de control. Sin embargo, generalmente las diferencias entre estos dos grupos son menores que en los casos de las vocales anteriormente analizadas. Cabe añadir que, en algunos casos, como cuando el sonido anterior es una aproximante, el número de realizaciones de /u/ es reducido.

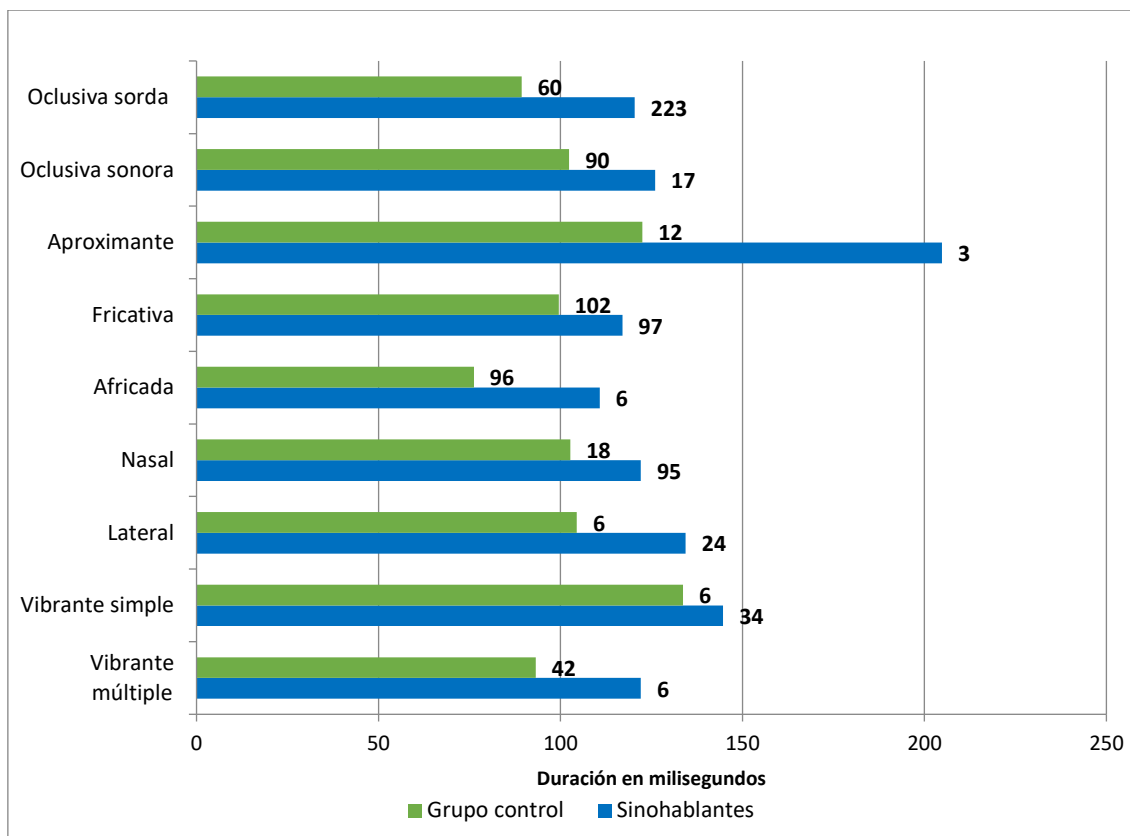


Figura 79. Comparación entre el grupo de control y el de informantes sinohablantes de la duración de la vocal /u/ según el entorno anterior.

Como se ha indicado, la duración media de /u/ producida por las hablantes que conforman el grupo de control es inferior a la de las informantes sinohablantes. Cuando esta es articulada por el primer grupo, en la tabla 65 se observa que /u/ es más breve cuando el sonido anterior es una consonante africada, y la duración más larga se da cuando dicho sonido es una vibrante simple. Sin embargo, el corpus se compone únicamente de 6 realizaciones de vocal después de estos dos sonidos vocálicos. En cuanto al resto de duraciones medias, esta oscila entre 93 y 122 ms. El orden, de más breves a más largas, es el siguiente: africada, oclusiva sorda, vibrante múltiple, fricativa, oclusiva sonora, nasal, lateral, aproximante y vibrante simple. Cuando /u/ es articulada por las hablantes cuya lengua materna es el chino, la duración media va desde 110 hasta 204 ms. Igual que en el grupo de control, hay pocas realizaciones de /u/ cuando es precedida por algunos sonidos consonánticos, como aproximante, africada o vibrante múltiple. Teniendo esto en cuenta, cabe señalar que el orden, de menor a mayor duración de esta vocal según su entorno anterior es el siguiente: africada, fricativa, oclusiva sorda, nasal, vibrante múltiple, oclusiva sonora, lateral y vibrante simple.



<b>Comparación de la duración de la vocal /u/ según el entorno anterior a la vocal</b>							
	<b>Grupo de control</b>			<b>Informantes sinohablantes</b>			<b>Nivel de significación</b>
	<b>Casos</b>	<b>DM</b>	<b>DS</b>	<b>Casos</b>	<b>DM</b>	<b>DS</b>	
<b>Oclusiva sorda</b>	60	89.3	21.7	223	120.4	40.2	< 0.001
<b>Oclusiva sonora</b>	90	102.4	22.3	17	126.0	45.0	0.052
<b>Aproximante</b>	12	122.5	31.6	3	204.8	104.6	0.033
<b>Fricativa</b>	102	99.5	25.0	97	117.0	38.8	< 0.001
<b>Africada</b>	6	76.2	21.4	6	110.8	27.3	0.055
<b>Nasal</b>	96	102.7	18.9	95	122.1	40.6	< 0.001
<b>Lateral</b>	18	104.5	18.7	24	134.4	41.8	0.088
<b>Vibrante simple</b>	6	133.7	26.8	34	144.7	48.9	0.010
<b>Vibrante múltiple</b>	42	93.2	15.8	6	122.1	31.1	< 0.001

*Tabla 65.* Duración media y desviación estándar de la vocal /u/ según el entorno consonántico anterior articulada por los dos grupos, y nivel de significación.

En cuanto a la desviación estándar, en el caso del grupo de control oscila entre 15 y 31, mientras que en el de las informantes sinohablantes, entre 27 y 104. Sin embargo, cabe señalar que esta desviación tan elevada corresponde a la realización de /u/ tras una consonante aproximante, de la que únicamente se han podido analizar tres casos. Si no se tienen en cuenta las realizaciones de /u/ tras este sonido, la desviación estándar oscila entre 27 y 48. Esto muestra que, si bien las realizaciones de las informantes sinohablantes son menos homogéneas que las del grupo de control, /u/ es la vocal articulada por las estudiantes chinas que menor variedad presenta en sus realizaciones.

En relación con el nivel de significación de los resultados obtenidos en el análisis de las vocales producidas por ambos grupos, se observa cuando el sonido anterior es una oclusiva sonora, una africada o lateral las diferencias entre ambos grupos no son significativas ( $>0.05$ ). Sin embargo, si el sonido anterior es una oclusiva sorda, aproximante, fricativa, nasal, vibrante simple, vibrante múltiple, el nivel de significación es menor que 0.05, lo que indica que existen diferencias significativas y que la duración de /u/ tras estos sonidos depende de la lengua materna de las informantes.

### 3.3.2 Comparación de la duración de las vocales según el entorno posterior

A continuación, se muestra la comparación de la duración de las vocales del español producidas por los dos grupos de estudio de la presente tesis doctoral, según el entorno consonántico posterior a las mismas.

#### 3.3.2.1 Vocal /i/

La figura 80 muestra la duración media de /i/ según el entorno posterior producida por el grupo de control y por las informantes sinohablantes. Tal y como se puede observar, la duración es mayor cuando esta vocal es producida por las estudiantes chinas que por las hablantes nativas de español. Sin embargo, cuando la vocal precede a una aproximante, la duración de la misma es similar en los dos grupos. Cabe añadir que no hay casos, ni en el grupo de control, ni en el de informantes sinohablantes, en los que /i/ es seguida por vibrante múltiple. En cuanto a esta vocal seguida por oclusiva sonora, no hay ningún caso realizado por las informantes sinohablantes. Esto se debe a que, como se ha señalado en la primera parte de este capítulo, las informantes sinohablantes hacen un cambio denominado resilabación, por el que hay cambios en la estructura de la palabra y en los sonidos consonánticos.

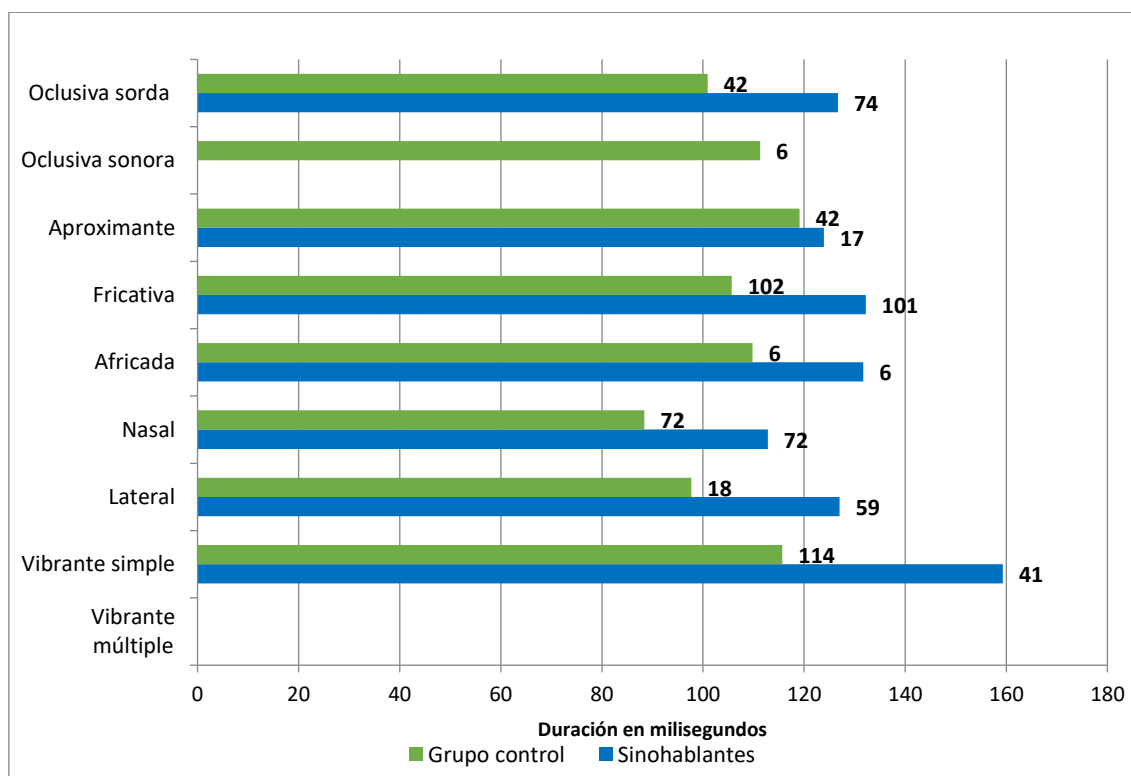


Figura 80. Comparación entre el grupo de control y el de informantes sinohablantes de la duración de la vocal /i/ según el entorno posterior.

Como se ha indicado, la duración media de las vocales del español producidas por las hablantes que conforman el grupo de control es inferior a la de las informantes sinohablantes. En el caso de las primeras, la media oscila entre 88 y 119 ms., mientras que en el de las segundas, entre 112 y 159 ms. En cuanto a la duración de /i/ articulada por las hablantes nativas de español, el orden de más breves a más largas según el entorno posterior es el siguiente: nasal, lateral, oclusiva sorda, fricativa, africada, oclusiva sonora, vibrante simple y aproximante. En cambio, este orden cambia cuando la vocal en cuestión es realizada por las informantes sinohablantes: nasal, aproximante, oclusiva sorda, lateral, africada, fricativa y vibrante múltiple.

<b>Comparación de la duración de la vocal /i/ según el entorno posterior a la vocal</b>							
	<b>Grupo de control</b>			<b>Informantes sinohablantes</b>			<b>Nivel de significación</b>
	<b>Casos</b>	<b>DM</b>	<b>DS</b>	<b>Casos</b>	<b>DM</b>	<b>DS</b>	
<b>Oclusiva sorda</b>	42	100.9	18.9	74	126.7	34.9	< 0.001
<b>Oclusiva sonora</b>	6	111.3	21.3	0	--	--	0.004
<b>Aproximante</b>	42	119.1	27.2	17	123.9	27.2	0.458
<b>Fricativa</b>	102	105.7	25.3	101	132.2	42.8	< 0.001
<b>Africada</b>	6	109.8	33.3	6	131.7	44.2	0.522
<b>Nasal</b>	72	88.4	14.1	72	112.8	33.3	< 0.001
<b>Lateral</b>	18	97.7	16.4	59	127.0	46.1	< 0.001
<b>Vibrante simple</b>	114	115.7	23.2	41	159.3	38.8	< 0.001
<b>Vibrante múltiple</b>	0	--	--	0	--	--	--

*Tabla 66.* Duración media y desviación estándar de las vocal /i/ según el entorno posterior articulada por los dos grupos, y nivel de significación.

La desviación estándar oscila entre 14 y 27 en las vocales producidas por el grupo de control, y entre 27 y 46 cuando /i/ es realizada por las informantes sinohablantes, Esto muestra que esta vocal tiene una duración más homogénea cuando es articulada por las hablantes nativas de español, que por las estudiantes chinas.

En cuanto al nivel de significación de los resultados obtenidos del análisis de la duración de /i/ según el entorno posterior, se observa que la diferencia es significativa cuando el entorno posterior a la vocal es una oclusiva sorda, oclusiva sonora, fricativa, nasal, lateral o vibrante simple. En cambio, la diferencia no es significativa cuando dicho sonido es una aproximante o una africada ( $> 0.05$ ). Por este motivo, se puede asumir que, en

general, existen diferencias significativas entre ambos grupos y que, por lo tanto, la lengua materna de las informantes sinohablantes interfiere en la duración de /e/, excepto cuando el sonido posterior es una aproximante o una consonante africada.

### 3.3.2.2 Vocal /e/

Tal y como muestra la figura 81, la duración media de /e/ según el entorno posterior es mayor cuando esta es producida por las informantes sinohablantes que cuando la articulan las hablantes que conforman el grupo de control. Cabe señalar que no hay casos de africada en posición posterior a /e/ en ninguno de los dos grupos, y que las informantes sinohablantes no han realizado esta vocal seguida por vibrante simple en ningún caso. Esto se debe, como se ha señalado anteriormente, al cambio denominado resilabación que hacen las estudiantes chinas en su producción oral en español, por el que hay cambios en la estructura de la palabra.

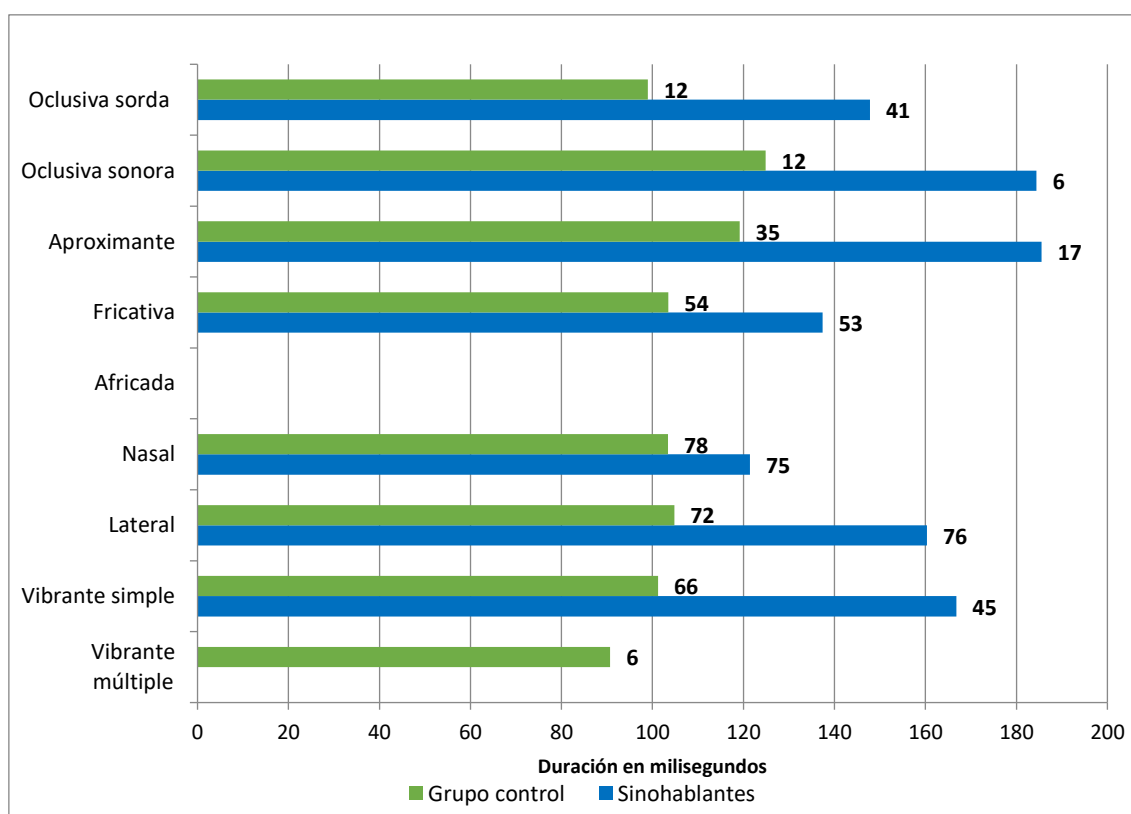


Figura 81. Comparación entre el grupo de control y el de informantes sinohablantes de la duración de la vocal /e/ según el entorno posterior.

Como se ha indicado, la duración media de la /e/ del español producida por las hablantes que conforman el grupo de control es menor que la de las informantes sinohablantes. En el caso de las primeras, la duración media oscila entre 90 y 124 ms., mientras que en el

de las segundas, entre 121 y 185 ms. La duración media de la vocal /e/ producida por el grupo de control de realizaciones más breves a más largas según el entorno posterior es la siguiente: vibrante múltiple, oclusiva sorda, vibrante simple, nasal, fricativa, lateral, aproximante y oclusiva sonora. Si esta vocal es realizada por las estudiantes chinas, este orden, también de más breves a más largas, cambia: nasal, fricativa, oclusiva sorda, lateral, vibrante simple, oclusiva sonora y aproximante.

<b>Comparación de la duración de la vocal /e/ según el entorno posterior a la vocal</b>							
	<b>Grupo de control</b>			<b>Informantes sinohablantes</b>			<b>Nivel de significación</b>
	<b>Casos</b>	<b>DM</b>	<b>DS</b>	<b>Casos</b>	<b>DM</b>	<b>DS</b>	
<b>Oclusiva sorda</b>	12	99.0	26.1	41	147.8	40.9	< 0.001
<b>Oclusiva sonora</b>	12	124.9	30.4	6	184.4	74.7	< 0.001
<b>Aproxi- mante</b>	35	119.2	31.4	17	185.5	91.1	0.005
<b>Fricativa</b>	54	103.5	26.9	53	137.4	35.4	< 0.001
<b>Africada</b>	0	--	--	0	--	--	--
<b>Nasal</b>	78	103.4	16.9	75	121.4	32.7	< 0.001
<b>Lateral</b>	72	104.8	16.9	76	160.3	45.2	< 0.001
<b>Vibrante simple</b>	66	101.2	16.8	45	166.8	44.2	< 0.001
<b>Vibrante múltiple</b>	6	90.7	--	0	--	--	0.078

*Tabla 67.* Duración media y desviación estándar de las vocal /e/ según el entorno posterior articulada por los dos grupos, y nivel de significación.

La desviación estándar oscila entre 16 y 31 en las vocales producidas por el grupo de control, y entre 32 y 91 cuando /e/ es realizada por las informantes sinohablantes. Esto muestra que la vocal /e/ tiene una duración considerablemente más homogénea cuando es articulada por las informantes hispanohablantes nativas que cuando lo es por las hablantes cuya lengua materna es el chino.

En cuanto al nivel de significación de los resultados obtenidos en el análisis de los dos grupos, se observa que las diferencias son significativas en todos los casos, excepto cuando el sonido posterior a la vocal /e/ es una vibrante múltiple ( $> 0.05$ ). Por este motivo, se puede asumir que existen diferencias significativas entre ambos grupos y que, por lo tanto, la lengua materna de las informantes sinohablantes interfiere en la duración de /e/, excepto cuando el sonido posterior es una vibrante múltiple.

### 3.3.2.3 Vocal /a/

La figura 82 muestra la duración media de /a/ según el entorno posterior producida por las informantes hispanohablantes y por las hablantes cuya lengua materna es el chino. Tal y como se puede observar, la duración es mayor cuando esta vocal es producida por las informantes sinohablantes que por las hablantes del grupo de control, aunque en el caso de /a/ seguida por oclusiva sonora la duración media de ambos grupos es bastante similar. Cabe añadir que no hay casos de /a/ seguida por africada en ninguno de los dos grupos, y que, en el caso de la vibrante múltiple, las informantes sinohablantes realizan este sonido consonántico tras la vocal /a/ en una única ocasión.

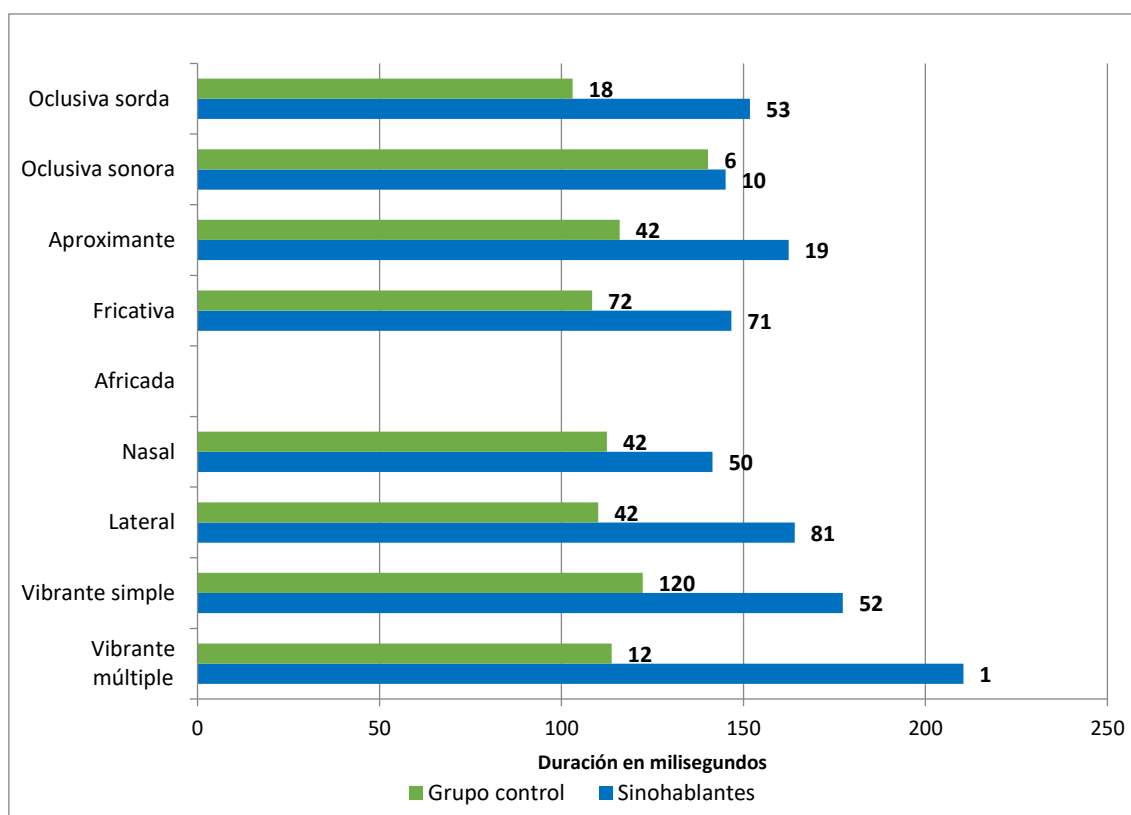


Figura 82. Comparación entre el grupo de control y el de informantes sinohablantes de la duración de la vocal /a/ según el entorno posterior.

Cuando la vocal /a/ es producida por las hablantes del grupo de control, la duración media oscila entre 103 y 140 ms., mientras que si es articulada por las informantes sinohablantes, dicha duración es mayor, y oscila entre 141 y 210 ms. En cuanto a la duración de esta vocal articulada por el primer grupo, el orden de realizaciones más breves a más largas según el entorno posterior es el siguiente: oclusiva sorda, fricativa, lateral, nasal, vibrante múltiple, aproximante, vibrante simple y oclusiva sonora. En el caso del segundo grupo, el orden es diferente: nasal, oclusiva sonora, fricativa, oclusiva sorda, aproximante, lateral, vibrante simple y vibrante múltiple.

<b>Comparación de la duración de la vocal /a/ según el entorno posterior a la vocal</b>							
	<b>Grupo de control</b>			<b>Informantes sinohablantes</b>			<b>Nivel de significación</b>
	<b>Casos</b>	<b>DM</b>	<b>DS</b>	<b>Casos</b>	<b>DM</b>	<b>DS</b>	
<b>Oclusiva sorda</b>	18	103.1	14.0	53	151.8	46.3	< 0.001
<b>Oclusiva sonora</b>	6	140.3	23.1	10	145.1	43.2	0.004
<b>Aproximante</b>	42	116.0	23.9	19	162.4	52.0	0.001
<b>Fricativa</b>	72	108.4	21.8	71	146.7	29.6	< 0.001
<b>Africada</b>	0	--	--	0	--	--	--
<b>Nasal</b>	42	112.5	17.0	50	141.5	35.0	< 0.001
<b>Lateral</b>	42	110.1	21.4	81	164.1	37.4	< 0.001
<b>Vibrante simple</b>	120	122.3	20.7	52	177.3	53.1	< 0.001
<b>Vibrante múltiple</b>	12	113.8	11.7	1	210.4	--	0.057

*Tabla 68.* Duración media y desviación estándar de las vocal /a/ según el entorno posterior articulada por los dos grupos, y nivel de significación.

La desviación estándar oscila entre 11 y 23 en las vocales realizadas por el grupo de control, y entre 29 y 53 cuando /a/ es producida por las estudiantes chinas. Esto muestra que esta vocal tiene una duración más homogénea cuando es articulada por el primer grupo que por el segundo.

En cuanto al nivel de significación de los resultados obtenidos en el análisis de los dos grupos, se observa que las diferencias son significativas en todos los casos excepto cuando el sonido posterior a /a/ es una vibrante múltiple. Sin embargo, cabe señalar que en este caso el nivel de significación es 0.057, es decir, muy próximo al nivel de significación establecido en la presente tesis doctoral. Por este motivo, se puede asumir que existen diferencias significativas entre ambos grupos y que, por lo tanto, la lengua materna de las informantes sinohablantes interfiere en la duración de /e/.

### **3.3.2.4 Vocal /o/**

De acuerdo con los datos mostrados en la figura 83, la duración de la vocal /o/ según el entorno posterior es mayor cuando esta es articulada por las informantes sinohablantes que cuando la producen las informantes del grupo de control. Sin embargo, se observa que cuando el sonido posterior es una vibrante múltiple, la duración de la vocal es superior cuando la producen las hablantes cuya lengua materna es el español, que

cuando la realizan las estudiantes chinas, aunque ha de tenerse en cuenta que estas últimas hablantes han producido la vibrante múltiple en posición posterior a la vocal /o/ en una única ocasión. Cabe señalar, además, que si bien en el grupo de control no hay casos de oclusiva sonora tras la vocal en cuestión, las informantes sinohablantes han realizado este sonido consonántico hasta en 11 ocasiones.

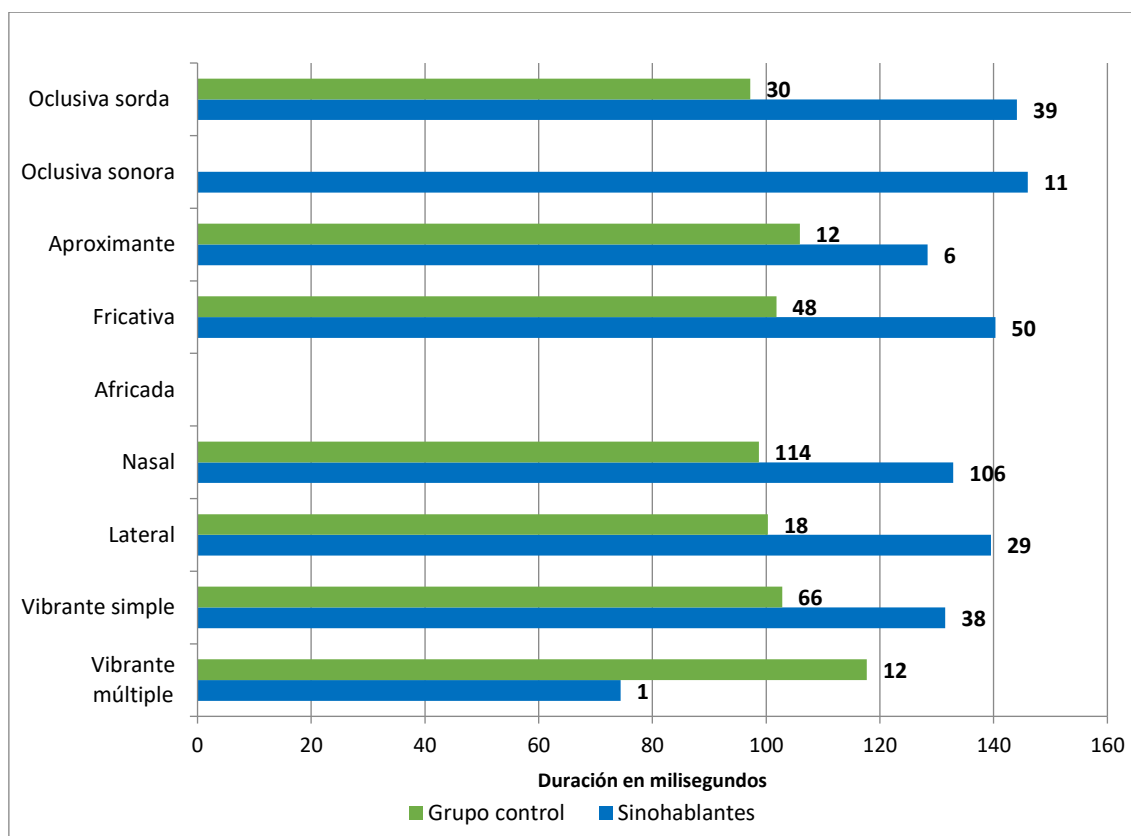


Figura 83. Comparación entre el grupo de control y el de informantes sinohablantes de la duración de la vocal /o/ según el entorno posterior.

Como se ha indicado, exceptuando el caso en el que la vocal /o/ precede a una vibrante múltiple, la duración media de este sonido vocálico producido por las informantes sinohablantes es mayor que cuando lo realiza el grupo de control. En el caso de las hablantes cuya lengua materna es el español, la duración media oscila entre 97 y 111 ms., mientras que en el de las estudiantes chinas, entre 74 y 146 ms.

En cuanto a la duración de /o/ articulada por el primer grupo, el orden de vocales más breves a más largas según el entorno posterior es el siguiente: oclusiva sorda, nasal, lateral, fricativa, vibrante simple, aproximante y vibrante múltiple. En cambio, este orden es diferente cuando la vocal en cuestión es realizada por las informantes sinohablantes: vibrante múltiple, aproximante, vibrante simple, nasal, lateral, fricativa, oclusiva sorda y oclusiva sonora.



<b>Comparación de la duración de la vocal /o/ según el entorno posterior a la vocal</b>							
	<b>Grupo de control</b>			<b>Informantes sinohablantes</b>			<b>Nivel de significación</b>
	<b>Casos</b>	<b>DM</b>	<b>DS</b>	<b>Casos</b>	<b>DM</b>	<b>DS</b>	
<b>Oclusiva sorda</b>	30	97.2	24.4	39	144.1	47.9	< 0.001
<b>Oclusiva sonora</b>	0	--	--	11	146.0	39.5	--
<b>Aproximante</b>	12	105.9	25.6	6	128.4	20.3	0.225
<b>Fricativa</b>	48	101.8	22.1	50	140.3	31.0	< 0.001
<b>Africada</b>	0	--	--	0	--	--	--
<b>Nasal</b>	114	98.7	17.9	106	132.9	39.8	< 0.001
<b>Lateral</b>	18	100.3	15.3	29	139.5	39.1	< 0.001
<b>Vibrante simple</b>	66	102.8	17.4	38	131.5	31.0	< 0.001
<b>Vibrante múltiple</b>	12	111.7	24.2	1	74.4	--	0.057

*Tabla 69.* Duración media y desviación estándar de las vocal /o/ según el entorno posterior articulada por los dos grupos, y nivel de significación.

La desviación estándar oscila entre 15 y 25 en las vocales realizadas por el grupo de control, y entre 20 y 47 cuando /o/ es producida por las estudiantes chinas. Esto muestra que esta vocal tiene una duración más homogénea cuando es articulada por el primer grupo que por el segundo.

En cuanto al nivel de significación de los resultados obtenidos en el análisis de los dos grupos, se observa que las diferencias son significativas en todos los casos excepto cuando el sonido posterior a /o/ es una aproximante (0.225) o una vibrante múltiple. Sin embargo, cabe señalar que en este caso el nivel de significación es 0.057, es decir, está en el límite del nivel de significación establecido en este trabajo para tomar las decisiones estadísticas. Por este motivo, se puede asumir que, en general, existen diferencias significativas entre ambos grupos y que, por lo tanto, la lengua materna de las informantes sinohablantes interfiere en la duración de /o/.

### 3.3.2.5 Vocal /u/

La figura 84 muestra la duración media de /e/ según el entorno posterior producida por el grupo de control y por las informantes sinohablantes. Tal y como se puede observar, la duración es mayor cuando esta vocal es producida por las hablantes del segundo grupo que por las del primero, aunque en el caso de /u/ seguida por oclusiva sonora la duración

media de ambos grupos es bastante similar. Además, como se puede observar, en el caso de esta vibrante múltiple posterior a esta vocal, mientras que hay 6 casos correspondientes al grupo de control, no hay ninguna realización por parte de las estudiantes chinas.

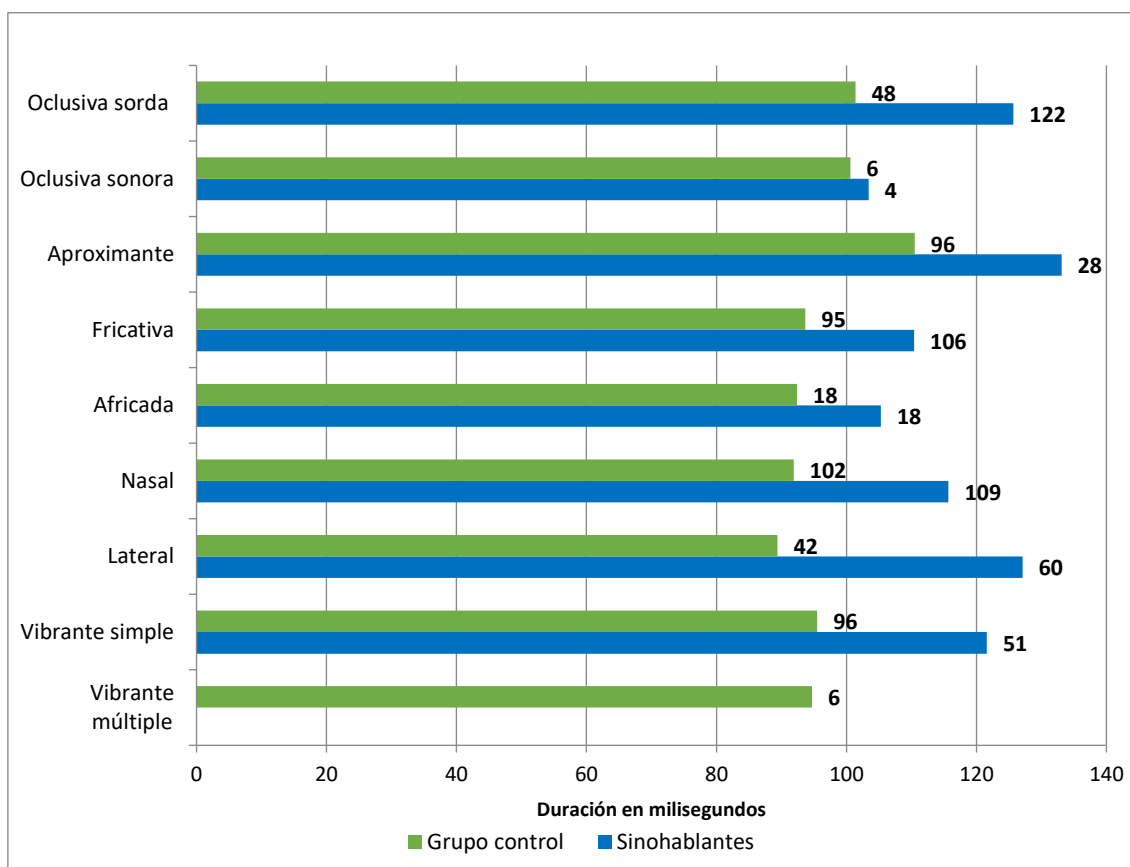


Figura 84. Comparación entre el grupo de control y el de informantes sinohablantes de la duración de la vocal /u/ según el entorno posterior.

Cuando la vocal /u/ es producida por las hablantes del grupo de control, la duración media oscila entre 89 y 110 ms., mientras que si es articulada por las informantes sinohablantes, dicha duración es mayor, y oscila entre 103 y 133 ms. En cuanto a la duración de esta vocal articulada por el primer grupo, el orden de vocales más breves a más largas según el entorno posterior es el siguiente: lateral, nasal, africada, fricativa, vibrante múltiple, vibrante simple, oclusiva sonora, oclusiva sorda y aproximante. En el caso del segundo grupo, el orden es diferente: oclusiva sonora, africada, fricativa, nasal, vibrante simple, oclusiva sorda, lateral y aproximante.

Comparación de la duración de la vocal /u/ según el entorno posterior a la vocal							
	Grupo de control			Informantes sinohablantes			Nivel de significación
	Casos	DM	DS	Casos	DM	DS	
Oclusiva sorda	48	101.4	24.0	122	125.7	38.4	<0.001
Oclusiva sonora	6	100.6	24.3	4	103.4	31.0	<0.001
Aproxi-mante	96	110.5	27.5	28	133.1	31.3	0.002
Fricativa	95	93.6	21.3	106	110.4	39.7	0.003
Africada	18	92.4	10.0	18	105.3	29.9	0.137
Nasal	102	91.9	19.0	109	115.7	34.6	<0.001
Lateral	42	89.4	12.5	60	127.1	41.2	<0.001
Vibrante simple	96	95.5	18.3	51	121.6	34.0	<0.001
Vibrante múltiple	6	94.7	--	0	--	--	0.150

Tabla 70. Duración media y desviación estándar de las vocal /u/ según el entorno posterior articulada por los dos grupos, y nivel de significación.

La desviación estándar oscila entre 10 y 27 en las vocales articuladas por las hablantes hispanohablantes nativas, y entre 29 y 41 cuando /u/ es articulada por las hablantes cuya lengua materna es el chino. Esto muestra que esta vocal tiene una duración más homogénea cuando es articulada por el primer grupo que por el segundo.

En cuanto al nivel de significación, los resultados muestran que las diferencias son significativas en todos los casos excepto cuando el sonido posterior a /u/ es una africada o una vibrante múltiple. Estos datos permiten asumir que existen diferencias significativas entre ambos grupos y que, por lo tanto, la lengua materna de las informantes sinohablantes interfiere en la duración de /u/, exceptuando aquellos casos en los que el sonido posterior es una africada o una vibrante múltiple.

### 3.3.3 Comparación de la duración de las vocales cuando se encuentran en posición final de palabra

La figura 85 muestra que, igual que se ha comprobado en los apartados en los que se analiza la duración de las vocales del español según el entorno anterior y el posterior, los sonidos vocálicos producidos por las informantes sinohablantes tienen una mayor duración que los realizados por las hablantes que conforman el grupo de control. De la misma manera que sucede en los apartados anteriores, en este también hay diferencias en lo que al número de casos se refiere, ya que las informantes sinohablantes en muchas

ocasiones eliden el sonido consonántico del final de palabra, de manera que hay más vocales en posición final de palabra aislada.

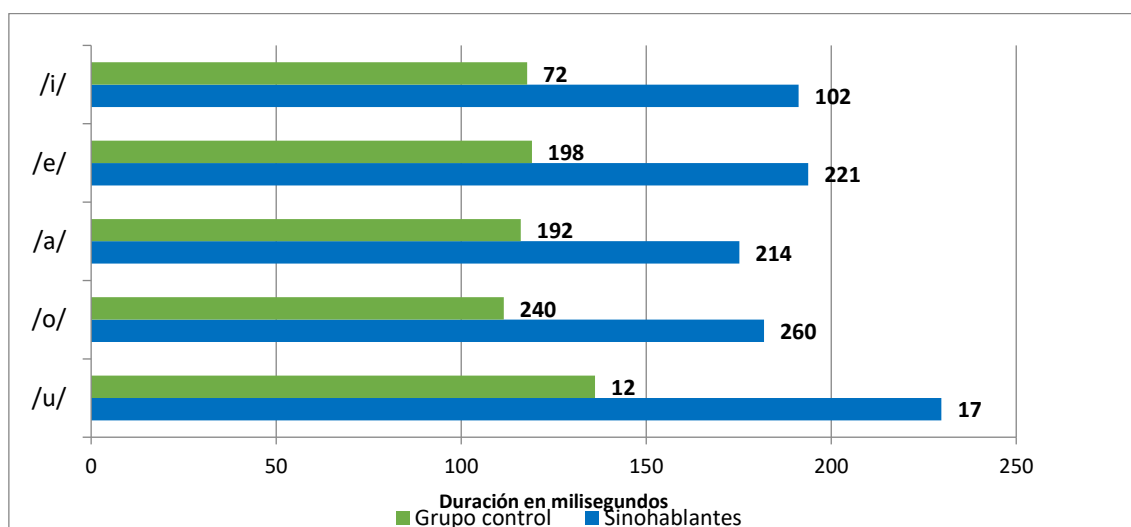


Figura 85. Comparación entre el grupo de control y el de informantes sinohablantes de la duración de las vocales del español en posición final de palabra aislada.

En el caso de las hablantes que conforman el grupo de control, la duración media oscila entre 111 y 136 ms., y la duración de las vocales sigue, de más breves a más largas, el siguiente orden: /o/, /a/, /i/, /e/, /u/. En el caso de las informantes sinohablantes, la duración media oscila entre 175 y 229 ms. Aquí, el orden, también de más breves a más largas, es el siguiente: /a/, /o/, /i/, /e/, /u/.

Comparación de la duración de las vocales en posición final de palabra							
	Grupo de control			Informantes sinohablantes			Nivel de significación
	Casos	DM	DS	Casos	DM	DS	
/i/	72	117.8	26.9	102	191.2	62.5	<0.001
/e/	198	119.1	24.1	221	193.8	66.3	<0.001
/a/	192	116.1	21.3	214	175.2	52.7	<0.001
/o/	240	111.5	24.2	260	181.8	62.7	<0.001
/u/	12	136.2	16.9	17	229.8	39.7	<0.001

Tabla 71. Duración media y desviación estándar de las vocales en posición de final de palabra aislada articuladas por los dos grupos, y nivel de significación.

La desviación estándar oscila entre 16 y 26 en las vocales realizadas por el grupo de control, y entre 39 y 66 cuando los sonidos vocálicos son producidos por las estudiantes chinas. Estos datos muestran que la duración de las distintas vocales del español es poco homogénea cuando estas son producidas por las informantes sinohablantes, mientras que hay menor variación cuando las articulan las hablantes del grupo de control.

En cuanto al nivel de significación de los datos obtenidos en el análisis de los dos grupos, se observa que las diferencias son significativas en todos los casos, por lo que se puede asumir que la lengua materna de las informantes sinohablantes interfiere en la duración de las vocales del español cuando estas se encuentran en posición final de palabra aislada.

### 3.4 Estructura silábica

En este apartado se comparan los resultados extraídos del análisis acústico de las grabaciones correspondientes a las informantes sinohablantes con los datos obtenidos en el estudio del grupo de control. Esta comparación se realiza en función de la estructura silábica. Como se ha mostrado en la primera parte de este capítulo, las informantes sinohablantes realizan una serie de estructuras que no son las propias de las palabras que conforman el corpus de estudio de esta investigación. Sin embargo, las estructuras que se comparan en este apartado son las propias del español presentes en este corpus: CV-CV, CV-CVC y CVC-CV.

#### 3.4.1 Duración de la vocal según la sílaba en la que se encuentra

Tal y como se puede observar en la figura 86, cuando la vocal analizada se encuentra en sílaba CV, la duración media de las vocales articuladas por las informantes sinohablantes es considerablemente superior a aquellas realizadas por el grupo de control, excepto en el caso de la vocal /u/, en el que se detecta que la duración media es igual en ambos grupos.

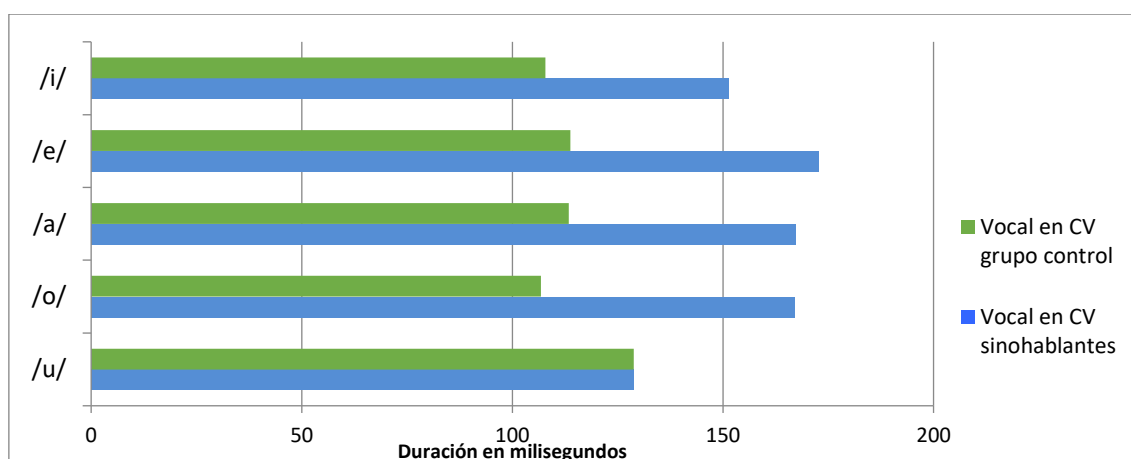


Figura 86. Comparación entre el grupo de control y el de informantes sinohablantes de la duración de la vocal en sílaba cuya estructura es CV.

En cuanto a la duración media de las vocales cuando estas se encuentran en sílaba CVC, las informantes sinohablantes producen sonidos vocálicos con una duración superior a los realizados por las hablantes nativas de español, tal y como se puede observar en la figura 87.

