

UNIVERSIDAD DE SEVILLA

FACULTAD DE FILOLOGÍA

Departamento de Lengua Inglesa



**APROXIMACIÓN A LA PERCEPCIÓN DE LAS
CONSONANTES INGLESAS POR APRENDICES DE L2:
BENEFICIOS DEL ENTRENAMIENTO
AUDITIVO INTENSIVO**

TESIS DOCTORAL

Beatriz Redondo Vigil

Dirigida por:

Dra. Teresa López Soto

Sevilla, 2015

Dra. Teresa López Soto

Beatriz Redondo Vigil

Tabla de contenidos generales

Índice de figuras	iv
Índice de tablas.....	v
Agradecimientos.....	vi
Dedicatoria	vii
Introducción.....	1
Capítulo 1. Antecedentes	4
1.1. Percepción del habla: indicios acústicos y el concepto de prototipo.....	3
1.2. Modelo del Imán de la Lengua Materna (NLM)	6
1.3. Modelo de Aprendizaje del Habla (SLM) y Modelo de Asimilación Perceptiva (PAM)	8
1.3.1. Modelo de Aprendizaje del Habla (SLM): la importancia de la transferencia de la lengua materna en la adquisición fonológica de lenguas extranjeras.....	8
1.3.2. Modelo de Asimilación Perceptiva (PAM): la convergencia entre la Teoría Motriz de la percepción del habla y el modelo auditivo (modelo psicoacústico)	11
1.4. La percepción de las consonantes inglesas por no nativos: antecedentes experimentales	16
Capítulo 2. Hipótesis y objetivos	25
2.1. Hipótesis.....	25
2.2. Objetivos	27
Capítulo 3. Materiales y métodos.....	29
3.1. Experimento I: Test general de identificación de consonantes inglesas	30
3.1.1. Población	30
3.1.2. Materiales	32
3.1.2.1. Estímulos	32
3.1.2.2. Tests de identificación	34
3.1.3. Procedimiento	35
3.2. Experimento II: Test de entrenamiento auditivo intensivo. Efectos del contexto adverso (ruido de fondo) y de la combinación de estímulos auditivos (sílabas vs. oraciones)	37
3.2.1. Población	37
3.2.2. Materiales	38

3.2.2.1. Estímulos	38
3.2.2.2. Tests de identificación	38
3.2.3. Procedimiento	41
Capítulo 4. Análisis de datos	44
4.1. Experimento I: Test general de consonantes inglesas	44
- Modo de articulación	46
- Lugar de articulación	48
- Sonoridad	50
- Consonantes oclusivas	52
- Consonantes fricativas	53
- Consonantes bilabiales	53
- Consonantes labiodentales	54
- Consonantes alveolares	54
- Consonantes velares	54
- Consonantes sonoras	55
- Consonantes aproximantes posición inicial	55
- Consonantes palatales posición inicial	56
- Consonantes africadas, nasales, líquidas en posición inicial y final	57
4.2. Experimento II: Entrenamiento auditivo intensivo. Efectos del contexto adverso (ruido de fondo) y de la combinación de estímulos auditivos (sílabas vs. oraciones)	58
4.2.1. Grupo de control: entrenamiento de sílabas a 15dB SNR	60
4.2.2. Grupo Experimental I: entrenamiento con sílabas a -5dB SNR	62
4.2.3. Grupo Experimental II: entrenamiento con palabras a 15dB SNR	63
Capítulo 5. Discusión.....	69
5.1. Cómo afecta la percepción del SNR en la muestra	74
5.2. Cómo afecta la percepción de la posición silábica en la muestra	78
5.3. Cómo afectan los rasgos articulatorios en la muestra	79
5.4. Cómo afecta el entrenamiento auditivo intensivo en relación a los rasgos articulatorios en la muestra, la posición en la sílaba y el ruido de fondo.....	82
Capítulo 6. Conclusión.....	90
Referencias bibliográficas.....	97
Anexos	
Anexo 1. Documentos	110
- Informe consentido.....	111
- Cuestionario demográfico	113

- Test de inglés.....	115
- Test auditivo.....	118
Anexo 2. Corpus estímulos silábicos	119
Anexo 3. Corpus de entrenamiento oracional	122
Anexo 4. Figuras de los mapas de etiquetas de los tests de entrenamiento	124
Anexo 5. Tablas de estadísticos descriptivos: modo de articulación y sonoridad en posición inicial y final	127

Índice de figuras

Figura 1. Noción de imán perceptivo	7
Figura 2. Medias atendiendo al modo de articulación en posición inicial	46
Figura 3. Medias atendiendo al modo de articulación en posición final	47
Figura 4. Medias atendiendo al lugar de articulación en posición inicial	48
Figura 5. Medias atendiendo al lugar de articulación en posición final	49
Figura 6. Medias atendiendo a la sonoridad en posición inicial	50
Figura 7. Medias atendiendo a la sonoridad en posición final	51
Figura 8. Porcentaje aciertos consonante inicial (Grupo Control)	60
Figura 9. Porcentaje aciertos consonante final (Grupo Control)	61
Figura 10. Porcentaje aciertos consonante inicial (Grupo experimental I)	62
Figura 11. Porcentaje aciertos consonante final (Grupo experimental I)	62
Figura 12. Porcentaje aciertos consonante inicial (Grupo experimental II)	64
Figura 13. Porcentaje aciertos consonante final (Grupo experimental II)	64
Figura 14. Comparativa Grupo de Control vs. Grupo Experimental I	66
Figura 15. Comparativa Grupo de Control vs. Grupo Experimental II	67
Figura 16. Comportamiento del SNR en las consonantes velares	74
Figura 17. Comportamiento del SNR en las consonantes aproximantes en posición inicial	75
Figura 18. Comportamiento del SNR en las consonantes palatales en posición inicial	76
Figura 19. Comportamiento del SNR en las consonantes líquidas en posición inicial	76
Figura 20. Comportamiento del SNR en las consonantes nasales en posición inicial	76
Figura 21. Comportamiento del SNR en las consonantes aspiradas en posición inicial	77
Figura 22. Comportamiento del SNR en las consonantes aspiradas en posición final	77
Figura 23. Mapa de etiquetas para el test de identificación de ataques (CV) 1	125
Figura 24. Mapa de etiquetas para el test de identificación de codas (VC) 1	125
Figura 25. Mapa de etiquetas para el test de identificación de ataques (CV) 2	125
Figura 26. Mapa de etiquetas para el test de identificación de codas (VC) 2	126
Figura 27. Mapa de etiquetas para el test de identificación de oraciones	127

Índice de tablas

Tabla 1. Clasificación de las consonantes atendiendo a su representación ortográfica y posición silábica	32, 120
Tabla 2. Medias atendiendo al modo de articulación en posición inicial	128
Tabla 3. Medias atendiendo al modo de articulación en posición final	128
Tabla 4. Tabla 2. Medias atendiendo a la sonoridad en posición inicial.....	129
Tabla 5. Tabla 2. Medias atendiendo a la sonoridad en posición final	129

Agradecimientos

Tras la finalización de este arduo trabajo debo mirar atrás y recordar a aquellas personas que han formado parte, directa e indirectamente, en este trabajo y sin su apoyo esto no hubiera sido posible. Para mí es necesario utilizar este espacio para ser consecuente y justa con ellas expresándole mis más profundos y sinceros agradecimientos.

Esta tesis doctoral no habría sido posible sin la valiosa ayuda de mis directora, doña Teresa López Soto. Le doy mi más profundo y sincero agradecimiento por la formación académica, la cual, sin duda, es muy valiosa, y también por la confianza depositada en mí y por aceptar la dirección de este proyecto. Especialmente quiero agradecerle su constante ayuda en las tareas, en cada problema experimental, en cada problema en la discusión de los resultados, pero más especialmente por su confianza brindada en mí, por hacer del trabajo diario un espacio para compartir alegrías, tristezas, problemas, proyectos, etc. No cabe duda que su dirección ha enriquecido el trabajo realizado y, además, ha supuesto el surgimiento de una buena amistad.

A los señores miembros del tribunal por haber asumido la importante tarea de evaluación de esta tesis.

Me gustaría agradecer a todos los profesores que han formado parte de mi educación, todos ellos son responsables de este trabajo. En especial al claustro del colegio Santo Domingo Savio de Madrid.

Por último, no puedo olvidar que la realización de esta tesis no habría sido posible sin el apoyo incondicional tanto económico como moral de mis padres. Sin su esfuerzo esto no habría sido posible. Además, su comprensión, paciencia y ánimo junto con la de mis amigos han sido fundamentales en la finalización de este trabajo.

Son muchas las personas que han formado parte de mi vida profesional y personal a las que me encantaría agradecerles su amistad, consejos, apoyo, ánimo y compañía en los momentos más difíciles de mi vida. Algunas están aquí conmigo y otras en mi recuerdo y mi corazón, sin importar en donde estén quiero darles las gracias por formar parte de mí y por todo lo que me han dado.

A todos, sinceramente, muchas gracias.

Dedicatoria

Dedico esta tesis a:

Mi abuela

que no la verá pero sé que desde donde esté estará orgullosa de mí.

Mis padres

por su apoyo y comprensión. Es a vosotros a quien debo la persona que soy.

Joaquín

por su apoyo y comprensión.

Os quiero.

Introducción

En los últimos años se ha ido afianzando la idea de que la lengua hablada se codifica en la mente del hablante a partir de prototipos previamente almacenados. La noción de prototipo se aplica en todo el campo de la ciencia cognitiva para explicar el procesamiento de información a partir de estímulos externos: los sentidos reciben los datos del exterior, pasan a través del sistema nervioso al cerebro, que va almacenándolos para crear patrones de identificación que se utilizarán posteriormente para la identificación de realidades similares. Junto a la idea de prototipo, los especialistas del campo trabajan con modelos inspirados en la teoría de cascada, que dentro de la ciencia cognitiva determina que la activación de uno de esos prototipos dispara al mismo tiempo la asociación de prototipos similares. En el campo del lenguaje, este concepto asociacionista se relaciona con la activación de redes de palabras que se activan a la vez que se percibe una unidad lingüística. Como ejemplo, la psicología cognitiva a partir del modelo de “cascada” afirma que el almacenamiento de unidades lingüísticas se procesa en términos de frecuencia de manera que para cada unidad que se recibe, otras tantas se activan en forma de expectativas: si el oyente escucha la sílaba “ca” en castellano, inmediatamente otras sílabas que con frecuencia han sido percibidas en secuencia, se activarán, como “ma” en “cama”, “sa” en “casa”, etc. pues tanto “cama” como “casa” se reconocen por su alta frecuencia en la lengua.