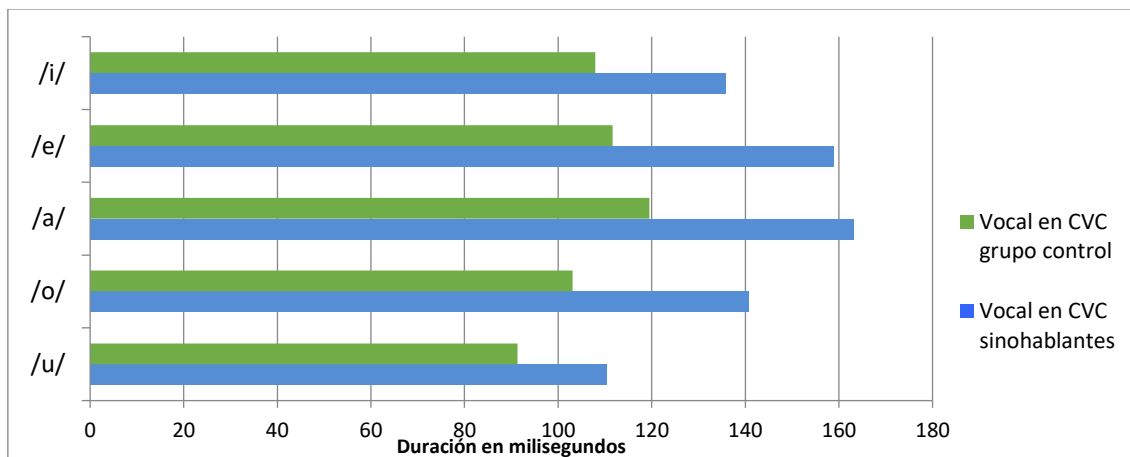


Figura 87. Comparación entre el grupo de control y el de informantes sinohablantes de la duración de la vocal en sílaba cuya estructura es CVC.

De acuerdo con los datos mostrados en la tabla 72, la duración media de las vocales producidas por el grupo de control va desde los 106 ms. hasta los 128 ms. cuando se encuentran en sílaba CV, y de 91 a 119 ms. en sílaba CVC. En cambio, tal y como se ha indicado, esta duración es superior cuando las vocales son realizadas por las informantes sinohablantes, ya que la duración media va desde 128 ms. hasta 172 ms. cuando la vocal se encuentra en sílaba CV, y desde 110 ms. hasta 163 ms. cuando se la estructura de la sílaba es CVC.

Comparación de la duración de la vocal según la sílaba en la que se encuentra								
Vocal	Estructura silábica	Grupo de control			Informantes sinohablantes			Nivel de significación
		Casos	DM	DS	Casos	DM	DS	
/i/	CV	306	107.8	22.3	306	151.3	62.9	< 0.001
	CVC	180	107.9	25.1	155	135.7	43.1	< 0.001
/e/	CV	360	113.7	24.2	359	172.8	67.2	< 0.001
	CVC	180	111.6	21.2	146	158.8	50.9	< 0.001
/a/	CV	360	113.4	21.1	381	167.3	53.7	< 0.001
	CVC	180	119.5	21.1	146	163.1	44.4	< 0.001
/o/	CV	360	106.7	23.6	390	166.9	60.2	< 0.001
	CVC	180	103.0	17.9	129	140.7	39.5	< 0.001
/u/	CV	324	102.3	20.0	300	128.8	42.7	< 0.001
	CVC	180	91.3	17.3	165	110.2	36.7	< 0.001

Tabla 72. Duración media y desviación estándar de las vocales en sílaba con estructura CV o CVC articuladas por los dos grupos, y nivel de significación.

La tabla 72 también permite observar que en todos los casos la desviación estándar es superior en las vocales producidas por las hablantes cuya lengua materna es el chino que en las de las hispanohablantes nativas. En el caso de las primeras, la desviación estándar puede oscilar entre 36 y 67, mientras que en el de las segundas entre 17 y 25, lo que indica que las vocales realizadas por las informantes sinohablantes presentan menor homogeneidad en la duración que las realizadas por el grupo de control.

El nivel de significación de estos datos es, en todos los casos  $< 0.001$ , lo que muestra que existen diferencias significativas en la realización de las vocales del español (tanto en sílaba CV como en sílaba CVC) entre el grupo de control y las informantes sinohablantes.

### 3.4.2 Duración de la vocal cuando está en sílaba inicial

La figura 88 muestra la diferencia entre el grupo de control y las informantes sinohablantes en la duración media de las vocales del español cuando se encuentran en sílaba inicial de palabra. Tal y como se puede observar, las hablantes cuya lengua materna es el chino producen vocales cuya duración es mayor que las hablantes del grupo de control.

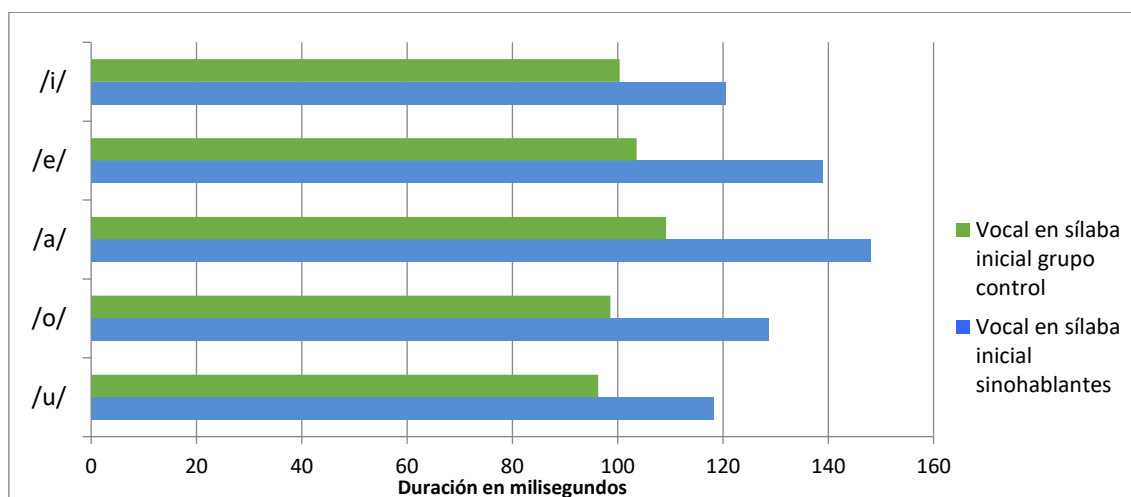


Figura 88. Comparación entre el grupo de control y el de informantes sinohablantes de la duración de la vocal cuando se encuentra en sílaba inicial.

De acuerdo con los datos mostrados en la tabla 73, cuando las vocales se encuentran en sílaba inicial de palabra la duración media va desde 96 ms. hasta 109 ms. en el caso del grupo de control, y desde 118 ms. hasta 148 ms. en el caso de las informantes sinohablantes.

Comparación de la duración de la vocal cuando está en sílaba inicial							
Vocal	Grupo de control			Informantes sinohablantes			Nivel de significación
	Casos	DM	DS	Casos	DM	DS	
/i/	270	100.4	18.1	306	120.4	35.7	< 0.001
/e/	270	103.5	18.6	248	139.0	48.7	< 0.001
/a/	270	109.2	15.9	272	148.0	38.0	< 0.001
/o/	204	98.6	17.3	207	128.7	35.2	< 0.001
/u/	270	96.3	18.2	480	118.2	36.6	< 0.001

Tabla 73. Duración media y desviación estándar de las vocales en sílaba inicial articuladas por los dos grupos, y nivel de significación.

En cuanto a la desviación estándar, esta es mayor en el caso de las informantes sinohablantes (entre 35 y 48.) que en el del grupo de control (entre 15 y 18), lo que indica que las vocales realizadas por las informantes chinas presentan menor homogeneidad que las de las hablantes nativas de español.

En relación con el nivel de significación, en todos los casos es < 0.001, lo que indica que las diferencias entre ambos grupos en la duración de las vocales cuando se estas se encuentran en sílaba inicial de palabra es significativa.

### 3.4.3 Duración de la vocal según la sílaba anterior

La comparación de la duración media de las vocales producidas por las hablantes de los dos grupos según la sílaba anterior, la cual aparece en las figuras 89 y 90, permite observar que las vocales producidas por las informantes sinohablantes tienen siempre una duración mayor que las realizadas por el grupo de control. Este fenómeno se da en ambos casos, tanto si la silaba anterior tiene estructura CV (figura 89), como si dicha sílaba tiene estructura CVC (figura 90).

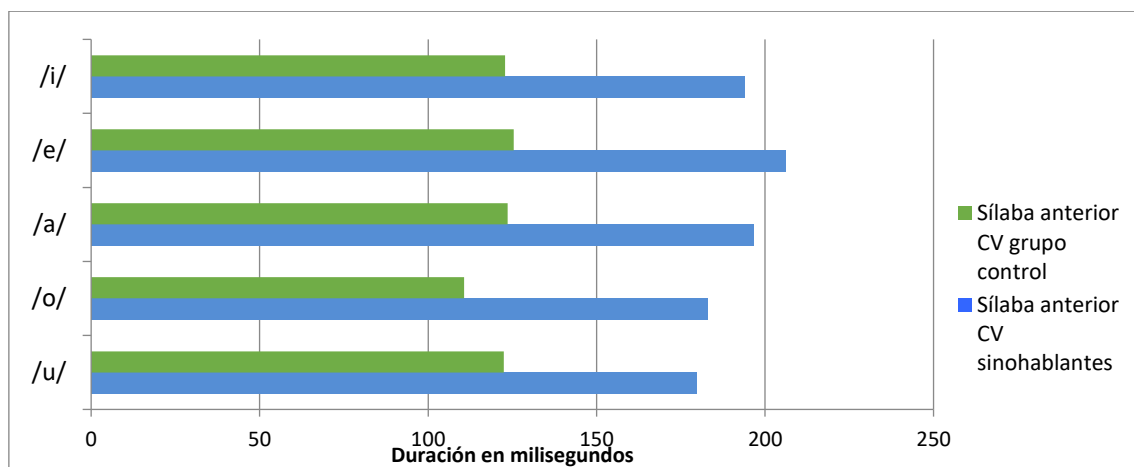


Figura 89. Comparación entre el grupo de control y el de informantes sinohablantes de la duración de la vocal cuando la sílaba anterior es CV.



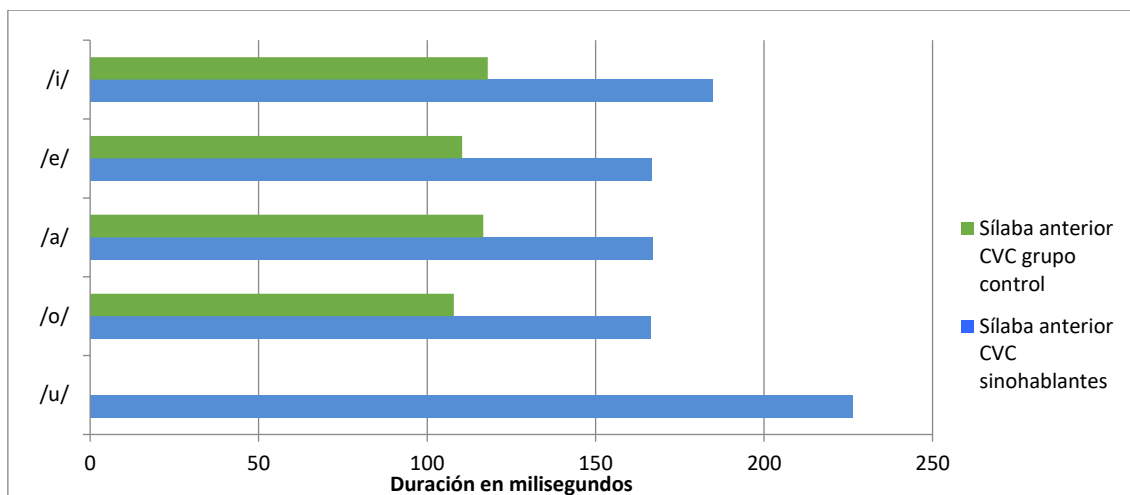


Figura 90. Comparación entre el grupo de control y el de informantes sinohablantes de la duración de la vocal cuando la sílaba anterior es CVC.

La tabla 74 muestra los datos detallados respecto a la comparación de los resultados del análisis acústico de ambos grupos. Cuando las vocales son realizadas por las hablantes del grupo de control, la duración media de estas oscila entre 110 y 123 ms. cuando la sílaba anterior es CV, y 107 y 118 si dicha sílaba tiene estructura CVC. En cambio, cuando las informantes realizan dichos sonidos, la duración media va desde 179 a 206 ms. cuando la estructura de la sílaba anterior es CV, y 166 y 226 cuando es CVC. Cabe señalar que en el corpus no hay casos de /u/ en posición posterior a sílaba CVC, pero las informantes sinohablantes han realizado dicha estructura anterior a /u/ en 2 ocasiones.

		Comparación de la duración de la vocal según la sílaba anterior						Nivel de significación
		Grupo de control			Informantes sinohablantes			
Vocal	Estructura silábica	Casos	DM	DS	Casos	DM	DS	
/i/	CV	132	122.8	23.8	118	193.8	61.7	< 0.001
	CVC	42	118.0	22.7	38	184.8	54.9	< 0.001
/e/	CV	180	125.4	22.0	212	206.0	63.7	< 0.001
	CVC	90	110.4	21.0	74	166.5	54.4	< 0.001
/a/	CV	180	123.5	22.3	177	196.5	56.5	< 0.001
	CVC	90	116.6	19.2	78	166.9	49.7	< 0.001
/o/	CV	180	110.6	24.4	221	183.0	59.2	< 0.001
	CVC	90	107.9	20.5	94	166.4	54.8	< 0.001
/u/	CV	36	122.4	21.8	34	179.6	58.0	< 0.001
	CVC	0	0	0	2	226.3	11.3	--

Tabla 74. Duración media y desviación estándar de las vocales según la estructura anterior articuladas por los dos grupos, y nivel de significación.

La desviación estándar mostrada en la tabla 74 es menor en el caso del grupo de control (entre 19 y 24) que en el de las informantes sinohablantes (entre 11 y 63), lo que indica que las realizaciones del grupo de control presentan mayor homogeneidad que las del grupo de estudiantes sinohablantes.

En cuanto al nivel de significación, este es  $< 0.001$  en todos los casos, lo que indica que las diferencias entre ambos grupos son significativas.

### 3.4.4 Duración de la vocal según la sílaba posterior

La comparación entre la duración media de las vocales según la estructura de la sílaba posterior mostrada en las figuras 91 y 92 muestra que las vocales producidas por el grupo de informantes sinohablantes tienen mayor duración que las realizadas por el grupo de control. Esta diferencia se puede observar tanto si la sílaba posterior tiene estructura CV (figura 91) como si tiene estructura CVC (figura 92).

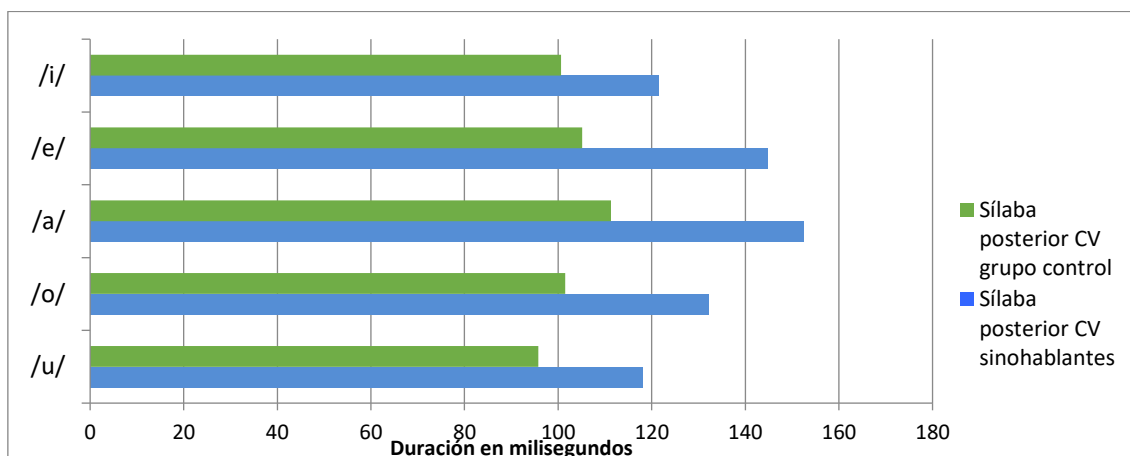


Figura 91. Comparación entre el grupo de control y el de informantes sinohablantes de la duración de la vocal cuando la sílaba posterior es CV.

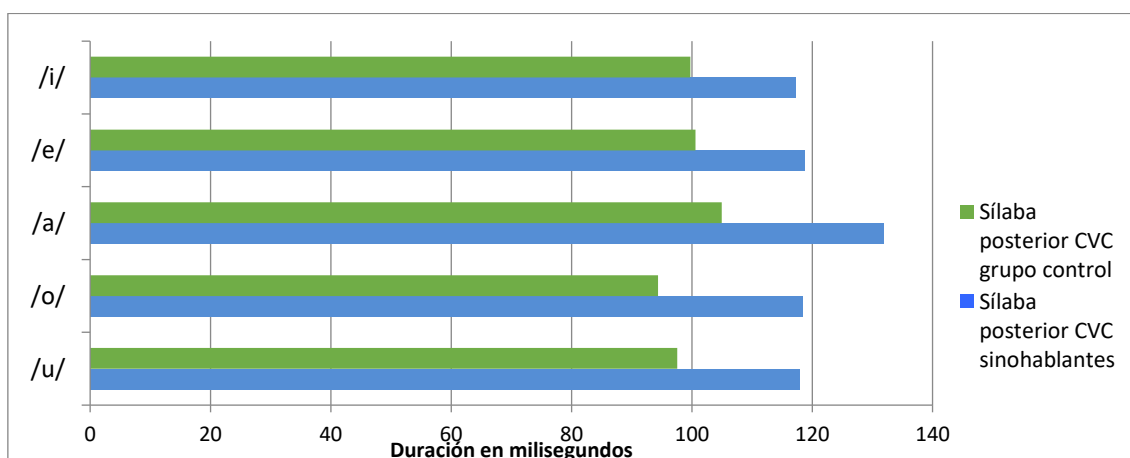


Figura 92. Comparación entre el grupo de control y el de informantes sinohablantes de la duración de la vocal cuando la sílaba posterior es CVC.

En el caso del grupo de control, tal y como se puede observar en la tabla 75, la duración media oscila entre 95 y 111 ms. si la sílaba posterior tiene estructura CV, y 94 y 104 ms. si dicha sílaba es CVC. En el caso de las informantes sinohablantes, dicha duración es mayor, y va desde 117 hasta 152 ms. cuando la sílaba posterior es CV, y 117 y 131 ms. si es CVC.

Comparación de la duración de la vocal según la sílaba posterior								
Vocal	Estructura silábica	Grupo de control			Informantes sinohablantes			Nivel de significación
		Casos	DM	DS	Casos	DM	DS	
/i/	CV	90	100.6	17.8	247	121.4	35.0	< 0.001
	CVC	78	99.7	18.0	56	117.1	38.6	0.004
/e/	CV	90	105.1	19.1	170	144.7	49.9	< 0.001
	CVC	108	100.5	16.5	66	118.7	23.2	< 0.001
/a/	CV	90	111.3	15.9	190	152.5	37.7	< 0.001
	CVC	108	104.9	14.6	64	131.9	32.5	< 0.001
/o/	CV	90	101.5	17.0	128	132.0	34.3	< 0.001
	CVC	96	94.4	15.5	65	118.4	35.3	< 0.001
/u/	CV	90	95.7	18.6	368	117.9	37.2	< 0.001
	CVC	132	97.5	17.2	100	117.9	33.0	< 0.001

Tabla 75. Duración media y desviación estándar de las vocales según la estructura posterior articuladas por los dos grupos, y nivel de significación.

La desviación estándar de las vocales producidas por el grupo de control no es elevada (entre 14 y 19), mientras que en el caso de las informantes sinohablantes es mayor (entre 23 y 49), lo que indica que las realizaciones del segundo grupo son menos homogéneas y que la variabilidad es mayor en caso de las hablantes chinas.

En cuanto al nivel de significación, en todos los casos es < 0.001, excepto cuando la estructura posterior a la vocal /i/ es CVC, en cuyo caso dicho nivel es de 0.004. No obstante, dado que las decisiones estadísticas se han realizado tomando como nivel de significación el valor 0.05 y en todos los casos el valor obtenido es menor, se considera que las diferencias son estadísticamente significativas.

### 3.4.5 Duración de la vocal cuando está en sílaba final

En último lugar, se compara la duración de las vocales producidas por ambos grupos cuando estas se encuentran en sílaba final de palabra. Tal y como se puede

observar en la figura 93, en todos los casos la duración media es mayor cuando las vocales son producidas por las informantes sinohablantes.

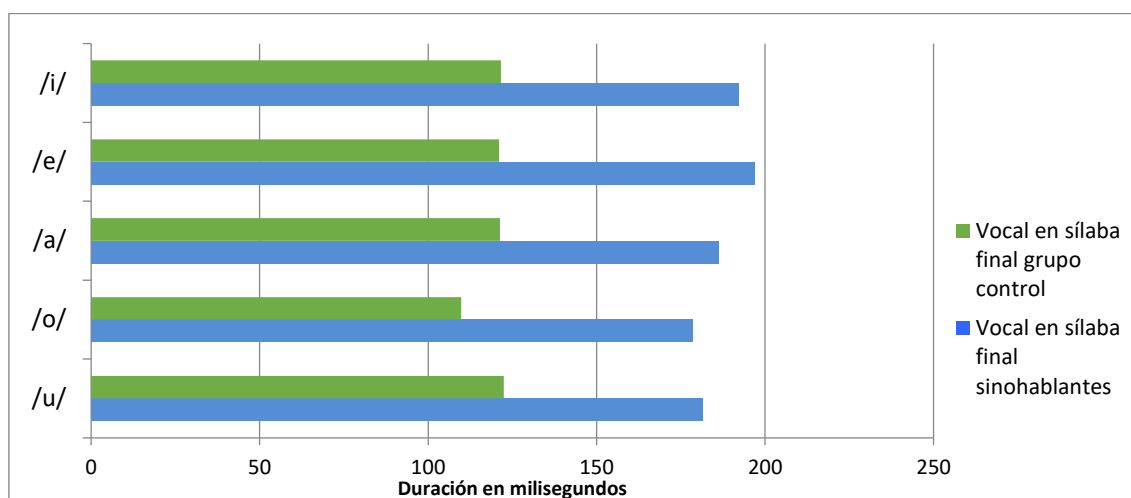


Figura 93. Comparación entre el grupo de control y el de informantes sinohablantes de la duración de la vocal cuando se encuentra en sílaba final.

En el caso del grupo de control, exceptuando la vocal /o/, que tiene una duración media menor, de 109 ms., la duración media de las otras vocales es muy similar, de 121-122 ms. Dicha duración es mayor en el caso de las informantes sinohablantes, ya que dicha duración va desde 178 ms. hasta 196 ms.

Comparación de la duración de la vocal cuando está en sílaba final							
Vocal	Grupo de control			Informantes sinohablantes			Nivel de significación
	Casos	DM	DS	Casos	DM	DS	
/i/	174	121.6	24.1	169	192.0	60.1	< 0.001
/e/	270	121.0	22.9	290	196.8	64.6	< 0.001
/a/	270	121.3	21.7	280	186.1	55.4	< 0.001
/o/	270	109.8	23.6	336	178.4	58.6	< 0.001
/u/	42	122.4	21.8	43	181.3	56.8	< 0.001

Tabla 76. Duración media y desviación estándar de las vocales en sílaba final articuladas por los dos grupos, y nivel de significación.

Como se puede observar en la tabla 76, la desviación estándar es mayor en el caso del grupo de estudiantes sinohablantes, lo que muestra que las realizaciones de estas hablantes son menos homogéneas y presentan mayor variabilidad que las de las hablantes cuya lengua materna es el español. En cuanto al nivel de significación, en todos los casos es < 0.001, lo que indica que las diferencias entre ambos grupos son estadísticamente significativas.

#### **4. RECAPITULACIÓN**

La comparación de las vocales realizadas por las hablantes que conforman el grupo de control y las informantes sinohablantes ha permitido observar que la duración de las vocales articuladas por las segundas es mayor en todos los casos, y dicha diferencia entre ambos grupos es estadísticamente significativa, lo que permite rechazar la hipótesis nula, y asumir que la lengua materna de las informantes sinohablantes interfiere en la duración de las vocales, analizadas según la estructura silábica, producidas por estas hablantes.

En este apartado dedicado a la influencia de las variables cualitativas independientes en la duración de las vocales del español producidas por sinohablantes, se han puesto de manifiesto dos fenómenos importantes:

Por un lado, como se ha indicado, se ha comprobado que las vocales producidas por las estudiantes chinas son más largas que las articuladas por las hablantes nativas de español.

Por otro lado, se ha mostrado que la desviación en la duración es superior en el caso de las informantes sinohablantes, lo que indica que sus realizaciones son menos homogéneas y presentan mayor variabilidad que las de las hablantes del grupo de control.

Ambos fenómenos están relacionados, sin duda, con las características de la lengua materna de las hablantes analizadas: los sonidos vocálicos de una lengua tonal como el chino presentan, forzosamente, una duración mayor que los propios de una lengua no tonal (como el español). La realización de los tonos, es decir, los cambios de timbre durante la realización de una vocal requieren que el sonido presente una duración que se corresponda con los movimientos articulatorios que producen los cambios de tono. Por tanto, los elementos vocálicos son más largos en una lengua tonal.

Por otra parte, la mayor variabilidad se explica por el hecho ya conocido de que el hablante de una lengua extranjera no se siente seguro de sus realizaciones en la misma de forma que, en su acercamiento a las realizaciones de la lengua meta, utiliza estrategias de producción diferentes para tratar de producir sonidos lo más parecidos posible a los que realizan los nativos y ello se refleja en la mayor variabilidad que presentan sus realizaciones.

**Así, una primera observación importante que arrojan los datos de duración analizados es que las vocales realizadas por los sinohablantes durante su proceso de aprendizaje del español difieren considerablemente de las realizadas por los hispanohablantes nativos en el hecho de que realizan transferencias de algunas características del mandarín como la mayor duración de las vocales (obligatoria en chino debido a la necesidad de realizar los tonos).**

**Como se verá en el siguiente apartado de resultados, la mayor duración conducirá a que los timbres de las vocales del español de las sinohablantes presenten variaciones en el seno de las mismas pues es muy difícil mantener la homogeneidad en el timbre cuando los valores de duración son tan elevados.**

## **TERCERA PARTE: EL TIMBRE VOCÁLICO**

Como se ha visto en el capítulo III, en el análisis auditivo de las vocales se detectó un timbre inestable en los sonidos vocálicos producidos por las informantes sinohablantes. A continuación se muestran los resultados del estudio del análisis acústico del timbre de las vocales del español articuladas por los dos grupos de informantes estudiados. Como se ha indicado en el capítulo III, se han analizado los dos primeros formantes de las vocales (F1 y F2) en tres puntos: el inicio, el centro y el final. El análisis se ha llevado a cabo de la misma forma tanto para el grupo de control como para el de informantes sinohablantes, de manera que se puedan comparar los resultados de ambos grupos. En primer lugar, se muestran los datos relativos al grupo de control, en segundo lugar, los de las informantes sinohablantes y, en tercer lugar, se comparan los resultados en ambos grupos.

### **1. GRUPO DE CONTROL**

Los datos que se presentan permiten, por un lado, mostrar el comportamiento de las vocales producidas por el grupo de control y compararlas con los valores relativos y absolutos señalados en otros trabajos que analizan las vocales del español. Y, por otro lado, se ha realizado la prueba de Friedman sobre los valores obtenidos, la cual indica si las diferencias entre los tres puntos de medición son significativas. Finalmente, se han analizado las áreas de dispersión en los tres puntos de medición de las vocales del español producidas por las informantes cuya lengua materna es el español.

#### **1.1 Características de las vocales**

Como se ha indicado, los formantes de las vocales se han analizado en tres puntos. Sin embargo, dado que la bibliografía que analiza las vocales del español toma como referencia el centro de la vocal, para poder realizar adecuadamente la comparación, en este apartado se emplean únicamente los datos obtenidos en el análisis del punto medio de los dos primeros formantes para caracterizar las vocales producidas por el grupo de control, los cuales aparecen en la tabla 77.

Vocal	Formante	Frecuencia media (Hz)
/i/	F1	461.2
	F2	2555.8
/e/	F1	612.8
	F2	2199.7
/a/	F1	853.5
	F2	1695.3
/o/	F1	542.2
	F2	1262.9
/u/	F1	451.5
	F2	1083.8

Tabla 77. Frecuencia media en el punto medio de las vocales producidas por el grupo de control.

A continuación, se analizan las características del primer formante (F1), el cual “se corresponde articulatoriamente con la abertura oral” (RAE, 2011: 86). La figura 94 permite apreciar que las vocales producidas por estas hablantes tienen el mismo comportamiento que el indicado en RAE (2011: 80). Las vocales /i/ y /u/ presentan los valores más bajos, lo que indica que son cerradas. La altura a la que aparecen las vocales /e/ y /o/ es mayor que la de /i/ y /u/, pero menor que /a/, lo que muestra que la abertura oral es media, mientras que /a/, la cual tiene las frecuencias medias más elevadas, se realiza como una vocal abierta.

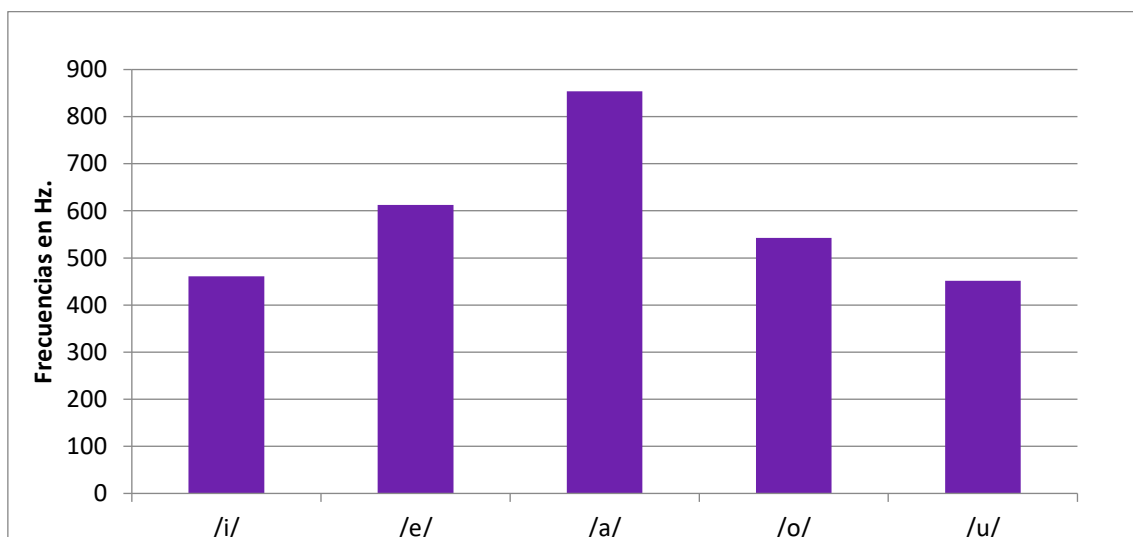


Figura 94. Frecuencias medias de F1 de las vocales producidas por el grupo de control.

El segundo formante (F2) permite conocer la posición de la lengua, y que los valores sean más altos indican que esta es más anterior (RAE, 2011: 87). Tal y como se puede observar en la figura 95 que muestra los valores correspondientes al F2, las vocales producidas por estas hablantes también tienen el mismo comportamiento que el indicado por RAE (2011: 87). Por un lado, los sonidos vocálicos que presentan las frecuencias más



bajas son /u/ y /o/, lo que indica que son vocales posteriores. Por otro lado, la frecuencia media en la que aparece /a/ es más elevada que en los casos de /u/ y /o/, por lo que esta vocal es más anterior. Finalmente, las vocales que aparecen en un rango de frecuencias más elevado son /e/ e /i/, lo que muestra que se trata de las vocales más anteriores del sistema.

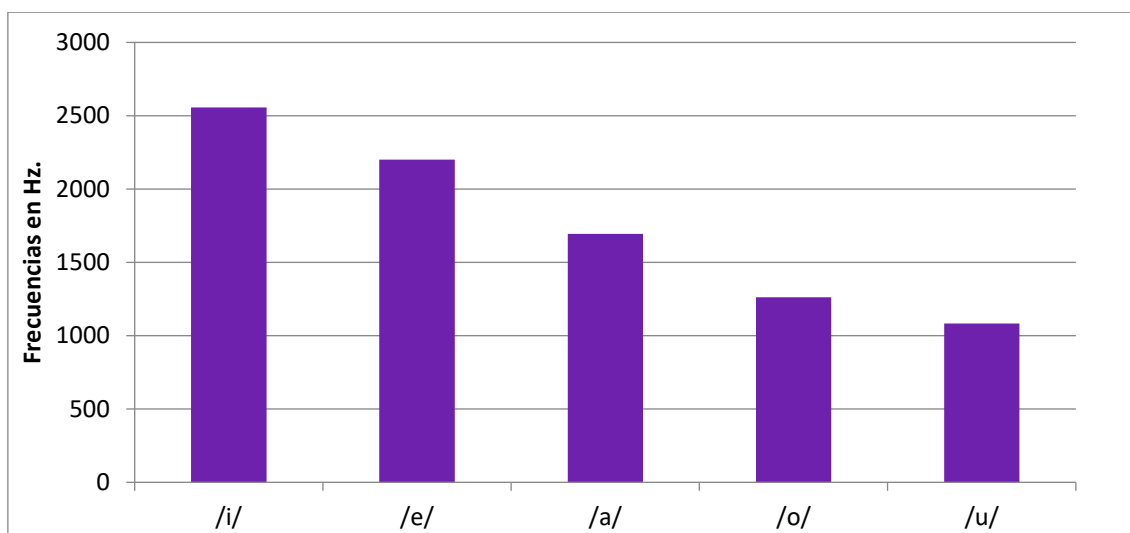


Figura 95. Frecuencias medias de F2 de las vocales producidas por el grupo de control.

Las vocales articuladas por las informantes cuya lengua materna es el español presentan los mismos valores relativos que los indicados por los trabajos de Martínez Celdrán (2007: 39) o RAE (2011: 86-87). Sin embargo, en las figuras 94 y 95 se puede apreciar que los valores absolutos obtenidos en esta investigación no coinciden con los de los citados trabajos y ello es debido al tipo de voz estudiada. Por este motivo, a continuación se comparan los valores relativos obtenidos en los trabajos de Martínez Celdrán (2007) y en RAE (2011), los cuales se centran su análisis en voces masculinas, con los de la investigación de Pérez García (2018) y con los de la presente tesis doctoral, los cuales estudian voces femeninas.

En relación con el primer formante de las vocales del español, la figura 96 muestra las frecuencias medias de las mismas según los trabajos de Martínez Celdrán (2007), en naranja, RAE (2011), en amarillo, Pérez García (2018), en morado, y la presente tesis doctoral, en rosa. En esta figura se observa que, en todos los casos, las vocales /i/ y /u/ son cerradas, ya que presentan los valores de F1 más bajos, /e/ y /o/ presentan una abertura oral media, y /a/ es la vocal más abierta. Sin embargo, las frecuencias medias en las que aparece el primer formante son siempre más altas cuando las voces son femeninas (Pérez

García, 2018 e Igarreta Fernández), que cuando se trata de voces masculinas (Martínez Celdrán, 2007 y RAE, 2011).

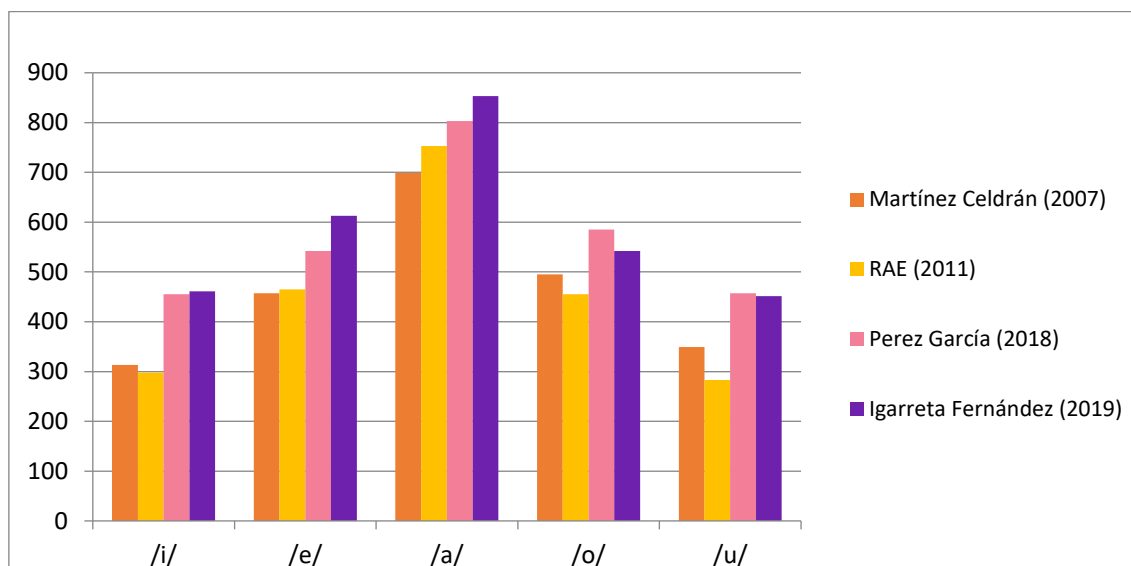


Figura 96. Valores medios de F1 de las vocales del español según los estudios de Martínez Celdrán (2007: 39), RAE (2011: 85), Pérez García (2018: 100) y el presente estudio.

En lo que al segundo formante se refiere, en la figura 96, la cual tiene la misma configuración que la figura 95, se observa que el comportamiento respecto a la posición de la lengua es el mismo tanto en voces masculinas como en voces femeninas, es decir, los valores relativos arrojan, cualitativamente, resultados similares. Sin embargo, igual que sucede con el primer formante, los valores absolutos no coinciden, ya que las frecuencias son más altas cuando las vocales son producidas por mujeres (Pérez García, 2018 e Igarreta Fernández).

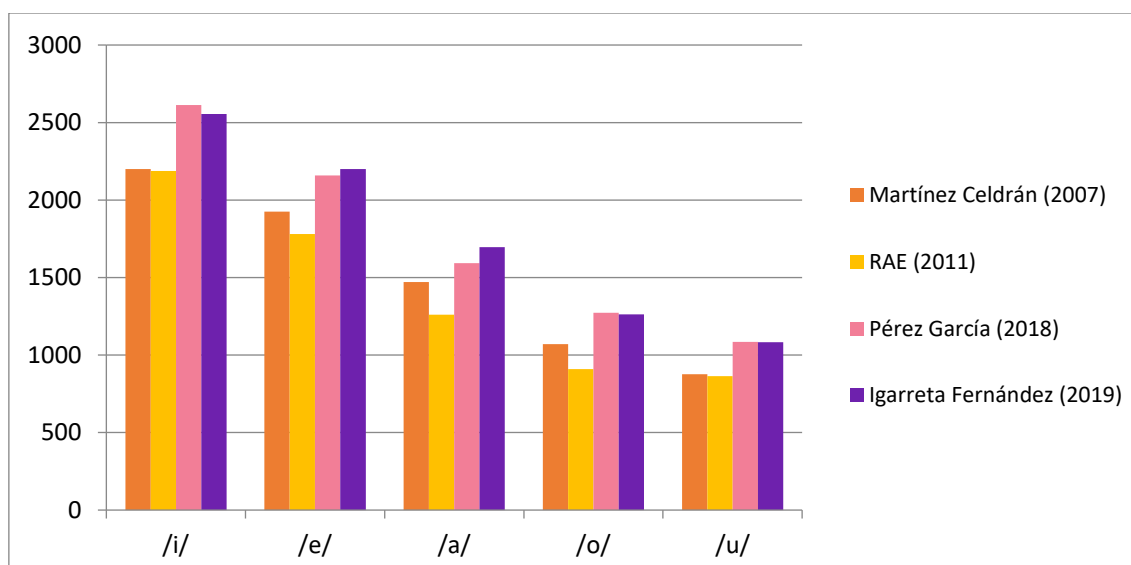


Figura 97. Valores medios de F2 de las vocales del español según los estudios de Martínez Celdrán (2007: 39), RAE (2011: 85), Pérez García (2018: 100) y el presente estudio.

Una vez analizados los valores relativos de los cuatro trabajos, la tabla 78 permite comparar los valores absolutos de los mismos, ya que, como se ha indicado, estos no coinciden cuando las voces son producidas por hombres o por mujeres. Las columnas ‘M.C’, ‘RAE’, ‘P.G.’ e ‘I.F’ corresponden a las obras de Martínez Celdrán (2007: 39), RAE (2011: 85), Pérez García (2018) y la presente tesis doctoral respectivamente.

Valores medios de frecuencias de los dos primeros formantes de las vocales del español según diferentes estudios																					
		[i]				[e]				[a]				[o]				[u]			
		M.C. (2007)	RAE (2011)	P.G. (2018)	I.F. (2019)	M.C. (2007)	RAE (2011)	P.G. (2018)	I.F. (2019)	M.C. (2007)	RAE (2011)	P.G. (2018)	I.F. (2019)	M.C. (2007)	RAE (2011)	P.G. (2018)	I.F. (2019)	M.C. (2007)	RAE (2011)	P.G. (2018)	I.F. (2019)
F <sub>1</sub>		313	298	454,8	461,2	457	465	541,9	612,8	699	753	803,5	853,5	495	455	585,3	542,2	349	283	456,7	451,5
F <sub>2</sub>		2200	2188	2613,4	2555,8	1926	1780	2159,1	2199,7	1471	1260	1591,8	1695,3	1070	910	1272,4	1262,9	877	865	1084,3	1083,8

Tabla 78. Comparación de valores medios de los formantes F1 y F2 de las vocales del español según los estudios de Martínez Celdrán (2007: 39), RAE (2011: 85), Pérez García (2018: 100) y el presente estudio.

Tal y como se puede observar, de acuerdo con los trabajos de Martínez Celdrán (2007: 39) y RAE (2011: 85), la frecuencia media en la que aparece el F1 de /i/ se sitúa en 313 y 298 Hz respectivamente, mientras que en el trabajo de Pérez García (2018: 100) y la presente tesis sitúan el primer formante de esta vocal en 454 y 461 Hz respectivamente. Tomando como referencia las medias indicadas por RAE (2011) y las obtenidas en esta investigación, para las voces femeninas las frecuencias en las que aparece el primer formante son un 54% más altas. La frecuencia media del segundo formante también es más elevada en el caso de los trabajos que se centran en voces femeninas, ya que en estos casos la media es de 2613 (Pérez García, 2018: 100) y 2555 Hz (presente trabajo), mientras que según Martínez Celdrán (2007: 39) la frecuencia media es de 2200 Hz y en el de RAE (2011: 85) es de 2188 Hz. Es decir, en lo que al segundo formante se refiere, y tomando como referencia los trabajos de RAE (2011) y la presente tesis doctoral, las frecuencias de las voces femeninas son un 16% más altas que las masculinas.

En cuanto a la vocal /e/, la frecuencia media del primer formante es de 457 y 465 Hz según Martínez Celdrán (2007: 39) y RAE (2011: 85), mientras que dichos valores son más elevados cuando las voces analizadas son femeninas, ya que ascienden a 541 Hz en el trabajo de Pérez García (2018: 100) y 612 Hz en esta investigación. Es decir, las voces femeninas presentan unas frecuencias medias un 31% superiores a las de las voces masculinas. El mismo fenómeno sucede en relación con el F2, ya que las frecuencias

medias son de 1926 y 1780 Hz en los dos primeros trabajos, y de 2159 y 2199 Hz en ellos segundos. En este caso, las frecuencias en las que aparecen las voces femeninas son un 23% superiores a las voces masculinas.

La vocal /a/ también presenta diferencias en función de si las vocales son producidas por voces masculinas o femeninas. En el caso de las primeras, la frecuencia media de F1 es de 699 Hz (Martínez Celdrán, 2007: 39) y 753 Hz (RAE, 2011: 85), mientras que en el de las segundas, la media asciende a 803 Hz (Pérez García, 2018: 100) y 853 Hz (presente investigación). Es decir, las vocales producidas por voces femeninas presentan unas frecuencias medias de F1 un 13% superiores a las de las voces masculinas. Lo mismo sucede con los valores en los que aparece el F2: 1471 y 1260 Hz en los trabajos de Martínez Celdrán (2007: 39) y RAE (2011: 85), y 1591 y 1695 Hz en los de Pérez García (2018: 100) y la presente tesis doctoral. Estos datos muestran que el segundo formante articulado por las mujeres presenta una frecuencias medias un 34% superiores a las de aquellas producidas por hombres.

De acuerdo con los trabajos de Martínez Celdrán (2007: 39) y RAE (2011: 85), la frecuencia en la que aparece el F1 de /o/ es 495 y 455 Hz respectivamente, y la del F2 es de 1070 y 910 Hz. Sin embargo, los valores son más elevados según lo indicado en el trabajo de Pérez García (2018: 100), ya que el F1 asciende a 585 Hz y el F2 a 1272, resultados corroborados por los extraídos en el análisis de la realización de las hablantes que conforman el grupo de control de la presente tesis doctoral: el F1 de /o/ aparece alrededor de 542 Hz, y el F2 en 1262 Hz. Es decir, la frecuencia media en la que aparece el primer formante de /o/ articulada por voces femeninas es un 19% superior que las masculinas, y el segundo formante un 38% superior cuando esta vocal es articulada por mujeres.

En cuanto a la /u/, Martínez Celdrán (2007: 39) y RAE (2011: 85) indican que el primer formante de dicha vocal aparece en 349 y 283 Hz respectivamente, mientras que en el caso de las voces femeninas aparece en 456 y 451 Hz (Pérez García, 2018: 100 y la presente tesis doctoral). Tomando como referencia los trabajos de RAE (2011) y esta investigación, se observa que las frecuencias medias en las que aparece el primer formante de /u/ son un 59% superiores cuando esta vocal es articulada por voces femeninas que por voces masculinas. También se observa que el segundo formante presenta diferentes medias, y es que este aparece en 877 y 865 Hz en voces masculinas (Martínez Celdrán, 2007: 39 y RAE, 2011: 85), y en 1084 y 1083 Hz en voces femeninas

(Pérez García, 2018: 100 e Igarreta, 2019). Es decir, el segundo formante es un 25% superior cuando es articulado por mujeres que por hombres.

Este apartado ha permitido comprobar que las vocales del español producidas por voces femeninas siguen el mismo comportamiento que los correspondientes a voces masculinas, es decir, los valores relativos son los mismos. Sin embargo, los valores absolutos no coinciden, ya que las frecuencias en las que aparecen el primer y el segundo formante son, en todas las vocales, más elevadas cuando estas son producidas por mujeres que por hombres, aspecto señalado en el trabajo Wang y Van Heuven (2006) en relación con las vocales del inglés, quienes indican que “(f)or female voices the formant frequencies are 10 to 15% higher due to the fact that the resonance cavities in the female vocal tract are smaller (shorter) by 10 to 15% than those of a male speaker” (Wang y Van Heuven, 2006: 238).

En relación con el F1, la /a/ es la vocal que presenta una diferencia menor entre las voces masculinas y las femeninas, ya que estas son un 13% más altas, las de /o/ un 19%, las de /e/ un 31%, las de /i/ un 54% y las de /u/ un 59% más altas. Es decir, en todos los casos la abertura oral es mayor cuando la vocal es articulada por voces femeninas. En cuanto al F2, la /i/ es la vocal que presenta menor diferencia entre las voces masculinas y las femeninas, ya que estas son un 16% más altas, las de /e/ un 23%, las de /u/ un 25%, las de /a/ un 34% y las de /o/ un 38%. Es decir, en todos los casos las vocales son más anteriores cuando son producidas por mujeres que por hombres.

Como se ha indicado anteriormente, la presente tesis doctoral estudia exclusivamente voces femeninas, por lo que el grupo de control proporciona el punto de referencia de los valores de las frecuencias de los formantes en voz femenina.