¿Cómo relacionamos estos conceptos de “prototipo” y “activación asociativa” a nivel cognitivo con el aprendizaje de segundas lenguas? Esta tesis se fundamenta en estas hipótesis científicas para presentar el diseño de un estudio experimental que busca comparar la adquisición de sonidos de una segunda lengua (L2) cuando se hace en distintos contextos, a saber, diferencia de ruido de fondo, diferencia del contexto lingüístico de presentación de las unidades (sílabas vs. oración).

El método de experimentación se nutre del entrenamiento perceptivo intensivo, que consiste en la presentación continuada, al azar, pero programada, de unidades lingüísticas en diferentes sesiones. La tesis ha sido posible gracias a la observación de diferentes grupos de aprendices, participantes voluntarios en el experimento, que siguieron una serie de sesiones de entrenamiento con estímulos (sílabas y palabras inglesas) presentados de manera diferente. Los grupos experimentales siguieron un entrenamiento en el que los estímulos se presentaban con un alto nivel de ruido de fondo, o bien, los mismos aparecían presentados con niveles de ruido que si bien afectaban a la señal, este no impedía la percepción de los estímulos.

Se realizó una experimentación basada en el entrenamiento auditivo intensivo para demostrar que este es necesario, no sólo en el campo de la enseñanza y aprendizaje de un L2, sino también en el campo de la rehabilitación auditiva. Algunos autores como Lively et al. (1993), Lively et al. (1994), Bradlow et al. (1997) o Bradlow et al. (1999) demostraron que el entrenamiento de estímulos que por sus características acústicas se consideran inherentemente difíciles es necesario para que se produzca una mejora de la percepción e identificación de estos estímulos, y, además, si el entrenamiento auditivo es eficaz, los resultados pueden prolongarse en el tiempo, es decir, los oyentes pueden ser capaces de identificar estos estímulos después de un tiempo sin entrenamiento. En esto coinciden investigadores Anderson y Kraus (2013), Burke y Humes (2007, 2008) o Schuman et al. (2015) centrados en el campo de la rehabilitación auditiva. Estos también encontraron que el entrenamiento es necesario para una mejora de la percepción de los sonidos. A diferencia de los primeros autores mencionados, estos últimos utilizaron el ruido de fondo durante el entrenamiento para que se asemejar más a las condiciones de vida real.

En uno de los estudios que presentamos en esta tesis hemos aplicado SNR (Relación Señal a Ruido – *Signal to Noise Ratio*) a la señal para demostrar la necesidad del mismo en experimentación de tipo acústica, ya que consideramos que al enmascarar la señal que reciben los oyentes podemos crear un contexto más cercano al que se puede dar en la vida real fuera de las condiciones óptimas que podemos encontrar en un laboratorio. Este experimento que demuestra la necesidad de aplicar SNR a la señal en estudios acústicos, el cual presentamos como Experimento I, hemos tomado como antecedentes estudios de la percepción de fonemas en situaciones de degradación, como por ejemplo

Nabelek y Donhaue (1984) o Cutler et al. (2004, 2007, 2008), hablan sobre el ruido de enmascaramiento o *masking noise* en percepción y cómo afecta a la percepción de los hablantes no nativos de inglés. Por lo tanto, teniendo en cuenta estos estudios previos, nuestra investigación confirmará o rechazará los resultados existentes hasta ahora. De manera indirecta, esta tesis también muestra resultados interesantes acerca de cómo la degradación de la señal por ruido de fondo consigue mayores o menores niveles de acierto, y, por consiguiente, de cuáles son los niveles más óptimos para que los sonidos en una L2 puedan alcanzar mejores cotas de categorización por aprendices no nativos.

Por tanto, los resultados que este estudio ofrezca darán nuevas perspectivas en la adquisición de segundas lenguas ya que nuestra investigación identificará y clasificará categorías consonánticas y descubrirá características inherentes de los sonidos ingleses. Esto demostrará que la adquisición de los fonemas ingleses no se basa sólo en la búsqueda de sonidos similares en la L1, sino que también hay características universales que afectan a la percepción de estos sonidos. Ese es un objetivo fundamental en nuestra investigación pues hemos intentado centrarnos en universales lingüísticos para comprender la percepción de las consonantes inglesas.

Además, todos estos descubrimientos podrían afectar al concepto del inglés como lengua franca (*English as Lingua Franca* – ELF en adelante) o del inglés como lengua internacional (*English as International Language* – EIL en adelante) y estos hallazgos permitirán también el diseño de nuevas estrategias metodológicas para enseñar inglés en contextos internacionales.

Para exponer los objetivos y contenidos detallados anteriormente, esta tesis doctoral se organiza de la siguiente manera: el capítulo uno presenta los antecedentes, es decir, una revisión de los modelos de percepción del habla y de la noción de prototipo: Modelo Imán de la Lengua Materna, Modelo de Aprendizaje del Habla y Modelo de Asimilación Perceptiva, se presenta también en este primer capítulo una revisión de los antecedentes para la percepción de las consonantes inglesas por hablantes no nativos. En el segundo capítulo se presentan nuestras hipótesis y objetivos. En el tercer capítulo se presentan y desarrollan los materiales que hemos utilizado en ambos experimentos y los métodos que hemos seguido para la obtención de los resultados. En el cuarto capítulo presentamos el análisis de los datos de los dos experimentos. El quinto capítulo

abarca la discusión de los datos, y por último, en el sexto capítulo se ofrecen nuestras conclusiones.

Nuestra población del Experimento I la componen 33 hablantes no nativos de inglés con diferentes L1 que han realizado una tarea de percepción basada en la identificación de todas las consonantes inglesas tanto en posición silábica inicial como final.

Para el análisis de datos se han agrupado las consonantes en tres grupos categoriales atendiendo al modo de articulación, lugar de articulación y sonoridad. Al no disponer de hablantes suficientes de cada una de las L1, no hemos considerado ésta como una variable en el análisis. Sin embargo, de los resultados obtenidos, tampoco podríamos deducir lógicamente que los mismos tengan una justificación debida a la L1 del participante en la prueba. Es por este motivo, que entendemos que este estudio se aleja de otros similares en tanto que la bidireccionalidad L1 – L2 no es evidente.

Nuestra población del Experimento II la componen 15 oyentes no nativos españoles que han realizado una tarea de percepción basada en la identificación de todas las consonantes inglesas en posición inicial y final y otra tarea de percepción basada en la identificación de estímulos en oraciones. En este segundo experimento se realizaron dos tareas de identificación para averiguar diferencias en la identificación de estímulos de forma aislada (sílabas) o en un contexto más amplio (oraciones).

Para el análisis de datos se analizaron los resultados de aquellos estímulos que en el Experimento I habían obtenido peores resultados en la identificación para así demostrar si un entrenamiento intensivo es productivo y necesario en los estímulos que producen errores en la identificación por parte de los oyentes.

No obstante, no hemos podido demostrar si los resultados obtenidos fruto de nuestro entrenamiento intensivo auditivo han permanecido a lo largo del tiempo, por lo que hemos dejado la puerta abierta a futuras investigaciones en este aspecto.

Antecedentes

1.1. Percepción del habla: indicios acústicos y el concepto de prototipo

Los estudios centrados en la percepción del habla han establecido que la identificación de los sonidos por parte del oyente parte de la teoría de la Ciencia Cognitiva que distingue entre indicios acústicos y prototipos.

La gran variabilidad en la señal sonora nos hace reflexionar sobre cómo el ser humano es capaz de identificar sonidos únicos entre variaciones acústicas tan grandes, debidas a cambio de hablante, cambio de situación, registro, posición del sonido en la cadena hablada, entre otras tantas. La decodificación de la onda sonora para la identificación de las unidades lingüísticas que la contienen se centra en el contexto en el que se produce la emisión de la información y la información acústica presente en la señal: nuestra mente identifica el contexto (registro, volumen, hablante, etc.) e inmediatamente relativiza la información acústica presente en el sonido.

Una vez contextualizada la señal, el oyente procede a extraer los indicios acústicos presentes en la misma. Estos indicios nos permiten agrupar los sonidos del lenguaje en diferentes categorías, lo que en Lingüística se conoce como categorías fonéticas, fonemas, etc. Indicios, por tanto, equivaldría a “asociación cognitiva”, mientras que categorías fonéticas puede entenderse como “prototipos cognitivos”. Así, el disparo de los indicios acústicos en la percepción encamina la activación de patrones cognitivos prototípicos que ayudan al oyente a identificar la señal sonora que recibe.

Por ejemplo, uno de los indicios más estudiados es el “voice onset time” (VOT en adelante). VOT es un indicio esencial que marca la diferencia entre las consonantes oclusivas sonoras y sordas, por ejemplo, /p/ y /b/. Otros indicios acústicos diferencian sonidos que se producen en diferentes lugares y modos de articulación, como la coarticulación o información acústica producida justo en la frontera entre sonidos, o la información frecuencial, como la que nos ayuda a interpretar la entonación en los enunciados.

Los indicios acústicos que un oyente normal procesa son muy numerosos. La variable cuantitativa se debe fundamentalmente a aspectos relacionados con la producción (diferentes hablantes), el contexto lingüístico (los sonidos vecinos) y el contexto acústico (efectos como la reverberación o resonancia afectan la variabilidad de los indicios acústicos). Esta ingente variabilidad no parece ser un problema, como decíamos anteriormente, en la percepción de los sonidos de la lengua materna: el niño, de hecho, adquiere de manera progresiva el lenguaje y lo maneja con gran soltura en la etapa de la primera madurez, en torno a los 6 años. Esta segmentación de los sonidos se pone de manifiesto en la creación de prototipos los cuales son modelos mentales que corresponden a patrones de nuevas señales tal y como expone el Modelo Imán de la Lengua Materna planteado y desarrollado por Kuhl (1992, 1993a, 1993b, 1994, 2000, 2004) y Kuhl e Iverson (1995). Los prototipos son los casos más representativos de una categoría fonética que comparte un máximo número de semejanzas con los miembros de la categoría fonética, pero a su vez, presenta también el máximo número de diferencias respecto a los componentes de otras categorías fonéticas. Cuanto mejor son los prototipos de una categoría fonética, mejor los identifica el oyente. En este punto, debemos considerar el concepto de “imán perceptivo”. El prototipo podría considerarse como el resultado del cálculo mental de las propiedades acústicas medias de los sonidos percibidos. De esta forma el niño se ve expuesto a multitud de estímulos sonoros. Mientras los va recibiendo, su cerebro va construyendo una base de datos que, llegado a un punto de madurez, se resumirá en lo que en estadística podríamos llamar las medias y la desviación estándar. Este es el concepto que defienden los autores del modelo de imán para explicar la construcción de patrones cognitivos. Una vez el niño se ha creado una imagen estable que permite identificar un estímulo externo, el mismo se mantiene de manera indefinida en su memoria asociativa y le servirá para seguir identificando estímulos sonoros que se sitúen en las zonas de contacto “magnético”, según