## **1.2 Áreas de dispersión**

En este apartado se analizan las áreas de dispersión correspondientes a cada una de las vocales, ya que el estudio de estas “pone de relieve que los valores de los formantes presentan ligeras variaciones debidas a la acción de factores como el contorno, la prominencia acentual o el tipo de habla” (RAE, 2011: 88). Como se ha indicado en el capítulo III, se han medido las frecuencias en las que aparecen los formantes de las vocales (inicio, medio y final). También se ha calculado la desviación estándar de cada uno de ellos, lo que proporciona un punto de referencia de la dispersión en las realizaciones. El análisis de las áreas correspondientes a los tres puntos analizados

permite realizar más adelante (en el apartado 3) una comparación entre las vocales producidas por los hablantes que conforman el grupo de control, y las realizadas por el grupo de informantes sinohablantes.

En el caso del grupo de control, con el fin de poder comprobar si las diferencias entre el inicio, el centro y el final de los formantes son significativas, y de comparar su comportamiento con el de las informantes del grupo meta, se ha realizado la prueba de Friedman, que es una prueba paramétrica que permite comparar más de dos variables relacionadas. Las decisiones estadísticas se han tomado adoptando como nivel de significación el valor 0.05; es decir, si los valores obtenidos son menores de 0.05, se considera que las diferencias son significativas, por lo que se rechaza la hipótesis nula, la cual postula que no existen diferencias significativas entre el inicio, el centro y el final de F1 y F2 de cada una de las vocales. Los resultados de esta prueba se muestran en el apartado correspondiente a cada vocal.

La tradición muestra que cuando se trata de áreas de dispersión de las vocales, se analizan globalmente en el interior del espacio vocálico conjunto. Sin embargo, en esta tesis se analizan las vocales de manera individual con el fin de realizar un análisis microscópico del comportamiento de estas. Así lo aconsejan los fenómenos de variabilidad detectados en el caso de las duraciones que influyen, sin duda, en las características del timbre.

### **1.2.1 Vocal /i/**

Los datos obtenidos en el análisis de las frecuencias medias de F1 y F2 de la vocal /i/ en el punto inicial, punto medio y punto final, y la desviación estándar de cada uno de ellos aparecen en la tabla 79. Como se puede observar, las frecuencias medias del F1 en los tres puntos son similares, aunque ligeramente superiores en el inicio que en el medio y en el final. En cambio, aunque las frecuencias del F2 también son similares en todos sus puntos, en el inicio estas son más bajas que en el centro. En cuanto a la desviación estándar, la homogeneidad que presenta el primer formante es menor en el inicio de la vocal que en el centro, y en el final es ligeramente superior que en el medio. Por otro lado, la desviación del segundo formante es mayor en el inicio que en el centro, y en el final la homogeneidad es menor que en el inicio y el punto medio, dado que en este punto la desviación estándar es mayor que en los anteriores.

Vocal	Formante	Punto inicial		Punto medio		Punto final	
		Frecuencia media	Desv. est.	Frecuencia media	Desv. est.	Frecuencia media	Desv. est.
/i/	F1	476.2	56.9	461.2	47.2	454.9	52.8
	F2	2432.3	157.6	2555.8	148.9	2510.3	176.9

Tabla 79. Frecuencia media y desviación estándar de los tres puntos de la vocal /i/ producida por el grupo de control.

De acuerdo con los resultados de la prueba de Friedman (tabla 80), las diferencias entre el punto inicial (en adelante, etiquetado como ‘1’), medio (en adelante, etiquetado como ‘2’) y final (en adelante, etiquetado como ‘3’) de /i/ realizada por el grupo de control son significativas.

Formante	Puntos comparados	Significación	Significación conjunta
F1	1 – 2	< 0.001	< 0.001
	1 – 3	< 0.001	
	2 – 3	0.0002	
F2	1 – 2	< 0.001	< 0.001
	1 – 3	< 0.001	
	2 – 3	< 0.001	

Tabla 80. Resultados de la prueba de Friedman para la vocal /i/ realizada por el grupo de control.

A continuación, se analizan las áreas de dispersión correspondientes a cada uno de los puntos analizados. La tabla 79 muestra que la frecuencia media en la que aparece el primer formante la vocal /i/ en su inicio es 476 Hz, y la del segundo de 2432 Hz. En el primero, la desviación estándar es de 56, y en el segundo, de 157. Estos datos muestran que la mayoría de las realizaciones del F1 de esta vocal se encuentran entre 400 y 550 Hz, y las de F2 entre 2200 y 2600 Hz. En la figura 98 se observa que hay casos en los que la vocal es más cerrada, ya que el F1 aparece en unas frecuencias más bajas que las señaladas; en otros, en cambio, la vocal es más abierta, dado que F1 aparece en un rango de frecuencias más altas. Cabe señalar que se observan algunas realizaciones de la informante 3, y en algún caso aislado de la informante 1, en las que el F1 es superior, por lo que la abertura oral es superior a la del resto de realizaciones. Este comportamiento puntual puede deberse a distintos factores, siendo uno de ellos el perfil lingüístico de las informantes.

En cuanto al F2, hay casos aislados en los que aparece en un rango de frecuencias ligeramente más bajas a las de la mayoría de realizaciones de esta vocal, alrededor de

2000 Hz, lo que indica que la posición de la lengua es más posterior que la del resto de realizaciones de /i/.

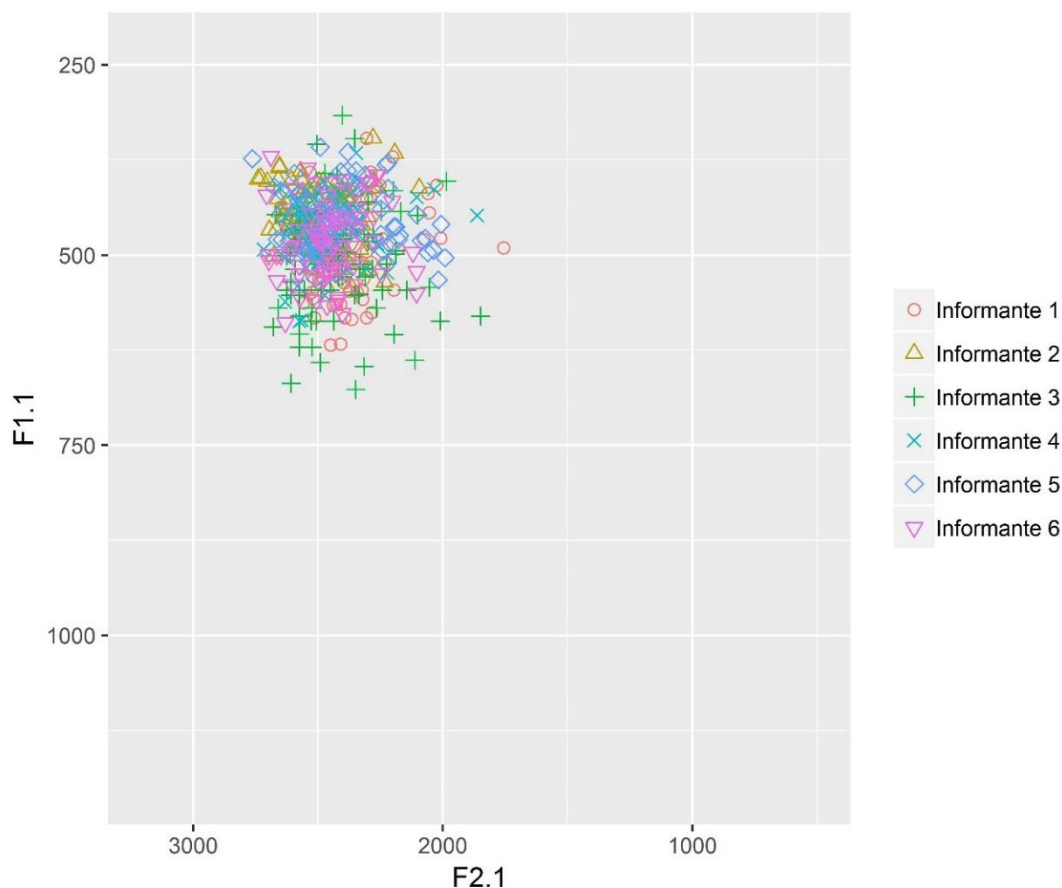
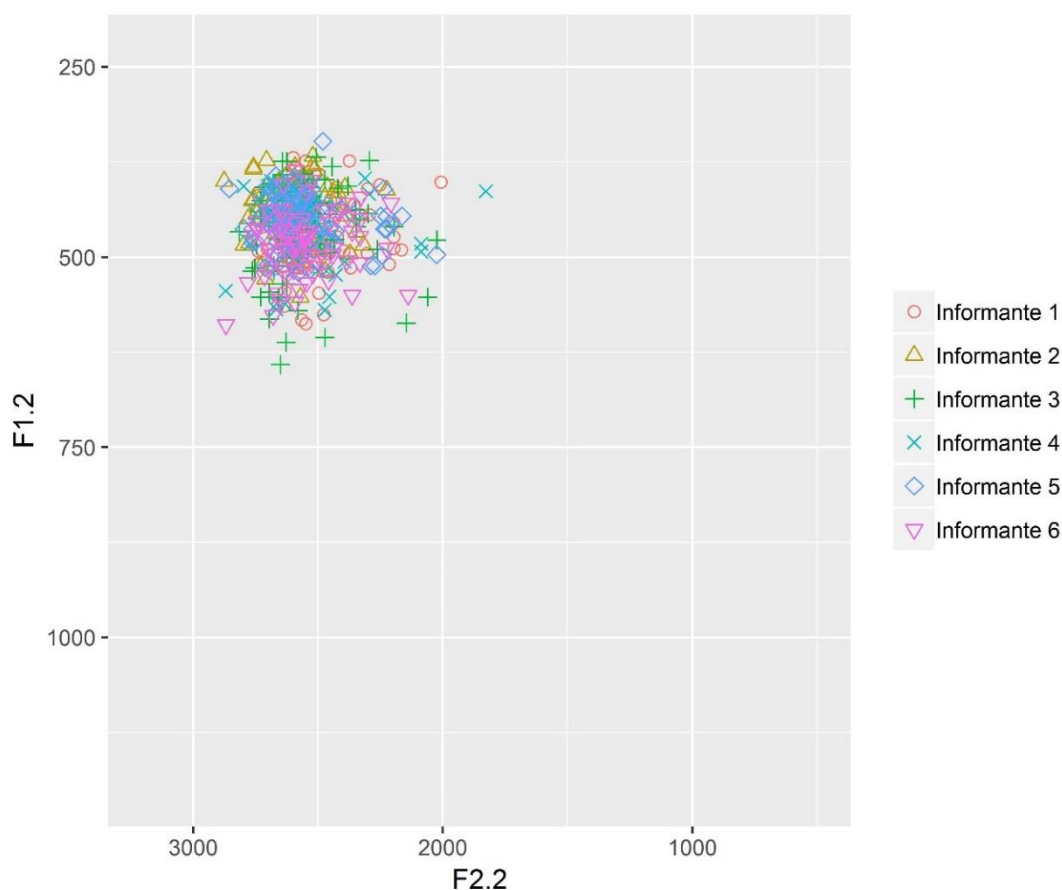


Figura 98. Área de dispersión de la vocal /i/ realizada por el grupo de control en el punto inicial.

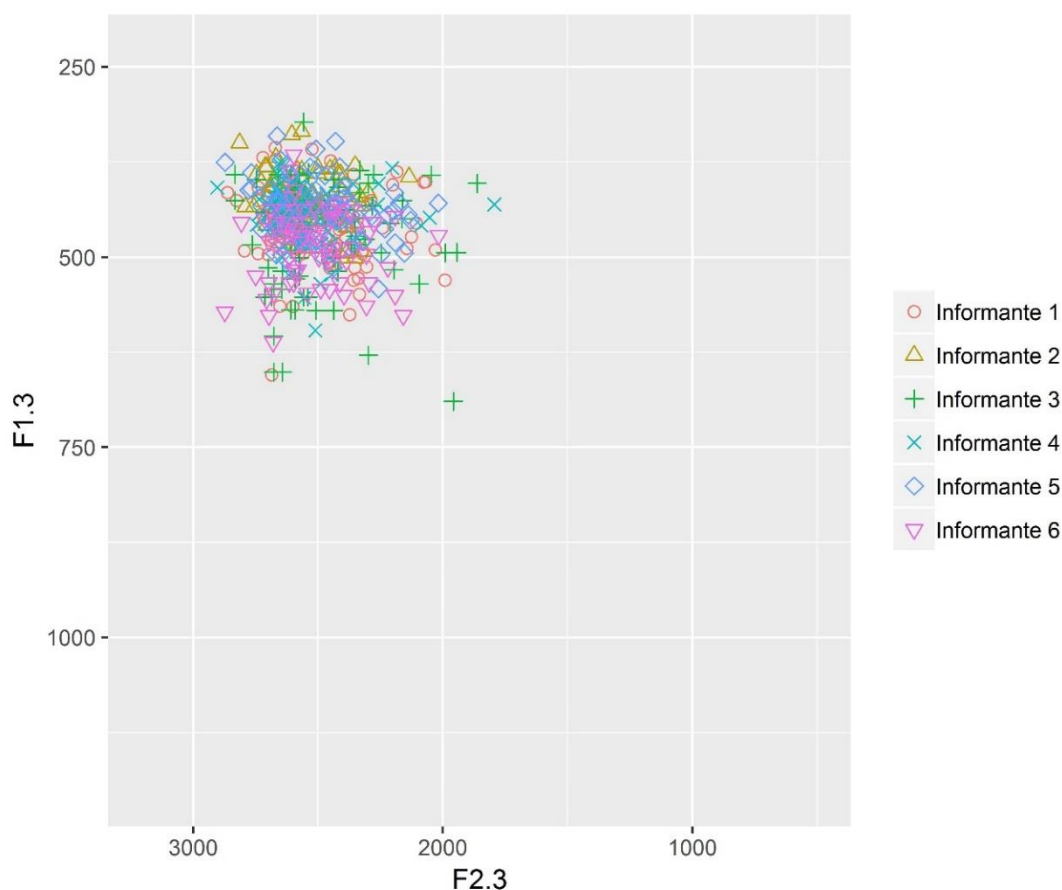
En el centro del formante, la media de F1 es de 461 Hz, y la de F2 de 2555 Hz, y la desviación estándar del primer formante es de 47, y la del segundo de 148, la cual es, en ambos casos, menor que en el inicio de la vocal. Estos datos muestran que la mayoría de realizaciones se concentran en un rango de frecuencias de entre 400 y 550 Hz en F1, y entre 2400 y 2700 Hz en F2, valores ligeramente superiores a los del inicio del formante. Como sucede en el punto inicial, se observa que hay realizaciones en las que la vocal es más cerrada, puesto que F1 presenta un valor más bajo, o más abierta, ya que las frecuencias en las que aparece son más altas; sin embargo, en este punto se puede observar que hay menos realizaciones de la informante 3 en las que la abertura oral es mayor que la de las demás informantes. Lo mismo sucede con el F2, en cuyo caso la posición de la lengua es más anterior o más posterior en función de si los valores son más o menos altos que los de la mayoría.





*Figura 99.* Área de dispersión de la vocal /i/ realizada por el grupo de control en el punto medio.

En cuanto al final de la vocal /i/ realizada por las informantes hispanohablantes, la frecuencia media en la que aparece el F1 es de 454 Hz, y la de F2 de 2510 Hz, siendo la desviación estándar de 52 y 176 respectivamente, la cual es ligeramente superior que la del centro de la vocal. Estos datos muestran que la mayoría de realizaciones de F1 aparecen entre 400 y 500 Hz, y las de F2 entre 2300 y 2600 Hz, aunque, igual que sucede en los anteriores puntos analizados, se dan casos en los que las realizaciones son más cerradas o más abiertas (F1 aparecen en otro punto inferior o superior al señalado), o más anteriores o más posteriores (F2). También en el final de la vocal se detectan algunos casos aislados en los que la frecuencia del primer formante es mayor que la del resto.



*Figura 100.* Área de dispersión de la vocal /i/ realizada por el grupo de control en el punto final.

La prueba de Friedman y el análisis de las áreas de dispersión de la vocal /i/ en su inicio, centro y final, han permitido comprobar que las diferencias entre estos tres puntos de las vocales producidas por las informantes que conforman el grupo de control son significativas. Como se ha indicado anteriormente, todas las vocales analizadas en el presente trabajo son precedidas por un sonido consonántico, los cuales “pueden modificar la abertura vocálica y provocar la aparición de variantes más abiertas o más cerradas” (RAE, 2011: 81), lo que explica que en el inicio de la vocal la dispersión sea ligeramente superior a la del centro, ya que los órganos fonatorios cambian de posición desde la consonante anterior hasta el centro de la vocal. Además, la dispersión aumenta levemente en el final respecto al centro de /i/, lo que se puede explicar por dos motivos: en primer lugar, porque si va seguida de un sonido consonántico, los órganos fonatorios se preparan para la realización del mismo y, en segundo lugar, si la vocal se encuentra en posición final de palabra, los órganos fonatorios se deslizan hacia la posición de reposo.

### 1.2.2 Vocal /e/

Los datos obtenidos en el análisis de las frecuencias de F1 y F2 de la vocal /e/ en el punto inicial, punto medio y punto final aparecen en la tabla 81. Como se puede observar, las frecuencias medias del F1 en los tres puntos son prácticamente iguales. En cambio, a pesar de que las del F2 también son similares entre sí, el punto inicial aparece en unas frecuencias menos elevadas que el centro y el final del formante. En cuanto a la desviación estándar, la homogeneidad que presenta el primer formante es muy similar en todos los puntos. Lo mismo sucede con el F2, el cual muestra una desviación ligeramente superior en el inicio que en el punto medio, y la dispersión del final es similar a la del comienzo del formante.

Vocal	Formante	Punto inicial		Punto medio		Punto final	
		Frecuencia media	Desv. est.	Frecuencia media	Desv. est.	Frecuencia media	Desv. est.
/e/	F1	609.7	59.1	612.8	52.1	610.9	56.2
	F2	2074.4	178.6	2199.7	162.2	2181.1	176.1

Tabla 81. Frecuencia media y desviación estándar de los tres puntos de la vocal /e/ producida por el grupo de control.

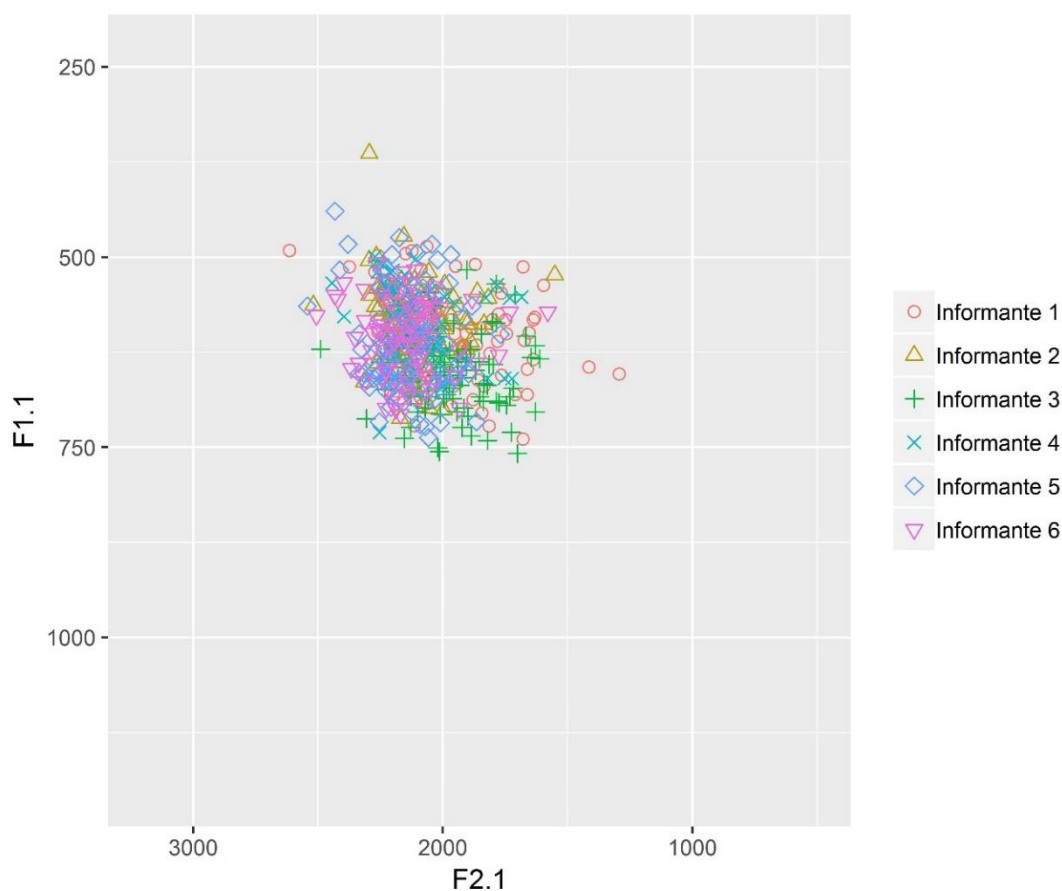
De acuerdo con los resultados de la prueba de Friedman (tabla 82), las diferencias entre el punto inicial, medio y final del F1 de /e/ realizada por las informantes hispanohablantes no son significativas en ninguno de los casos, mientras que las del F2 son significativas en todos los casos.

Formante	Puntos comparados	Significación	Significación conjunta
F1	1 – 2	1	0.2315
	1 – 3	1	
	2 – 3	0.1988	
F2	1 – 2	< 0.001	< 0.001
	1 – 3	< 0.001	
	2 – 3	0.0006	

Tabla 82. Resultados de la prueba de Friedman para la vocal /e/ realizada por el grupo de control.

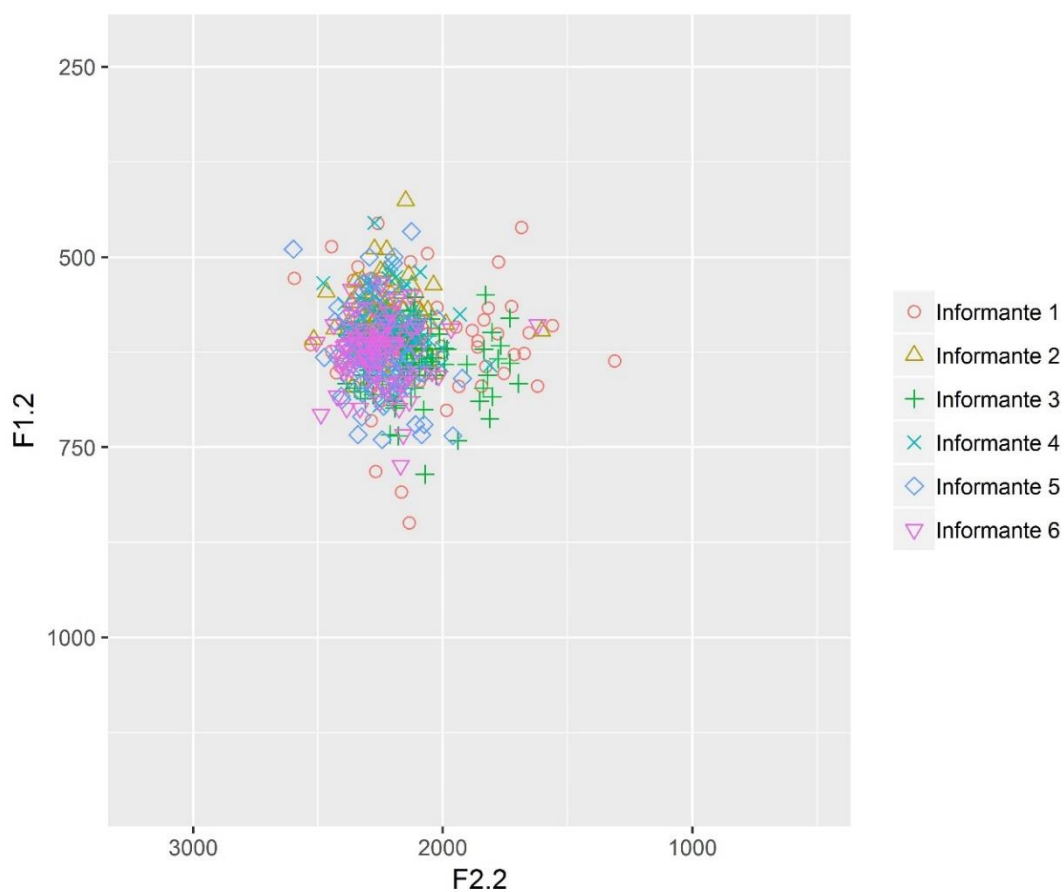
A continuación, se analizan las áreas de dispersión correspondientes a cada uno de los puntos analizados. La frecuencia media de F1 en el inicio de la vocal es de 609 Hz, y la de F2 de 2074 Hz. En el primer formante, la desviación estándar es de 59, y en el segundo, de 178. Estos datos indican que, en la mayoría de casos, el F1 aparece entre 550 y 680 Hz, y el F2 entre 1900 y 2250 Hz. A pesar de que se pueden detectar realizaciones en los que los valores de F1 o F2 difieren de la mayoría, la dispersión es baja. En la figura 101 se observa que, en algunos casos puntuales, los valores de F1 son más bajos que los de la

mayoría, por lo que la realización de /e/ es ligeramente más cerrada. Sin embargo, se ha detectado un mayor número de realizaciones en las que la vocal es más abierta, ya que el primer formante alcanza los 700 Hz.



*Figura 101.* Área de dispersión de la vocal /e/ realizada por el grupo de control en el punto inicial.

En cuanto al centro de la vocal, F1 y F2 tienen una frecuencia media de 612 y 2199 Hz respectivamente. La desviación estándar del primer formante es de 52, y la del segundo, de 162, valores muy similares a los del inicio de la vocal. En este caso, la mayoría de realizaciones de F1 aparecen entre 550 y 670 Hz, y las de F2 entre 2000 y 2350 Hz. Igual que sucede en el inicio, se observa que hay casos en los que F1 y F2 presentan valores ligeramente más bajos o más altos que los de la mayoría, lo que indica que la abertura oral y la colocación de la lengua varían ligeramente, y que en ocasiones la realización de /e/ es ligeramente más cerrada y anterior, o levemente más abierta y más posterior. Destacan las realizaciones aisladas en las que el F1 supera los 750 Hz, lo que indica que la vocal es muy abierta.



*Figura 102.* Área de dispersión de la vocal /e/ realizada por el grupo de control en el punto medio.

En cuanto al final de la /e/ realizada por las hablantes cuya lengua materna es el español, la frecuencia media de F1 es de 610 Hz, y la de F2, de 2181 Hz y la desviación estándar es de 56 en el caso del primero, y de 176 en el del segundo. Estos datos son muy similares a los del inicio y el centro de la vocal. Además, también se dan casos en los que el final del formante presenta valores más bajos o más altos que los de la mayoría.

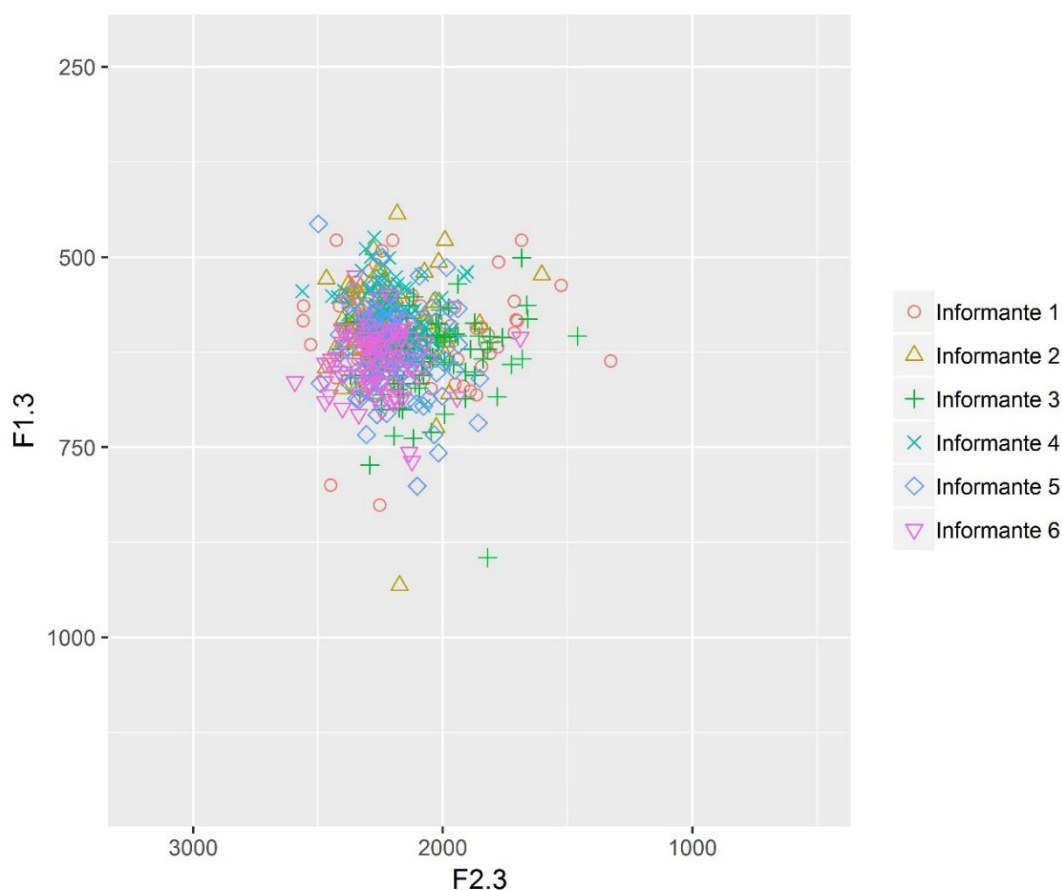


Figura 103. Área de dispersión de la vocal /e/ realizada por el grupo de control en el punto final.

La prueba de Friedman y el análisis de las áreas de dispersión de la vocal /e/ en su inicio, centro y final, han permitido comprobar que las diferencias entre estos tres puntos del primer formante no son significativas, mientras que las del segundo sí lo son. La mayoría de realizaciones comparten las mismas características, aunque se han detectado realizaciones aisladas en las que /e/ es ligeramente más cerrada y anterior, o más abierta y más posterior. La desviación estándar es similar en todos los casos, lo que indica que las realizaciones presentan una homogeneidad similar. Sin embargo, cabe señalar que, tanto en el F1 como en el F2, la desviación del centro de la vocal es ligeramente más baja que la del inicio y la del final. Como se ha señalado en el análisis de /i/, la dispersión en el inicio de la vocal se debe a que las vocales son precedidas por un sonido consonántico, por lo que los órganos fonatorios cambian de posición desde la consonante anterior hasta el centro de la vocal. En cuanto a la dispersión del final de /e/, igual que en el caso de /i/ se puede explicar por dos motivos: el primero, es que si la vocal va seguida por un sonido consonántico, los órganos fonatorios se preparan para la realización de este; el segundo,

es que si la vocal se encuentra en posición final de palabra aislada, los órganos fonatorios se deslizan hacia la posición de reposo.

### 1.2.3 Vocal /a/

En la tabla 83 se observa que la frecuencia media en la que aparece el F1 en el inicio es más baja que la del punto medio y el punto final, las cuales son muy similares. En cambio, las frecuencias medias de los tres puntos de /a/ se asemejan mucho. La desviación estándar de F1 es ligeramente superior en el inicio que en el centro y el final de la vocal, lo que indica que la homogeneidad es mayor a partir del punto medio de /a/. En el F2, la desviación es considerablemente superior en el inicio que en el medio y el final, puntos que presentan una homogeneidad muy similar.

Vocal	Formante	Punto inicial		Punto medio		Punto final	
		Frecuencia media	Desv. est.	Frecuencia media	Desv. est.	Frecuencia media	Desv. est.
/a/	F1	786.5	77.3	853.5	69.4	856.4	68.5
	F2	1692.6	172.4	1695.3	114.3	1687.1	119.6

Tabla 83. Frecuencia media y desviación estándar de los tres puntos de la vocal /a/ producida por el grupo de control.

De acuerdo con los resultados de la prueba de Friedman (tabla 84), en el F1 las diferencias entre el inicio y el centro ( $< 0.001$ ), y entre el inicio y el final ( $< 0.001$ ) son significativas, mientras que las diferencias entre el centro y el final (0.9079) de este formante no lo son. En cuanto al F2, las diferencias entre el inicio y el centro (1), y entre el inicio y el final (0.7083) no son significativas, mientras que las diferencias entre el centro y el final (0.0453) sí lo son. Estos datos dan como resultado que la significación conjunta del primer formante es  $< 0.001$ , por lo tanto, significativa, mientras que la del segundo es 0.275, es decir, no hay una diferencia significativa en lo que a los tres puntos del segundo formante se refiere.

Formante	Puntos comparados	Significación	Significación conjunta
F1	1 – 2	$< 0.001$	$< 0.001$
	1 – 3	$< 0.001$	
	2 – 3	0.9079	
F2	1 – 2	1	0.275
	1 – 3	0.7083	
	2 – 3	0.0453	

Tabla 84. Resultados de la prueba de Friedman para la vocal /a/ realizada por el grupo de control.

A continuación, se analizan las áreas de dispersión correspondientes a cada uno de los puntos analizados. Los datos muestran que el punto inicial de F1 presenta una frecuencia media de 786 Hz, y el de F2 de 1692 Hz. La desviación estándar del primer formante es de 77, y la del segundo, de 172. Estos datos muestran que, en la mayoría de los casos, las vocales realizadas por las informantes hispanohablantes la frecuencia en la que aparece el F1 oscila entre 700 y 850 Hz, mientras que la del F2 aparece entre 1500 y 1900 Hz. Igual que sucede con las vocales anteriormente analizadas, se dan casos en los que las frecuencias de F1, F2, o ambos, difieren de las de la mayoría. Así, se observa que algunas realizaciones de /a/ presentan una abertura oral mayor (F1 más alto). También hay algunas realizaciones de esta vocal que presentan unas frecuencias más bajas en el segundo formante, lo que indica que la posición de la lengua es más posterior y más central.

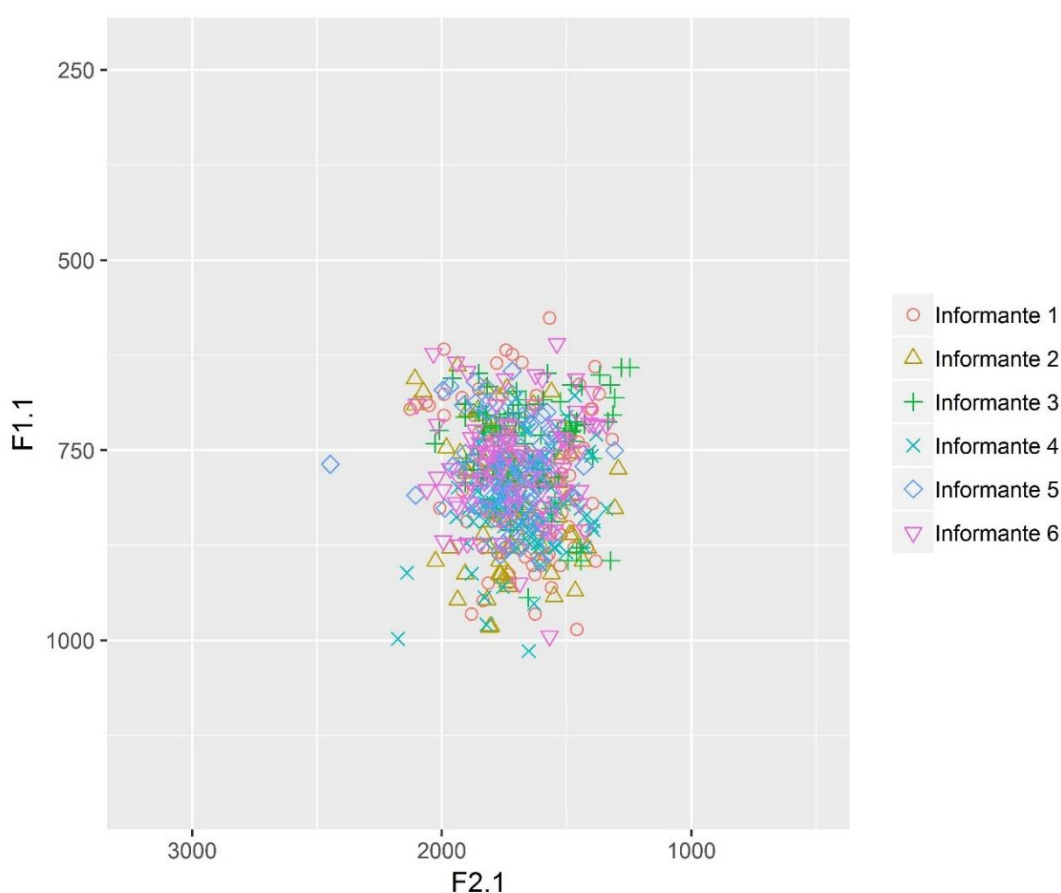
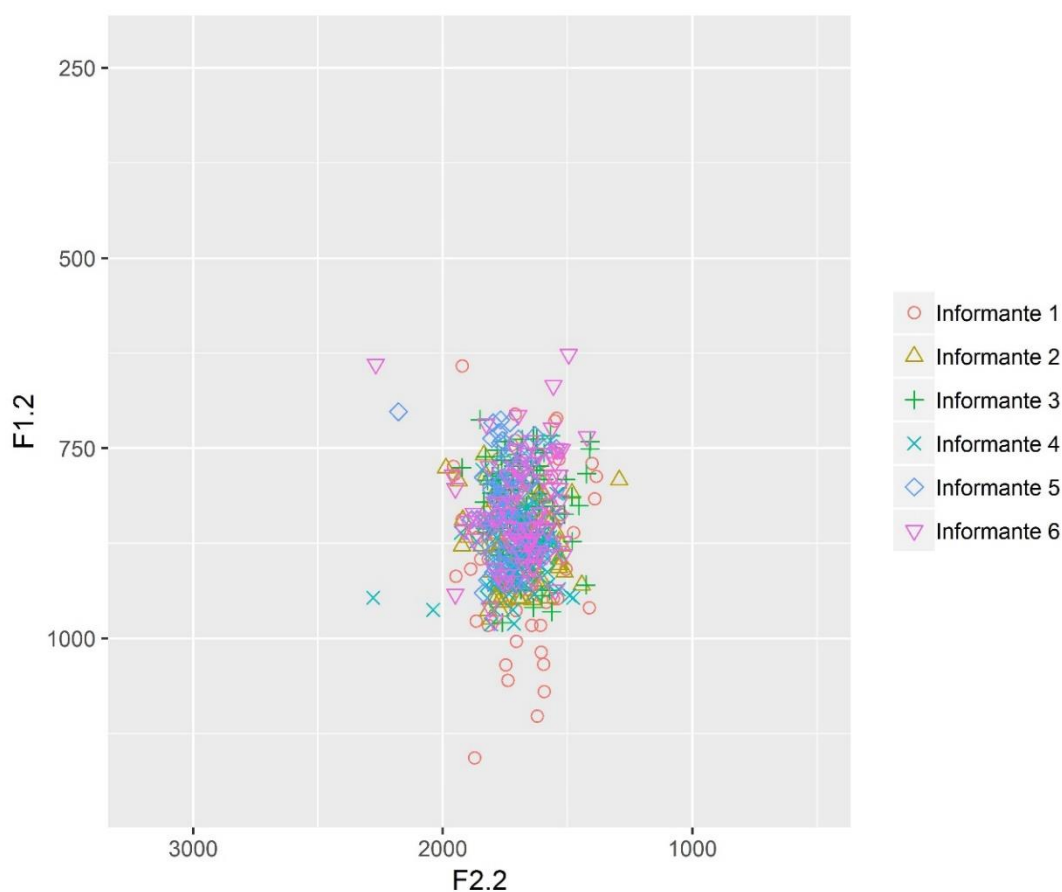


Figura 104. Área de dispersión de la vocal /a/ realizada por el grupo de control en el punto inicial.

En cuanto al punto medio de /a/, el F1 tiene una frecuencia media de 853 Hz, y el F2 de 1695 Hz. La desviación estándar del primer formante es de 69, y la del segundo de 114, la cual es ligeramente inferior a la del inicio de la vocal. Estos datos muestran que, en el punto medio de la vocal, el F1 de la mayoría de las realizaciones aparece entre 750 y 900



Hz, mientras que el F2 se sitúa entre 1500 y 1800 Hz. Los valores de F1, en este caso, son ligeramente superiores a los del punto inicial del formante, lo que indica que, en general, la abertura oral es ligeramente mayor que en el inicio. En este punto también se detectan casos en los que la abertura oral es menor, o en los que la posición de la lengua es más central. Sin embargo, el área de dispersión muestra que las realizaciones de /a/ son homogéneas.



*Figura 105.* Área de dispersión de la vocal /a/ realizada por el grupo de control en el punto medio.

En cuanto al final de la vocal, la frecuencia media del F1 de /a/ es de 856 Hz, y la del F2 de 1687 Hz, y la desviación estándar es muy similar a la del punto medio, ya que la del F1 es de 68 y la del F2 de 119. En la mayoría de los casos el primer formante de la vocal tiene una frecuencia que va desde los 750 hasta los 900 Hz, y el segundo, desde 1500 hasta 1800 Hz. Igual que en los puntos anteriores, hay realizaciones en las que la vocal presenta una abertura oral o una posición de la lengua diferente a la del resto.

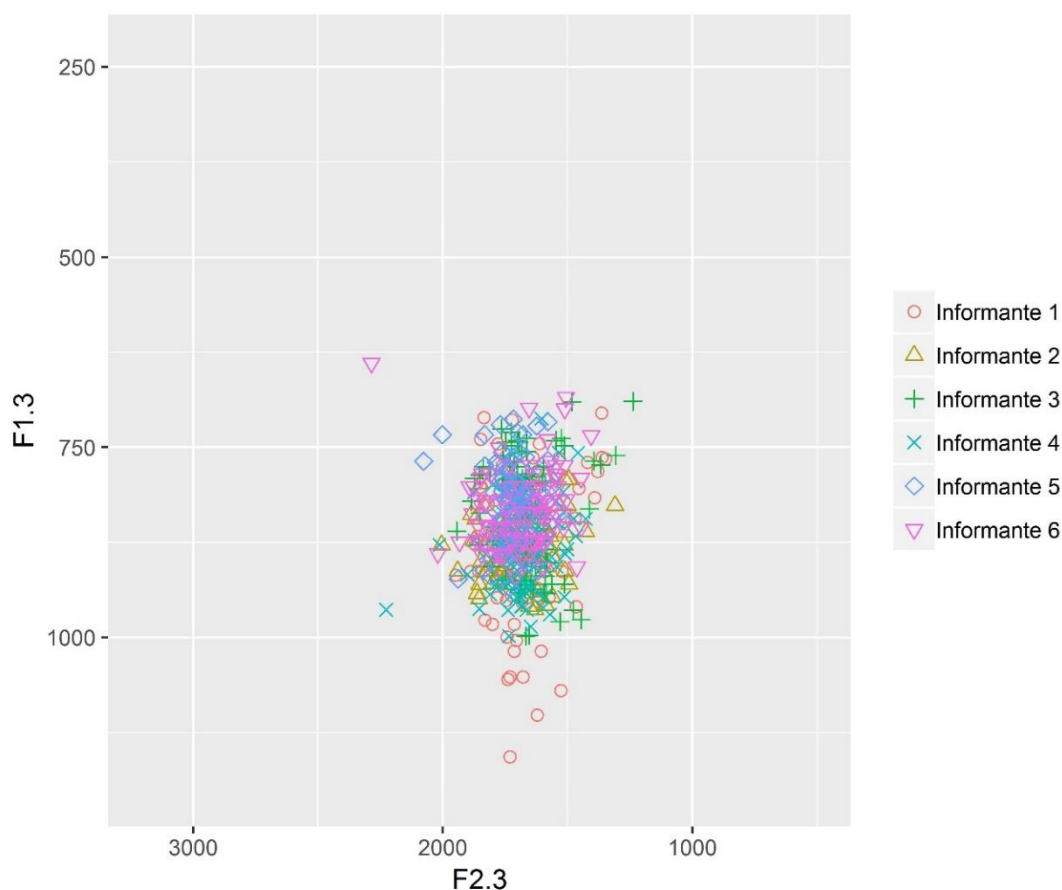


Figura 106. Área de dispersión de la vocal /a/ realizada por el grupo de control en el punto final.

La prueba de Friedman y el análisis de las áreas de dispersión de la vocal /a/ en su inicio, centro y final, han permitido comprobar que las diferencias a nivel global del primer formante son significativas, mientras que las del segundo formante no lo son. A pesar de que se encuentran casos en los que la realización de /a/ tiene una abertura oral menor, la articulación de esta vocal es homogénea puesto que la mayoría de realizaciones de la misma por parte de las informantes del grupo de control presentan las mismas características. Como sucede con las vocales /i/ y /e/, estudiadas anteriormente, los casos en los que la abertura oral o la posición de la lengua difieren del resto se explican, probablemente, porque en el inicio de la vocal los órganos fonatorios cambian de posición desde la consonante anterior hasta el centro de /a/, y, en el final, porque los órganos fonatorios se preparan para la articulación de otra consonante, o se deslizan hacia la posición de reposo.

### 1.2.4 Vocal /o/

Los datos obtenidos en el análisis de las frecuencias de F1 y F2 de la vocal /o/ en el punto inicial, punto medio y punto final aparecen en la tabla 85. Como se puede observar, las frecuencias medias de los tres puntos del F1 son muy similares, especialmente el punto central y el final. Lo mismo sucede con el F2, el cual presenta unas frecuencias semejantes, y las del inicio son ligeramente superiores a las de los otros dos puntos analizados. En cuanto a la desviación estándar, la homogeneidad que presentan los tres puntos del F1 es igual en todos ellos, mientras que en el caso de F2 esta es mayor en el centro y el final del formante que en el inicio, el cual presenta una mayor variabilidad ya que la desviación es mayor.

Vocal	Formante	Punto inicial		Punto medio		Punto final	
		Frecuencia media	Desv. est.	Frecuencia media	Desv. est.	Frecuencia media	Desv. est.
/o/	F1	551.4	65.0	542.2	65.5	540.9	62.4
	F2	1310.5	239.0	1262.9	169.6	1241.0	164.2

Tabla 85. Frecuencia media y desviación estándar de los tres puntos de la vocal /o/ producida por el grupo de control.