terminología de esta teoría, del prototipo. El centro del prototipo es lo que llamamos “categoría fónica”.

Hay otras diferentes teorías relacionadas con la percepción del habla que se centran en explicar cómo los oyentes de una L2 procesan los sonidos del habla. Usar el aprendizaje de una L2 ha ayudado a los investigadores a comprender la percepción de la L1, puesto que estos modelos teóricos siempre se basan en experimentación de naturaleza psicoacústica que normalmente utiliza métodos basados en la presentación de estímulos a oyentes que devuelven sus respuestas. Esta metodología es más potente cuando se usan oyentes colaborativos por lo que la mayor parte de las investigaciones realizadas previas a la definición de modelos teóricos se han hecho con oyentes adultos que estudian una L2.

Aparte del Modelo del Imán Perceptivo de la Lengua Materna (Kuhl 1992, 1993a, 1993b, 1994, 2000, 2004 y Kuhl e Iverson 1995), otros modelos que deberíamos mencionar son el Modelo de Asimilación Perceptiva (Best, 1993, 1994a, 1994b, 1995) y el Modelo de Aprendizaje del Habla (Flege 1995, 1999, 2002).

Estos tres modelos comparten la idea de que el sistema fonológico de la L1 del oyente influye en la percepción de los sonidos de la L2. En estos tres modelos de percepción se considera que tanto los sonidos de la L1 como los sonidos de la L2 comparten un espacio fonológico común. Es decir, el patrón, prototipo o categoría fónica adquirida en la L1 determinar cómo se procesan los sonidos de la L2. Debido a esto, la similitud fonética de la L1 con la L2 tiene un papel fundamental sobre cómo se categorizan los segmentos de la L2 en el espacio fonético. Además, estas teorías comparten la creencia de que el contacto continuo con los estímulos, o experiencia, determina ineludiblemente la creación de nuevas categorías.

1.2. El Modelo Imán de la Lengua Materna (NLM)

El modelo NLM (*Native Language Magnet*) desarrollado por Kuhl (1992, 1993a, 1993b, 1994, 2000, 2004) y Kuhl e Iverson (1995) establece que los niños reciben de los adultos y del entorno que les rodea la información fonológica de su L1. Algunos autores como Fernald y Simon (1984) determinaron que la vocalización rítmica y continua de los adultos a los bebés determina la adquisición fonológica. Esto significa que es la repetición constante y aislada de unidades lo que facilita la adquisición de patrones cognitivos de naturaleza estadística. Los niños usan estrategias computacionales para detectar patrones estadísticos y entonativos que les ayudarán a aprender sonidos y palabras. La capacidad neuronal en edades tempranas para activar estos procesos determina el éxito de la adquisición en niños vs. adultos (Kuhl, 2004).

La teoría del NLM surgió principalmente como un modelo para explicar el desarrollo de la percepción auditiva y de la adquisición de los sonidos de la L1 en el niño. No obstante, los conceptos establecidos por ella como el concepto de prototipo y el de efecto imán se han extrapolado y aplicado al estudio del aprendizaje de los sonidos de una L2 por parte de hablantes adultos.

Estos patrones jugarán un papel determinante e imprescindible en la percepción de sonidos de la L1 pero también en la L2. El modelo NLM se basa en la idea de que el niño está dotado genéticamente e inherentemente de la capacidad establecer las diferencias acústicas entre sonidos que pertenecen a diferentes categorías fonéticas. Esta capacidad se hace incluso más notable en la percepción de aquellos sonidos que se encuentran en la frontera entre dos categorías. En el caso de la adquisición de una L2 tendríamos que admitir que el aprendiz ha de reorganizar el espacio perceptivo según las categorías desarrolladas para los sonidos de la L1 y la información fonética que recibe de la L2.

La adquisición de los prototipos de la L1 produce que se disminuyan las fronteras cercanas al prototipo y al mismo tiempo se reconfigura el espacio auditivo para incorporar un imán nuevo o más específico (como en el contraste de las vocales del

español y el inglés). Esto produce que se disminuyan ciertas fronteras (las que están cercanas al prototipo), al mismo tiempo que se reconfigura el espacio auditivo para incorporar un determinado imán. Los estudios en sujetos adultos demuestran que las fronteras cuestan de desaparecer, quizás porque la plasticidad neuronal va disminuyendo con el paso del tiempo. Para afianzar la creación necesaria de nuevos prototipos según se incorporan nuevas lenguas se ha probado posible a partir de un entrenamiento adecuado. El concepto de entrenamiento intensivo parte de la idea de que es la exposición continuada a estímulos lo que permite la reconfiguración del espacio cognitivo.