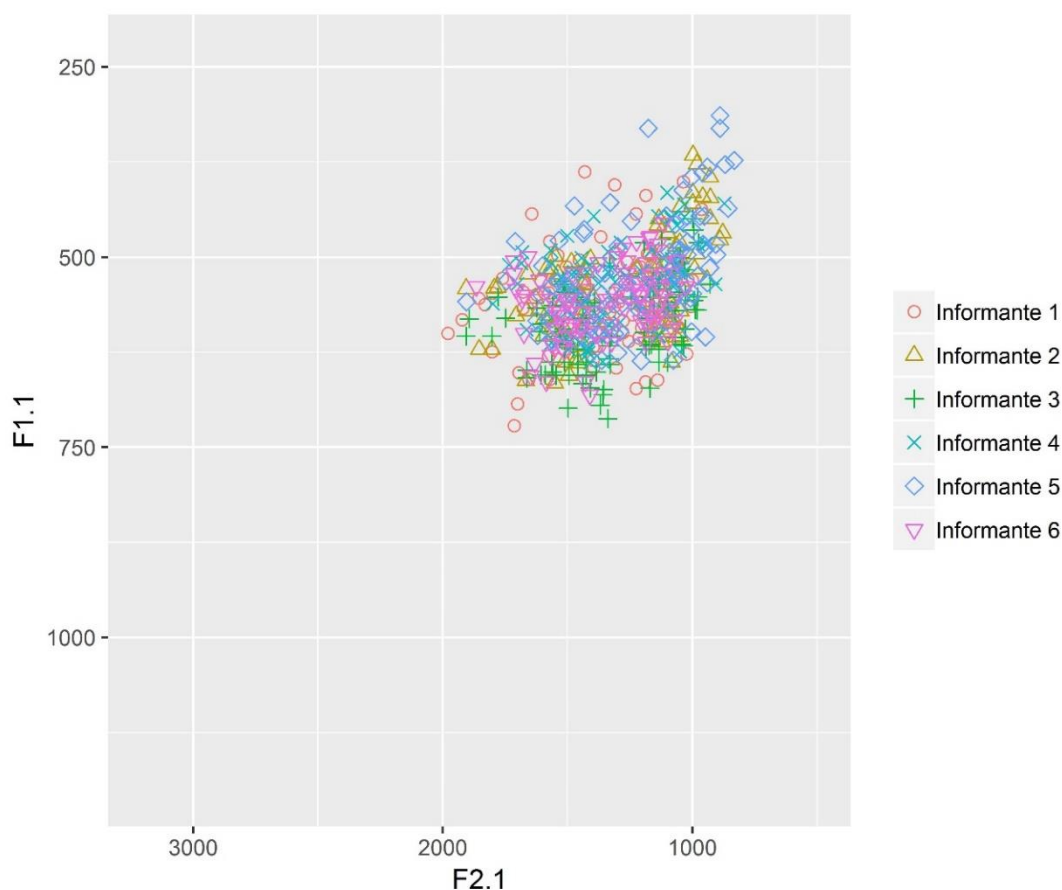
De acuerdo con los resultados de la prueba de Friedman (tabla 86), las diferencias entre el punto inicial y el punto medio (0.001), y entre el inicio y el final (0.0002) del primer formante son significativas, mientras que entre el centro y el final (1) no lo son. Sin embargo, en conjunto, estas diferencias son significativas ( $< 0.001$ ). En cuanto al segundo formante, las diferencias entre el punto inicial, el medio y el final de esta vocal son significativas en todos los casos, también en conjunto ( $< 0.001$ ).

Formante	Puntos comparados	Significación	Significación conjunta
F1	1 – 2	0.0001	$< 0.001$
	1 – 3	0.0002	
	2 – 3	1	
F2	1 – 2	$< 0.001$	$< 0.001$
	1 – 3	$< 0.001$	
	2 – 3	$< 0.001$	

Tabla 86. Resultados de la prueba de Friedman para la vocal /o/ realizada por el grupo de control.

La frecuencia media de F1 en el inicio de la vocal es de 551 Hz, y la del F2 de 1310 Hz. En el primer formante, la desviación estándar es de 65, y en el segundo, de 239. Como muestra la tabla 85, la desviación en el F2 es mucho mayor en el inicio de la vocal que en el punto medio o el final de la misma, lo que muestra que la dispersión en este punto es

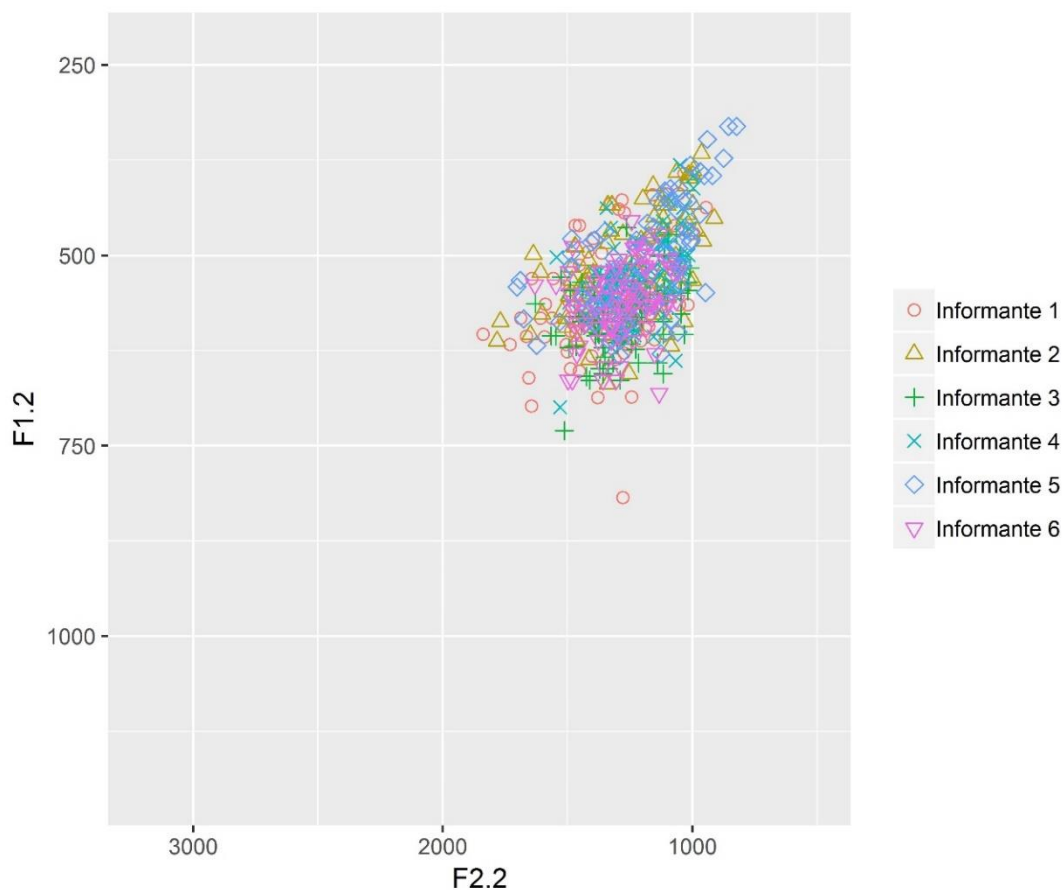
mayor, tal y como se puede observar en la figura 10. Estos datos muestran que, en la mayoría de casos, F1 aparece entre 500 y 620 Hz, y el F2 entre 1100 y 1600 Hz. Además, se observa que existen un número considerable de realizaciones en las que la articulación es más cerrada, ya que el F1 aparece en unas frecuencias inferiores a 500 Hz, y, además, más posterior, dado que la frecuencia de F2 también es inferior a 1000 Hz. También se detectan casos en los que la realización de la vocal presenta una abertura oral mayor que la mayoría.



*Figura 107.* Área de dispersión de la vocal /o/ realizada por el grupo de control en el punto inicial

El F1 de /o/ producida por el grupo de control tiene en su centro una frecuencia media de 542 Hz, y el F2 de 1262 Hz. La desviación estándar del primero es de 65, y la del segundo, de 169. Como se ha indicado, la desviación es inferior que en el punto inicial, lo cual indica que la realización es más homogénea en el centro del formante que en el inicio. En cuanto a la mayoría de casos en los que las informantes han producido esta vocal, se observa que la frecuencia de F1 va desde 500 hasta 620 Hz, y la de F2 desde 1100 hasta 1400 Hz. Cabe señalar que, de la misma manera que sucede en el inicio de la vocal, aquí también se observa que hay casos en los que la realización de /o/ es más cerrada y más

posterior ya que tanto el F1 como el F2 presentan valores inferiores a los indicados. En cuanto a la realización más anterior de /o/, el número de casos en los que sucede esto es más reducido que en el inicio de la vocal, ya que en este punto no hay realizaciones en las que F2 sea próximo a 2000 Hz.



*Figura 108.* Área de dispersión de la vocal /o/ realizada por el grupo de control en el punto medio.

En lo que respecta al final de la vocal, la frecuencia media en la que aparece F1 es de 540 Hz, y la de F2 de 1241 Hz. La desviación estándar del primer formante es de 62, y la del segundo, de 164. Como se puede observar en la figura 109, la mayoría de casos tienen un F1 que presenta unos valores de entre 500 y 600 Hz, y un F2 de entre 1100 y 1400 Hz. Si bien en este punto también se dan casos en los que la abertura oral es menor, o la posición de la lengua más posterior que la media, la realización de esta vocal presenta gran homogeneidad.

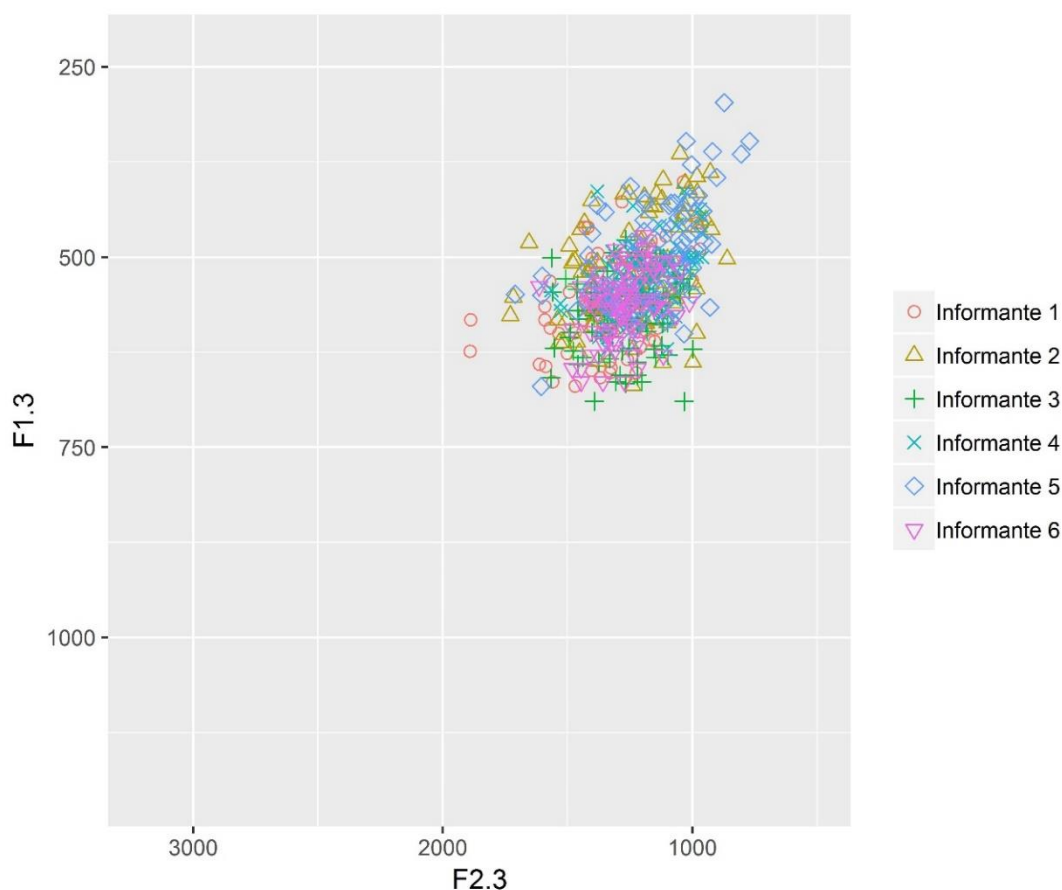


Figura 109. Área de dispersión de la vocal /o/ realizada por el grupo de control en el punto final.

La prueba de Friedman y el estudio de las áreas de dispersión de /o/ en su inicio, centro y final, han permitido comprobar que las diferencias entre estos tres puntos de las vocales producidas por el grupo de control son significativas. Cabe señalar que, mientras que en /i/, /e/ y /a/ aparecen vocales que en el inicio o final son más abiertas o cerradas que en el centro, debido probablemente al entorno consonántico, en el caso de /o/ la variabilidad es la misma en los tres puntos analizados del primer formante. Sin embargo, el inicio del segundo formante presenta una dispersión superior a la de los otros dos puntos de la vocal, y también mayor que en el resto de vocales analizadas. Igual que se ha indicado en las vocales anteriormente analizadas, esta mayor dispersión en el inicio se debe a que, puesto que en todos los casos la /o/ es precedida por una consonante, y los órganos fonatorios cambian de posición desde la consonante anterior hasta el centro de la vocal.

### 1.2.5 Vocal /u/

La tabla 87 muestra los datos obtenidos en el análisis de las frecuencias de F1 y F2 de la vocal /u/ en el punto inicial, punto medio y punto final. Como se puede observar, las frecuencias medias en las que aparece el inicio de F1 son ligeramente superiores a las del centro y el final. También las frecuencias medias de los tres puntos de F2 son muy similares. En cuanto a la desviación estándar, la homogeneidad que presenta el primer formante es muy similar en todos los puntos, mientras que el segundo formante presenta mayor variabilidad en el inicio que en el centro, y en el final la homogeneidad es ligeramente inferior que en el punto medio de la vocal, ya que la desviación es algo superior.

Vocal	Formante	Punto inicial		Punto medio		Punto final	
		Frecuencia media	Desv. est.	Frecuencia media	Desv. est.	Frecuencia media	Desv. est.
/u/	F1	480.3	64.4	451.5	52.1	458.0	51.7
	F2	1100.7	182.2	1083.8	109.7	1116.8	129.9

Tabla 87. Frecuencia media y desviación estándar de los tres puntos de la vocal /u/ producida por el grupo de control.

De acuerdo con los resultados de la prueba de Friedman (tabla 88), las diferencias entre el punto inicial, medio y final del primer formante son significativas. En cuanto al segundo formante, las diferencias entre el inicio y el centro (0.1943) de este no son significativas, mientras que entre el inicio y el final (0.0443), y entre el punto medio y el punto final ( $< 0.001$ ), sí lo son. La significación conjunta de los tres puntos de estos dos formantes indica que las diferencias son significativas.

Formante	Puntos comparados	Significación	Significación conjunta
F1	1 – 2	$< 0.001$	$< 0.001$
	1 – 3	$< 0.001$	
	2 – 3	0.0040	
F2	1 – 2	0.1943	0.0001
	1 – 3	0.0443	
	2 – 3	$< 0.001$	

Tabla 88. Resultados de la prueba de Friedman para la vocal /u/ realizada por el grupo de control.

La frecuencia media de F1 en el inicio de la vocal es de 480 Hz, y la de F2 de 1100 Hz. En el primer formante, la desviación estándar es de 64, y en el segundo, de 182. Estos datos muestran que, en la mayoría de los casos, F1 aparece entre 400 y 550 Hz, y F2 aparece entre 900 y 1300 Hz. Igual que sucede en las vocales anteriormente analizadas,

la figura 110 muestra que se dan casos en los que la vocal es ligeramente más abierta (F1 más alto) y la posición de la lengua menos posterior (frecuencias de F2 más elevadas).

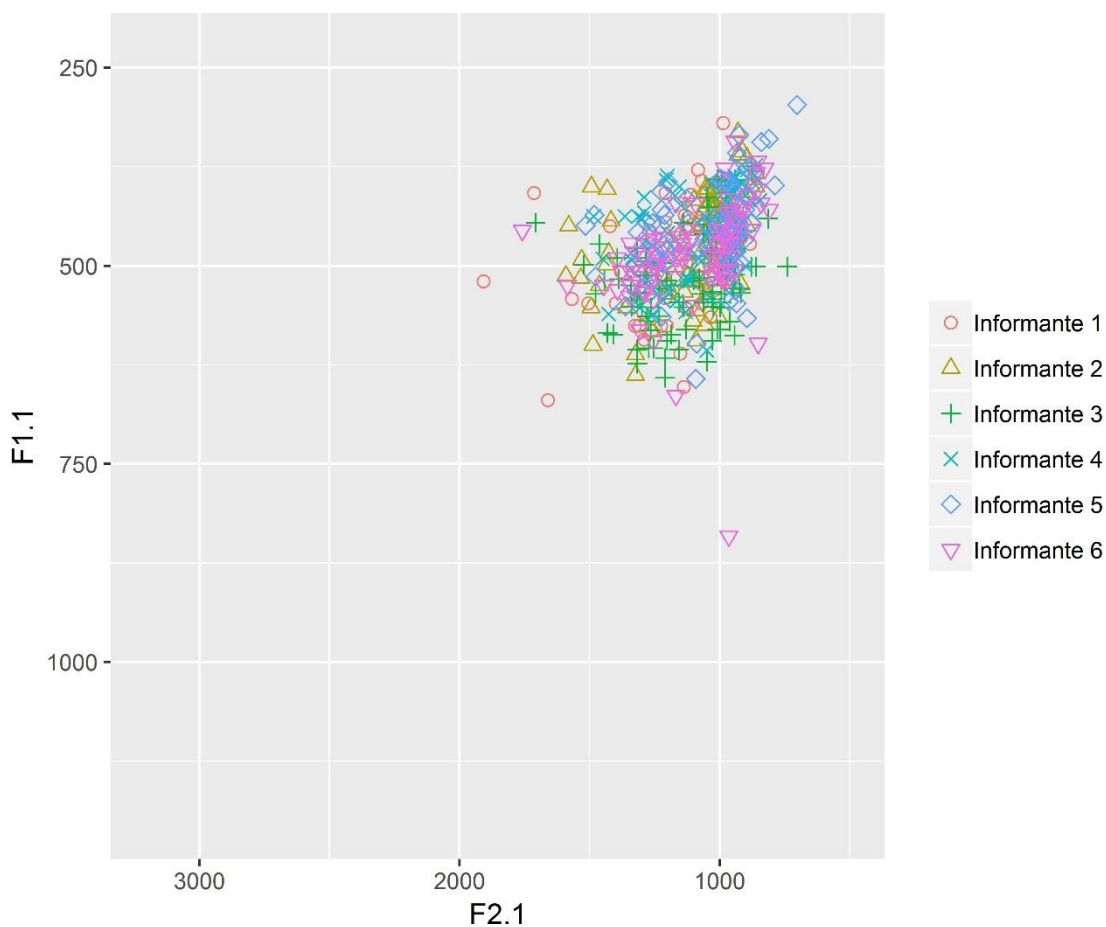
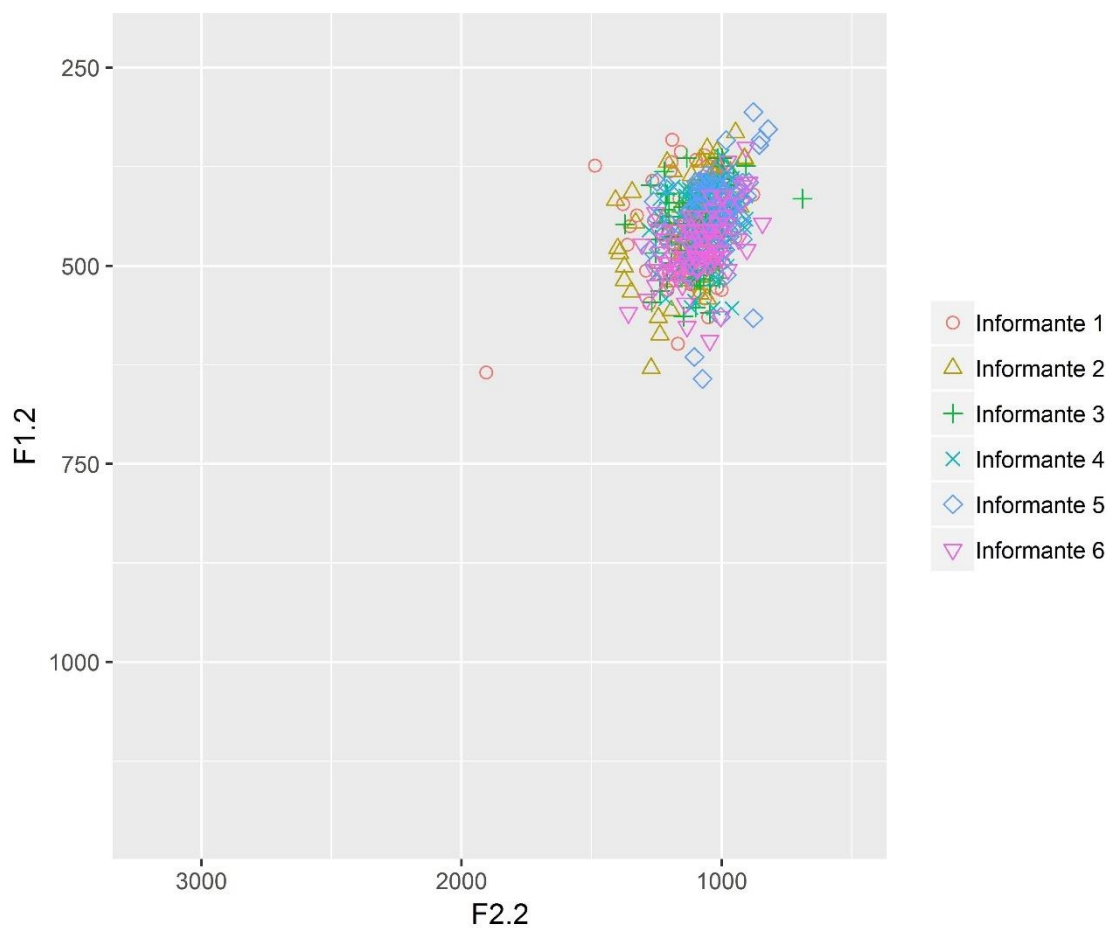


Figura 110. Área de dispersión de la vocal /u/ realizada por el grupo de control en el punto inicial.

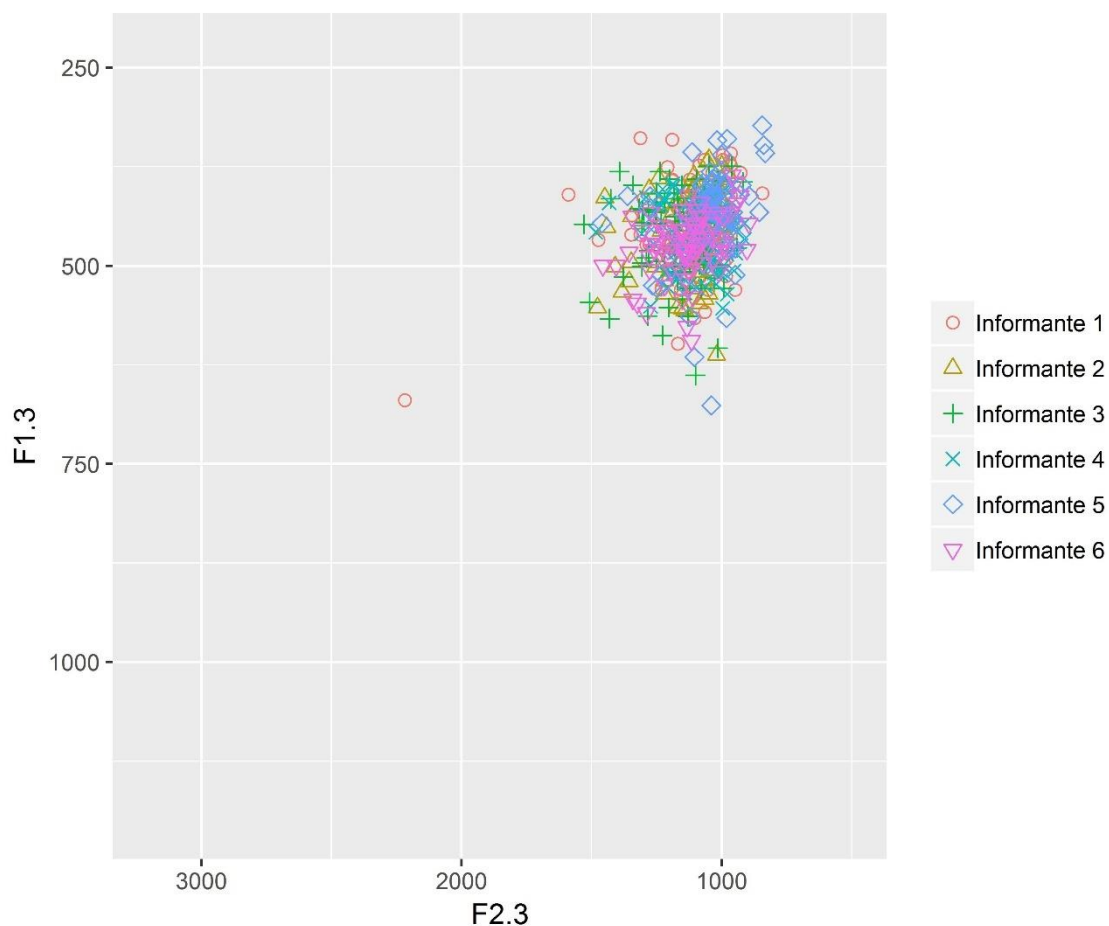
El centro de /u/ tiene una frecuencia media de 451 Hz en su F1, y de 1083 Hz en F2. La desviación estándar del primer formante es de 52, y la del segundo, de 109. La figura 111 muestra que la mayoría de realizaciones del primer formante aparecen entre 400 y 500 Hz, y las del segundo, entre 950 y 1200 Hz. Igual que sucede en el inicio de la vocal, algunas de las realizaciones son más abiertas, aunque el número de casos es menor que en el comienzo de /u/.





*Figura 111.* Área de dispersión de la vocal /u/ realizada por el grupo de control en el punto medio.

En cuanto al final de la vocal, la media de F1 es de 458 Hz, y la de F2 de 1116 Hz. La desviación estándar, de 51 en F1 y 129 en F2, es muy similar a la del centro de /u/, lo que indica que no hay variación en la dispersión. También en este caso se detectan realizaciones más abiertas y en las que la posición de la lengua es más anterior.



*Figura 112.* Área de dispersión de la vocal /u/ realizada por el grupo de control en el punto final.

La prueba de Friedman y el análisis de las áreas de dispersión de la vocal /u/ en su inicio, centro y final, han permitido comprobar que las diferencias entre estos tres puntos de los dos formantes analizados son significativas. A pesar de que se encuentran casos en los que /u/ es ligeramente menos cerrada, la articulación de esta vocal es homogénea, ya que en la mayoría de realizaciones de esta por parte de las informantes del grupo de control presentan las mismas características. Como se ha observado en /i/, /e/ y /a/, hay casos en los que el inicio de la vocal es más abierto o más cerrado que el centro y el final, y en ocasiones, este último punto también presenta mayor variabilidad que el punto medio de la vocal. Esto se debe a que, como se ha indicado anteriormente, los órganos fonatorios cambian la posición desde la consonante anterior hasta el centro de la vocal, y desde el centro hasta el final, en este caso porque, o se preparan para la realización de la siguiente consonante, o porque se deslizan hacia la posición de reposo si se encuentran en posición final de palabra aislada.

### 1.3 Recapitulación

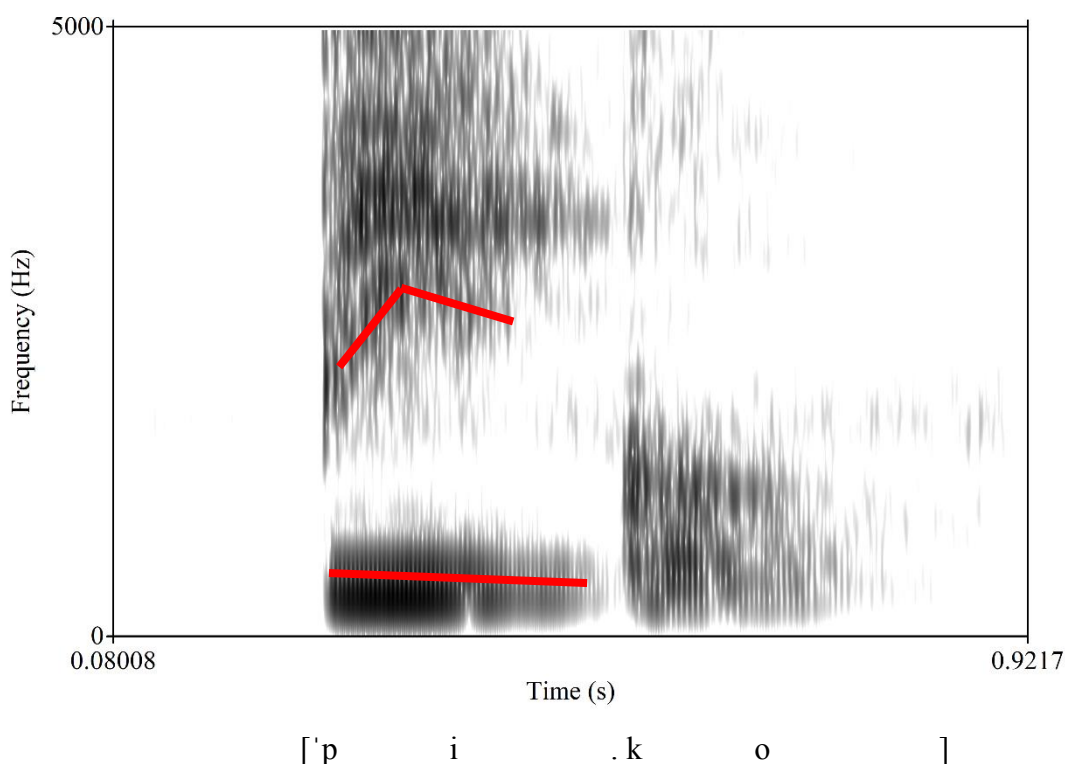
El análisis de las vocales del español realizadas por las informantes que conforman el grupo de control ha permitido mostrar que cuando estas son producidas por voces femeninas el comportamiento es, cualitativamente, el mismo que el de las voces masculinas, es decir, los valores relativos son los mismos. Sin embargo, los valores absolutos no coinciden. En el caso del primer formante, las vocales articuladas por mujeres aparecen en unas frecuencias entre un 13% y un 59% más altas que las de los hombres, es decir, la abertura oral es mayor en todos los casos. En cuanto al segundo formante, las frecuencias son entre un 16% y un 38% más altas, lo que indica que la posición de la lengua es más anterior cuando las vocales son producidas por mujeres que cuando son articuladas por hombres.

El análisis en tres puntos del primer y segundo formante ha permitido comprobar que, generalmente, las diferencias entre los mismos son significativas. Cabe señalar que, si bien el número de vocales analizadas en la presente tesis doctoral es equilibrado, los casos en los que estas aparecen precedidas por un sonido consonántico u otro varían de una vocal a otra y, como se ha indicado anteriormente, el entorno puede modificar la abertura vocálica. Por este motivo, debe tenerse en cuenta que el número de ocasiones en las que las vocales son precedidas por un sonido consonántico u otro puede influir en que las diferencias entre unos puntos y otros sean, o no, significativas.

En cuanto al estudio de la dispersión, se ha mostrado que, en la mayoría de casos, la realización de las vocales es más homogénea en el centro que en el inicio y el final de las mismas, ya que la desviación es menor en este punto. Como se ha indicado anteriormente, todos los sonidos vocálicos analizados en esta investigación son precedidos por una consonante, por lo que los órganos fonatorios cambian de posición desde este sonido hasta el centro de la vocal, lo que explica que en el inicio de las vocales la dispersión sea ligeramente superior a la del centro. En cuanto a la dispersión superior en el final del formante respecto al punto medio del mismo, se ha indicado que este fenómeno se puede explicar por dos motivos: en primer lugar, en aquellos casos en los que la vocal va seguida por un sonido consonántico, los órganos fonatorios se preparan para la realización del mismo; en segundo lugar, si la vocal se encuentra en posición final de palabra aislada, los órganos fonatorios se deslizan hacia la posición de reposo.

## 2. INFORMANTES SINOABLANTES

Como se ha indicado en el capítulo III, al realizar el análisis acústico de las vocales producidas por el grupo de informantes sinohablantes se ha detectado que muchas de ellas presentan un timbre inestable. Por este motivo, se ha considerado necesario realizar un análisis acústico exhaustivo de los formantes y medir, además del punto medio o centro, los puntos inicial y final del F1 y F2 de las vocales producidas por las informantes y, como se acaba de presentar, así se ha hecho también en el caso del grupo de control para poder establecer adecuadamente la comparación del comportamiento de los dos grupos de informantes. En la fase de realización del análisis acústico con el programa *Praat* se ha podido comprobar que, efectivamente, el timbre de algunas vocales de las informantes sinohablantes es inestable. Ejemplo de ello es la realización de la vocal /i/ en la palabra *pico* realizada por una de las informantes sinohablantes que aparece en el espectrograma 12.



*Espectrograma 12.* Timbre inestable de la vocal /i/ en la palabra *pico* producida por una de las informantes sinohablantes.

Se observa que, si bien la vocal /i/ presenta un primer formante con timbre estable, la energía del segundo formante comienza en torno a los 2200 Hz, asciende a 2500 Hz en el centro, y desciende a una altura de 2400 Hz en el punto final del formante.

Tras realizar la medición de los formantes de las vocales en los tres puntos señalados, se ha procedido a calcular, igual que en el grupo de control, el valor medio en

cada uno de los puntos de medición de los formantes (en Hz) y la desviación estándar de cada uno de ellos. Estos datos han permitido realizar la prueba de Friedman con el fin de comprobar si las diferencias entre el punto inicial, el medio y el final de cada vocal son significativas o no. Igual que en el análisis de las vocales producidas por el grupo de control, las decisiones estadísticas han tomado como nivel de significación el valor 0.05; es decir, si el resultado de la prueba es menor a 0.05 las diferencias son significativas, por lo que se rechaza la hipótesis nula, la cual dice que no existen diferencias significativas entre el inicio, el centro y el final de F1 y F2 de cada una de las vocales producidas por las informantes sinohablantes. Estos datos permiten comparar, más adelante, las realizaciones del grupo de control con las de las informantes sinohablantes; sin embargo, es necesario realizar el análisis de las áreas de dispersión de cada una de las vocales para poner en relieve los diferentes tipos de articulación de los sonidos vocálicos del español detectados en el análisis auditivo.

## 2.1 Vocal /i/

Los datos obtenidos en el análisis de las frecuencias de F1 y F2 de la vocal /i/ en el punto inicial, punto medio y punto final aparecen en la tabla 89. En esta se observa que la frecuencia media en la que aparecen los tres puntos del primer formante son muy similares. En el caso del segundo formante, el inicio aparece en un rango de frecuencias inferior al del punto medio y el final, los cuales son muy similares. En cuanto a la desviación estándar, en el F1 es ligeramente superior en el inicio, y en el centro y el final es similar, y lo mismo sucede con el F2, lo que significa que en ambos formantes el centro y el final de la vocal presentan mayor homogeneidad que el inicio de los mismos.

Vocal	Formante	Punto inicial		Punto medio		Punto final	
		Frecuencia media	Desv. est.	Frecuencia media	Desv. est.	Frecuencia media	Desv. est.
/i/	F1	459.2	65.5	451.2	58.7	446.2	57.8
	F2	2436.4	276.2	2531.8	267.8	2516.0	268.7

Tabla 89. Frecuencia media y desviación estándar de los tres puntos de F1 y F2 de /i/ producida por las informantes sinohablantes.

De acuerdo con los resultados de la prueba de Friedman (tabla 90), tanto en el primer formante, como en el segundo, el nivel de significación es menor que 0.05, lo que indica que las diferencias entre los tres puntos analizados de la vocal /i/ realizada por las informantes sinohablantes son significativas. Cabe señalar que, mientras que en el caso de las hablantes del grupo de control el índice de significación era  $<0.001$ , en el de las

informantes sinohablantes se puede apreciar que es  $<0.0001$ , lo que indica que el grado de significación es mayor en el caso de las informantes sinohablantes.

Formante	Puntos comparados	Significación	Significación conjunta
F1	1 – 2	$< 0.0001$	$< 0.0001$
	1 – 3	$< 0.0001$	
	2 – 3	0.005	
F2	1 – 2	$< 0.0001$	$< 0.0001$
	1 – 3	$< 0.0001$	
	2 – 3	0.017	

Tabla 90. Resultados de la prueba de Friedman para la vocal /i/ realizada por las informantes sinohablantes.

De acuerdo con los datos mostrados en la tabla 90, la frecuencia media en la que aparece el primer formante de la vocal /i/ en su inicio es de 459 Hz, y la del segundo de 2436 Hz. En el primero, la desviación estándar es de 65, y en el segundo, de 276. Estos datos permiten saber que, en la mayoría de casos, las realizaciones del F1 de esta vocal por parte de las informantes sinohablantes se encuentran, aproximadamente, entre 350 y 550 Hz, y las de F2 entre 2100 y 2700 Hz. Sin embargo, en la figura 113 se observa que hay casos en los que la realización de la vocal difiere del comportamiento general. Ejemplo de ello es la realización más cerrada de esta vocal, como muestran las realizaciones en las que el F1 es menor de 300 Hz. También hay muchas realizaciones en las que sucede lo contrario, y es que el primer formante aparece en un rango de frecuencias más elevadas, cercanas a 600 Hz, lo que indica que la abertura oral es similar a la de la /e/ del español. En cuanto a la posición de la lengua, hay casos en los que la realización es más posterior, ya que el F2 aparece en un rango de frecuencias inferior a 2000 Hz, o realizaciones más anteriores de /i/, en las que el F2 es cercano a 3000 Hz. Además, hay casos aislados en los que el primer formante aparece en unas frecuencias próximas a 300 Hz, y el segundo cerca de los 1000 Hz, lo que indica que esta realización de /i/ por parte de las informantes sinohablantes es semejante a la de la /u/ del español.

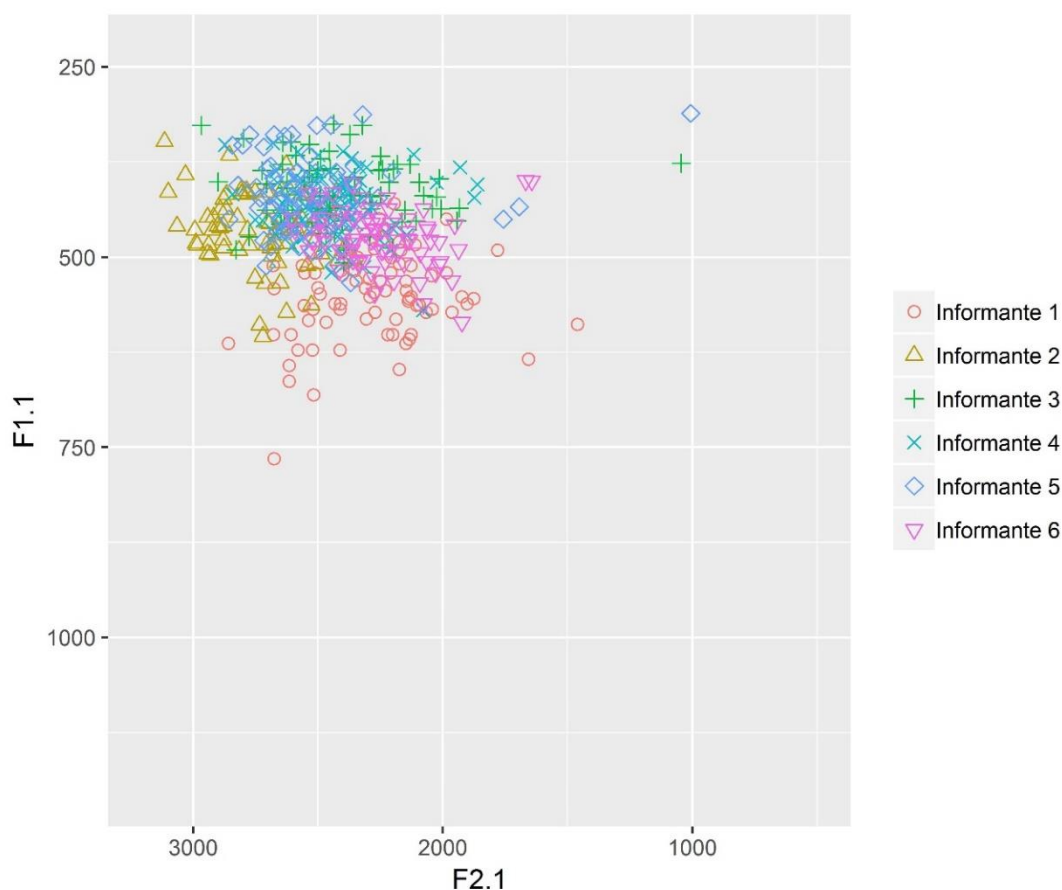
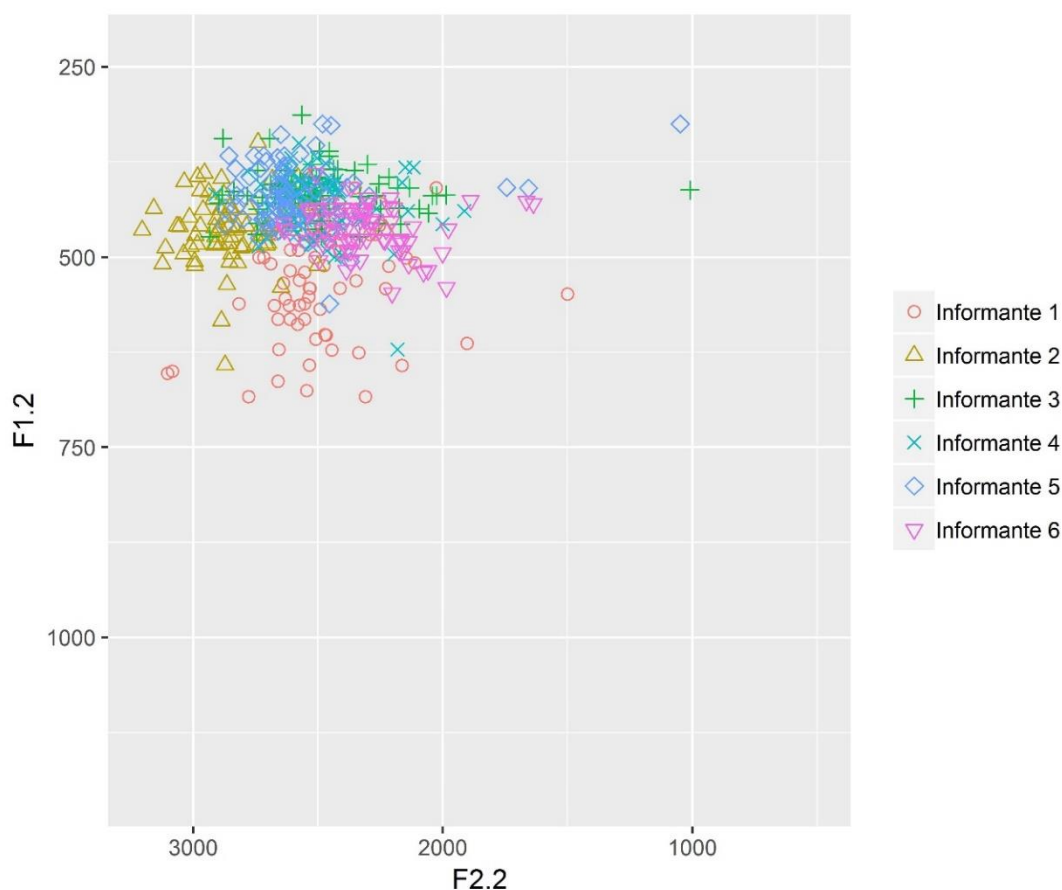


Figura 113. Área de dispersión de la vocal /i/ realizada por informantes sinohablantes en el punto inicial.

En el centro del formante, la media de F1 es de 451 Hz, y la de F2 de 2531 Hz. En cuanto a la desviación estándar, esta es de 58 en F1, y de 267 en F2, lo que indica que la realización de esta vocal es más homogénea en el centro que en el inicio. En este punto, la mayoría de realizaciones de /i/ se concentran en un rango de frecuencias entre 400 y 500 Hz en F1, y entre 2300 y 2800 Hz en F2. Igual que sucede en el punto inicial de /i/ producida por las informantes sinohablantes, en el punto medio de la vocal se detectan casos que difieren del resto. En el centro del formante también hay realizaciones en las que el F1 aparece en unas frecuencias más bajas, lo que indica que la articulación de la vocal es más cerrada. Por el contrario, e igual que sucede en el inicio, hay realizaciones en las que la abertura es mayor. En cuanto a la posición de la lengua, se observa que hay menos casos en los que el segundo formante aparece en una franja más baja que 2000 Hz, por lo que las realizaciones son ligeramente más anteriores que en el inicio. En cuanto a los casos aislados en los que la articulación de la vocal se asemeja a la de la /u/ del español, las informantes sinohablantes mantienen dicha articulación.



*Figura 114.* Área de dispersión de la vocal /i/ realizada por informantes sinohablantes en el punto medio.

En el final de la vocal, de acuerdo con los datos mostrados en la tabla 89, la frecuencia media del primer formante es de 446 Hz, y la del segundo, de 2516 Hz. En cuanto a la desviación estándar, la de F1 es de 57, y la de F2 de 268, lo que muestra que la homogeneidad de las realizaciones de esta vocal es similar tanto en el centro como en el final de la misma. Aquí, igual que en centro del formante, la mayoría de realizaciones de F1 se sitúan entre 400 y 500 Hz, y las de F2 entre 2300 y 2800 Hz. Como se puede observar en la figura 115, las realizaciones de esta vocal se asemejan mucho a las del centro de la misma. Cabe añadir a los fenómenos señalados anteriormente que, en este caso, el número de casos en que la frecuencia del primer formante es superior a 600 Hz es menor.



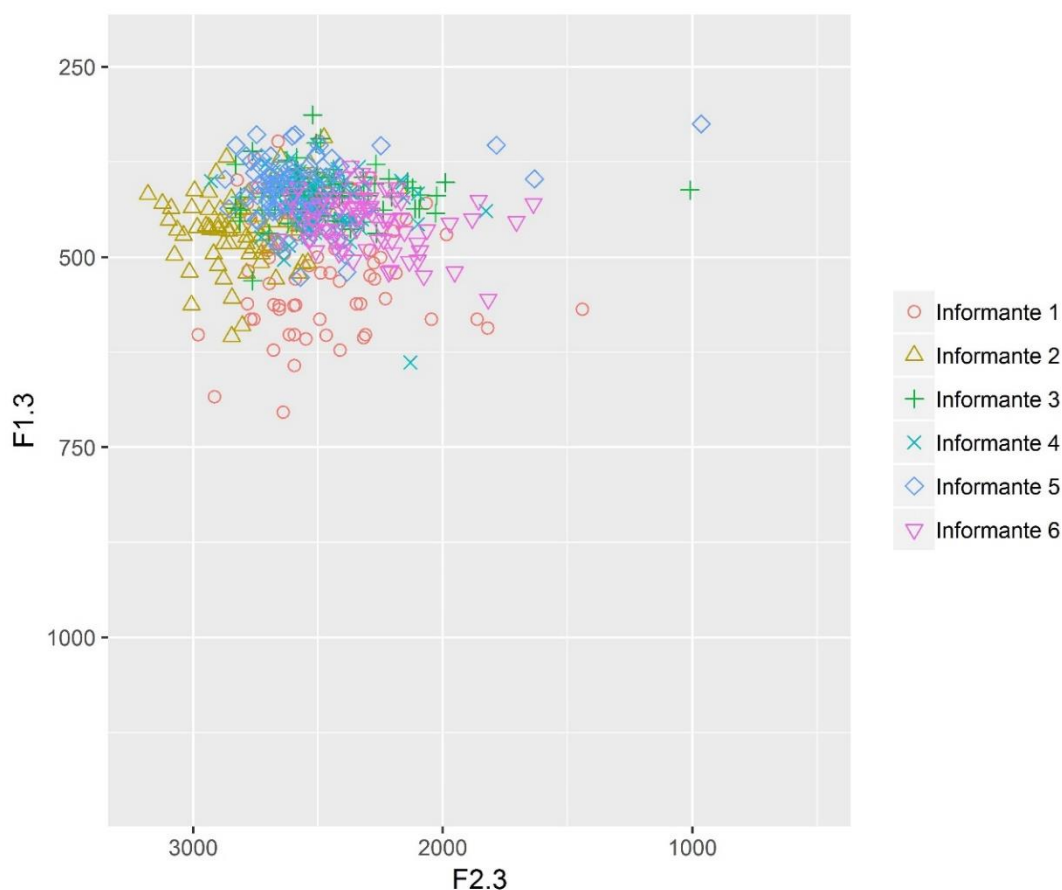


Figura 115. Área de dispersión de la vocal /i/ realizada por informantes sinohablantes en el punto final.

La prueba de Friedman y el análisis de las áreas de dispersión de la vocal /i/ en su inicio, centro y final, han permitido comprobar que las diferencias entre estos tres puntos de las vocales producidas por las informantes sinohablantes son significativas. Como se ha podido observar, la desviación estándar de los dos formantes es ligeramente superior en el inicio del formante que en el final. Esto muestra que, tanto la abertura oral, como la posición de la lengua, tienden a ser más homogéneas en el centro y el final de la vocal, que en el inicio de la misma. Se han detectado casos en los que las realizaciones difieren de las de la mayoría y, en algunos casos aislados, presentan características similares a las de otras vocales del español, como la /e/ y la /u/.

## 2.2 Vocal /e/

Los datos obtenidos en el análisis de las frecuencias de F1 y F2 de la vocal /e/ del español en el punto inicial, punto medio y punto final aparecen en la tabla 91. En esta se observa que la frecuencia media en la que aparece el inicio de F1 es ligeramente inferior

a la del centro, y este punto tiene una media similar a la del final del formante. En el caso del F2, el inicio presenta una frecuencia media inferior a la del centro, y esta es ligeramente superior a la del final. En cuanto a la desviación estándar del primer formante, los tres puntos presentan valores prácticamente iguales, por lo que la homogeneidad es similar, y en el caso del segundo formante, la desviación aumenta levemente desde el inicio, pasando por el centro y hasta el final.

Vocal	Formante	Punto inicial		Punto medio		Punto final	
		Frecuencia media	Desv. est.	Frecuencia media	Desv. est.	Frecuencia media	Desv. est.
/e/	F1	608.2	67.4	634.9	72.4	641.5	73.0
	F2	2039.7	227.3	2098.0	233.9	2081.3	240.5

Tabla 91. Frecuencia media y desviación estándar de los tres puntos de F1 y F2 de /e/ producida por las informantes sinohablantes.

Los resultados de la prueba de Friedman (tabla 92) han permitido demostrar que las diferencias entre el inicio, el medio y el final del primer y segundo formante de la vocal /e/ producida por las informantes sinohablantes son significativas en todos los casos.

Formante	Puntos comparados	Significación	Significación conjunta
F1	1 – 2	< 0.0001	< 0.0001
	1 – 3	< 0.0001	
	2 – 3	0.011	
F2	1 – 2	< 0.0001	< 0.0001
	1 – 3	< 0.001	
	2 – 3	0.0001	

Tabla 92. Resultados de la prueba de Friedman para la vocal /e/ realizada por las informantes sinohablantes.

De acuerdo con los datos mostrados en la tabla 91, la frecuencia media de F1 en el inicio de la vocal es de 608 Hz, y la de F2 de 2039 Hz. En el primer formante, la desviación estándar es de 67, y en el segundo, de 227. Estos datos permiten conocer que, en la mayoría de los casos, las realizaciones del F1 de /e/ por parte de las estudiantes chinas se encuentran, aproximadamente, entre 500 y 700 Hz, y las de F2 entre 1800 y 2300 Hz. Sin embargo, en la figura 116 se observa que hay casos en los que la realización de la vocal difiere del comportamiento general.

En primer lugar, se observa que hay realizaciones en las que el F1 aparece en un rango de frecuencias cercano, e incluso superior a 750 Hz, lo que indica que la abertura oral es mayor que en la mayoría de casos. Además, en algunas de estas realizaciones el F2 presenta una frecuencia inferior, cercana a 1500 Hz. En estos casos, la realización de /e/ por parte de las hablantes cuya lengua materna es el chino se caracteriza porque la

abertura oral y la posición de la lengua se asemejan a las propias de la /a/ del español. En segundo lugar, hay casos en los que el F1 es menor, y aparece en un rango de frecuencias inferior a 500 Hz, por lo que la abertura oral se asemeja a la de /i/ del español, aunque la posición de la lengua aparece en unas frecuencias similares a las de la mayoría de realizaciones de /e/ articulada por las informantes sinohablantes (entre 1800 y 2300 Hz). En tercer lugar, se puede observar que hay ocasiones en las que la abertura oral es la propia de /e/, pero el F2 tiene un valor inferior, cercano a 1500 Hz, por lo que la posición de la lengua es más posterior, similar a la de /a/ del español. Por último, cabe mencionar que se detectan casos aislados en los que el F2 de /e/ articulado por la informante 2 aparece cerca de 1000 Hz, lo que indica que la posición de la lengua es posterior, más cercana a la de /u/ del español que a la de /e/.

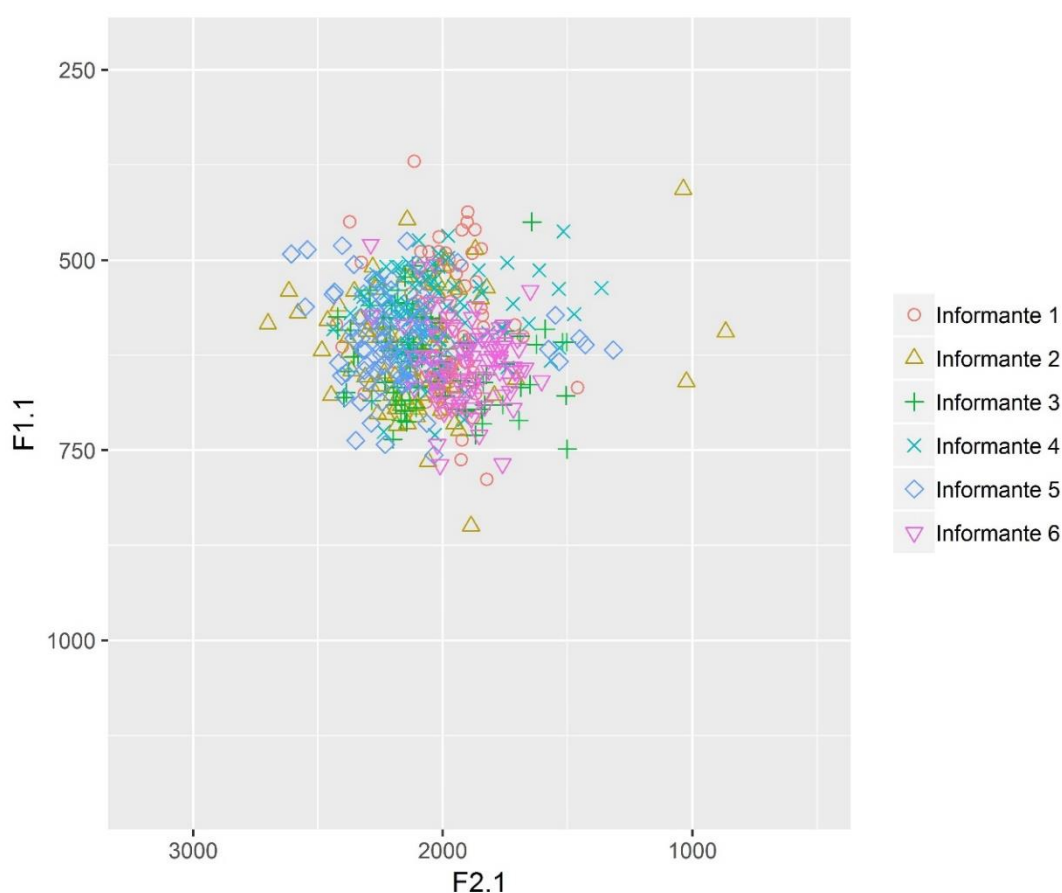


Figura 116. Área de dispersión de la vocal /e/ realizada por informantes sinohablantes en el punto inicial.

En cuanto al centro de la vocal, F1 y F2 tienen una frecuencia media de 634 y 2098 Hz respectivamente. La desviación estándar del primer formante es de 72, y la del segundo, de 233, valores similares a los del inicio de la vocal, aunque ligeramente superiores. En este caso, los datos indican que la mayoría de realizaciones de F1 aparecen entre 550 y

750 Hz, y las de F2 entre 1800 y 2500 Hz. Igual que sucede en el inicio, en la figura 117 se observa que hay casos en los que F1 y F2 presentan valores que se alejan de los propios de la /e/ del español.

En primer lugar, cabe señalar que, si bien la dispersión del primer formante es similar a la del inicio, las frecuencias han aumentado levemente, lo que indica que la abertura oral es algo mayor que en el punto inicial. En segundo lugar, se observa que ha aumentado el número de realizaciones de F1 que superan los 750 Hz, es decir, la abertura oral es mayor en más ocasiones. En tercer lugar, se mantienen las realizaciones en las que la frecuencia del primer formante es menor de 500 Hz. En cuanto al segundo formante, también en el centro de la vocal hay casos en los que este aparece en frecuencias cercanas a 1500 Hz, y, además, se mantienen las realizaciones aisladas en las que dicho formante tiene un valor cercano a 1000 Hz.

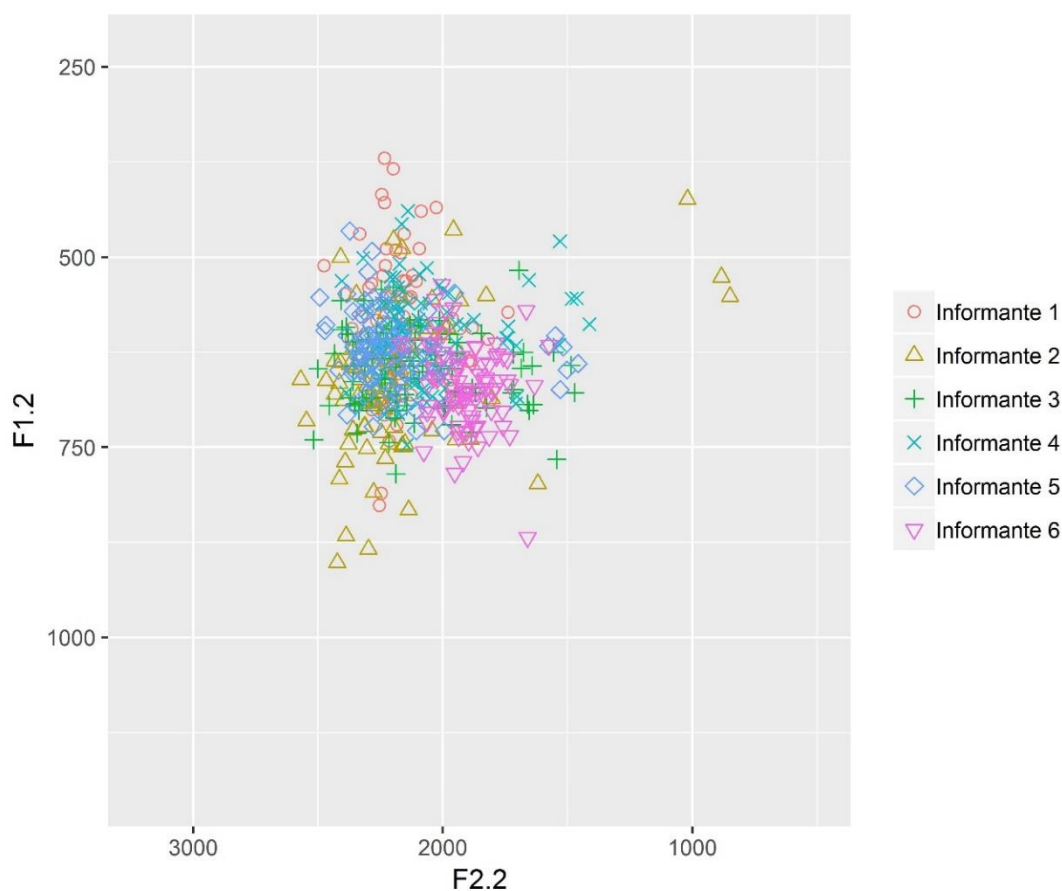


Figura 117. Área de dispersión de la vocal /e/ realizada por informantes sinohablantes en el punto medio.

En cuanto al final de /e/ realizada por las informantes sinohablantes, los datos obtenidos en el análisis son muy similares a los del centro de la vocal, ya que la frecuencia media de F1 es de 641 Hz y la de F2, de 2081 Hz. También la desviación estándar es muy similar

a la del punto medio, ya que la del formante es de 73, y la del segundo, de 240. La figura 118 muestra que, si bien hay pequeñas diferencias entre el centro del formante y el final, en ambos casos la dispersión y las realizaciones son muy similares.

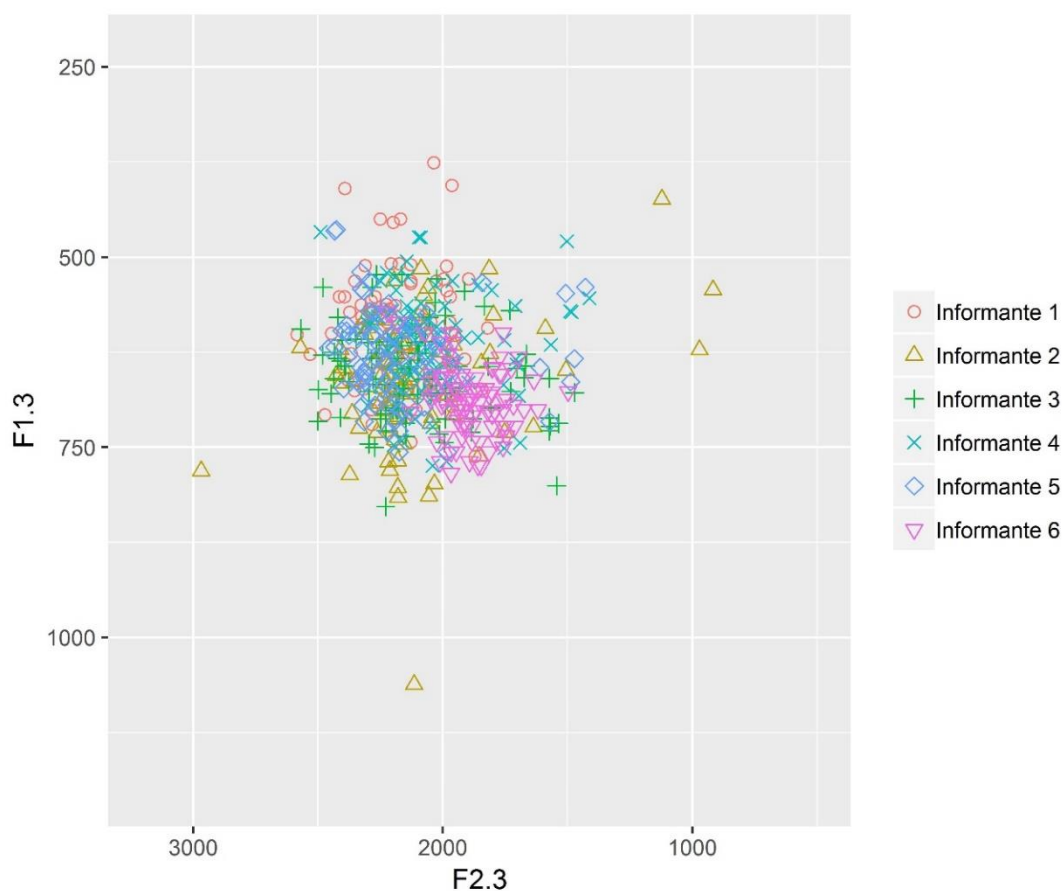


Figura 118. Área de dispersión de la vocal /e/ realizada por informantes sinohablantes en el punto final.

La prueba de Friedman y el análisis de las áreas de dispersión de la vocal /e/ en su inicio, centro y final, han permitido comprobar que las diferencias entre estos tres puntos de las vocales producidas por las informantes sinohablantes son significativas. Se ha podido observar que la desviación estándar de F1 y F2 es menor en el inicio que en el centro y el final, lo que indica que la dispersión es mayor en estos dos puntos. Las áreas de dispersión muestran que, si bien muchas de las realizaciones de /e/ por las hablantes chinas se asemejan a la /e/ del español, se han detectado vocales más abiertas o más cerradas, con valores de F1 semejantes a los de la /a/ y la /i/ del español. En relación con el F2, se han detectado algunos casos en los que los valores de dicho formante son menores que los de la mayoría, por lo que en algunos casos la posición de la lengua es similar a la de la /a/ o la /u/ del español.

### 2.3 Vocal /a/

Los datos obtenidos en el análisis de las frecuencias de F1 y F2 de la vocal /a/ en el punto inicial, punto medio y punto final aparecen en la tabla 93. En esta se observa que la frecuencia media en la que aparece el inicio de F1 es inferior que la del centro y el final, las cuales son similares. En el caso del F2, el inicio presenta una frecuencia media más elevada que la del punto medio o el punto final, las cuales también son semejantes. En cuanto a la desviación estándar del primer formante, esta es muy similar en los tres puntos de la vocal analizados, mientras que, en el caso del segundo formante, el inicio presenta una mayor desviación que el punto medio, y la de este es ligeramente superior que la del final, por lo que la homogeneidad de F2 es mayor en el final que en el inicio.

Vocal	Formante	Punto inicial		Punto medio		Punto final	
		Frecuencia media	Desv. est.	Frecuencia media	Desv. est.	Frecuencia media	Desv. est.
/a/	F1	761.7	84.3	824.1	89.4	833.5	81.5
	F2	1631.7	173.8	1589.8	144.1	1582.6	134.1

Tabla 93. Frecuencia media y desviación estándar de los tres puntos de F1 y F2 de /a/ producida por las informantes sinohablantes.

De acuerdo con los resultados de la prueba de Friedman (tabla 94), las diferencias entre el punto inicial, medio y final de /a/ realizada por las informantes sinohablantes son significativas en todos los casos.

Formante	Puntos comparados	Significación	Significación conjunta
F1	1 – 2	< 0.0001	< 0.0001
	1 – 3	< 0.0001	
	2 – 3	< 0.0001	
F2	1 – 2	< 0.0001	< 0.0001
	1 – 3	< 0.0001	
	2 – 3	0.002	

Tabla 94. Resultados de la prueba de Friedman para la vocal /a/ realizada por las informantes sinohablantes.

Los datos muestran que el punto inicial de F1 presenta una frecuencia media de 761 Hz, y el de F2 de 1631 Hz. La desviación estándar del primer formante es de 84, y la del segundo, de 173. Estos datos permiten saber que, en la mayoría de casos, las realizaciones del F1 de /a/ por parte de las informantes sinohablantes se encuentran, aproximadamente, entre 650 y 850 Hz, y las de F2 entre 1400 y 1800 Hz, valores similares a los del español. Sin embargo, es necesario estudiar todo el conjunto para conocer las diferentes realizaciones de esta vocal por parte de las estudiantes chinas. Como se puede observar

en la figura 119, hay realizaciones en las que el F1 aparece en una frecuencia inferior a 650 Hz, por lo que la abertura oral es menor que la de la mayoría y se asemeja a la de /e/ del español. Hay casos aislados producidos por la informante 2 en los que sucede lo contrario y la abertura oral es mayor, es decir, F1 tiene un valor superior a 1000 Hz. En cuanto al segundo formante, aquellos casos en los que /a/ aparece entre 1600 y 1800 Hz se asemeja a la /a/ del español, por lo que la posición de la lengua es central. Sin embargo, cuando F2 aparece en un rango de frecuencias inferior, e incluso cercano a 1000 Hz, la posición de la lengua es más posterior, semejante en algunos casos a la propia de la /o/ del español, y en otros a la de /u/. Cabe señalar, además, que hay un caso aislado en el que el F1 es cercano a 500 Hz, y el F2 supera los 2500 Hz, valores propios de la /i/ del español.

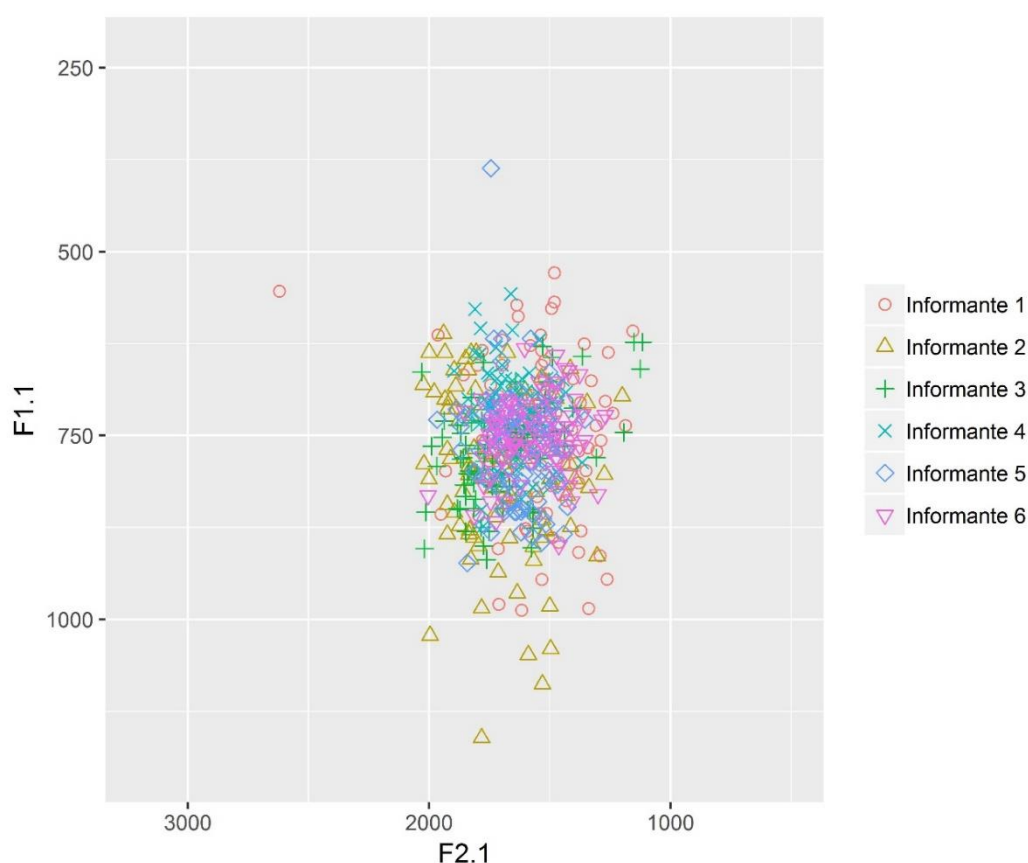


Figura 119. Área de dispersión de la vocal /a/ realizada por informantes sinohablantes en el punto inicial.

En cuanto al punto medio de /a/, el F1 tiene una frecuencia media de 824 Hz, y el F2 de 1589 Hz. La desviación estándar del primer formante es de 89, y la del segundo de 144. Esto muestra que la desviación del segundo formante es menor en el centro de este que en el punto inicial. Estos datos indican que, generalmente, F1 aparece entre 700 y 900 Hz, y F2 entre 1400 y 1700 Hz. Esto muestra que la abertura oral es ligeramente superior

que la del inicio de la vocal, y que la dispersión del segundo formante es menor que en el punto inicial de /a/ articulada por las estudiantes chinas. En cuanto a las diferencias respecto al inicio de la vocal, en primer lugar, cabe señalar que la más notable es que, como se puede observar en la figura 8, hay un número considerable de casos en los que la abertura oral es mayor que la del resto, ya que el F1 supera los 1000 Hz. En segundo lugar, otra diferencia notoria es la menor dispersión del F2, lo que indica que hay mayor homogeneidad en lo que a la posición de la lengua se refiere que en el inicio de la vocal. En cuanto al caso aislado en el que la realización de /a/ tiene valores próximos a los de la /i/ del español, se puede observar que la abertura oral es más cerrada, ya que ahora la realización se encuentra en la banda de 500 Hz, y la posición de la lengua se mantiene.

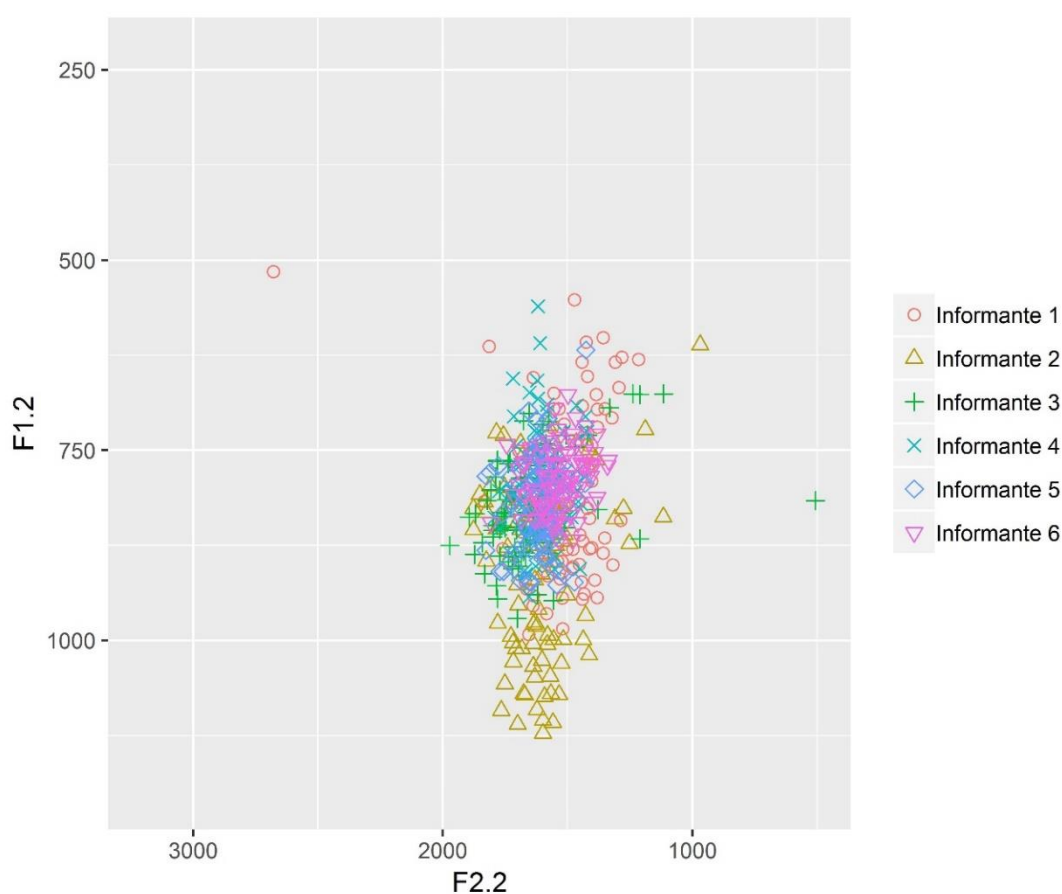


Figura 120. Área de dispersión de la vocal /a/ realizada por informantes sinohablantes en el punto medio.

En lo que respecta al final de la vocal, la frecuencia media de F1 de /a/ es de 833 Hz, y la del F2 de 1582 Hz. La desviación estándar del primer formante es de 81, y la del segundo formante es de 134, es decir, inferior tanto a la del inicio como a la del centro de la vocal. De acuerdo con los datos, generalmente el F1 de /a/ aparece en un rango de frecuencias entre 750 y 900 Hz, y el F2 entre 1400 y 1600 Hz. Como se observa en la figura 121, el



final de /a/ presenta un área de dispersión con unas características similares a las del centro de dicha vocal. A los fenómenos señalados en el punto anterior, hay que añadir que, en el final de la vocal la abertura es ligeramente superior a los puntos anteriores, y que la posición de la lengua es más homogénea, ya que hay menos variedad de realizaciones.

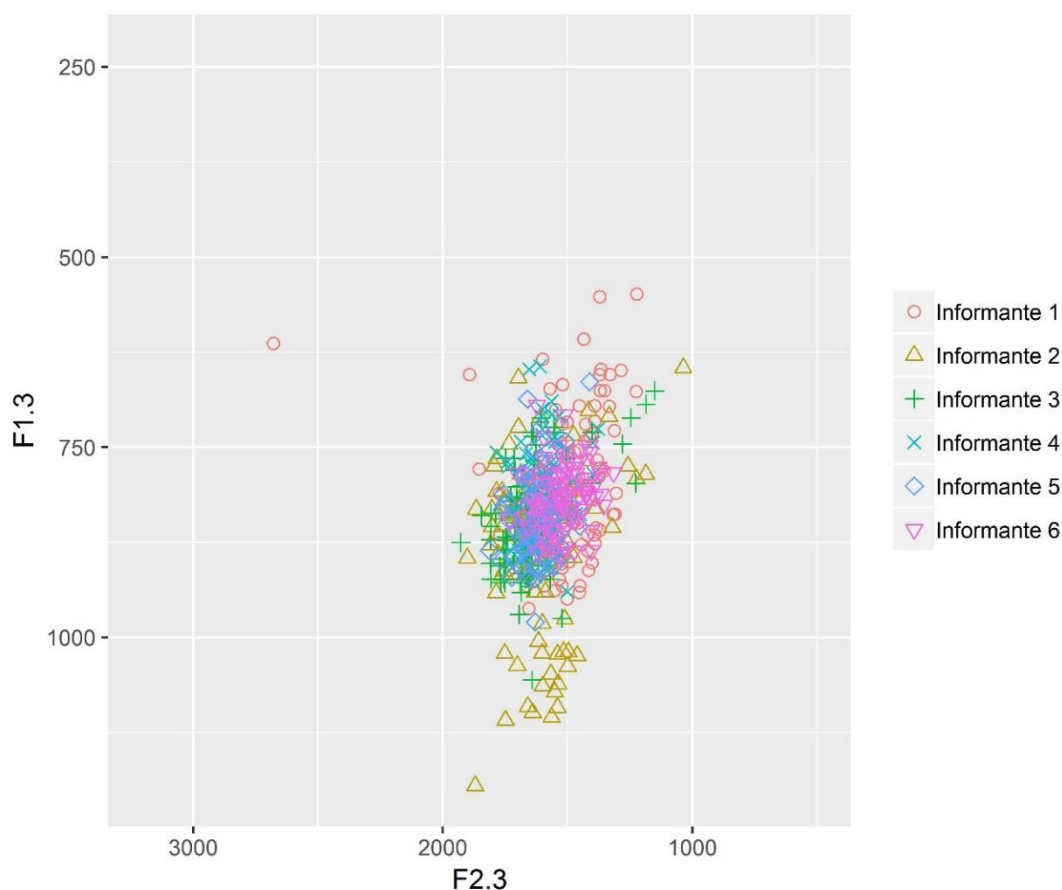


Figura 121. Área de dispersión de la vocal /a/ realizada por informantes sinohablantes en el punto final.

La prueba de Friedman y el análisis de las áreas de dispersión de la vocal /a/ en su inicio, centro y final, han permitido comprobar que las diferencias entre estos tres puntos de las vocales producidas por las estudiantes sinohablantes son significativas. Se ha podido observar que, si bien muchas de las realizaciones presentan características propias de la /a/, hay casos en los que la abertura oral es menor, siendo estos valores similares a los de la /e/ del español. También se han detectado realizaciones en las que las frecuencias en las que aparece F1 superan los 1000 Hz, por lo que la articulación de esta vocal es muy abierta. Además, hay realizaciones de /a/ cuyo F2 se aparece en un rango de frecuencias menor, por lo que la posición de la lengua se asemeja a la de /o/ y /u/ del español. Anteriormente se ha mencionado la existencia de un caso aislado en el que la articulación

de /a/ es semejante a la de la /i/ en español. En el final de la vocal se puede observar que esta es más abierta, ya que el F1 aparece en una franja superior, cercana a los 600 Hz, y la posición de la lengua sigue siendo la misma. En cuanto a la desviación estándar, se ha podido observar que en el primer formante esta aumenta en el centro respecto al inicio, pero disminuye al final; sin embargo, la del segundo formante disminuye considerablemente desde el inicio hasta el final de la vocal, lo que indica que la posición de la lengua resulta menos homogénea en el principio de /a/ que en el final.

## 2.4 Vocal /o/

Los datos obtenidos del análisis de las frecuencias de F1 y F2 de la vocal /o/ en el punto inicial, punto medio y punto final aparecen en la tabla 95. En esta se observa que la frecuencia media en la que aparece el inicio de F1 es similar en los tres puntos de la vocal. En el caso del F2, el punto inicial del formante presenta una frecuencia media superior a la del medio, y esta es ligeramente superior a la del final. En cuanto a la desviación estándar del primer formante, todos los puntos presentan la misma homogeneidad, mientras que, en el caso del segundo formante, la desviación es superior en el inicio que en el centro, y en el final es ligeramente mayor que en el centro, pero inferior también a la del inicio.

Vocal	Formante	Punto inicial		Punto medio		Punto final	
		Frecuencia media	Desv. est.	Frecuencia media	Desv. est.	Frecuencia media	Desv. est.
/o/	F1	526.9	69.7	518.8	69.4	522.2	69.2
	F2	1229.6	177.2	1160.5	125.0	1135.0	139.0

Tabla 95. Frecuencia media y desviación estándar de los tres puntos de F1 y F2 de /o/ producida por las informantes sinohablantes.

Los resultados de la prueba de Friedman (tabla 96) muestran que, en el caso del primer formante, la diferencia entre el inicio y el centro (0.0004) es significativa, mientras que entre el inicio y el final (0.6503) y entre el centro y el final (0.1028) no son significativas. Sin embargo, la significación conjunta de este formante muestra que existen diferencias significativas. En cuanto al segundo formante, las diferencias entre el punto inicial, el medio y el final de la vocal /o/ producida por las informantes sinohablantes son significativas en todos los casos ( $< 0.0001$ ), incluida la significación conjunta ( $< 0.0001$ ).

Formante	Puntos comparados	Significación	Significación conjunta
F1	1 – 2	0.0004	0.006
	1 – 3	0.6503	
	2 – 3	0.1028	
F2	1 – 2	< 0.0001	< 0.0001
	1 – 3	< 0.0001	
	2 – 3	< 0.0001	

Tabla 96. Resultados de la prueba de Friedman para la vocal /o/ realizada por las informantes sinohablantes.

Como muestra la tabla 95, la frecuencia media de F1 en el inicio de la vocal es de 526 Hz, y la del F2 de 1229 Hz. En el primer formante, la desviación estándar es de 69, y en el segundo, de 177. Estos datos permiten conocer que, en la mayoría de los casos, las realizaciones del F1 de /o/ por parte de las informantes sinohablantes se encuentran, aproximadamente, entre 450 y 600 Hz, y las de F2 entre 1000 y 1400 Hz. Sin embargo, en la figura 122 se observa que hay casos en los que la realización de la vocal difiere del comportamiento general. En primer lugar, hay numerosas realizaciones de /o/ cuyo F1 aparece en unas frecuencias más bajas, más semejantes a las propias de la /u/ del español que a la /o/, lo que indica que en esos casos la abertura oral es menor. En segundo lugar, también se detectan realizaciones de /o/ cuyo F1 es mayor que el de la mayoría, lo que indica que la abertura oral es mayor, y se asemeja a la de /e/ del español. En tercer lugar, hay casos en los que el F2 de /o/ supera los 1500 Hz, lo que muestra que la posición de la lengua es más anterior, similar a la de la /a/ del español. Por último, cabe añadir que aparecen realizaciones aisladas en las que la abertura oral es considerablemente superior a la del resto (F1 cercano a 750 Hz), y la posición de la lengua más anterior (F2 cercano a 1500 Hz), características similares a las de la /a/ del español.

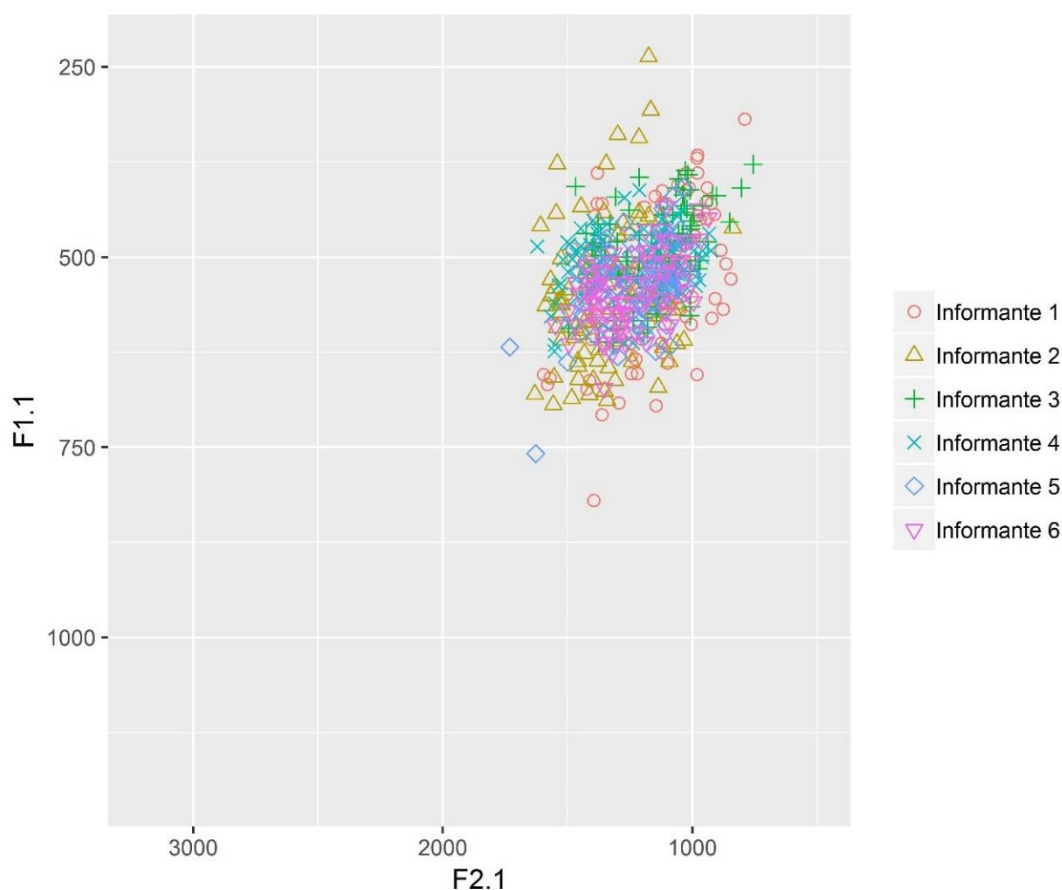
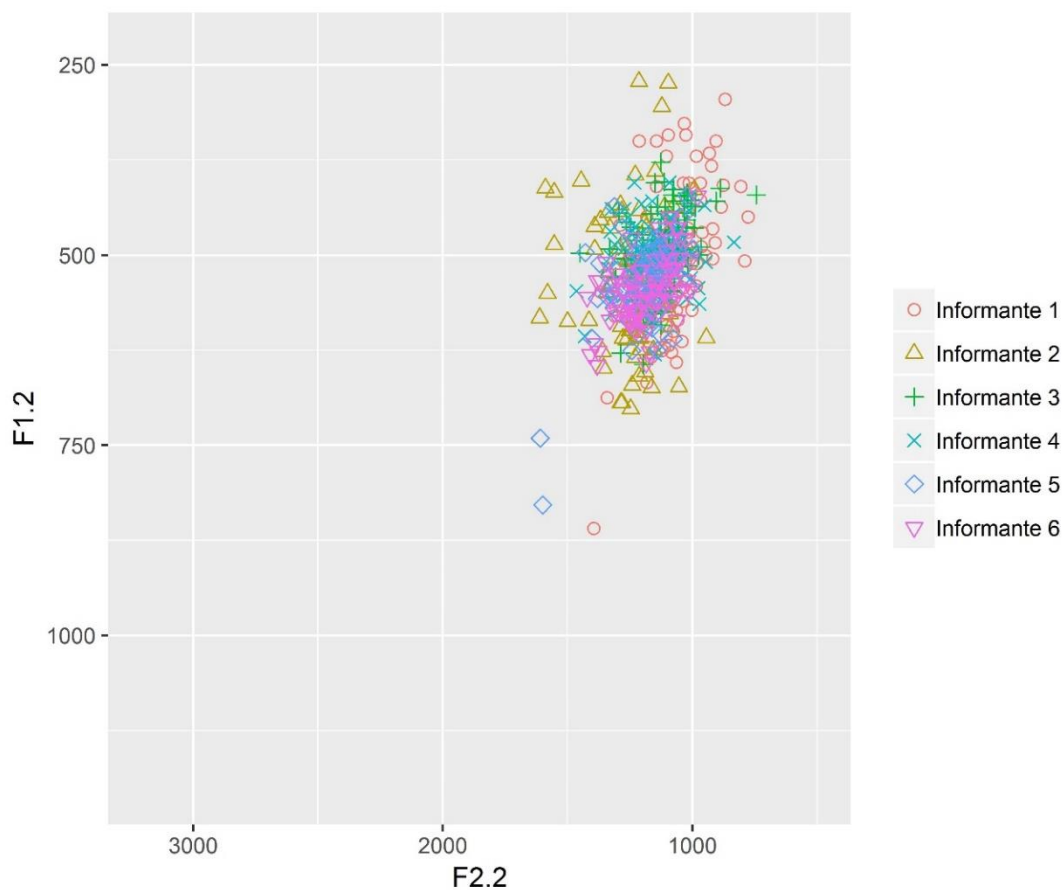


Figura 122. Área de dispersión de la vocal /o/ realizada por informantes sinohablantes en el punto inicial.

El F1 de /o/ producida por las informantes sinohablantes tiene en su centro una frecuencia media de 518 Hz, y el F2 de 1160 Hz. La desviación estándar del primero es de 69, y la del segundo, de 125, lo que indica que, si bien se mantiene en el F1, la del F2 disminuye respecto al inicio de la vocal. Estos datos muestran que, en este punto, la mayoría de realizaciones de F1 aparecen entre 450 y 600 Hz, y las de F2 entre 1000 y 1300 Hz, por lo que la dispersión del segundo formante es más reducida, como se puede observar en la figura 123. A las realizaciones descritas en el punto inicial de la /o/, cabe añadir que en este punto el número de casos en los que el F1 es más bajo es más numeroso, por lo que hay más realizaciones más cerradas que en el inicio, algunas de ellas articuladas con una abertura oral más semejante a la de la /u/ del español.



*Figura 123.* Área de dispersión de la vocal /o/ realizada por informantes sinohablantes en el punto medio.

En lo que respecta al final de la vocal, la frecuencia media en la que aparece F1 es de 522 Hz, y la de F2 de 1135 Hz. La desviación estándar del primer formante es de 69, y la del segundo, de 139. Esto muestra que en F1 la desviación es mantiene, mientras que en F2 aumenta respecto al centro, aunque es menor que en el inicio de /o/. En este punto, F1 aparece, generalmente, entre 450 y 600 Hz, y F2 entre 900 y 1300 Hz. Como se puede observar en la figura 124, la mayor diferencia respecto al centro de /o/ es que disminuye el número de casos en los que F1 es menor, la posición de la lengua presenta menor homogeneidad, ya que las realizaciones están distribuidas en una franja más amplia del segundo formante.

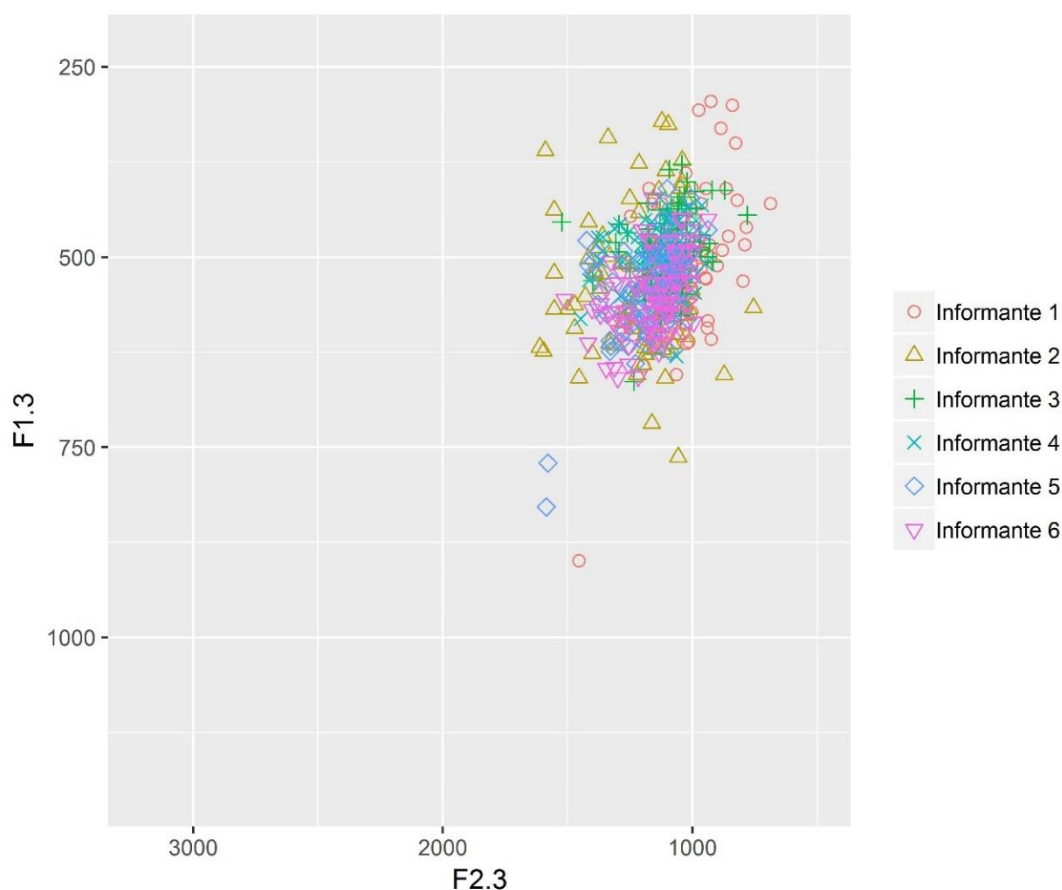


Figura 124. Área de dispersión de la vocal /o/ realizada por informantes sinohablantes en el punto final.

La prueba de Friedman y el estudio de las áreas de dispersión de /o/ en su inicio, centro y final han permitido comprobar que las diferencias entre estos tres puntos de las vocales producidas por las estudiantes sinohablantes son significativas a nivel global. La realización de este sonido vocálico no es homogénea, como se ha podido observar. Por un lado, se han detectado realizaciones en las que la abertura oral es menor y similar a la de la /u/ del español, mientras que en otros casos la abertura oral es mayor, y se asemeja a la de /e/. También se han podido detectar vocales en las que la realización es considerablemente más cerrada que el resto, cercanas a la /u/ del español. Además, en algunas ocasiones la abertura oral y la posición de la lengua se asemejan, también, a las de /a/. En cuanto a la desviación estándar, se ha observado que esta se mantiene en todos los puntos de F1, mientras que en el caso de F2 desciende en el centro respecto al inicio, pero en el final es mayor que en el centro, lo que indica que la abertura oral presenta, generalmente, la misma homogeneidad en las realizaciones, mientras que la posición de la lengua presenta variaciones.

## 2.5 Vocal /u/

Los datos obtenidos en el análisis de las frecuencias de F1 y F2 de la vocal /u/ en el punto inicial, punto medio y punto final aparecen en la tabla 97. En esta se observa que la frecuencia media en la que aparece el F1 es similar en todos los puntos, y lo mismo sucede con el F2. En cuanto a la desviación estándar del primer formante, esta es muy similar en todos los puntos analizados, mientras que, en el caso del segundo formante, el inicio presenta una homogeneidad menor que el centro y el final, y el punto medio es el que presenta menor desviación, por lo tanto, es el punto más homogéneo del formante.

Vocal	Formante	Punto inicial		Punto medio		Punto final	
		Frecuencia media	Desv. est.	Frecuencia media	Desv. est.	Frecuencia media	Desv. est.
/u/	F1	456.9	55.6	444.5	53.6	449.4	53.2
	F2	1116.8	187.3	1097.7	151.2	1104.7	160.4

Tabla 97. Frecuencia media y desviación estándar de los tres puntos de F1 y F2 de /u/ producida por las informantes sinohablantes.