En el adulto, la dificultad para aprender un nuevo contraste fonético viene determinada en gran medida por la proximidad entre el contraste y un imán de la L1 ya que este imán distorsiona el espacio auditivo subyacente lo que conlleva la atracción de sonidos similares. La teoría del NLM establece que un sonido de la L2 será más difícil de discriminar y, por tanto, de adquirir cuanto más próximo esté a un imán de la L1. Esta teoría se denomina “principio de proximidad” (Kuhl e Iverson, 1995; Best, 1993).

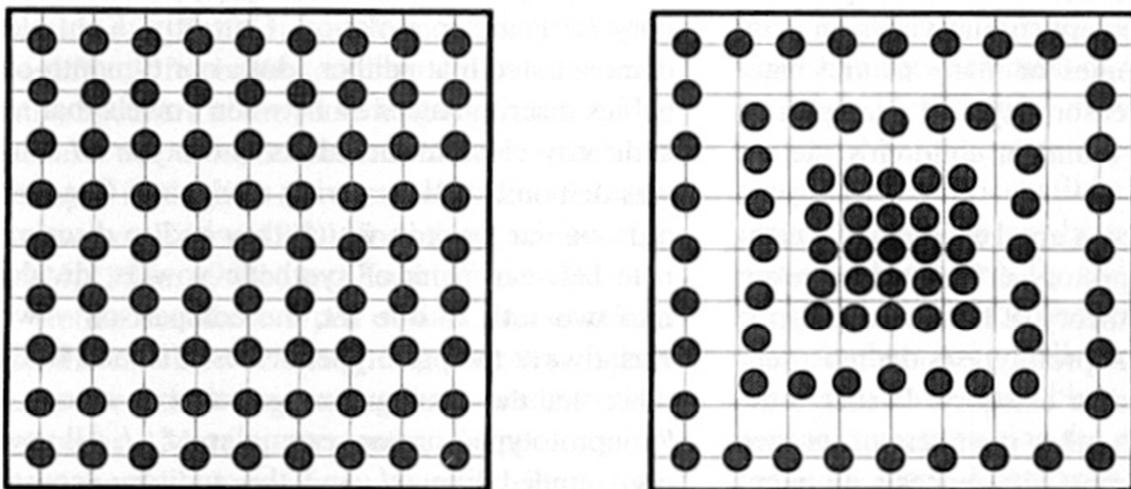


Figura 1. Noción de imán perceptivo

La figura 1 representa a la izquierda el patrón cognitivo sin activar mientras que a la derecha se representa el patrón cognitivo activado. La zona central en la imagen de la derecha corresponde a la zona de acción magnética en la que el prototipo atrae los sonidos similares produciendo un efecto de “imán”.

1.3. Modelo de Aprendizaje del Habla (SLM) y Modelo de Asimilación Perceptiva (PAM)

1.3.1. Modelo de Aprendizaje del Habla (SLM): la importancia de la transferencia de la lengua materna en la adquisición fonológica de lenguas extranjeras.

Flege (1995) desarrolló su modelo SLM (*Speech Learning Model*) fundamentalmente a partir de la observación del aprendizaje del inglés por no nativos. Sus estudios sirvieron de modelo a multitud de investigadores y ha tenido gran repercusión en modelos y métodos centrados en la producción y pronunciación. De hecho, la perspectiva de estas investigaciones estuvo siempre centrada en aspectos más productivos que perceptivos.

Flege (1981) descubrió que cuando las habilidades sensomotoras de los niños se están desarrollando, parece que pierden la habilidad para el aprendizaje de vocales y consonantes de una segunda lengua. Las investigaciones que se llevaron a cabo a nivel segmental durante los años 70 se centraron en el aprendizaje en el entorno del aula de idiomas o en etapas tempranas en la adquisición de la L2. Se consideraba la interferencia producida por la L1 en la L2 como la causa fonológica principal de la existencia de acentos extranjeros. Un punto de vista común era, que por un lado, un sonido de la L2 es identificado como un sonido de la L1 y automáticamente el sonido de la L2 es reemplazado por el sonido de la L1. Por otro lado, los sonidos de los contrastes de la L2 no se tienen en cuenta cuando estos no existen en la L1. Por último, otro punto de vista común era que ciertos sonidos contrastivos de la L1 se transfieren a la L2 pero éstos no existen en la L1.

Según Flege (1995) el punto de vista que afirma que la existencia de acentos extranjeros se produce por una dificultad motora está en consonancia con observaciones hechas por diferentes autores. Por ejemplo, Neufeld (1979) demostró que un grupo de sujetos adultos produjeron frases en un idioma diferente de su L1 sin tener acento extranjero, lo que demuestra que los adultos son capaces de imitar sonidos extranjeros de la misma forma que lo hacen los niños. Por otro lado, Snow y Hoefnagel-Höhle (1979)

descubrieron que tras un proceso de aprendizaje e imitación, la producción espontánea de sonidos en adultos y niños era similar.