De acuerdo con los resultados de la prueba de Friedman, tanto las diferencias entre los distintos puntos del primer formante de /u/ realizada por las informantes sinohablantes, como la significación conjunta son significativas ( $< 0.0001$ ). En cambio, la diferencia entre el inicio y el centro del segundo formante es significativa (0.02), mientras que entre el inicio y el final (0.495), y entre el centro y el final (1). Como consecuencia, las diferencias entre los distintos puntos del segundo formante no son significativas a nivel global.

Formante	Puntos comparados	Significación	Significación conjunta
F1	1 – 2	$< 0.0001$	$< 0.0001$
	1 – 3	0.001	
	2 – 3	0.004	
F2	1 – 2	0.020	0.593
	1 – 3	0.495	
	2 – 3	1	

Tabla 98. Resultados de la prueba de Friedman para la vocal /u/ realizada por las informantes sinohablantes.

De acuerdo con los datos mostrados en la tabla 97, la frecuencia media de F1 en el inicio de la vocal es de 456 Hz, y la de F2 de 1116 Hz. En el primer formante, la desviación estándar es de 55, y en el segundo, de 187. Esta información permite saber que, en la mayoría de los casos, las realizaciones del F1 de /u/ se encuentran, aproximadamente, entre 400 y 550 Hz, y las de F2 entre 900 y 1300 Hz. Sin embargo, igual que sucede con

las vocales anteriormente analizadas, en la figura 125 se observa que hay muchos casos en los que la realización de /u/ por parte de las estudiantes chinas no se asemeja a la del español.

En primer lugar, se detecta que hay numerosos casos en los que F1 aparece en un rango de frecuencias cercana a 300 Hz, lo que indica que la abertura oral es menor que en el resto de realizaciones. En segundo lugar, aparecen casos en los que la abertura oral es mayor que la de la mayoría de realizaciones, más cercana a 600 Hz. Este valor se asemeja al del primer formante de /e/ del español. En tercer lugar, se observa que, si bien el F1 presenta los mismos valores que la mayoría (entre 400 y 550 Hz), el F2 es algo menor de 900 Hz, lo que indica que la realización es más posterior.

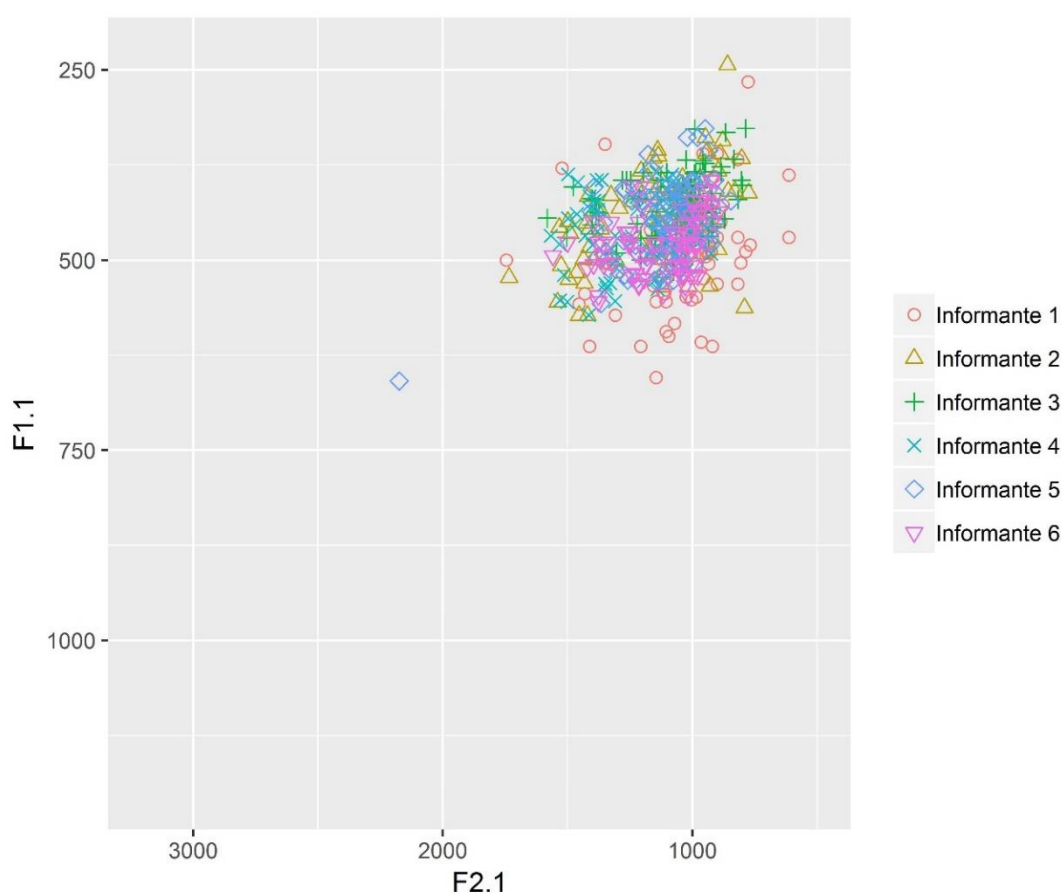


Figura 125. Área de dispersión de la vocal /u/ realizada por informantes sinohablantes en el punto inicial.

El centro de /u/ tiene una frecuencia media de 444 Hz en su F1, y de 1097 en su F2. La desviación estándar del primer formante es de 53, y la del segundo, de 151, lo que indica que en el centro de F2 hay menos dispersión que en el inicio. En este punto de la vocal, los datos muestran que la mayoría de realizaciones del primer formante se concentran en un rango de frecuencias entre 350 y 500 Hz, y las del segundo, entre 900 y 1200 Hz, lo



que muestra que en este punto las realizaciones son, generalmente, más cerradas, y hay mayor homogeneidad en lo que a la posición de la lengua se refiere. Igual que sucede en el punto inicial de /u/ producida por las informantes sinohablantes, y a las realizaciones señaladas en el punto inicial hay que añadir que hay casos aislados en los que el primer formante aparece alrededor de 700 Hz, lo que indica que la abertura oral es mayor.

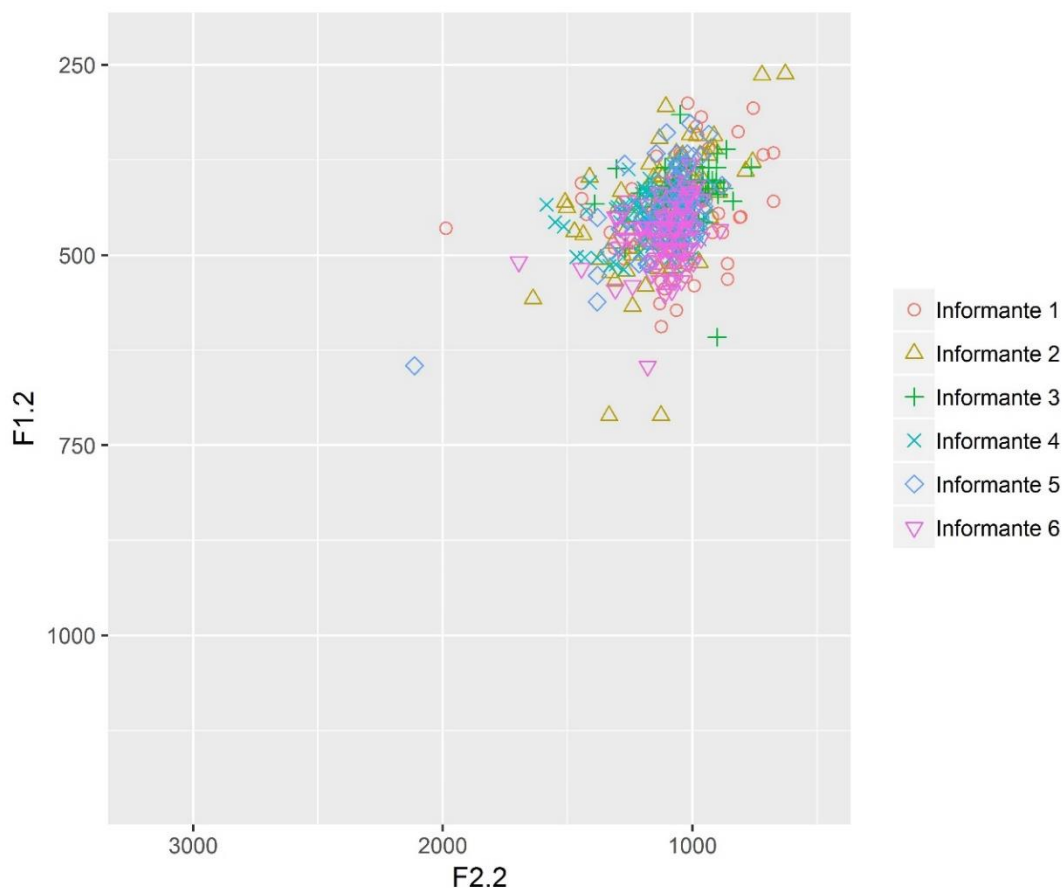


Figura 126. Área de dispersión de la vocal /u/ realizada por informantes sinohablantes en el punto medio.

En cuanto al final de la vocal, la media de F1 es de 449 Hz, y la de F2 de 1104 Hz. La desviación estándar, de 53 en F1 y de 160 en F2, por lo que en este punto es ligeramente superior respecto a la del centro del segundo formante. De acuerdo con estos datos, la mayoría de realizaciones de F1 aparecen entre 350 y 500 Hz, y las de F2 entre 900 y 1300 Hz, lo que indica que la dispersión del segundo formante en el final del mismo es superior que la del centro. Además de los casos señalados en los puntos anteriores, en la figura 127 se observa que en el final de /u/ aparecen más realizaciones en las que F1 es similar al de la mayoría de realizaciones, y el F2 es menor de 1000 Hz. Esto indica que hay una articulación más posterior de dichas vocales.

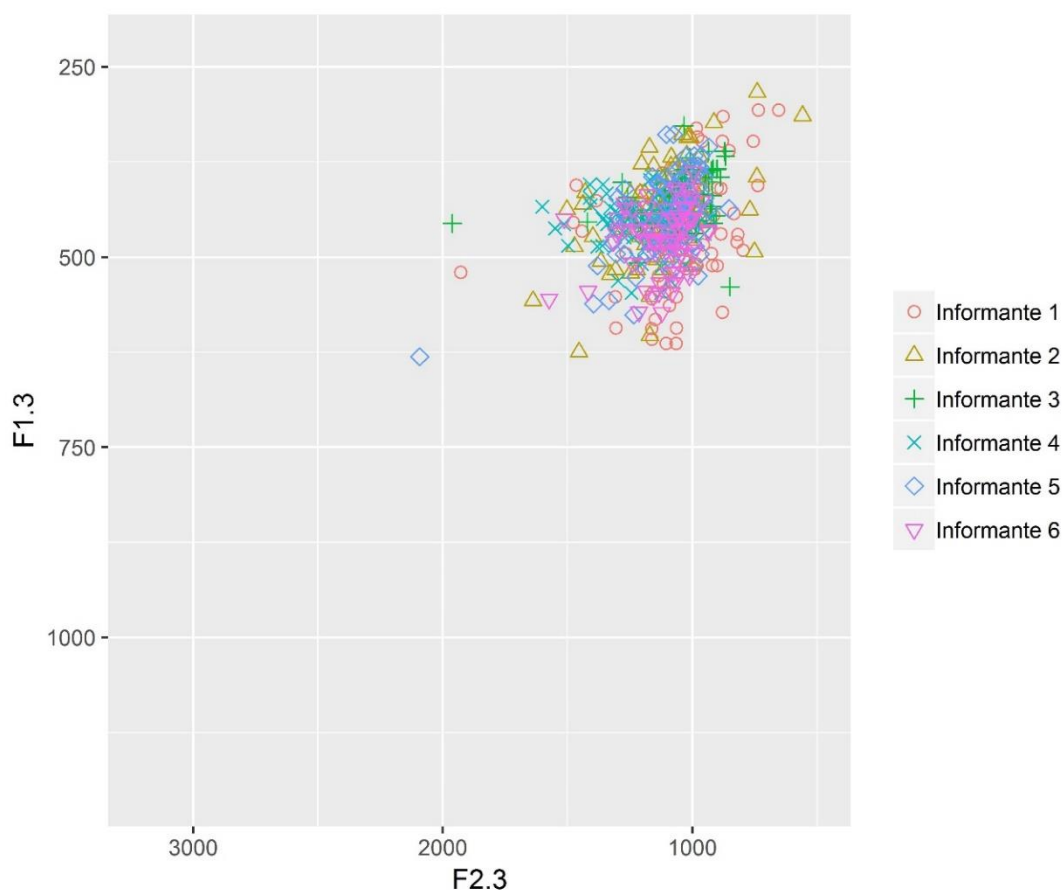


Figura 127. Área de dispersión de la vocal /u/ realizada por informantes sinohablantes en el punto final.

La prueba de Friedman y el análisis de las áreas de dispersión de la vocal /u/ en su inicio, centro y final, han permitido comprobar que las diferencias entre los tres puntos de F1 son significativas, mientras que en el caso de F2 no son significativas. Además, se ha podido detectar que, si bien hay un amplio número de casos en los que la realización es similar a la de /u/ del español, hay casos en los que dicha vocal es articulada de distintas maneras. Por un lado, hay casos en los que F1 y F2 tienen valores inferiores al resto, por lo que la abertura oral es menor, y la posición de la lengua más posterior. Además, se han detectado vocales cuyo F1 se asemeja al propio de la /e/ del español. En cuanto a la desviación estándar, esta es similar en los tres puntos del primer formante, mientras que en el segundo es mayor al inicio, descende en el centro, y es ligeramente superior al anterior en el final, lo que indica que, si bien la abertura oral se mantiene similar en todos los puntos, la posición de la lengua sufre más variaciones.

## 2.6 Recapitulación

El análisis de las vocales producidas por las informantes sinohablantes en tres puntos, y el estudio de las áreas de dispersión correspondientes a cada uno de ellos ha permitido describir el comportamiento de las estudiantes sinohablantes durante el proceso de aprendizaje del español. En las realizaciones de dichas hablantes se han detectado numerosos casos en los que los formantes presentan valores más altos o más bajos, por lo que la realización es más abierta o más cerrada, y más anterior o más posterior que la de las hablantes nativas. A continuación se resumen brevemente los aspectos más relevantes de estas diferencias relativos a cada una de las vocales.

En relación con la vocal /i/, se ha comprobado que los tres puntos del primer formante aparecen en el mismo rango de frecuencias; en el caso del segundo formante, las frecuencias en las que aparece el inicio son inferiores que las del punto medio y el final, las cuales son semejantes. En cuanto a la desviación estándar, esta es ligeramente superior en el inicio que en el centro y el final, lo que indica que la homogeneidad en las realizaciones es mayor a partir del punto medio de la vocal. Cabe señalar, además, que las diferencias entre los tres puntos de ambos formantes son altamente significativas.

Si bien hay un amplio número de realizaciones que presentan las características propias de la /i/ del español en lo que abertura oral y posición de la lengua se refiere, se han detectado casos en los que hay realizaciones más abiertas o más cerradas, y más posteriores o más anteriores. Además, se han detectado otro tipo de realizaciones, como aquellas en las que los valores del primer formante se aproximan a los de la /e/ del español, o los casos aislados en los que el segundo formante es muy posterior, por lo que la posición de la lengua es semejante a la de la /u/ del español.

En cuanto a la vocal /e/, igual que sucede en /i/, la frecuencia media en la que aparece el inicio del primer formante es ligeramente inferior que la del centro y el final del mismo. En el segundo formante el inicio también aparece en un rango de frecuencias inferior que el centro, el cual, a su vez, aparece en unas frecuencias ligeramente superiores que el final. La realización del primer formante presenta la misma variación en los tres puntos, mientras que en el segundo formante, la

variación aumenta desde el inicio hasta el final, lo que indica que este último punto de la vocal presenta más variación que el principio. Cabe añadir que las diferencias entre los distintos puntos, tanto del primer formante, como del segundo, son muy significativas.

Si bien muchas de las realizaciones de esta vocal por parte de las informantes sinohablantes presentan las características propias de la /e/ del español en lo que a la abertura oral y la posición de la lengua se refiere, se dan numerosos casos en los que los valores de los formantes se aproximan a los de otras vocales del español. Ejemplo de ello es que hay casos en los que tanto el primer formante, como el segundo, aparecen en unas frecuencias propias de la /a/ del español; o realizaciones de /e/ en las que el primer formante aparece en frecuencias más bajas, propias de la /i/ del español.

En lo que concierne a la vocal /a/ realizada por las informantes sinohablantes, se ha mostrado que la frecuencia media en la que aparece el inicio de F1 es inferior que la del centro y el final, y en el caso del F2, el inicio presenta una frecuencia media más elevada que los puntos medio y final, los cuales aparecen en el mismo rango de frecuencias. Mientras que la desviación estándar del primer formante muestra que la variación de los tres puntos analizados es similar, la del segundo formante es mayor en el inicio que en el centro, y la del centro es superior que la del final. En cuanto a las diferencias entre los distintos puntos de F1 y F2, estas son altamente significativas.

Igual que sucede en /i/ y /e/, en numerosas realizaciones de /a/ la abertura oral, la posición de la lengua, o ambas, no son las características de esta vocal en español. Así, hay casos en los que el primer formante aparece en un rango de frecuencias propio de la /e/ del español, y otros en el que el segundo aparece en unas frecuencias similares a las de la /o/ y la /u/ del español.

En cuanto a la realización de /o/ por las estudiantes chinas, la frecuencia media en la que aparece F1 es similar en los tres puntos, mientras que en el caso del F2 el inicio tiene una frecuencia media superior que el centro, y este es, a su vez, más alto que el final. Mientras que los tres puntos del primer formante presentan la misma desviación estándar, el centro presenta la menor variación, mientras que el inicio y el final presentan una mayor desviación respecto a la media que el punto

medio. La significación conjunta del primer y segundo formante muestra que las diferencias entre los distintos puntos son muy significativas.

En cuanto a las realizaciones que se alejan de las características propias de la /o/ del español, se han detectado casos en los que el primer formante es más bajo y las frecuencias son semejantes a las de la /u/ del español; también hay casos en los que la abertura oral es mayor y las frecuencias cercanas a las de la /e/ del español. Además, hay realizaciones de /o/ en las que la posición de la lengua es más anterior, semejante a la de la /a/ del español.

Finalmente, en relación con la vocal /u/, la frecuencia media en la que aparece el primer formante es similar en todos los puntos, y lo mismo sucede con el segundo formante. En cuanto a la desviación estándar, esta es semejante en los tres puntos del F1, mientras que en el F2 es mayor en el inicio que en el centro y el final, por lo que la realización en este punto es menos homogénea. En cuanto a las diferencias entre los distintos puntos analizados, se ha comprobado que en el caso del primer formante son significativas, mientras que en el del segundo no lo son.

Igual que sucede en las demás vocales, aquí también se han detectado casos en los que las realizaciones de las informantes sinohablantes se alejan de las características propias del primer y segundo formante de la /u/ del español. Ejemplo de ello es que se ha podido observar que en ocasiones el F1 aparece en unas altas, propias de la /e/ del español, lo que indica que la abertura oral es considerablemente mayor; y también hay casos en los que el F1 o el F2 aparecen en unas frecuencias medias más bajas que las propias de la /u/ del español, lo que indica que la realización de la misma por parte de las estudiantes chinas es, por un lado, más cerrada, y, por otro lado, más posterior.

Teniendo en cuenta los datos aquí presentados respecto a la realización de las vocales del español por las informantes sinohablantes cabe señalar que, tal como se esperaba, el comportamiento fónico de las hablantes estudiadas es diferente del de las informantes nativas por lo que los timbres de las vocales producidas por las estudiantes de español poseen características distintas de los de las hablantes del grupo de control, como se explicará en el apartado siguiente.

### **3. DIFERENCIAS ENTRE EL GRUPO DE CONTROL Y EL DE INFORMANTES SINOHABLANTES**

En el presente apartado se realiza una comparación de los resultados obtenidos en el análisis del timbre de las vocales articuladas por el grupo de control y por el de informantes sinohablantes. En primer lugar, se comparan las frecuencias medias y la desviación estándar de ambos grupos. En segundo lugar, se indica si las diferencias entre ambos grupos son significativas. Para ello, se ha realizado la prueba de Mann Whitney, una prueba paramétrica aplicada a dos muestras independientes. Las decisiones estadísticas se han realizado tomando como nivel de significación el valor 0.05. Es decir, si el valor obtenido es mayor que 0.05, no existen diferencias significativas entre el grupo de control y las informantes sinohablantes; si dicho valor es menor que 0.05, las diferencias sí son significativas, por lo que se asume que la lengua materna interfiere en la abertura oral y/o en la posición de la lengua de las informantes sinohablantes cuando articulan las vocales del español. En tercer lugar, se comparan las áreas de dispersión de las vocales producidas por ambos grupo, de manera que se pueden analizar las diferencias en las realizaciones de unas y otras, la dispersión, y aquellas articulaciones que se alejan de las características de las vocales del español.

#### **3.1 Vocal /i/**

En la tabla 99 se muestra, por un lado, la frecuencia media en la que aparecen el punto inicial, el central y el final del primer y segundo formante de la vocal /i/ articulada por el grupo de control y por las informantes sinohablantes, además de la desviación estándar de cada uno de ellos. En la columna de la derecha ('significación') se muestra el valor obtenido en la prueba de Mann Whitney, es decir, el nivel de significación de las diferencias entre un grupo y el otro.

En relación con el primer formante, los datos indican que, si bien las frecuencias en las que aparecen los distintos puntos de F1 son similares, las frecuencias son ligeramente más altas cuando /i/ es articulada por las hablantes del grupo de control, ya que en el caso de estas la media del punto inicial es de 476 Hz, la del medio de 461 Hz y la del final de 454 Hz. En cuanto a las informantes sinohablantes, el punto inicial tiene una frecuencia media de 459 Hz, el medio 451 Hz, y el final 446 Hz. Sin embargo, la desviación estándar es ligeramente superior cuando /i/ es articulada por las estudiantes chinas (65, 58 y 57) que cuando las producen las hablantes nativas de español (56, 47 y

52), lo que muestra que la realización de esta vocal por parte de las hablantes que conforman el grupo de control es más homogénea.

La frecuencia media en la que aparecen los tres puntos del F2 de /i/ articulada por las hablantes del grupo de control es de 2432, 2555 y 2510 Hz respectivamente, mientras que en el caso de las informantes sinohablantes es de 2436, 2531 y 2516 Hz respectivamente. Como se puede observar, las medias son muy similares; sin embargo, la desviación estándar es mayor en el caso de las estudiantes chinas (276, 267 y 268) que en el de las mujeres hispanohablantes nativas (157, 148 y 176), lo que indica que las realizaciones de estas últimas son más homogéneas que las de las aprendices de español.

			<b>Control</b>	<b>Sinohablantes</b>	<b>Significación</b>
<b>Punto inicial</b>	<b>F1</b>	Media	476.2	459.2	< 0.0001
		DS	56.9	65.5	
	<b>F2</b>	Media	2432.3	2436.4	< 0.0001
		DS	157.6	276.2	
<b>Punto medio</b>	<b>F1</b>	Media	461.2	451.2	< 0.0001
		DS	47.2	58.7	
	<b>F2</b>	Media	2555.8	2531.8	< 0.0001
		DS	148.9	267.8	
<b>Punto final</b>	<b>F1</b>	Media	454.9	446.2	< 0.0001
		DS	52.8	57.8	
	<b>F2</b>	Media	2510.3	2516.0	< 0.0001
		DS	176.9	268.7	

*Tabla 99.* Comparación de frecuencias medias y desviación estándar de la vocal /i/ producida por el grupo de control y las informantes sinohablantes.

Los resultados obtenidos en la prueba de Mann Whitney muestran que, tanto en los tres puntos del primer formante, como en los del segundo, las diferencias entre ambos grupos son significativas, dado que el valor obtenido en esta prueba es menor que 0.05. Esto indica que la lengua materna de las informantes interfiere en su realización de la vocal /i/, tanto en lo que a la abertura oral se refiere, como en lo que a la posición de la lengua se refiere.

A continuación se analizan las áreas de dispersión con el fin de mostrar las diferencias entre ambos grupos de estudio que, como ya se ha indicado, son significativas, y señalar aquellas realizaciones que se alejan de las características de la /i/ del español en lo que a abertura oral y/o posición de la lengua se refiere.

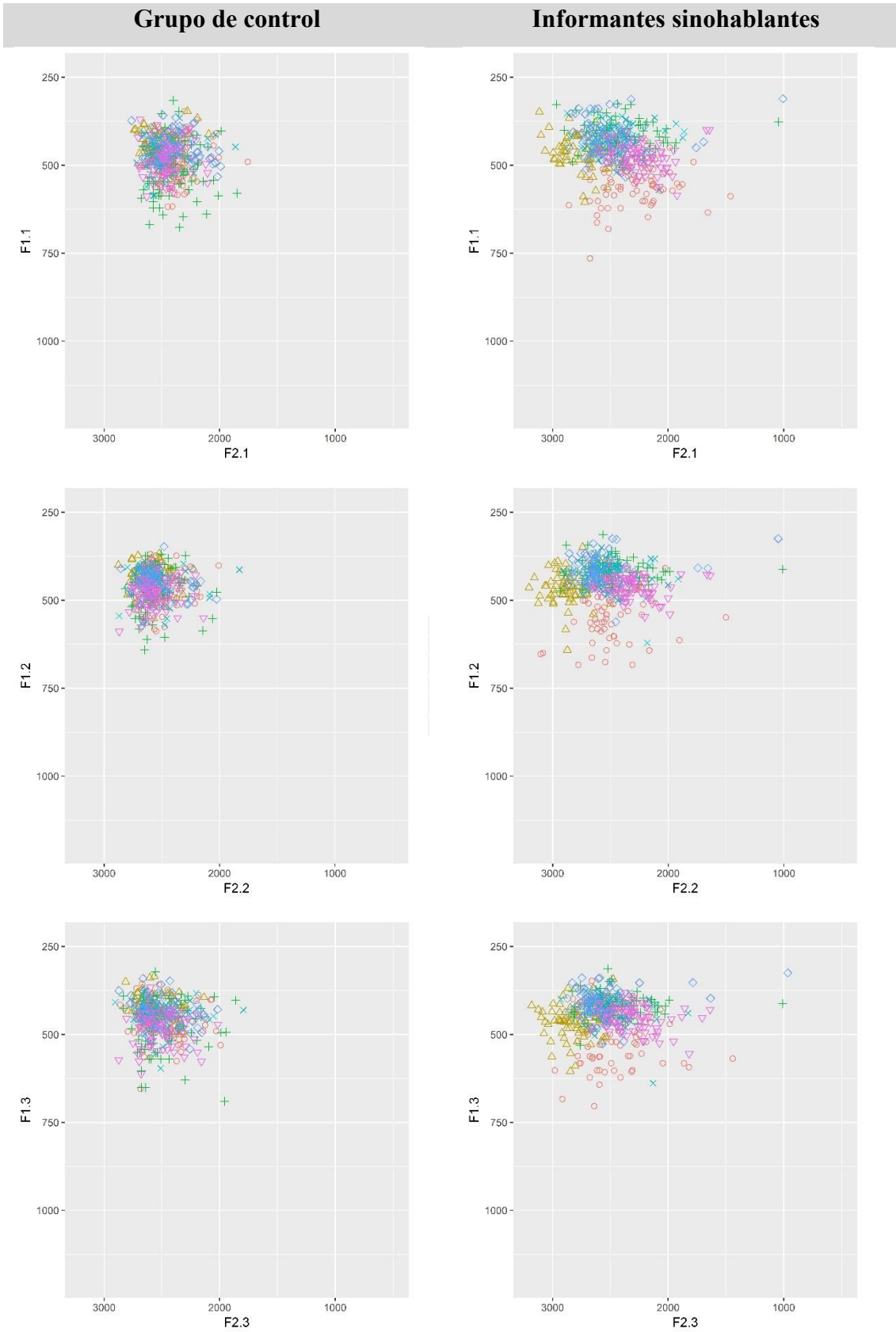


Figura 128. Comparación de las áreas de dispersión de la vocal /i/ en su inicio (F1.1 y F2.1), centro (F1.2 y F2.2) y final (F1.3 y F2.3).



Como se ha indicado anteriormente, las realizaciones del grupo de control son más homogéneas que las de las informantes sinohablantes, aspecto que se aprecia en los tres puntos de la vocal /i/ mostrados en la figura 128.

De acuerdo con los datos mostrados en la tabla 99, en el inicio de la vocal la mayoría de realizaciones del F1 del grupo de control aparecen entre 400 y 550 Hz, mientras que las de las informantes sinohablantes se encuentran entre 350 y 550 Hz. En cuanto al F2, en el caso del primer grupo, las realizaciones se concentran entre 2200 y 2600 Hz, mientras que las del segundo entre 2100 y 2700 Hz. Esto muestra que hay mayor variación en las realizaciones de las hablantes chinas. En el área correspondiente al grupo de control se puede observar que hay casos en los que F1 aparece en frecuencias más bajas, por lo que la vocal es más cerrada, y otros en los que /i/ es más abierta, y las frecuencias más altas; además, en algunos casos F2 aparece en un rango de frecuencias inferior al de la mayoría, lo que indica que la posición de la lengua es menos anterior que el resto. En cuanto a las realizaciones de las informantes sinohablantes, se observa que, igual que sucede en las del grupo de control, hay casos en los que la abertura oral es mayor o menor que la de la mayoría de realizaciones, y la posición de la lengua es más anterior o más posterior. Sin embargo, hay casos en los que el primer formante, el segundo, o ambos, aparecen en un rango de frecuencias que no son los característicos de la /i/ del español. Ejemplo de ello son los casos en los que las frecuencias del F1 son más elevadas, y se asemejan a las de la /e/ del español. También hay casos aislados en los que el F2 es cercano a 1000 Hz, por lo que la posición de la lengua es más posterior, similar a la de la /u/ del español.

En el centro, la mayoría de realizaciones de /i/ articulada por las informantes que conforman el grupo de control son similares a las del inicio de la vocal. En la mayoría de casos, el F1 se mantiene en el mismo rango de frecuencias, aunque el de F2 es ligeramente superior (entre 2400 y 2700 Hz), lo que indica que la posición de la lengua levemente más anterior que la del inicio y, también, que la dispersión es menor. En cuanto a las informantes sinohablantes, el rango de frecuencias de F1 en el que aparecen la mayoría de las realizaciones es de 400 - 500 Hz, mientras que el de F2 es ligeramente superior, entre 2300 y 2800 Hz. En cuanto a las realizaciones que presentan características diferentes de las de la /i/ del español, se observa que se mantienen las mismas articulaciones que en el inicio de la vocal.

En el final de la vocal, el rango de frecuencias en el que aparecen la mayoría de realizaciones del F1 y el F2 de /i/ articulada, tanto por el grupo de control, como por el

de las informantes sinohablantes, es similar al del centro de dicho sonido vocálico, aunque en ambos formantes se detecta una dispersión ligeramente superior que en el centro. Cabe añadir que en este punto se mantienen las realizaciones cuyas características no son las propias de la /i/ del español.

### **3.2 Vocal /e/**

Los datos relativos a la realización de la vocal /e/ por ambos grupos aparecen en la tabla 100. En relación con el primer formante, se observa que la frecuencia media en la que aparece el punto inicial es igual para el grupo de control y para las informantes sinohablantes. En cuanto al punto central y el final, la frecuencia media es ligeramente más alta cuando /e/ es articulada por las informantes sinohablantes (634 y 641 Hz respectivamente) que cuando es producida por las hablantes del grupo de control (612 y 610 Hz respectivamente). La desviación estándar es mayor en la producción de las estudiantes chinas (64, 72 y 73 en cada uno de los puntos), lo que indica que la abertura oral es más homogénea cuando /e/ es realizada por las hablantes que conforman el grupo de control.

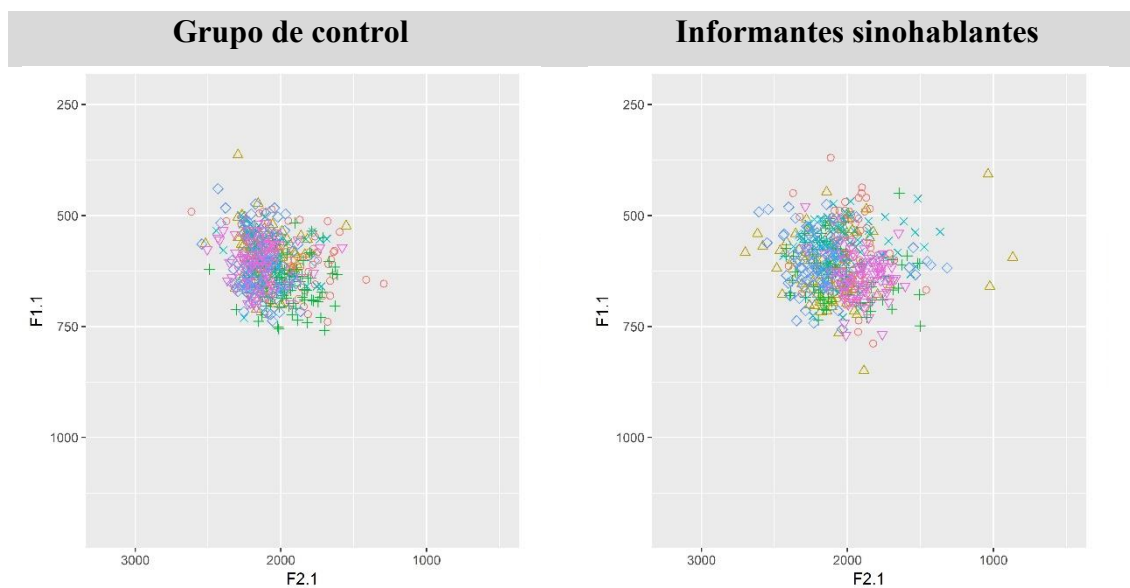
En relación con el segundo formante, este presenta una frecuencia media inferior en todos sus puntos cuando /e/ es articulada por las informantes sinohablantes (2039, 2098 y 2081 Hz respectivamente) que cuando la producen las hablantes del grupo de control (2074, 2199 y 2181 Hz respectivamente). En cuanto a la desviación estándar, esta es mayor en el caso de las hablantes chinas (227, 233 y 240 respectivamente) que en el de las hispanohablantes nativas (178, 162 y 176 respectivamente). Además, en el caso de las primeras se observa que la desviación aumenta ligeramente desde el punto inicial hasta el punto final del formante, mientras que en el de las segundas disminuye del inicio al centro, y aumenta del centro al final. Esto muestra que la realización de F2 es más homogénea por parte del grupo de control que por la de las estudiantes chinas, igual que sucede con el primer formante, y que, en el caso de las informantes sinohablantes, la dispersión es mayor en el final de F2 que en el inicio.

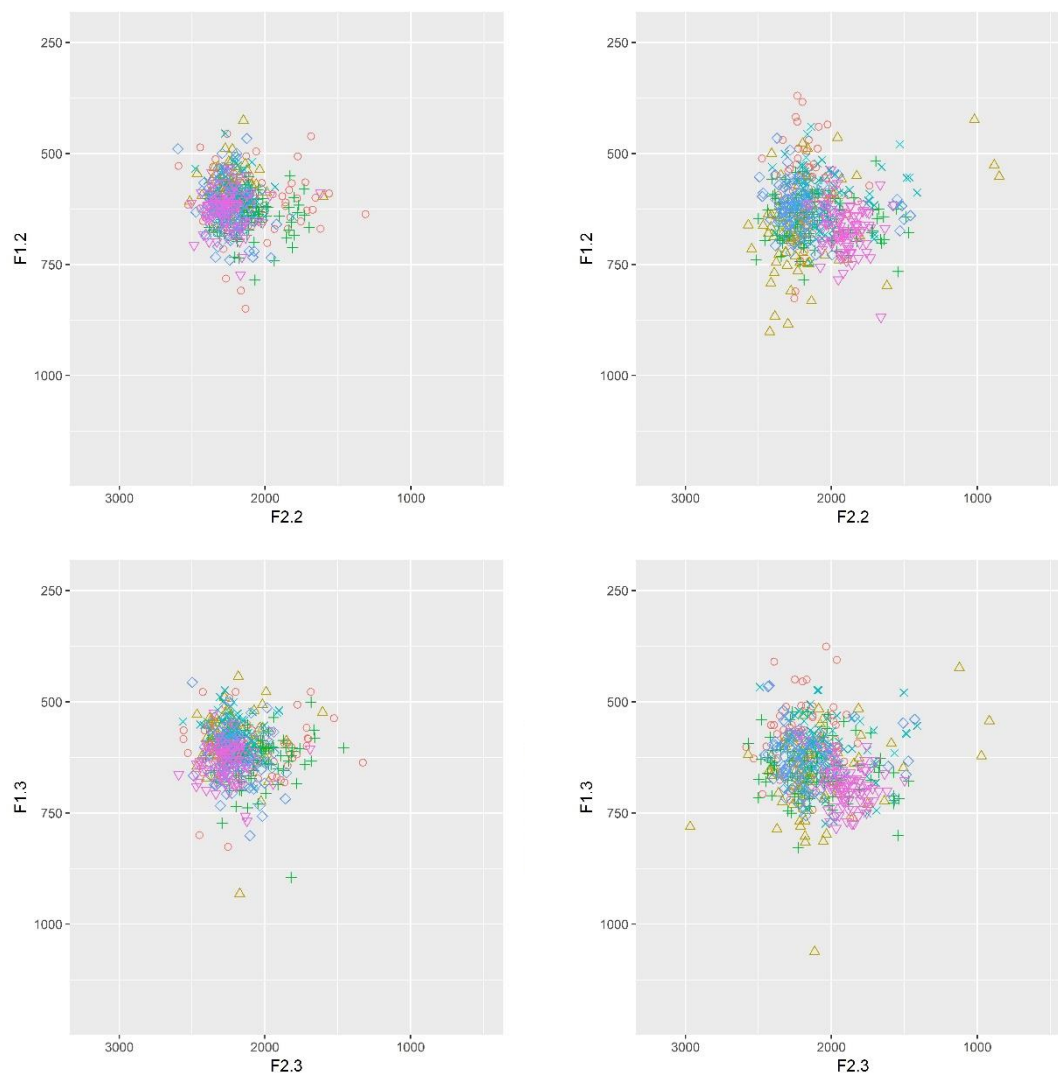
			Control	Sinohablantes	Significación
Punto inicial	F1	Media	609.7	608.2	0.9452
		DS	59.1	67.4	
	F2	Media	2074.4	2039.7	0.0024
		DS	178.6	227.3	
Punto medio	F1	Media	612.8	634.9	< 0.0001
		DS	52.1	72.4	
	F2	Media	2199.7	2098.0	< 0.0001
		DS	162.2	233.9	
Punto final	F1	Media	610.9	641.5	< 0.0001
		DS	56.2	73.0	
	F2	Media	2181.1	2081.3	< 0.0001
		DS	176.1	240.5	

Tabla 100. Comparación de frecuencias medias y desviación estándar de la vocal /e/ producida por el grupo de control y las informantes sinohablantes.

Los resultados de la prueba de Mann Whitney muestran que las diferencias entre la vocal /e/ realizada por el grupo de control y por las informantes sinohablantes son significativas, dado que el valor obtenido es menor que 0.05, excepto el punto inicial del F1, en el cual el índice de significación es de 0.9452, por lo que la diferencia entre ambos grupos en este caso no es significativa. Esto indica que, a excepción del punto inicial del F1, la lengua materna de las informantes interfiere en su realización de la vocal /e/.

A continuación se analizan las áreas de dispersión con el fin de mostrar las diferencias entre ambos grupos de estudio y señalar aquellas realizaciones que se alejan de las características de la /e/ del español en el caso de las informantes sinohablantes.





*Figura 129.* Comparación de las áreas de dispersión de la vocal /e/ en su inicio (F1.1 y F2.1), centro (F1.2 y F2.2) y final (F1.3 y F2.3).

En la figura 129, en la cual aparecen las áreas de dispersión correspondientes a cada uno de los puntos de la vocal /e/ realizada por los dos grupos de estudio, se observa que, tal y como se ha indicado anteriormente, las realizaciones de las hablantes nativas de español son más homogéneas que las de las estudiantes chinas, igual que sucede con la vocal /i/.

De acuerdo con los datos mostrados en la tabla 100, en el inicio de la vocal la mayoría de las realizaciones del F1 del grupo de control aparecen entre 550 y 680 Hz, mientras que las de las informantes sinohablantes se encuentran entre 500 y 700 Hz. En cuanto al F2, en el caso del primer grupo, las realizaciones se concentran entre 1900 y 2250 Hz, mientras que las del segundo entre 1800 y 2300 Hz. En el caso del grupo de control, hay casos en los que la realización de /e/ es más abierta (mayor F1), y un número reducido de ocasiones en las que la articulación de esta vocal es más cerrada (menor F1).

En cuanto a las informantes sinohablantes, además de los casos en los que, tanto el F1, como el F2 presentan los mismos valores que la /e/ del español articulada por las hablantes del grupo de control, hay realizaciones en las que los formantes aparecen en un rango de frecuencias diferente. En primer lugar, se observa que el número de casos en el que la frecuencia del primer formante es inferior a 500 Hz es mayor que en el de las hablantes del grupo de control, lo que indica que las estudiantes chinas realizan la /e/ más cerrada en más ocasiones, y en algunos casos la abertura oral es más cercana a la propia de la /i/ del español que a la de /e/. Además, hay casos aislados en los que la posición de la lengua es posterior, ya que el F2 aparece en torno a los 1000 Hz, valor propio de la /u/ del español, y alejado de la posición de la lengua característica de la /e/.

En el centro de la vocal, la mayoría de realizaciones del F1 del grupo de control aparecen entre 550 y 670 Hz, mientras que las de las informantes sinohablantes se encuentran entre 550 y 750 Hz. En cuanto al F2, en el caso del primer grupo, las realizaciones se concentran entre 2000 y 2350 Hz, y las del segundo entre 1800 y 2500 Hz. Las informantes hispanohablantes nativas mantienen las realizaciones más abiertas /e/, y también las más cerradas. En cuanto a las estudiantes chinas, como se ha indicado, el primer formante aparece en un rango de frecuencias ligeramente superior al del inicio de la vocal, por lo que, en general, todas las realizaciones son más abiertas en este punto. Aquí se mantienen las mismas realizaciones que en el inicio de la vocal se alejan de las características propias de la /e/ del español, aunque cabe destacar un aumento de casos en los que las frecuencias del primer formante son más altas, próximas a las de la /a/ del español.

En el final de /e/, las articulaciones del grupo de control son similares a las del centro, aunque aquí la dispersión es ligeramente superior, tal y como indica la desviación estándar que aparece en la tabla 100. También las realizaciones de las informantes sinohablantes mantienen las mismas características que las del centro de la vocal, aunque en este punto se reduce el número de casos en los que el primer formante presenta valores próximos a los de la /a/ del español.

### **3.3 Vocal /a/**

Los datos relativos a la realización de la vocal /a/ por ambos grupos aparecen en la tabla 101. En esta se observa que la frecuencia media en la que aparecen los tres puntos del primer formante de /a/ es ligeramente superior cuando esta es producida por el grupo

de control (786, 853 y 856 Hz) que cuando la realizan las estudiantes chinas (761, 824 y 833 Hz). Sin embargo, la desviación estándar es mayor cuando esta vocal es realizada por el segundo grupo (84, 91 y 83 respectivamente) que por el primero (77, 69 y 68), lo que muestra que las realizaciones de las informantes sinohablantes son menos homogéneas que las del grupo de control y que no hay una gran diferencia en la dispersión de ambos grupos entre los tres puntos analizados.

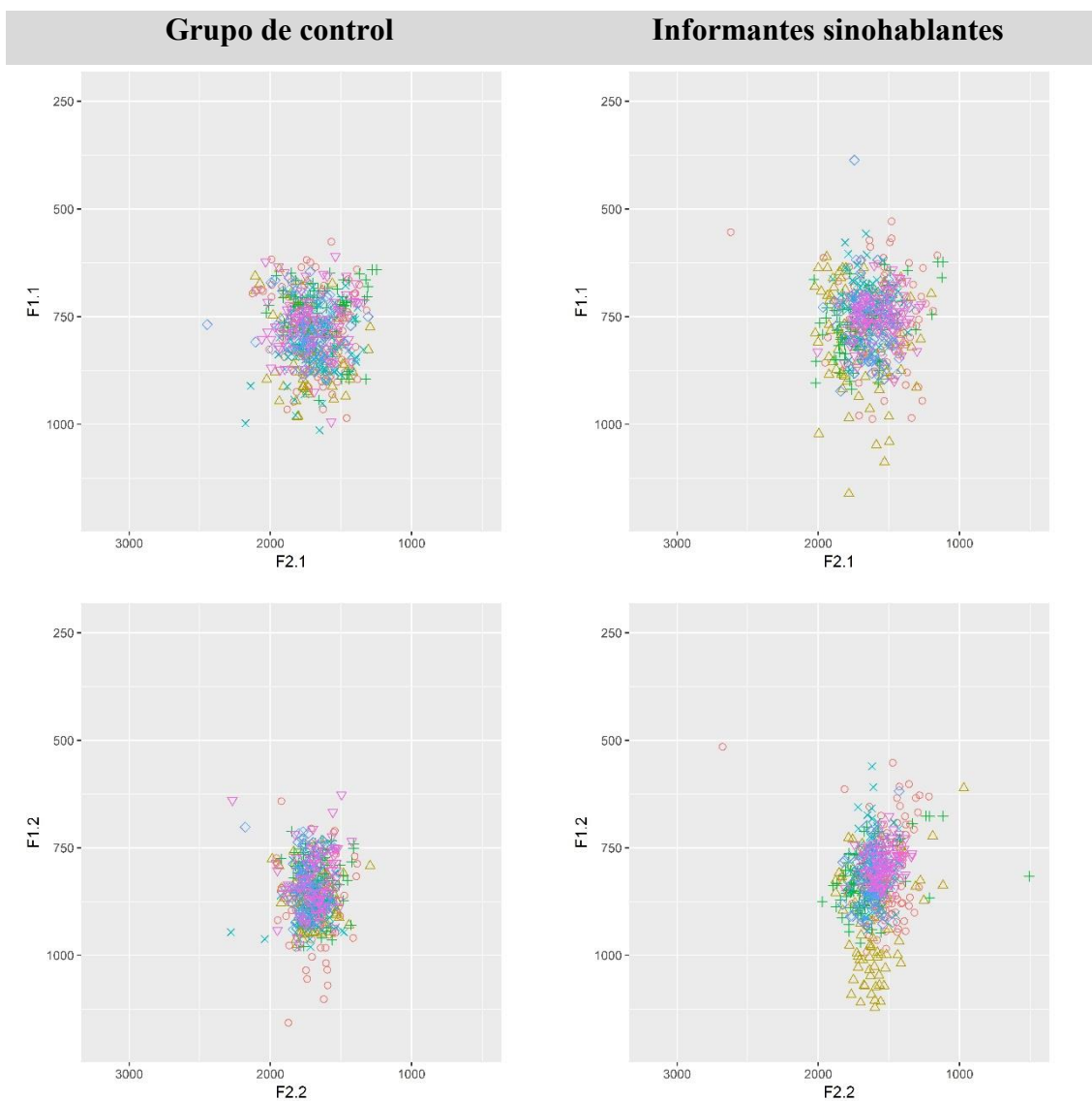
En cuanto a las frecuencias medias en las que aparecen los tres puntos del segundo formante, estas también son superiores cuando la vocal es realizada por las hablantes que conforman el grupo de control (1692, 1695 y 1687 Hz) que cuando las realizan las estudiantes chinas (1631, 1589 y 1582 Hz). La desviación estándar es la misma en el punto inicial para los dos grupos, aunque es superior en el centro y el final del formante cuando /a/ es articulada por las informantes sinohablantes (144 y 134) que por las hablantes cuya lengua materna es el español (114 y 119). Cabe señalar que dado que la desviación es mayor en los puntos iniciales, la dispersión es mayor en este punto, y disminuye en el centro y el final del formante, lo que indica que en estos puntos la posición de la lengua es más homogénea.

			Control	Sinohablantes	Significación
<b>Punto inicial</b>	<b>F1</b>	Media	786.5	761.7	< 0.0001
		DS	77.3	84.3	
	<b>F2</b>	Media	1692.6	1631.7	< 0.0001
		DS	172.4	173.8	
<b>Punto medio</b>	<b>F1</b>	Media	853.5	824.1	< 0.0001
		DS	69.4	89.4	
	<b>F2</b>	Media	1695.3	1589.8	< 0.0001
		DS	114.3	144.1	
<b>Punto final</b>	<b>F1</b>	Media	856.4	833.5	< 0.0001
		DS	68.5	81.5	
	<b>F2</b>	Media	1687.1	1582.6	< 0.0001
		DS	119.6	134.1	

*Tabla 101.* Comparación de frecuencias medias y desviación estándar de la vocal /a/ producida por el grupo de control y las informantes sinohablantes.

Los resultados obtenidos en la prueba de Mann Whitney muestran que la diferencia entre las realizaciones del grupo de control y las de las informantes sinohablantes son significativas, dado que el nivel de significación en todos los casos es menor que 0.05. Esto indica que la lengua materna de las informantes interfiere en su realización de la vocal /a/, tanto en la abertura oral como en la posición de la lengua.

A continuación, se muestran y se comparan las áreas de dispersión de la vocal /a/ producida por el grupo de control y por las informantes sinohablantes para mostrar los distintos tipos de realizaciones que producen estas hablantes.



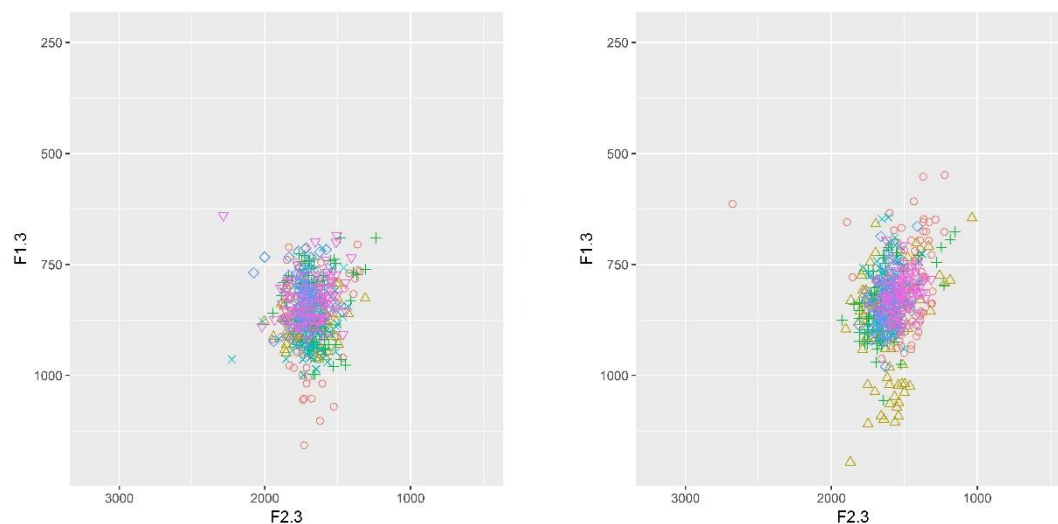


Figura 130. Comparación de las áreas de dispersión de la vocal /a/ en su inicio (F1.1 y F2.1), centro (F1.2 y F2.2) y final (F1.3 y F2.3).

Igual que sucede con las vocales /i/ y /e/, las áreas de dispersión de /a/ corroboran que las realizaciones de las hablantes que conforman el grupo de control son más homogéneas que las de las hablantes chinas, ya que la dispersión es mayor en el caso de las segundas. A continuación, se realiza un análisis de cada uno de los puntos de /a/.

De acuerdo con los datos mostrados en la tabla 101, en el inicio de la vocal, la mayoría de las realizaciones del F1 del grupo de control aparecen entre 700 y 850 Hz, mientras que las de las informantes sinohablantes se encuentran entre 650 y 850 Hz. En cuanto al F2, en el caso del primer grupo, las realizaciones se concentran entre 1500 y 1900 Hz, mientras que las del segundo entre 1400 y 1800 Hz. En el grupo de control, hay realizaciones de /a/ en las que la abertura oral es menor (F1 más bajo), y también hay casos en los que la posición de la lengua es más central (F2 más bajo). En cuanto a las informantes sinohablantes, se observa que muchas de las realizaciones se asemejan a las de /a/ producida por el grupo de control. Sin embargo, hay casos en los que la articulación de esta vocal es diferente. Ejemplo de ello es que, en ocasiones, la abertura oral es considerablemente mayor ya que el F1 alcanza los 1000 Hz, e incluso los supera en algunos casos; además, hay más casos en los que el F1 es más bajo, próximo a 500 Hz, por lo que la abertura oral es menor y semejante a la de la /e/ del español.

En el centro de la vocal, la mayoría de realizaciones del F1 del grupo de control se centran entre 750 y 900 Hz, valores ligeramente superiores a los del inicio, y estas frecuencias también aumentan en la articulación de /a/ por las informantes sinohablantes, las cuales se encuentran entre 700 y 900 Hz. En cuanto al F2, en el caso de las primeras, la mayoría de realizaciones se concentran entre 1500 y 1800 Hz, mientras que las de las



segundas aparecen en un rango de frecuencias de 1400-1700 Hz. En este punto, las realizaciones de las informantes del grupo de control son similares a las del inicio de la vocal y, aunque en general son ligeramente más abiertas, tanto el área de dispersión como la desviación estándar indican que las realizaciones son más homogéneas. En cuanto a las estudiantes chinas, como se ha indicado, las realizaciones son, en general, más abiertas que en el inicio de la vocal, y hay un aumento del número de casos en los que el primer formante aparece en un rango de frecuencias superior a 1000 Hz. Sin embargo, el número de casos en los que el F1 aparece en un rango de frecuencias más bajo es mayor que en el área correspondiente al grupo de control. Además, se puede apreciar que la dispersión del segundo formante es considerablemente menor que en el inicio de /a/.

En el final de la vocal, gran parte de las realizaciones del F1 por el grupo de control aparecen en el mismo rango de frecuencias que en el punto anterior, mientras que las de las informantes sinohablantes aparecen entre 750 y 900 Hz. En cuanto al F2, el rango de frecuencias correspondientes a las informantes hispanohablantes se mantiene igual que en el centro, y en el de las estudiantes chinas el rango de frecuencias es más limitado, entre 1400 y 1600 Hz. En el caso de las hablantes nativas de español, las realizaciones del final de la vocal se asemejan a las del centro de la misma. En cuanto a la realización de las informantes sinohablantes, a los fenómenos señalados en el punto anterior hay que añadir que, en este punto, las realizaciones del primer formante son más homogéneas, como muestra la desviación estándar que aparece en la tabla 101.

### **3.4 Vocal /o/**

Los datos relativos a la realización de la vocal /o/ por ambos grupos aparecen en la tabla 102. La comparación del primer formante producido por ambos grupos muestra que la frecuencia media es ligeramente superior cuando el grupo de control articula /o/, que cuando lo hacen las informantes sinohablantes. En el caso de las primeras, las frecuencias medias de los tres puntos analizados se sitúan en 551, 542 y 540 Hz respectivamente, y en el de las segundas, en 526, 518 y 522 Hz. En cuanto a la desviación estándar, es muy similar entre ambos grupos, ya que para el grupo de control es de 65, 65 y 62 en el inicio, medio y final del formante respectivamente, y para las hablantes chinas de 69 en los tres puntos de F1. Esto muestra que la homogeneidad de las realizaciones de ambos grupos es similar, que la variabilidad de un punto a otro del formante es

prácticamente la misma y que las informantes sinohablantes presentan el mismo comportamiento que las hablantes nativas de español en lo que a la abertura oral se refiere.

En cuanto al segundo formante de /o/, cuando esta vocal es articulada por el grupo de control las frecuencias medias en las que aparece (1310, 1262 y 1241 Hz) son más altas que las correspondientes a las informantes sinohablantes (1229, 1160, 1135 Hz). En cuanto a la desviación estándar, esta es de 239, 169 y 164 en el caso de las primeras, y de 177, 125 y 139 en el de las segundas, lo que muestra que el F2 realizado por las estudiantes chinas es más homogéneo que el realizado por las hablantes cuya lengua materna es el español. Este aspecto contrasta con las realizaciones de F1 y F2 de /i/, /e/ y /a/, y de F1 de /o/, en las que las hablantes del grupo de control presentan mayor homogeneidad que las informantes sinohablantes. Cabe señalar que, en ambos grupos, la dispersión es mayor en el inicio de la vocal que en el centro y el final de esta.

			Control	Sinohablantes	Significación
<b>Punto inicial</b>	<b>F1</b>	Media	551.4	526.9	< 0.001
		DS	65.0	69.7	
	<b>F2</b>	Media	1310.5	1229.6	< 0.001
		DS	239.0	177.2	
<b>Punto medio</b>	<b>F1</b>	Media	542.2	518.8	< 0.001
		DS	65.5	69.4	
	<b>F2</b>	Media	1262.9	1160.5	< 0.001
		DS	169.6	125.0	
<b>Punto final</b>	<b>F1</b>	Media	540.9	522.2	< 0.001
		DS	62.4	69.2	
	<b>F2</b>	Media	1241.0	1135.0	< 0.001
		DS	164.2	139.0	

*Tabla 102.* Comparación de frecuencias medias y desviación estándar de la vocal /o/ producida por el grupo de control y las informantes sinohablantes.

Como se ha visto, la realización del primer formante de /o/ es muy similar en los dos grupos, y en el caso del segundo formante se ha comprobado que las realizaciones de las informantes sinohablantes son más homogéneas. El valor obtenido en la prueba de Mann Whitney es, en todos los casos, menor que 0.05, lo que indica que las diferencias entre ambos grupos son significativas y que la lengua materna de las informantes interfiere en la articulación de /o/.

A continuación, se presentan y comparan las áreas de dispersión de la vocal /o/ producida por los dos grupos de estudio con el fin de mostrar las diferencias y similitudes entre las hablantes del grupo de control y las informantes sinohablantes en la realización de esta vocal.

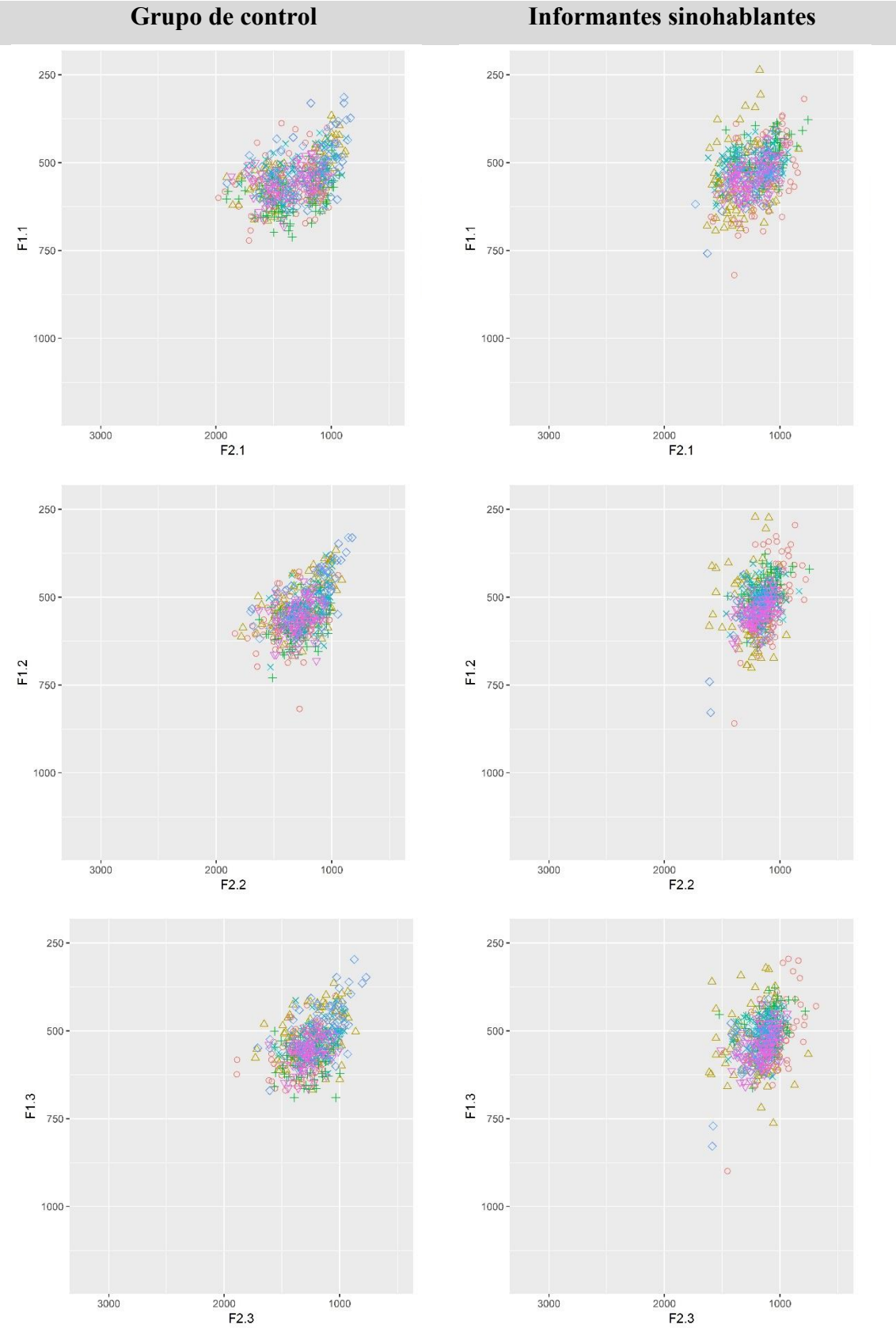


Figura 131. Comparación de las áreas de dispersión de la vocal /o/ en su inicio (F1.1 y F2.1), centro (F1.2 y F2.2) y final (F1.3 y F2.3).

Anteriormente se ha mostrado que las realizaciones de /i/, /e/ y /a/ por parte de las hablantes del grupo de control son más homogéneas que las de las informantes sinohablantes. En cambio, los datos mostrados en la tabla 102 indican que, en el caso de la vocal /o/, el primer formante articulado presenta características similares en ambos grupos en lo que a la frecuencia media y la desviación estándar se refiere, mientras que en el caso del segundo formante las realizaciones de las informantes sinohablantes son más homogéneas que las de las hablantes nativas de español. Una comparación general de todas las áreas permite observar que, efectivamente, las realizaciones del segundo formante presentan más variación cuando la vocal es articulada por las informantes del grupo de control.

De acuerdo con los datos mostrados en la tabla 102, en el inicio de la vocal la mayoría de las realizaciones del F1 de las informantes hispanohablantes aparecen entre 500 y 620 Hz, mientras que las de las estudiantes chinas aparecen entre 450 y 600 Hz. En cuanto al F2, en el caso del primer grupo las realizaciones se concentran entre 1100 y 1600 Hz, mientras que las del segundo entre 1000 y 1400 Hz. Como se puede observar en el área correspondiente al grupo de control, hay casos en los que el F1 y el F2 aparecen en un rango de frecuencias más bajas, por lo que la realización es más cerrada y más posterior. En cuanto a las estudiantes chinas, hay casos aislados en los que el F1 aparece en un rango de frecuencias más bajas, cercanas a 250 Hz, por lo que la abertura oral es menor y se asemeja a la de la /u/ del español; en cambio, también hay casos aislados en los que la abertura oral es mayor, y el F1 supera los 750 Hz, valor propio del primer formante de la /a/ del español. Anteriormente se ha indicado que la desviación estándar del segundo formante de /o/ producida por el grupo de control es de 239, mientras que la de las informantes sinohablantes es de 177. Esta diferencia es notable en las áreas correspondientes al inicio de la vocal, ya que, como se puede observar, hay más realizaciones de F2 que aparecen en un rango de frecuencias entre 1500 y 2000 Hz en el caso de las hablantes nativas de español, que en el de las estudiantes chinas.

En el centro de la vocal, la mayoría de las realizaciones del primer formante se mantienen en el mismo rango de frecuencias en los dos grupos. Sin embargo, estas cambian en el segundo formante. En el caso del grupo de control, estas se sitúan entre 1100 y 1400 Hz, mientras que las de las informantes sinohablantes aparecen entre 1000 y 1300 Hz, lo que muestra que hay menor dispersión en el centro de la vocal que en el inicio. En cuanto a las realizaciones de las informantes del grupo de control, además de aquellos casos en los que la abertura oral es menor dado que F1 aparece en un rango de

frecuencias inferior, se observa que el número de realizaciones más abiertas es menor que en el inicio de la vocal. En cuanto a las informantes sinohablantes, a las realizaciones descritas en el punto anterior cabe añadir, por un lado, que hay más casos en los que la abertura oral es más cerrada (F1 más bajo), y, por otro, que hay menos realizaciones en las que el segundo formante aparece en un rango de frecuencias superior a 1500 Hz.

En el final de la vocal la mayoría de realizaciones del F1 del grupo de control aparecen entre 500 y 600 Hz, mientras que las de las estudiantes chinas se mantienen igual que en el centro de la vocal. En cuanto al F2, en el caso del grupo de control también se mantiene igual que en el centro, pero no sucede lo mismo con las informantes sinohablantes, cuyas realizaciones aparecen entre 900 y 1300 Hz. En este punto aparecen los mismos fenómenos que en el inicio y el centro de la vocal cuando es articulada por las hablantes que conforman el grupo de control. En lo que a las informantes sinohablantes se refiere, la mayor diferencia respecto al centro del sonido vocálico es que la dispersión del segundo formante aumenta respecto al centro del mismo.

### **3.5 Vocal /u/**

Los datos relativos a la realización de la vocal /u/ por ambos grupos aparecen en la tabla 103. En esta se observa que la frecuencia media en la que se sitúa el primer formante realizado por las hablantes del grupo de control (480, 451 y 458 Hz) es ligeramente superior que la correspondiente a las informantes sinohablantes (456, 444 y 449 Hz). En cuanto a la desviación estándar, en el caso del primer grupo es de 64, 52 y 51 respectivamente, mientras que en el caso del segundo es de 55, 53 y 53, lo que indica que las realizaciones de las informantes sinohablantes presentan, ligeramente, una mayor homogeneidad que las del grupo de control.

En cuanto al segundo formante, este es muy similar cuando las hablantes del grupo de control articulan la /u/ (1100, 1083 y 1116 Hz) y cuando lo realizan las informantes sinohablantes (1116, 1097 y 1104 Hz).

El grupo de control presenta una desviación estándar de 182, 109 y 129 en los tres puntos del formante, mientras que la de las informantes sinohablantes es de 187, 151 y 160. Como se puede observar, la dispersión en ambos casos es mayor en el inicio del formante, disminuye en el centro, y aumenta en el final de la vocal, aunque en el caso de las estudiantes chinas la desviación es mayor, lo que indica que en estos puntos la realización es menos homogénea en lo que a la posición de la lengua se refiere.

			<b>Control</b>	<b>Sinohablantes</b>	<b>Significación</b>
<b>Punto inicial</b>	<b>F1</b>	Media	480.3	456.9	< 0.0001
		DS	64.4	55.6	
	<b>F2</b>	Media	1100.7	1116.8	0.8808
		DS	182.2	187.3	
<b>Punto medio</b>	<b>F1</b>	Media	451.5	444.5	0.0105
		DS	52.1	53.6	
	<b>F2</b>	Media	1083.8	1097.7	0.4551
		DS	109.7	151.2	
<b>Punto final</b>	<b>F1</b>	Media	458.0	449.4	0.0037
		DS	51.7	53.2	
	<b>F2</b>	Media	1116.8	1104.7	0.0002
		DS	129.9	160.4	

*Tabla 103.* Comparación de frecuencias medias y desviación estándar de la vocal /u/ producida por el grupo de control y las informantes sinohablantes.

De acuerdo con los resultados obtenidos al realizar la prueba de Mann Whitney, las diferencias entre el primer formante de la vocal /u/ articulada por el grupo de control y por las informantes sinohablantes son significativas en los tres puntos analizados. En cambio, en el segundo formante estas diferencias son significativas únicamente en el punto final de la vocal, y no en los anteriores, ya que en estos casos el nivel de significación es de 0.8808 en el punto inicial, y de 0.4551 en el punto medio. Por este motivo, se considera que la lengua materna de las informantes influye en la realización de /u/ en lo que a la abertura oral se refiere, pero no en la posición de la lengua.

A continuación, se presentan y analizan las áreas de dispersión de la vocal /u/ producida por los dos grupos de estudio con el fin de mostrar las diferencias y similitudes en la realizaciones de las hablantes.

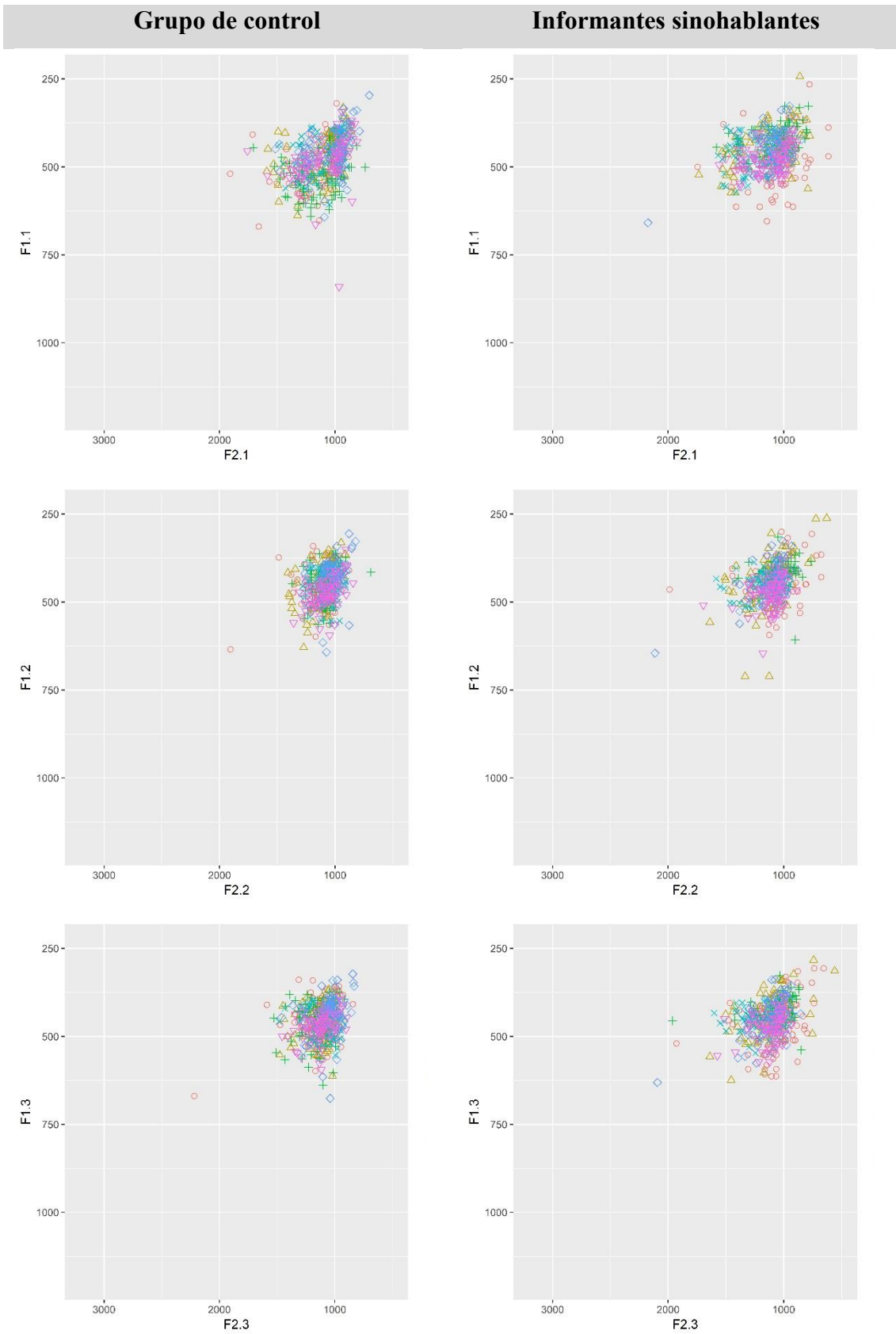


Figura 132. Comparación de las áreas de dispersión de la vocal /u/ en su inicio (F1.1 y F2.1), centro (F1.2 y F2.2) y final (F1.3 y F2.3).

Anteriormente se ha indicado que, en lo que concierne al primer formante de /u/, las realizaciones son más homogéneas en el inicio del mismo cuando la vocal es articulada por las informantes sinohablantes que por el grupo de control, y en el centro y el final la desviación es similar. En cambio, en el segundo formante la homogeneidad es mayor en todos los casos cuando /u/ es articulada por las hablantes del grupo de control. Estos aspectos se pueden observar en las áreas de dispersión que aparecen en la figura 132.

Según los datos de la tabla 103, en el inicio de la vocal la mayoría de las realizaciones del F1 del grupo de control aparecen, aproximadamente, entre 400 y 550 Hz, el mismo rango de frecuencias en el que se sitúan gran parte de las realizaciones de las estudiantes chinas. En cuanto al F2, el rango de frecuencias en el que aparecen la mayoría de las realizaciones del grupo de control es 900 – 1300 Hz, el mismo que en el que aparecen las de las informantes sinohablantes. Igual que sucede en las vocales anteriormente analizadas, en el grupo de control hay casos en los que las realizaciones pueden ser más abiertas o más cerradas, o más anteriores o más posteriores. que el F1 es más alto, por lo que la vocal es ligeramente más abierta, o el F2 más elevado, por lo que la posición de la lengua es menos posterior. Las realizaciones de las informantes sinohablantes se asemejan a las de las hablantes del grupo de control, aunque se puede observar que hay más casos en los que la frecuencia del F1 es más baja, por lo que la realización es más cerrada.

En el centro de la vocal, la mayoría de realizaciones del F1 del grupo de control aparecen entre 400 y 500 Hz, y las de las informantes sinohablantes entre 350 y 500 Hz. En cuanto al F2, en el caso de las primeras, la mayoría de realizaciones se concentran entre 950 y 1200 Hz, mientras que en el de las segundas aparecen en un rango de frecuencias de 900 – 1200 Hz. En el caso de las hablantes nativas de español, la posición de la lengua es más posterior en este punto de la vocal, ya que las frecuencias en las que aparece el F2 son más bajas, y únicamente superan los 1500 Hz en dos realizaciones aislados. En el caso de las informantes sinohablantes, hay articulaciones más abiertas, en las que la abertura oral es similar a la de la /o/ del español, y casos en los que la /u/ es más cerrada y más posterior que la mayoría.

En el final de la vocal, gran parte de las realizaciones del F1 por los dos grupos aparecen en el mismo rango de frecuencias que en el punto anterior, mientras que en el F2 las frecuencias correspondientes a las informantes hispanohablantes se mantienen igual que en el centro, pero no así las de las estudiantes chinas, las cuales aparecen entre 900 y 1300 Hz. En ambos grupos las realizaciones son similares a las del anterior punto,



aunque en el caso de las informantes sinohablantes las realizaciones abiertas, similares a /o/, no aparecen, lo que indica que en son más cerradas.

### **3.6 Recapitulación**

**En este apartado se realiza una breve recapitulación de los datos obtenidos en la comparación de las realizaciones de las vocales por parte de los dos grupos de estudio de la presente tesis doctoral.**

**Tanto el primer formante como el segundo de la vocal /i/ aparecen en un rango de frecuencias medias similar en ambos grupos, aunque las realizaciones de las hablantes del grupo de control son más homogéneas que las de las estudiantes chinas en lo que a la abertura oral y la posición de la lengua se refiere. No obstante, el análisis estadístico de los datos indica que las diferencias entre ambos grupos son significativas. Cabe señalar que, en el caso de las informantes sinohablantes, se han detectado casos en los que las características del F1 se asemejan a las de la /e/ del español, o las del F2 a las de la /u/ de esta lengua.**

**En el caso de la vocal /e/, pese a que ambos grupos tienen una frecuencia media similar en el inicio del primer formante, en el caso de las estudiantes chinas esta es más alta en el centro y en el final del mismo. En cambio, las frecuencias medias en las que aparecen los tres puntos del segundo formante son más bajas cuando /e/ es articulada por las informantes sinohablantes. Igual que sucede con la /i/, las realizaciones de los dos formantes son más homogéneas en la articulación de /e/ de las hablantes del grupo de control. El análisis estadístico de los datos muestra que las diferencias entre ambos grupos son, en general, significativas. Igual que en la vocal anterior, también en la realización de /e/ hay casos que presentan características propias de otras vocales del español, como una abertura oral más cerrada, similar a la de la /i/, realizaciones más abiertas en las que el F1 aparece en un rango de frecuencias característico de /a/, o realizaciones en las que el F2 tiene unas frecuencias más bajas, propias de la /u/ del español.**

**Tanto el primer formante como el segundo de /a/ tienen unas frecuencias medias ligeramente superiores cuando esta vocal es producida por las hablantes del grupo de control, y sus realizaciones son, además, más homogéneas que las de las informantes sinohablantes. El análisis estadístico de estos datos ha permitido**

mostrar que las diferencias entre ambos grupos son significativas. Igual que en articulación de /i/ y /e/ por las estudiantes chinas, se han detectado casos en los que el primer formante aparece en un rango de frecuencias más altas que la mayoría de realizaciones, y casos en los que las frecuencias son más bajas, próximas a las de la /e/ del español en lugar de las de la /a/.

En cuanto a la vocal /o/, las frecuencias medias en las que aparecen el primer y el segundo formante son más altas cuando esta vocal es producida por las hablantes que conforman el grupo de control. En cuanto a la homogeneidad de las realizaciones, el F1 presenta una variabilidad similar en ambos grupos, es decir, la abertura oral es similar, mientras que el F2 es más homogéneo en la producción oral de las estudiantes chinas, es decir, hay menor variabilidad en lo que a la posición de la lengua se refiere, lo que contrasta con los datos obtenidos en las vocales /i/, /e/ y /a/. El resultado del análisis estadístico de los datos indica que las diferencias entre ambos grupos son significativas. Igual que en las vocales anteriores, aquí también se han detectado casos en los que la articulación de /o/ por parte de las estudiantes chinas tiene características propias de otras vocales del español, como las realizaciones más cerradas en las que el primer formante aparece en un rango de frecuencias propio de la /u/, o el caso contrario, en las que la realización es abierta y el primer formante aparece en unas frecuencias propias de la /a/.

En cuanto a la realización de /u/ por parte de los dos grupos, el primer formante aparece en un rango de frecuencias más altas cuando es producido por las hablantes cuya lengua materna es el español, mientras que el segundo formante aparece en un rango de frecuencias similares en ambos grupos. Las realizaciones del F1 son más homogéneas en la producción de las estudiantes chinas, lo contrario que las del F2, las cuales son más homogéneas cuando /u/ es articulada por el grupo de control. El estudio estadístico de los datos ha revelado que las diferencias entre los dos grupos son significativas en el caso del primer formante, y en el segundo únicamente en el final de la vocal. Igual que en los casos anteriores, se han detectado realizaciones más abiertas, similares a las de la /o/ del español, o articulaciones más cerradas en las que las frecuencias medias son más bajas que las de la propia /u/.

En el análisis de las producciones de las hablantes que conforman el grupo de control se ha detectado que, generalmente, la desviación es mayor en el inicio y el final de los formantes. En el caso del inicio, se ha indicado que hay mayor

**variabilidad porque la vocal es precedida por un sonido consonántico, por lo que los órganos fonatorios cambian de posición desde dicha consonante hasta el centro de la vocal. En cuanto al final, se ha indicado que hay dos razones que explican este aumento en la dispersión: la primera, cuando la vocal va seguida de un sonido consonántico, los órganos fonatorios se preparan para la realización de este; la segunda, si la vocal se encuentra en posición final de palabra aislada, los órganos fonatorios se deslizan hacia la posición de reposo.**

**En cambio, este fenómeno no se ha detectado en la realización de las vocales del español por parte de las informantes sinohablantes, ya que en algunos casos la dispersión es similar en los tres puntos analizados, en otros aumenta desde el principio hasta el final, y en otros disminuye la desviación desde el inicio del formante hasta el final.**

**Además, cabe destacar que, mientras que las realizaciones de /i/, /e/ y /a/ son más homogéneas cuando son producidas por las hablantes del grupo de control, en el caso de /o/ se ha mostrado que la posición de la lengua (F2) presenta mayor homogeneidad cuando es articulada por las estudiantes chinas, y en el de /u/, la abertura oral (F1) es más homogénea cuando es producida por estas aprendices que por las hablantes cuya lengua materna es el español**

**CAPÍTULO V**  
**DISCUSIÓN DE LOS RESULTADOS**



## **V. DISCUSION DE LOS RESULTADOS**

En el presente capítulo se contrastan los resultados de esta investigación con los obtenidos en los trabajos previos recogidos en el capítulo referentes tanto a la adquisición de los sistemas fonológicos de lenguas extranjeras en general como a la adquisición del español por sinohablantes. El orden de exposición es el mismo que se ha utilizado en el capítulo IV, dedicado a la presentación de los resultados.

### **1. INFLUENCIA DE LAS VARIABLES INDEPENDIENTES SOBRE LA REALIZACIÓN DE LAS VOCALES Y PROCESOS DE RESILABACIÓN**

#### **1.1 Acentuación**

La bibliografía analizada en el capítulo II muestra que el número de trabajos que analizan la acentuación de los sonidos vocálicos del español por parte de informantes sinohablantes es muy reducido. Cortés Moreno (2005 y 2009b) indica que las estudiantes chinas no presentan dificultades remarcables en la realización de las distintas vocales del español, tanto tónicas como átonas, en palabras bisílabas. Blanco y Nogueroles (2013), por su parte, indican que existe un desplazamiento de patrones acentuales en palabras trisílabas, según el cual los estudiantes cuya lengua materna es el chino realizan las palabras llanas como esdrújulas.

En esta investigación se ha analizado la acentuación de palabras bisílabas por parte de las informantes sinohablantes. Los resultados muestran que, a pesar de que hay casos en los que la acentuación no ha sido la correcta, generalmente las informantes sinohablantes no presentan dificultades en la acentuación correcta de las vocales del español en palabras de dos sílabas, lo que confirma lo indicado por Cortés Moreno en sus trabajos. De cara a investigaciones futuras convendría analizar la acentuación de las palabras trisílabas con el fin de contrastar la afirmación realizada por Blanco y Nogueroles (2013).

#### **1.2 Entorno consonántico**

Como se ha indicado, aunque el estudio de las realizaciones del entorno consonántico no ha constituido un objetivo de esta tesis, en la medida en que dicho

entorno influye en la realización de los sonidos vocálicos, se han analizado también las consonantes anterior y siguiente a las vocales. Este planteamiento no figura en los trabajos publicados hasta la fecha dado que la mayoría de los autores consideran que la realización de las vocales por alumnos sinohablantes es un problema menor y que, por lo tanto, son los sonidos consonánticos los que presentan mayores dificultades.

Así, tal y como se ha podido observar en el capítulo II, la mayoría de trabajos que estudian específicamente los problemas consonánticos coinciden en señalar las mismas dificultades para estos hablantes: confusión entre las vibrantes [r], [r] y la lateral [l]; confusión entre [s] y [θ]; y confusión entre las oclusivas sordas [p], [t], [k] y las oclusivas sonoras [b], [d], [g]. Sin embargo, los trabajos de Zhi (2011), López (2012), Blanco y Nogueroles (2013), Poch e Igarreta (2014), Igarreta Fernández (2015b) y Pose (2017) ofrecen un análisis más exhaustivo de estos fenómenos, y realizan una descripción de los mismos. A continuación, se comparan los resultados de estos trabajos con los obtenidos en la presente tesis doctoral.

### **1.2.1 Oclusivas sordas y sonoras**

En relación con la articulación de las oclusivas sordas y sonoras, los autores coinciden en que los estudiantes sinohablantes articulan las oclusivas sonoras como sordas. Sin embargo, Zhi (2011: 55) indica que esta articulación sucede en menor medida, y que, en los informantes analizados en su estudio, la tendencia más generalizada es a realizar las oclusivas sordas como sonoras.

Los datos obtenidos en la presente tesis doctoral muestran que las oclusivas sordas se encuentran entre los sonidos que resultan menos problemáticos para las informantes sinohablantes (más del 95% de realizaciones correctas). En aquellos casos en los que no las han articulado como tal, las han realizado, en la mayoría de los casos, como oclusiva sonora, lo que suscribiría lo indicado por Zhi (2011), pero cabe señalar que también se han detectado casos en los que dicho sonido se ha elidido y no se ha articulado otro, y también hay un número reducido de ocasiones en las que las estudiantes han realizado la oclusiva sorda como fricativa, vibrante simple o nasal.

En cuanto a las consonantes oclusivas sonoras, los datos muestran que las informantes sinohablantes tienen problemas para articularlas como tal, ya que la realización ha sido correcta en menos del 20% de los casos. Se ha detectado una clara tendencia a articularlas como oclusiva sorda, y, además, igual que sucede con las

oclusivas sordas, hay casos aislados en los que la oclusiva sonora ha sido realizada como fricativa, vibrante simple o nasal.

### **1.2.2 Aproximantes**

En relación con las consonantes aproximantes, Yufei (2007: 92-94), señala que, al no existir éstas en chino, los estudiantes sinohablantes tienen dificultades para articularlas correctamente, y López (2012: 49-50) indica que la tendencia es a realizarlas como oclusivas sordas.

Los datos obtenidos en la presente tesis doctoral muestran que, efectivamente, las consonantes aproximantes suponen un gran escollo para las informantes chinas, ya que el porcentaje de realizaciones correctas no supera el 39%. La tendencia general, como se ha podido observar, es a articularlas como oclusivas sordas, realización que corrobora lo indicado por López (2012). Además, se han detectado casos aislados en los que las estudiantes chinas han articulado las aproximantes como nasal o vocal, o la han elidido.

### **1.2.3 Fricativas**

En cuanto a los sonidos fricativos, como se ha visto en el capítulo II, la bibliografía señala que las informantes sinohablantes presentan confusión entre [s] y [θ].

Los datos analizados en esta investigación muestran que las consonantes fricativas son menos problemáticas que otros sonidos, ya que las informantes sinohablantes las han articulado correctamente en el 97% de los casos. En aquellas ocasiones en las que no han producido este sonido correctamente, las estudiantes chinas han elidido la consonante, o la han realizado como oclusiva sorda, oclusiva sonora, lateral o como vocal.

### **1.2.4 Africadas**

Los trabajos mencionados en el capítulo II no señalan dificultades para articular las consonantes africadas del español. Los datos obtenidos en esta investigación muestran que la articulación de estas no es problemática para las informantes sinohablantes, ya que la realización de las mismas ha sido correcta en el 100% de los casos. Sin embargo, cabe señalar que el número de casos en los que este sonido está presente en el corpus de esta investigación es reducido (como ocurre en español), por lo que convendría hacer un análisis más exhaustivo estudiando un número más elevado de estas realizaciones.



### **1.2.5 Nasaes**

La bibliografía analizada no menciona dificultades por parte de los sinohablantes para articular correctamente las consonantes nasales del español. No obstante, en esta investigación se ha comprobado que, si bien estas consonantes no plantean grandes dificultades a estos hablantes, ya que las han realizado correctamente en el 94% de las ocasiones, se han detectado algunas realizaciones incorrectas de las mismas. Estas han consistido en elidir la consonante nasal, en realizarla como oclusiva sorda y, en ocasiones aisladas, en articularla como lateral.

### **1.2.6 Laterales**

En cuanto a la articulación de las consonantes laterales, Poch e Igarreta (2014: 153) e Igarreta Fernández (2015b) detectan realizaciones de [l] como [n], aspecto que señalan que puede deberse a la variedad de chino que hablan los informantes. Igarreta Fernández (2015b), además, señala que hay casos en los que los estudiantes realizan [l] como [r].

Los resultados obtenidos muestran también que, cuando el sonido lateral es anterior a la vocal, este plantea menos problemas a las informantes sinohablantes, ya que la han articulado correctamente en el 96% de los casos. Sin embargo, si dicho sonido es posterior a la vocal, las estudiantes chinas han tenido más problemas, y únicamente la han realizado correctamente en el 64% de los casos. En cuanto a las realizaciones incorrectas de la lateral, las informantes sinohablantes han mostrado cierta tendencia a realizarla como vibrante simple o a elidirla, aunque también se han detectado articulaciones de esta como fricativa, oclusiva sorda, oclusiva sonora, nasal y vocal.

### **1.2.7 Vibrantes**

En lo que concierne a la realización de la vibrante simple, la bibliografía generalmente señala que los estudiantes chinos articulan dicha consonante como [l]. Blanco y Nogueroles (2013: 217-218), además, indican que hay casos en los que se inserta la [d] epentética ante [r], y que también hay ocasiones en las que la vibrante simple es realizada como la [ɹ] del inglés, aspecto que también señalan Poch e Igarreta (2014: 154). Además, Igarreta Fernández (2015b) indica que hay casos en los que las informantes sinohablantes articulan la nasal [n] en lugar de la vibrante simple.

Los datos obtenidos muestran que las informantes sinohablantes tienen problemas para articular correctamente la vibrante simple tanto si se encuentra en posición anterior, como en posición posterior a la vocal, aunque esta es menos problemática si se encuentra después de la vocal. En los casos en los que la articulación no ha sido la correcta, las estudiantes chinas, generalmente, la han articulado como lateral, lo que corrobora lo indicado por los distintos trabajos mencionados. Sin embargo, cabe señalar que en muchas ocasiones la han elidido, y menos frecuentemente la han realizado como oclusiva sorda, fricativa, nasal, y, en algunos casos aislados, como oclusiva sonora o vibrante múltiple.

En cuanto a la vibrante múltiple, la bibliografía señala que, generalmente, esta es realizada como lateral. Los datos muestran que la realización de la vibrante múltiple es un escollo prácticamente insalvable, ya que la han realizado correctamente en un número muy reducido de ocasiones. Las estudiantes chinas han mostrado una clara tendencia a articular [r] como vibrante simple, y también hay realizaciones de esta como oclusiva sonora o lateral, además de casos aislados en los que ha sido articulada como oclusiva sorda, lateral o africada. Cabe señalar que en muchos de los casos en los que la vibrante múltiple se ha articulado como vibrante simple, se ha añadido antes de [r] la [d] epentética señalada por Blanco y Nogueroles (2013: 217-218).

La reflexión más importante que se deduce de los datos obtenidos conduce a pensar en la necesidad de replantear el enfoque tradicional hasta ahora seguido en los estudios sobre la adquisición del sistema fonológico de una lengua extranjera cuando ésta mantiene una considerable distancia lingüística con respecto a la lengua base de los alumnos. Dicha distancia, como se ha visto (y como se apunta ya en Poch – Igarreta, 2014), conduce a situaciones como las siguientes: si en el corpus aparece la palabra *duda* con el objetivo de estudiar las realizaciones vocálicas de los sinohablantes en contacto con oclusivas sonoras, la realidad de la pronunciación conducirá a que los alumnos pronuncien ['tu.ta], con oclusivas sordas, de forma que el corpus ya no funciona según las expectativas de pronunciación de los hispanohablantes nativos o de los extranjeros cuya lengua base mantiene una pequeña distancia lingüística con respecto a la lengua meta, como sería el caso de alumnos cuya lengua materna sea indoeuropea.

Como se expondrá a continuación, la estructura silábica constituye otro factor de variación en la realización de las palabras que integran el corpus de trabajo.

### 1.3 Estructura silábica

La bibliografía analizada en el capítulo II muestra que son pocos los investigadores que han analizado los problemas que presentan los estudiantes sinohablantes en la correcta articulación de los distintos tipos de estructura silábica del español. No obstante, todos coinciden en señalar que la principal dificultad para los alumnos sinohablantes es la pronunciación de los grupos consonánticos, ya que este tipo de estructura silábica no existe en chino. Algunas de las publicaciones se limitan a señalar dónde se encuentran las dificultades, mientras que otras indican cómo realizan las estructuras problemáticas los estudiantes cuya lengua materna es el chino.

Ortí Mateu (1990) indica que estos hablantes tienen, en general, dificultades con las estructuras silábicas en las que existe una consonante trabada, con aquellas que tienen una consonante final (exceptuando los casos en los que dicho sonido es [n]), o con otras estructuras menos frecuentes del español, como VCC, CVCC, CCVCC, CVVC, CCVV. Según el trabajo de Cortés Moreno (2009b: 7), en chino “no hay grupos consonánticos iniciales, lo que plantea dificultades a los estudiantes sinohablantes de ELE, aunque no solo en esta posición, si no en cualquier otra”. Además, el investigador indica que estos hablantes tienen tendencia a pronunciar sílaba por sílaba, lo que “impregna su discurso de un característico ritmo staccato” (Cortés Moreno, 2009a: 186). Blanco y Nogueroles (2013: 218), por su parte, indican que existe una tendencia entre los estudiantes sinohablantes de español a simplificar los grupos consonánticos, tanto en una misma sílaba, como en sílabas contiguas. Alcoholado (2013a: 34) indica que los hablantes chinos tienen dificultad para articular correctamente las combinaciones de consonantes en posición marginal y, también, las combinaciones vocálicas, además de mostrar cierta tendencia a pronunciar sílaba por sílaba. Miranda Márquez (2014), por su parte, apunta que estos estudiantes separan en dos o más partes una sílaba para adaptar el sonido al más cercano en su lengua materna. En cuanto a Poch e Igarreta (2014), las investigadoras indican que la tendencia general es la reducción de los grupos consonánticos a una sola consonante y una vocal. Pose (2017), además de señalar esta reducción, indica que las estudiantes sinohablantes articulan de manera débil o eliden las consonantes en posición implosiva. La suma de estos fenómenos se debe, según el autor, a las diferencias entre el chino y el español y a la transferencia de rasgos de la primera lengua a la segunda. El autor señala que estos procesos de transferencia únicamente afectan a una sílaba, y no traspasan a las sílabas contiguas, lo que da lugar a cortes entre sílabas y al ritmo staccato

mencionado por Cortés Moreno (2009a). Por estos motivos, estos hablantes no realizan la resilabación propia del español y hay cortes glóticos en ataque duro, además de otros fenómenos relacionados con el ritmo y la entonación, lo que da lugar a una pronunciación entrecortada que afecta a segmentos menores, como consonantes y vocales.

Como se ha indicado, en la presente tesis doctoral no se han incluido silabas que contengan grupos consonánticos, y los tres tipos de estructuras de palabra analizadas han sido: CV-CV, CV-CVC y CVC-CV. No obstante, a pesar de que el corpus no contiene las combinaciones de sonidos consideradas como más problemáticas, se ha podido comprobar que estas hablantes tienen dificultades para articular los tres tipos de estructuras más simples y frecuentes en español.

La estructura menos problemática es CV-CV (la más parecida a la propia del chino mandarín), y en aquellos casos en los que la articulación no ha sido la propia del español, la tendencia general ha sido a incluir otro sonido consonántico en la primera o en la segunda sílaba, dando lugar a realizaciones como CCV-CV, CV-CCV o CVC-CV. Estas modificaciones implican la introducción de un proceso de resilabación en las palabras que contiene el corpus.

La estructura más problemática es CV-CVC. En la mayoría de casos en los que la realización no ha sido la correcta, las informantes sinohablantes han producido palabras cuya estructura es CV-CV, CCV-CVC o CCV-CV. Es decir, en el primer caso, se ha elidido la consonante final, en el segundo se ha añadido una consonante en la primera sílaba de la palabra, y en la última, además de añadir un sonido consonántico en la primera sílaba, se ha eliminado el que se encuentra en posición final de palabra con lo que, de nuevo, se producen resilabaciones en la estructura del corpus.