Esto plantea por qué los adultos no son capaces de explotar las habilidades motoras completamente. Flege (1995) afirma que esto podría deberse a que los adultos no perciben de manera adecuada los sonidos de la L2. Normalmente, los hablantes no nativos necesitan una mayor porción de la palabra para poder identificarla, aunque esto no ocurre con los hablantes nativos. Además, las diferencias perceptuales se perciben cuando los hablantes no nativos tienen que enfrentarse a frases anómalas o versiones de oraciones en la L2 en las que el habla está distorsionada o es sintética.

Sin embargo, Flege también afirma que la relación entre los sonidos de la L1 y la L2 pueden cambiar en un entorno de aprendizaje natural debido a la modificación que puede producirse en los patrones auditivos que son específicos de un idioma. Como consecuencia de esto la percepción de sonidos en la L2 no sería constante.

El SLM fue un modelo creado a partir de estudios hechos con hablantes bilingües experimentados. Entre sus posicionamientos se afirma que existen unos contrastes fonológicos que perdurarán toda la vida y que claramente están basados en la L1. Según este modelo, se pueden cometer errores en la identificación de las diferencias fónicas entre los pares de sonidos de la L2 o entre la L2 y la L1 ya que los sonidos se asimilan a una única categoría de la L1. Esto se produce porque la L1 filtra las características de los sonidos de la L2 que son únicamente importantes a nivel fonético pero no a nivel fonológico. A partir de esta hipótesis Flege (1995) establece que (1) se deben aplicar los procesos de aprendizaje de la L1 en el proceso de aprendizaje de la L2 porque los procesos de la L1 no varían a lo largo del tiempo; (2) las categorías fonéticas están establecidas en las representaciones de la memoria a largo plazo ya que son los factores específicos de un idioma; (3) es más probable que los niños creen nuevas categorías que los adultos aunque son los adultos los que son capaces de conservar la capacidad de creación de nuevas categorías. Además, las categorías creadas en la infancia, evolucionan con el paso del tiempo para mostrar las propiedades de todos los fonemas de la L1 y L2 los cuales se asocian a la realización de las categorías fonéticas; (4) los

hablantes bilingües intentan mantener los contrastes entre la L1 y la L2 de aquellos sonidos que comparten un espacio fonológico común.

Además, establece una serie de hipótesis (H) básicas en la adquisición de una L2 en las que se basa el modelo SLM:

- H1: los sonidos de la L1 y la L2 están relacionados de manera auditiva en un determinado nivel contextual. Es decir, dos sonidos de la L2 se simplifican en un sonido de la L1 pero difieren en su percepción dependiendo de la posición, ya sea inicial o final. Por ejemplo, se demostró que se favorece la percepción de las consonantes líquidas en posición final.
- H2: Se pueden crear nuevas categorías para sonidos de la L2 que se diferencian fonéticamente de los sonidos de la L1.
- H3: Cuanto más se diferencia un sonido de la L2 de un sonido de la L1, mejor se adquiere el sonido de la L2.
- H4: La cuarta hipótesis está relacionada con la edad de adquisición del aprendiz. Cuanto más mayor es el aprendiz, es menos probable que se produzcan distinciones fonéticas entre los sonidos de la L1 y de la L2 y entre los sonidos de la L2 que no tienen un contraste directo en la L1.
- H5: Cuando no existe una creación de categorías debido a que el mecanismo encargado de clasificar las equivalencias impide la formación de las mismas, es probable que los sonidos de la L2 se asimilen a una categoría existente de la L1, creando la fusión de una categoría L1-L2 y, por tanto, la producción de estos sonidos es similar.
- H6: Es posible que las categorías se diferencien para mantener un contraste entre ellas. En base a esto, las categorías fonéticas creadas para los sonidos de la L2 en un hablante bilingüe pueden ser diferentes de las categorías creadas por un hablante bilingüe.

1.3.2. Modelo de Asimilación Perceptiva (PAM): la convergencia entre la “Teoría Motriz de la percepción del habla” y el modelo auditivo (modelo psicoacústico)

Best (1993, 1994a, 1994b, 1995) se basa en el modelo psicoacústico, en la teoría de la percepción auditiva y en la teoría motriz de la percepción del habla para establecer y desarrollar las bases de la percepción del habla.

La percepción del habla se produce de manera diferente en las tres teorías mencionadas anteriormente. El modelo psicoacústico establece que el habla se percibe mediante un estímulo próximo a la periferia auditiva (sistema auditivo periférico, aparato auditivo). La teoría motriz afirma que la percepción del habla se basa en un módulo innato neurológico que sintetiza los gestos articulatorios. Por último, según la teoría de la percepción del habla se basa en los movimientos reales del tracto vocal, o gestos, y los fonemas no abstractos o gestos previstos. Los oyentes perciben estos gestos a través de la información de la señal acústica específica de los gestos que lo forman.