La estructura CVC-CV también ha creado problemas a las informantes sinohablantes (aunque menos que CV-CVC), y en aquellos casos en los que la articulación no ha sido la adecuada, la tendencia general ha sido a elidir la segunda consonante de la primera sílaba (CV-CV), o a insertar una nueva consonante en dicha sílaba y articular la palabra como CCVC-CV.

Además, en los tres tipos de estructuras se han detectado casos en los que ha habido una reducción de sonidos considerable, y la palabra ha pasado de ser bisilábica a monosilábica.

Así pues, las informantes sinohablantes presentan dos tendencias generales en lo que a la realización de las estructuras de las sílabas CV y CVC del español se refiere: la primera de ellas es la elisión de la segunda consonante en las sílabas cuya estructura es CVC; la segunda, la integración de un sonido consonántico nuevo, especialmente en casos en los que la consonante anterior o posterior es especialmente problemática para las estudiantes sinohablantes.

Mientras que la bibliografía únicamente señala los grupos consonánticos y vocálicos como problemáticos en lo que a la estructura silábica se refiere, en esta tesis doctoral se ha comprobado que las informantes sinohablantes también tienen dificultades para realizar correctamente las estructuras CV y CVC. Además, se ha observado que, generalmente, dichos problemas surgen cuando el sonido anterior y posterior es problemático para las informantes sinohablantes. Estas dificultades conducen, como se ha indicado, a la aparición de importantes fenómenos de resilabación.

Los fenómenos hasta ahora resumidos proceden del estudio auditivo del corpus que ha puesto de manifiesto que, en el caso del aprendizaje del sistema fonológico de una lengua que mantiene una importante distancia lingüística con respecto a la lengua base de los alumnos, deben ser tenidos en cuenta estos fenómenos. El análisis contrastivo tradicional, tal como se ha practicado en los estudios revisados y comentados, que considera los problemas separados por categorías de sonidos, no parece suficiente para explicar las dificultades de aprendizaje. Los datos parecen apuntar a la necesidad de abordar estos problemas desde una perspectiva holística que considere los problemas en su globalidad de tal forma que los modelos de adquisición de los sistemas fonológicos puedan integrar estos comportamientos.

A continuación se comentan, desde este punto de vista, los resultados del análisis acústico de las vocales del español realizadas por sinohablantes. Como se resumirá, la duración de los sonidos vocálicos y la cualidad de sus timbres vocálicos se verá influida por los fenómenos que se acaban de comentar.

## **2. DURACIÓN DE LAS VOCALES**

Como se ha mostrado en el capítulo II, Pérez García (2018: 163) analiza la duración de las vocales del español por estudiantes sinohablantes y las compara con las realizaciones del grupo de control. El investigador concluye que las realizaciones de las estudiantes chinas son más largas que las de las hablantes nativas de español. Según sus datos, la [i] tiene una duración más breve, con una media de 119 ms., [e], [o] y [u] tienen la misma duración media (122 ms.) y [a] es la más larga de todas las vocales, con una media de 133 ms. Jiménez y Tang (2018) también analizan la duración de las vocales del español producidas por informantes sinohablantes. Los resultados muestran que /i/ tiene la duración más breve, con una media de 102 ms., /u/ y /e/ son más largas, con una duración media de 117 y 122 ms. cada una, y /a/ y /o/ son las vocales más largas y tienen la misma duración media, de 129 ms.

En este trabajo se ha comprobado que las vocales producidas por las informantes sinohablantes siguen el orden /u/, /i/, /o/, /a/, /e/, siendo /u/ la más breve, y /e/ la más larga, y la duración media va desde 122 hasta 169 ms. cuando se analizan independientemente de la acentuación, el entorno o la estructura silábica. En todos los casos las vocales producidas por las estudiantes chinas son más largas que las articuladas por las hablantes del grupo de control, aunque las realizaciones de estas últimas son más homogéneas, y se ha comprobado que las diferencias entre ambos grupos son estadísticamente significativas.

### **2.1. Duración de las vocales en función del acento**

El análisis de la duración de las vocales realizadas por las informantes sinohablantes según la tonicidad ha mostrado que las vocales tónicas son más largas que las átonas. Cuando estas son articuladas por las hablantes del grupo de control, la duración es más breve y más homogénea que la de las estudiantes chinas, diferencias que, como se ha mostrado, son significativas.

### **2.2. Duración de las vocales en función del entorno**

Los resultados del análisis de la duración de las vocales según el entorno han mostrado que, si bien no todas las vocales presentan la misma duración en función del sonido consonántico anterior o posterior, se ha podido detectar que, generalmente, las

vocales son más breves cuando son precedidas por una fricativa, una africada o una oclusiva sorda; la duración es más larga cuando el sonido anterior es una oclusiva sonora, una nasal, una lateral o una vibrante múltiple, y cuando la consonante que precede a la vocal es una vibrante simple o una aproximante, las vocales son más largas que en los casos anteriores. También se ha podido observar que, generalmente, las vocales son más largas cuando el sonido posterior a la vocal es una oclusiva sonora, una aproximante, una lateral, vibrante simple o una vibrante múltiple. La comparación de estos datos con los de la realización del grupo de control ha permitido mostrar que, generalmente, las vocales son más largas y menos homogéneas cuando son producidas por las informantes sinohablantes, y, de acuerdo con los resultados del análisis estadístico, las diferencias entre ambos grupos son, generalmente, significativas.

En cuanto a la realización de las vocales en final de palabra aislada, estas son más largas cuando son articuladas por las informantes sinohablantes, y las de las hablantes del grupo de control son más homogéneas. Estas diferencias son estadísticamente significativas.

Estos datos han permitido concluir que, dado que las diferencias entre ambos grupos son significativas, la lengua materna de las informantes, es decir, el entorno consonántico en este caso, influye en la duración de las vocales del español, lo que corrobora los resultados obtenidos por Pérez García (2018) y por Jiménez y Tang (2018).

### **2.3. Duración de las vocales en función de la estructura silábica**

En cuanto a la duración de las vocales según la estructura silábica en la que se encuentran, se ha comprobado que cuando la sílaba está compuesta por CV, las vocales son más largas que cuando la estructura es CVC. En relación con las estructuras más repetidas por las informantes sinohablantes que no son las propias de las palabras que conforman el corpus, como CCV o CCVC, se ha comprobado que las vocales son más largas que cuando se encuentran en las estructuras propias de las palabras del corpus. La comparación de estos resultados con la realización del grupo de control ha mostrado que en todos los casos los sonidos vocálicos son más largos y menos homogéneos cuando son articulados por las estudiantes sinohablantes, diferencias que se ha mostrado que son estadísticamente significativas.

Como se aprecia, pues, los factores de variación cualitativos señalados anteriormente, poseen una influencia importante sobre los datos acústicos de duración de las vocales, lo que confirma la interdependencia entre las características de todos los elementos de la cadena fónica.

### **3. TIMBRE DE LAS VOCALES**

Como se ha visto en el capítulo II, existe una falta de consenso en lo que a la descripción del sistema vocálico del chino se refiere. Los trabajos de Ortí Mateu (1990), Wang (2001) y Cortés Moreno (2009a, 2009b) coinciden en señalar que tanto el chino como el español presentan los cinco fonemas vocálicos /i, /e/, /a/, /o/, /u/, y añaden dos sonidos más al sistema vocálico chino, aunque con diferentes matices según el autor. Según Cortés Moreno (2009a: 4), del análisis contrastivo de ambas lenguas se desprende que los estudiantes sinohablantes no tienen dificultades para percibir y pronunciar los sonidos vocálicos del español. En cambio, los trabajos señalados a continuación indican lo contrario.

Planas Morales indica que estos estudiantes sí tienen dificultades, especialmente derivadas de las equivalencias entre la grafía y el sonido, por lo que, según la autora, los estudiantes chinos articulan la /e/ del español más abierta y más posterior (2008: 501), el timbre de la vocal [a] es inestable, sustituyen la [e] por la neutra y pronuncian /o/ como [u] (Planas Morales, 2013).

En lo que concierne a la pronunciación de las vocales del español por mujeres sinohablantes, el análisis acústico realizado por López (2012: 43) muestra que las vocales cerradas tienen mayor abertura, la /a/ también es más abierta, las vocales medias presentan, además de mayor abertura, adelantamiento, la /i/ presenta un retroceso y hay producciones de [e] cercanas a [ɛ] y de /o/ similar a [ɤ].

Del análisis de Poch e Igarreta (2014: 150) se concluye que hay una gran inestabilidad en el timbre de las vocales, y en los casos en los que la articulación no es la adecuada, las realizaciones de los alumnos son más abiertas, realizan el sonido [ə] en lugar de [a] o [e], articulan la vocal labializada [y] en lugar de [u], y en ocasiones producen un timbre distinto del esperado.

Pérez García (2018), por su parte, indica que las oscilaciones internas de los formantes son lo que diferencian las vocales producidas por las estudiantes chinas de las de las hablantes nativas de español. En cuanto al primer formante, el investigador indica



que la [a] es la que presenta mayor oscilación, [e] y [o] presentan una tendencia hacia la apertura inicial y posterior cierre, y [u] e [i] muestran una mayor estabilidad. En cuanto al segundo formante, las vocales [i] y [e] presentan un mayor grado de variabilidad, ya que en el inicio los valores ascienden y posteriormente descienden, [o] y [u] presentan mayor variabilidad que las vocales anteriores, y en este caso en el inicio hay un descenso y posteriormente un ascenso, y la [a] es la más estable de todas.

Los resultados del estudio de Jiménez y Tang (2018) muestran que el rango de frecuencias en el que aparecen las vocales articuladas por estudiantes sinohablantes es mayor que el de las propias del español, las vocales medias y /a/ presentan mayor abertura, y /a/, /e/ y /o/ presentan mayor variabilidad.

El análisis acústico sistemático y el tratamiento estadístico de los datos han permitido realizar una descripción exhaustiva de las características de las vocales del español producidas por las estudiantes chinas que se resumen a continuación:

- En primer lugar, se ha comprobado que, si bien las frecuencias medias de las vocales articuladas por las informantes sinohablantes son similares a las del español, igual que indica Pérez García (2018), la gran variabilidad en los valores de frecuencia de los formantes es lo que distingue las realizaciones de las estudiantes chinas de las de las hablantes que conforman el grupo de control. Se ha probado que, generalmente, las realizaciones de las informantes hispanohablantes nativas son más homogéneas que las de las aprendices de español. No obstante, cabe señalar, por un lado, que en el caso de la vocal /o/ ambos grupos presentan una homogeneidad similar en lo que a la abertura oral se refiere (F1), y la posición de la lengua (F2) es más homogénea en el caso de las informantes sinohablantes. Por otro lado, en el caso de /u/ las realizaciones del F1 son más homogéneas en la producción oral de las estudiantes chinas.
- En segundo lugar, se ha comprobado que, en general, las realizaciones de las vocales del español son más abiertas, lo que corrobora lo indicado en los trabajos de Planas Morales (2008, 2013), López (2012), Poch e Igarreta (2014) y Jiménez y Tang (2018). No obstante, la variabilidad ya señalada ha conducido a que algunas realizaciones sean más cerradas que las producidas por las informantes hispanohablantes.

- En tercer lugar, se ha mostrado que también hay adelantamiento de las vocales del español producidas por las informantes sinohablantes, tal y como indica López (2012). Sin embargo, de nuevo por cuestiones de variabilidad, también se han detectado casos en los que hay un retroceso en la posición de la lengua y las frecuencias en las que aparece el segundo formante son más bajas.
- En cuarto lugar, se han detectado realizaciones de las distintas vocales del español en las que el timbre es distinto del esperado y el rango de frecuencias en el que aparecen el primer formante, el segundo o ambos, es propio de otra vocal del español. Se han detectado casos en los que el primer formante de /i/ tiene valores propios de la /e/ del español, es decir, es más abierta, o de la de /u/. Hay realizaciones de /e/ en las que la vocal es más cerrada y el F1 aparece en un rango de frecuencias propio de /i/, o más abierta, y el rango de frecuencias en el que aparece es el de /a/, o casos en los que la posición de la lengua es posterior, y el segundo formante aparece en unas frecuencias propias de la /u/. También se han detectado realizaciones de /a/ como /e/, es decir, la vocal es más cerrada. En ocasiones, la abertura de /o/ es más cerrada y el primer formante aparece en un rango de frecuencias propio de /u/, o más abierta, y aparece en el rango de /a/. En cuanto a la vocal /u/, se han detectado realizaciones de la misma en las que la abertura oral es mayor, propia de la /o/ del español.
- En quinto lugar, se ha comprobado que las diferencias entre los tres puntos de las vocales del español articuladas por las estudiantes chinas son, en general, estadísticamente significativas, y los datos han mostrado que también son significativas las diferencias entre los dos grupos de estudio.

En relación con las diferencias entre los tres puntos de medición de los formantes, en el caso del grupo de control se ha mostrado que, en la mayoría de los casos, la desviación estándar es mayor en el inicio y en el final del formante, y menor en el centro, lo que indica que el punto medio de la vocal es más homogéneo que los otros dos. Estas diferencias son significativas, y, como se ha indicado, están relacionadas con el entorno de la vocal, tanto el anterior, como el posterior. En el caso del inicio, los órganos fonatorios cambian de posición desde el sonido consonántico anterior hasta el centro de la vocal, y en el final, estos órganos, o bien se preparan para la realización del otro sonido consonántico, o se deslizan hacia la posición de reposo si se encuentran en final de palabra aislada. Es decir,

los datos obtenidos en esta investigación siguen las consideraciones de RAE (2011), la cual indica que “(e)l contorno y otras variables, como el tipo de habla, pueden modificar la abertura vocálica y provocar la aparición de variantes más abiertas o más cerradas” (RAE, 2011: 82). En cambio, en la producción oral de las informantes sinohablantes no se ha detectado un comportamiento generalizado en lo que a las diferencias entre los distintos puntos se refiere.

Como se ha indicado en los capítulos I y II, hay falta de acuerdo entre los sinólogos chinos en lo que a los fonemas vocálicos que componen el chino estándar se refiere. De acuerdo con Duanmu (2007: 35), son los siguientes: [i], [y], [u], [ə], [a]. No obstante, el investigador realiza un análisis de las propuestas realizadas por otros autores, e indica que, en general hay poca controversia en relación con las vocales altas [i] y [u], las cuales indica el autor que son similares a las del inglés. En cuanto a [y], Duanmu (2007: 37) señala que algunos investigadores proponen que puede ser sustituida por [iu]. En cambio, la vocal central [ə] tiene varias variantes, y hay desacuerdo en cuanto a la manera de transcribirlas. Las variantes indicadas son [o], [E], [ɤ], [e], [ə], [ʌ], y algunos incluso escriben [ou]. Duanmu (2007: 37) indica que las diferencias en la transcripción pueden, en parte, reflejar la variación entre los hablantes de chino estándar, y concluye que la vocal central puede cambiar su anterioridad y, posiblemente, ser redonda o no, pero no su altura. Como se ha visto en esta tesis, una de las vocales que mayor variedad de realizaciones presenta es la /e/, lo que indica que las características de los sonidos vocálicos de la lengua materna de las informantes sinohablantes se reflejan en su producción oral en español.

- En último lugar, y en relación con los resultados del análisis del timbre de las vocales producidas por las hablantes que conforman el grupo de control, cabe señalar que en la presente tesis doctoral se ha mostrado que cuando los sonidos vocálicos son producidos por voces femeninas, el primer formante aparece en un rango de frecuencias entre un 13 y 59% más altas que las voces masculinas, y en el segundo, las voces son entre un 16 y 38% más agudas. Por ello, se concluye que, dado que la mayoría de trabajos que analizan las vocales del español se centran en voces masculinas, es necesario tener referencias de voces femeninas que permitan realizar un análisis adecuado de las vocales del español.

Así pues, del análisis acústico del timbre de las vocales del español producidas por estudiantes sinohablantes se deduce que esta clase de sonidos, acústicamente, presentan importantes diferencias con respecto a los timbres de las vocales producidas por las informantes hispanohablantes. Y estas diferencias no son debidas exclusivamente a las características intrínsecas de los sonidos vocálicos del español, sino que están relacionadas con la forma en la que han sido realizados los sonidos del entorno, y a la estructura silábica de las palabras tal como realmente la han pronunciado las estudiantes sinohablantes.



**CAPÍTULO VI**  
**CONCLUSIONES GENERALES**



## VI. CONCLUSIONES GENERALES

Tal como se ha indicado en los primeros capítulos, la presente investigación analiza el proceso de adquisición del sistema vocálico del español por sinohablantes. El mandarín y el español son dos lenguas que pertenecen a familias muy distintas y que, por lo tanto, poseen características fonéticas, morfológicas y sintácticas muy diferentes. Dicho en otras palabras, guardan una distancia lingüística muy grande entre ellas, por lo que cabe pensar que el proceso de adquisición del español por hablantes de mandarín presenta unas características distintas de las que presenta la adquisición de una lengua indoeuropea por parte de hablantes de otra lengua indoeuropea.

En el capítulo II se ha mostrado que la mayoría de trabajos que analizan la adquisición del sistema fonético-fonológico de segundas lenguas o lenguas extranjeras sigue la *Hipótesis del Análisis Contrastivo* o el *Speech Learning Model*. En ambos casos, se trata de modelos elaborados a partir de datos procedentes de lenguas emparentadas en un grado mayor o menor, pero no se han contemplado en dichos modelos datos procedentes del proceso de adquisición de lenguas tan alejadas entre sí como el chino y el español.

Según el análisis contrastivo, el estudiante de una lengua extranjera encuentra siempre en ella sonidos similares a los de su lengua nativa, por lo que el aprendizaje de estos se da por simple transferencia, sin dificultad alguna. Por otro lado, se enfrenta también con sonidos que no forman parte del sistema fonético de la lengua materna, cuya estructura o distribución es diferente. El aprendizaje de estos sonidos se realiza más lentamente que el de los primeros y presenta mayores dificultades a los estudiantes (Lado, 1957: 12).

En cambio, el *Speech Learning Model* propuesto por Flege (1995) afirma que los sonidos similares en ambas lenguas son difíciles de percibir y producir, mientras que los nuevos resultan más fáciles. Es decir, ambos modelos coinciden en subrayar que la influencia de la lengua materna del estudiante es fundamental en la adquisición del sistema fonético de la lengua meta, pero difieren en lo que a los sonidos problemáticos se refiere.



Los trabajos que analizan las dificultades de los estudiantes sinohablantes durante la adquisición del sistema fonético del español son muy recientes: la difusión de la enseñanza de español en China data de los años 1990 y, por tanto, los primeros títulos de la bibliografía que estudian estas cuestiones se publican entre 1990 y 1995 de la mano de profesores que enseñan español en China y se enfrentan a enormes dificultades de todo tipo. Estos profesores cuentan, como instrumentos conceptuales, con las herramientas de estudio, descripción y clasificación empleados en la enseñanza de lenguas extranjeras en Europa y, sobre todo, en Rusia, como se ha apuntado en la introducción, y adoptan, por tanto, la perspectiva contrastiva para describir los problemas y proponer estrategias didácticas para resolverlos.

El Análisis Contrastivo proporciona, aparentemente, un sistema útil para predecir los posibles errores y dificultades de los sinohablantes frente al sistema del español. Generalmente estos problemas se clasifican en errores que impiden la comunicación y en errores que la dificultan, y aquellos sonidos que no dificultan la inteligibilidad del mensaje pasan a considerarse como no problemáticos. Además, cabe señalar que este tipo de estudios toman en consideración los fonemas como si se pronunciaran de manera aislada, y no tienen en cuenta que están sometidos a factores que modifican la pronunciación. Así, las vocales del español son clasificadas, con mucha frecuencia, en la tercera categoría, la de los sonidos no problemáticos, puesto que los profesores de español detectan que la pronunciación de los sonidos consonánticos del español por parte de sus estudiantes chinos plantea graves problemas de inteligibilidad mientras que no es el caso de las vocales. Por ejemplo, los alumnos sinohablantes pueden pronunciar *gamba* como ['kamba], *velas* como ['pelas] o *duda* como ['tuta]. Frente a estos fenómenos poca importancia tienen para los profesores las realizaciones más abiertas o más cerradas o nasalizadas de una vocal determinada porque son rasgos que no alteran el mensaje. No obstante, el estudio del proceso de adquisición del sistema vocálico sí debe tener en cuenta estas cuestiones y estudiar las características articulatorias y acústicas de las realizaciones de los estudiantes.

Fenómenos como los tonos, que tienen valor fonológico en chino y no existen en español, ejercen una importante influencia en la pronunciación de las vocales del español en lo referente a su duración y a las oscilaciones en los valores de frecuencia de los formantes que imprimen los sinohablantes a sus producciones en español. Otros aspectos del mandarín como la estructura silábica que, en chino, es casi siempre CV, ejercen

también una importante influencia en los timbres de las vocales españolas realizadas por estos estudiantes. Todas estas cuestiones están relacionadas con la enorme distancia lingüística que mantienen, entre sí, las dos lenguas.

En esta tesis, se ha intentado poner de manifiesto todos estos aspectos del proceso de adquisición mediante la primera fase del trabajo, el análisis auditivo, que ha permitido establecer, cualitativamente, cuáles son las influencias ejercidas por estos rasgos sobre el timbre de las realizaciones vocálicas.

El análisis de la influencia de estas variables cualitativas independientes en la pronunciación de las vocales del español ha permitido comprobar que, si bien los sonidos consonánticos que no existen en el sistema fonético chino resultan, en general, problemáticos, estos estudiantes también presentan dificultades al producir aquellos sonidos que son similares en ambas lenguas. De la misma manera, las estudiantes chinas han mostrado tener dificultades para realizar correctamente las estructuras silábicas del español, tan distintas de las de su lengua, lo que ha conducido a la aparición de fenómenos de resilabación no señalados en la bibliografía existente. Además, se ha podido observar que muchos de los procesos de resilabación tienen lugar cuando el entorno anterior y/o posterior de la vocal es un sonido consonántico que resulta especialmente problemático para las informantes sinohablantes. Producto de estos fenómenos, los timbres de las vocales del español producidas por las hablantes chinas se han visto alterados, tal y como se ha podido comprobar.

En cuanto al análisis de la influencia de las variables cualitativas en la duración de las vocales del español producidas por las informantes sinohablantes, se ha comprobado que los sonidos vocálicos articulados por estas hablantes son más largos y menos homogéneos que aquellos realizados por los hablantes cuya lengua materna es el español. Como se ha indicado, estos fenómenos están relacionados con las características de la lengua materna de las estudiantes analizadas, y es que los sonidos vocálicos de las lenguas tonales, como el chino, tienen una duración mayor que los propios de las lenguas no tonales, como el español, ya que la realización de los tonos requiere que el sonido presente una mayor duración. Dado que es difícil mantener la homogeneidad en el timbre cuando las vocales son más largas, este fenómeno también conduce a que los timbres de las vocales del español producidas por estas hablantes presenten variaciones en su seno.

Finalmente, el análisis de las características del timbre de las vocales producidas por las informantes sinohablantes ha permitido comprobar que las realizaciones de las hablantes chinas son, en general, menos homogéneas y presentan mayor variación que las de las del grupo de control. Además, se han podido detectar realizaciones alejadas de los valores propios de las vocales del español.

A modo de conclusión de carácter general, la presente tesis doctoral ha mostrado que el estudio de la adquisición de una clase de sonidos no depende exclusivamente de las características intrínsecas de los mismos, sino que para comprender las razones que conducen a los estudiantes sinohablantes a pronunciar, en este caso, las vocales del español de forma diferente a la de los nativos, deben considerarse todos los factores que influyen en la realización de los mismos. Es decir, para que la enseñanza del sistema fonético del español llegue a buen puerto, es necesario tener un conocimiento lingüístico que permita definir estrategias para que la adquisición sea efectiva y no adoptar un enfoque reduccionista.

Así, en realidad, esta tesis se propone constituir una aportación al conocimiento del proceso de adquisición de una lengua que mantiene una distancia lingüística muy grande con respecto a la lengua base de los estudiantes de forma que los datos obtenidos puedan ser integrados en los modelos de adquisición de lenguas extranjeras. Desde el punto de vista práctico, los datos obtenidos, que permiten conocer mejor el proceso de adquisición del español por sinohablantes, pueden constituir una base importante para la preparación de actividades de enseñanza de la pronunciación adaptadas a las necesidades de estos estudiantes.

## **BIBLIOGRAFÍA**



## BIBLIOGRAFÍA

Para citar las referencias bibliográficas se han seguido las normas de citación indicadas en el manual de publicaciones de la *American Psychological Association* (APA), sexta edición.

Alcoholado Feltstrom, A. (2012). La sílaba como unidad estructural: propuesta de actividades para principiantes sinófonos de ELE. *SinoELE* 6, 71-83. Recuperado de [https://www.sinoele.org/images/Revista/6/alcoholado\\_71-83.pdf](https://www.sinoele.org/images/Revista/6/alcoholado_71-83.pdf)

Alcoholado Feltstrom, A. (2013a). Vocales en contacto en la expresión oral del español. Planteamientos teóricos y perspectivas de trabajo con estudiantes universitarios chinos de español como lengua extranjera”. *Suplementos SinoELE* 8. Recuperado de [https://www.sinoele.org/images/Revista/8/Suplementos/antonio\\_alcoholado.pdf](https://www.sinoele.org/images/Revista/8/Suplementos/antonio_alcoholado.pdf)

- - - (2013b) Las combinaciones vocálicas del español en la expresión oral de universitarios chinos. *SinoELE* 9, 1-16. Recuperado de: [http://www.sinoele.org/images/Revista/9/Articulos/alcoholado\\_1-16.pdf](http://www.sinoele.org/images/Revista/9/Articulos/alcoholado_1-16.pdf)

Archangeli, D. (1988). Aspects of underspecification theory. *Phonology*, 5 (2), 183-207. doi: 10.1017/S0952675700002268

Best, C. (1993). Emergence of language-specific constraints in perception of non-native speech: A window on early phonological development. En B. de Boysson-Bardies (Ed.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 289-304). Dordrecht, Países Bajos: Kluwer.

- - - (1994). The emergence of native-language phonological influence in infants: A perceptual assimilation model. En H. Nussbaum, J. Goodman, & C. Howard (Eds.), *The transition from speech to spoken words: The development of speech perception* (pp. 167-224). Cambridge, Estados Unidos: MIT Press.

- - - (1995). A direct realist view of cross-language speech perception. En W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171-204). Timonium, Estados Unidos: York Press.
- Best, C.T., McRoberts, G.W. y Sithole, N.M. (1988). Examination of perceptual reorganization for non-native speech contrasts: Zulu click discrimination by English speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* 14(3), 345-360.
- Best, C.T., Faber, A. y Levitt, A. (1996). Assimilation of non-native vowel contrasts to the American English vowel system. *The Journal of the Acoustical Society of America*. doi: 10.1121/1.415316
- Blanco Canales, A. y Nogueroles López, M. (2013). Descripción y Categorización de Errores Fónicos en Estudiantes De Español/L2. Validación de la Taxonomía de Errores AACFELE. *Logos: Revista de Lingüística, Filosofía y Literatura* 23 (2), 196-225. Recuperado de <https://revistas.userena.cl/index.php/logos/article/view/365/360>
- Broselow, E., Chen, S., & Wang, C. (1998). The emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition*, 20, 261-280. Recuperado de: <https://pdfs.semanticscholar.org/af9f/db0cea0a6abe834b19fe566cd9c2b268c703.pdf>
- Brown, C. (1998). The role of the L1 grammar in the acquisition of segmental structure. *Second Language Research*, 14, 139-193. Recuperado de: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.860.5090&rep=rep1&type=pdf>
- - - (2000). The interrelation between speech perception and phonological acquisition from infant to adult. En J. Archibald (Ed.), *Second language acquisition and linguistic theory* (pp. 4-63). Oxford, Reino Unido: Blackwell.
- Chaoju, T. y Van Heuven, V.J. (2007). Predicting mutual intelligibility in Chinese dialects from subjective and objective linguistic similarity. *Interlingüística*, 17, 1019-1028.

- - - (2008). Mutual intelligibility of Chinese dialects tested functionally. *Linguistics in the Netherlands*. doi: 10.1075/avt.25.16tan
- - - (2009). Mutual intelligibility of Chinese dialects experimentally tested. *Lingua*, 119. doi: 10.1016/j.lingua.2008.10.001
- Chih, M. T. (2013). E/LE en Taiwán: Problemas de apreciación fonética en estudiantes universitarios de grado. *SinoELE* 9 (julio-diciembre) 17-32. Recuperado de: [http://www.sinoele.org/images/Revista/9/Articulos/mchih\\_17-32.pdf](http://www.sinoele.org/images/Revista/9/Articulos/mchih_17-32.pdf)
- Chomsky, N. & Halle, M. (1968). *The sound pattern of English*. Nueva York, Estados Unidos: Harper & Row Publishers.
- Consejo de Europa (2002). *Marco Europeo de Referencia para las lenguas: aprendizaje, enseñanza, evaluación*. Madrid, España: Anaya/Instituto Cervantes.
- Corder, S. P. (1971). Idiosyncratic dialects and error análisis. *IRAL*, 9, 149-159.
- Cortés Moreno, M. (1998). Sobre la percepción y adquisición de la entonación española por parte de hablantes nativos de chino. *Estudios de Fonética Experimental*, IX, 67-134. Recuperado de [http://stel.ub.edu/labfon/sites/default/files/EFE-IX-MCortes-percepcion\\_y\\_adquisicion\\_entonacion\\_esp%C2%A7aola\\_por\\_parte\\_de\\_nativos\\_de\\_chino.pdf](http://stel.ub.edu/labfon/sites/default/files/EFE-IX-MCortes-percepcion_y_adquisicion_entonacion_esp%C2%A7aola_por_parte_de_nativos_de_chino.pdf)
- - - (2001a). Percepción y adquisición de la entonación española en enunciados de habla espontánea: el caso de los estudiantes taiwaneses. *Estudios de Fonética Experimental*, XI, 89-119. Recuperado de <http://www.raco.cat/index.php/EFE/article/viewFile/143935/195623>
- - - (2001b). Producción de la entonación española en enunciados de habla espontánea: el caso de los estudiantes taiwaneses. *Estudios de Fonética Experimental*, XI, 191-209. <https://www.raco.cat/index.php/EFE/article/view/144006/195706>
- - - (2003a). Percepción y adquisición de la acentuación española en la lectura: el caso de los estudiantes taiwaneses. *Estudios de Fonética Experimental*, XII, 61-89. Recuperado de: <https://www.raco.cat/index.php/EFE/article/view/140003/194684>
- - - (2003b). Producción y adquisición de la acentuación española en habla espontánea: el caso de los estudiantes taiwaneses. *Estudios de Fonética Experimental*, XII, 91-



103. Recuperado de <https://www.raco.cat/index.php/EFE/article/view/140004/194685>
- - - (2004). Análisis acústico de la producción de la entonación española por parte de sinohablantes. *Estudios de Fonética Experimental*, XIII, 79-110. Recuperado de <https://core.ac.uk/download/pdf/39111113.pdf>
- - - (2005). Análisis experimental del aprendizaje de la acentuación y la entonación españolas por parte de hablantes nativos de chino. *PHONICA*, 1, 1-25. Recuperado de [http://www.publicacions.ub.es/revistes/phonica1/PDF/articulo\\_04.pdf](http://www.publicacions.ub.es/revistes/phonica1/PDF/articulo_04.pdf)
- - - (2006). Análisis acústico de la transferencia de rasgos del sistema tonal chino al habla en español como lengua extranjera. *Estudios de Fonética Experimental*, XV, 43-65. Recuperado de <https://www.raco.cat/index.php/EFE/article/view/140039/219174>
- - - (2009a). De la fonología a la enseñanza de la pronunciación mediante juegos. *Suplementos MarcoELE*, 8, 1-18. Recuperado de [https://marcoele.com/descargas/china/cortes\\_pronunciacion.pdf](https://marcoele.com/descargas/china/cortes_pronunciacion.pdf)
- - - (2009b). Chino y español: un análisis contrastivo. En Sánchez Griñán, A. y Melo, M. (comp.) *Qué saber para enseñar a estudiantes chinos* (pp. 183-210). Buenos Aires, Argentina: Ediciones Voces del Sur.
- Dong, Y y Liu, J (2006). *El Español Moderno, volumen 1*. Pekín, China: Editorial Enseñanza de lenguas extranjeras y prensa de investigación.
- Duanmu, S. (2007). *The phonology of standard Chinese* (2da edic.). Oxford, Reino Unido: Oxford University Press.
- Eckman, F. R. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning* 27, 315-330. doi: 10.1111/j.1467-1770.1977.tb00124.x
- - - (1991). The Structural Conformity Hypothesis and the acquisition of consonant clusters in the interlanguage of ESL learners. *Studies in Second Language Acquisition* 13, 23-41. doi: 10.1017/S0272263100009700
- - - (2004). From phonemic differences to constraint rankings. *Studies in Second Language Acquisition*, 26, 513-549. doi: 10.1017/S027226310404001X

- Eckman, F. R. e Iverson, G. K. (1997). Structure preservation in interlanguage phonology. En Hannahs, S.J. & Young-Scholten, Martha (eds.) *Focus on phonological acquisition*, (pp. 183-207). Amsterdam/Philadelphia, Países Bajos/Estados Unidos: John Benjamins Publishing Company. doi: 10.1075/lald.16
- Fisac, T. (2000). La enseñanza del español en Asia Oriental. En Instituto Cervantes (Ed.) *El español en el mundo. Anuario del Instituto Cervantes*. Madrid, España: Instituto Cervantes, Círculo de Lectores, Plaza & Janés.
- Flege, J. E. (1980). Phonetic approximation in second language acquisition. *Language Learning*, 30 (1), 117-134. doi: 10.1111/j.1467-1770.1980.tb00154.x
- - - (1987). The production of “new” and “similar” pones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics* 15, 47-65.
- - - (1995). Second language speech learning: theory, findings and problems. En Strange, W. (Ed.), *Speech perception and linguistic experience: issues in cross-language research* (pp. 233-277). Timonium, Estados Unidos: York Press.
- - - (2002). Interactions between the native and second-language phonetic systems. En Burmeister, P., Piske, T. & Rohde, A. (Eds.), *An integrated view of language development: papers in honor of Henning Wode* (pp. 217-244). Trier, Alemania: Wissenschaftlicher Verlag.
- Gong, Y., Chow, I.H. y Ahlstrom, D. (2010). Cultural diversity in China: Dialect, job embeddedness, and turnover. *Asia Pac J. Manag*, 28. doi: 10.1007/s10490-010-9232-6
- Hammerly, H. (1982). Contrastive phonology and error analysis. *International Review of Applied Linguistics*, 20, 17-32. doi: 10.1515/iral.1982.20.1-4.17
- Hancin-Bhatt, B. (1994). Segment transfer: a consequence of a dynamic system. *Second Language Research*, 10, 241-269.
- - - (2008). Second language phonology in optimality theory. En Edwards, H., Jette G. y Zampini, M. L. (Eds.), *Phonology and Second Language Acquisition* (pp. 117-151). Amsterdam/Philadelphia, Países Bajos/Estados Unidos: John Benjamins Publishing Company.
- Hancin-Bhatt, B. & Bhatt, R. (1997). Optimal L2 syllables: interaction of transfer and developmental effects. *Studies in Second Language Acquisition*, 19, 331-378.

- Henly, E. y Sheldon, A. (1986). Duration and context effects on the perception of English /r/ and /l/: A comparison of Cantonese and Japanese speakers. *Lang Learn* 36: 505-521.
- Igarreta Fernández, A. (2015a). Los problemas de pronunciación de la lengua española por parte de los estudiantes sinohablantes. *Actas del I Congreso Internacional de la Asociación Europea de Profesores*, 258-269. Recuperado de [https://cvc.cervantes.es/ensenanza/biblioteca\\_ele/aepe/pdf/congreso\\_50/congreso\\_50\\_29.pdf](https://cvc.cervantes.es/ensenanza/biblioteca_ele/aepe/pdf/congreso_50/congreso_50_29.pdf)
- - - (2015b). La corrección de la pronunciación de los estudiantes sinohablantes en el aula de E/LE. *Foro de Profesores de E/LE*, 11, 189-196. Recuperado de <https://ojs.uv.es/index.php/foroэле/article/view/7115/9419>
- Instituto Cervantes (Ed.) (2000). *El español en el mundo. Anuario del Instituto Cervantes*. Madrid, España: Instituto Cervantes, Círculo de Lectores, Plaza & Janés.
- Jakobson, R. (1941). *Child language, aphasia, and phonological universals*. The Hague, Países Bajos: Mouton Publishers.
- Jiménez, J. y Tang, A. (2018). Producción del sistema vocálico del español por hablantes de chino. *Revista de Lingüística Teórica y Aplicada*, 56 (1), 13-34. Recuperado de <https://scielo.conicyt.cl/pdf/rla/v56n1/0718-4883-rla-56-01-13.pdf>
- Jingsheng, L. (2012). Dificultades del sistema consonántico para sinohablantes desde el enfoque contrastivo. *Monográficos SINOELE*, 10, 7-16. Recuperado de [https://www.sinoele.org/images/Revista/10/Monograficos/EPES\\_2012/EPES%20III%20-%20LU\\_7-16.pdf](https://www.sinoele.org/images/Revista/10/Monograficos/EPES_2012/EPES%20III%20-%20LU_7-16.pdf)
- - - (2014). La génesis y el desarrollo de los estudios de español en China. En Instituto Cervantes (Ed.) *El español en el mundo. Anuario del Instituto Cervantes*. Recuperado de: [https://cvc.cervantes.es/lengua/anuario/anuario\\_14/jingsheng/p01.htm](https://cvc.cervantes.es/lengua/anuario/anuario_14/jingsheng/p01.htm)
- Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor, Estados Unidos: University of Michigan Press.
- Landis JR. y Koch GG. (1977). The measurements of observer agreement for categorical data. *Biometrics* 33, 159-174.

- Lehiste, I. (1988). *Lectures on Language Contact*. Cambridge, Estados Unidos: MIT Press.
- Lenneberg, E. (1967). *Biological Foundations of Language*. New York, Estados Unidos: Wiley.
- Lin, Y.-H. (2007). *The Sounds of Chinese*. Cambridge, Reino Unido: Cambridge University Press.
- Liu, N. (2012). *Análisis de errores y evolución de la Interlengua de los estudiantes sinohablantes: esbozo de un estudio longitudinal* (Trabajo de fin de máster). Recuperado de <https://www.recercat.cat/bitstream/handle/2072/209907/Trabajo%20de%20investigaci%C3%B3n-Liu%20Nian.pdf?sequence=1>
- Logan, J., Lively, S., y Pisoni, D. (1991). Training Japanese listeners to identify /r/ and /l/: A first report. *Journal of the Acoustical Society of America* 89, 874-86.
- López, M. (2012). *Estudio de los problemas de pronunciación de los estudiantes sinohablantes que aprenden español* (Trabajo de fin de máster). Universidad de Cantabria – Fundación Comillas, España.
- Major, R. C. (2008). Transfer in second language phonology. En Hansen Edwards, J. G. y Zampini, M. L. (Eds.), *Phonology and Second Language Acquisition*. Amsterdam/Philadelphia, Países Bajos/Estados Unidos: John Benjamins Publishing Company.
- Martínez Celdrán, E. (2007). *Análisis espectrográfico de los sonidos del habla* (2ª edición). Barcelona, España: Ariel.
- Méndez Marassa, E. (2009). Problemas de los estudiantes chinos de español. Ejercicios específicos. *Suplementoc SinoELE*, 1, 1-99. Recuperado de <http://www.sinoele.org/images/Revista/1/mendez.pdf>
- Miranda Márquez, G. (2014). Distancia lingüística, a nivel fonético-fonológico, entre las lenguas china y española. *Philologia Hispalentis* 28, 51-68. doi: 10.12795/PH.2014.v28.i01.04
- Muñoz Torres, M. (2012). Dificultades fonético-fonológicas de los estudiantes chinos. *Monográficos Sinoele*, 19, 208-227. Recuperado de

[http://www.sinoele.org/images/Revista/10/Monograficos/EPES\\_2012/EPES%20III%20-%20MUNOZ\\_208-227.pdf](http://www.sinoele.org/images/Revista/10/Monograficos/EPES_2012/EPES%20III%20-%20MUNOZ_208-227.pdf)

- Nemser, W. (1971). Approximative systems of foreign language learners. *IRAL*, 9, 115-123. doi: 10.1515/iral.1971.9.2.115
- Oller, J. & Ziahosseiny, S. (1970). The contrastive analysis hypothesis and spelling errors. *Language Learning*, 20, 183-189. doi: 10.1111/j.1467-1770.1970.tb00475.x
- Ortí Mateu, R. (1990). *Comparación fonética, diagnóstico y tratamiento de las dificultades de los estudiantes chinos para aprender español* (Tesis doctoral). Recuperado de <https://digitum.um.es/digitum/handle/10201/5769>
- Pan, Y. (2000). Code-switching and social change in Guangzhou and Hong Kong. *Internationa Journal of the Sociology of Language* 146, 21-41.
- Penfield, W. (1965). Conditioning the uncommitted cortex for language learning. *Brain* 88, 787-98. doi: 10.1093/brain/88.4.787
- Pérez García, R. (2018). *La adquisición del sistema vocálico del español por hablantes con lengua maternal alemán y chino* (Tesis doctoral). Recuperado de <https://www.tesisenred.net/bitstream/handle/10803/666059/rpg1de1.pdf?sequence=1&isAllowed=y>
- Planas Morales, S. (2008). Enseñanza de la pronunciación del español en estudiantes chinos: la importancia de las destrezas y los contenidos prosódicos. *XVIII Congreso internacional de la asociación para la enseñanza del español como lengua extranjera*, 3, 497-503. Recuperado de [https://cvc.cervantes.es/ensenanza/biblioteca\\_ele/asele/pdf/18/18\\_0497.pdf](https://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/pdf/18/18_0497.pdf)
- - - (2013). Análisis contrastivo de las vocales del español y el chino mandarín: reflexiones para el profesor de ELE. *Revista Nebrija de Lingüística Aplicada* 13, pp.
- - - (2016). Enseñanza y evaluación de la pronunciación de E/LE en alumnos chinos. *redeLE*, 17, 70-86. Recuperado de <https://www.educacionyfp.gob.es/dam/jcr:df5c6459-1273-436b-966a-5e749d27e570/2009-redele-17-04planas-pdf.pdf>

- Planas Morales, S. y Villalba, X. (2007). Las similitudes en el contorno entonativo de las palabras del chino mandarín y de los grupos fónicos del español. *XVII Congreso Internacional de la Asociación del Español como lengua extranjera*, 2, 901-914. Recuperado de [https://cvc.cervantes.es/ensenanza/biblioteca\\_ele/asele/pdf/17/17\\_0901.pdf](https://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/pdf/17/17_0901.pdf)
- Poch, D. e Igarreta, A. (2014). Tender puentes: distancia lingüística y pronunciación. En Ferrús, B. y Poch, D. (Eds.), *El español entre dos mundos. Estudios de ELE en Lengua y Literatura* (pp. 139-157). Madrid/Frankfurt, España/Alemania: Iberoamericana – Vervuert.
- Polivanov, E. (1931). La perception des sons d'une langue étrangère. *Travaux du Cercle Linguistique de Prague* 4, 111-114. París, Francia.
- Pose, R. (2017). Fonética del español para sinohablantes: de la teoría a la práctica. *SIGNOS ELE*, 11, 1-19. Recuperado de <https://p3.usal.edu.ar/index.php/ele/article/view/4120>
- Prince, A. & Smolensky, P. (1993). *Optimality Theory: Constraint interaction in a generative grammar*. New Brunswick, Estados Unidos: Rutgers University Center for Cognitive Science.
- (2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Oxford, Reino Unido: Blackwell.
- Ramsey, R. (1987). *The languages of China*. Princeton, Estados Unidos: Princeton University Press.
- Real Academia Española y Asociación de la Lengua Española (2011). *Nueva gramática de la lengua española. Fonética y fonología*. Madrid, España: Espasa.
- Real Instituto Elcano (2018). *Relaciones España-China*. Madrid, España: Real Instituto Elcano.
- Sagey, E. (1986). *The representation of features in non-linear phonology* (Tesis doctoral). Massachusetts Institute of Technology, Cambridge.
- Selinker, L. (1972). Interlanguage. *IRAL*, 10, 209-231.
- Stockwell, R., & Bowen, J. (1965). *The sounds of English and Spanish*. Chicago, Estados Unidos: University of Chicago Press.

- Strange, W. (1992). Learning non-native phoneme contrasts: Interactions among subject, stimulus, and task variables. En E. Tohkura, E., Vatikiotis-Bateson, e. y Sagisaka, Y. (Eds.), *Speech perception, production, and linguistic structure* (pp. 197-219). Tokyo, Japón: Ohmsha.
- Strange, W. y Dittmann, S. (1984). Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. *Perception & Psychophysics*, 36(2), 131-145.
- Trubetzkoy, N. (1939). *Principios de fonología*. Versión traducida de 1976. España: Cincel.
- Wang, H-Y. (2001). Estudio fónico del chino mandarín y del español. *Encuentros en Catay*, 15, 70-118. Recuperado de [http://www.sinoele.org/images/Encuentros/wang%20\(2001\).pdf](http://www.sinoele.org/images/Encuentros/wang%20(2001).pdf)
- Wang, H. y Van Heuven, V.J. (2006). Acoustical análisis of English vowels produced by Chinese, Dutch and American speakers. En J. M Van de Weijer y B. Los (Eds.), *Linguistic in the Netherlands* (págs. 237-248). Amsterdam/Philadelphia: John Benjamins.
- Weinreich, U. (1957). On the description of phonic interference. *Word* 13: 1-11. doi: 10.1080/00437956.1957.11659624
- - - (1968). *Languages in Contact. Findings and problems*. The Hague-Paris, Países Bajos, Francia: Mouton.
- Wen, D. (2008). *Voz en español estándar*. China: Prensa de Enseñanza e Investigación en Lenguas Extranjeras.
- Wode, H. (1976). Developmental sequences in naturalistic second language acquisition: The question of transfer. *Studies in Second Language Acquisition*, 9, 261-286.
- - - (1978). The beginning of non-school room L2 phonological acquisition. *International Review of Applied Linguistics in Language Learning*, 16, 109-125.
- - - (1983a). Contrastive analysis and language learning. En H. Wode (Ed.), *Papers on language acquisition, language learning, and language teaching* (pp. 202-212). Heidelberg, Alemania: Groos.

- - - (1983b). Phonology in L2 acquisition. En Wode, H. (Ed.) *Papers on language acquisition, language learning, and language teaching* (pp. 175-187). Heidelberg, Alemania: Groos.
- - - (1994). On Positive Evidence and Ultimate Attainment in L2 Phonology. *Second Language Research* 10, 193-214.
- Young-Scholten, M. (1985). Interference reconsidered: The role of similarity in second language acquisition. *Selecta* 6, 6-12.
- Young-Scholten, M. & Archibald, J. (2004). Second Language Syllable Structure. En Archibald, J. (Ed.) *Second Language Acquisition and Linguistic Theory*, Oxford, Reino Unido: Blackwell Publishers.
- Yufei, C. (2007). Un estudio contrastivo de los fonemas oclusivos entre español y chino. Reflexiones sobre la enseñanza de la fonética español a estudiantes chinos. *México y la cuenca del Pacífico*, 10, (28), 91-98. Recuperado de <http://sincronia.cucsh.udg.mx/caoyufei06.htm>
- Zhi, C. (2011). Errores articulatorios de los estudiantes chinos en la pronunciación de las consonantes españolas. *SinoELE* 4, 54-67. Recuperado de <http://www.sinoele.org/images/Revista/4/chen.pdf>
- Zhou, M. (1995). *Estudio comparativo del chino y el español. Aspectos lingüísticos y culturales* (Tesis doctoral). Recuperado de <https://www.tesisenred.net/bitstream/handle/10803/5277/TMZ2de2.pdf?sequence=2&isAllowed=y>