Respecto a la transformación de la percepción primitiva en elementos fonológicos y fonéticos, la teoría motriz y la teoría de la percepción del habla coinciden en afirmar que los niños aprenden por sí mismos qué conjunto de gestos articulatorios se utilizan en su L1. No obstante, el modelo psicoacústico difiere de esta afirmación y establece que los niños aprenden a asociar combinaciones de características que inherentemente no tienen significado y que no son lingüísticas con unidades como oraciones y morfemas, y con estructuras fonéticas como sílabas y segmentos. Por consiguiente, los niños construyen y memorizan indicaciones auditivas asociadas a unidades lingüísticas abstractas a través de la experiencia. Sin embargo, la teoría motora y la teoría del realismo directo asumen que los niños tienen que aprender por sí mismos qué grupos de gestos articulatorios se utilizan en su lengua nativa.

Teniendo en cuenta las teorías expuestas anteriormente, Best (1995) establece que es fundamental tener en cuenta las características fonéticas específicas de la L1. Para ello es necesario entender cómo el niño reconoce los principios que rigen el sistema de sonidos en su L1. Al principio, el niño sólo tiene acceso a los detalles fonéticos superficiales de su L1 y debe percibir la relación entre estos detalles y las estructuras

fonológicas que son más abstractas. Según Best (1995), el ser humano nace sabiendo todas las estructuras fonológicas ya que estas son innatas pero por falta de uso se olvidan y sólo perviven las que se mantienen a través de la experiencia. Por ello, es importante que los niños detecten a una edad temprana un amplio rango de posibles gestos articulatorios en el habla para poder descubrir patrones y características específicas que se emplean en el idioma y que utilizan características fonéticas para realizar funciones fonológicas.

Por tanto, cuando los hablantes se adaptan a la L1 se produce un descubrimiento de los niveles más complejos que están unidos a los gestos que componen el inventario fonológico. Esto significa que a la vez que los hablantes se ajustan a esta información específica del idioma, se vuelven más eficientes para distinguir las propiedades de esos gestos. Como consecuencia de esto, los hablantes reúnen información que es más específica y compacta, lo cual es muy importante para comprender cómo la experiencia que se tienen en la L1 afecta a la percepción de la L2. Esto lleva a afirmar que es fundamental una mayor exposición a las lenguas para que el oyente reconozca un mayor número de gestos que están relacionados con los elementos fonológicos que son específicos de una lengua, como las sílabas y sus constituyentes y las unidades rítmicas.

Aunque haya diferencias en los grupos gestuales de diferentes lenguas, no debería afirmarse que estas lenguas tienen poco o nada en común ya que están compuestas de todas las posibles realizaciones gestuales del tracto vocal. Además, normalmente hay una gran cantidad de información que coexiste al mismo tiempo en los gestos y características que se encuentran en sus espacios fonológicos individuales, especialmente a nivel segmental. Tradicionalmente se ha considerado que existe un gran número de segmentos compartidos por diferentes lenguas y, en ocasiones, estos segmentos son iguales o casi iguales. Estas similitudes han generado un gran interés por parte de los lingüistas que se interesan por las propiedades universales del lenguaje. Sin embargo, éstos no han establecido cómo se perciben los segmentos de una L2. Ésto estaría intrínsecamente relacionado con nuestra hipótesis ya que creemos que hay ciertos segmentos fonéticos del inglés que son universales y se perciben de la misma forma por diferentes hablantes con distintas L1.

La afirmación más importante que establece el modelo PAM es que los oyentes tienden a asimilar sonidos de la L2 que son parecidos a sonidos de la L1 o a categorías de fonemas de la L1, relacionándoles con la categoría a la cual fonéticamente son más parecidos. Los sonidos que se encuentran entre categorías de la L1 son sonidos “no categorizados”. Por su parte, los sonidos que se asimilan como sonidos no específicos del lenguaje, como por ejemplo el sonido que produce un grito de socorro, no se asimilan dentro del espacio fonológico.

El modelo PAM establece una serie de predicciones relacionados con los contrastes de la L2: (1) la discriminación de los sonidos es óptima cuando cada fonema de la se asimila a categorías diferentes de la L1 (Asimilación a dos categorías - *Two-Category Assimilation* (TC)); (2) La discriminación varía entre moderada y buena cuando dos sonidos de la L2 se asimilan a la misma categoría de la L1, aunque difieran en el grado del cual en el que se desvían del ejemplo nativo ideal (Diferencia de la bondad de la categoría – *Category Goodness Difference* (CG)); (3) la discriminación es pobre cuando dos sonidos de la L2 se asimilan a la misma categoría en la L1, pero se alejan del ideal nativo (Asimilación a una sola categoría - *Single Category Assimilation* (SC)); (4) la discriminación varía de pobre a buena cuando los contrastes de la L2 se encuentran dentro del mismo espacio fonético pero fuera de las categorías nativas. En este caso, el éxito en la discriminación depende de la similitud de los sonidos de la L2 respecto a las categorías de la L1 (Ambos categorizables - *Both Uncategorizable* (UU)); (5) la discriminación debería ser buena cuando un sonido un par L2 se asimila a una categoría de la L1, pero el otro sonido del mismo par no puede asimilarse a ninguna categoría de la L1 (No categorizado versus categorizado - *Uncategorized versus Categorized* (UC)); (6) por último, la discriminación debería ser buena o muy buena cuando ambos sonidos de la L2 son percibidos como sonidos no lingüísticos por lo que recaen fuera de la esfera del habla y no puede asimilarse a ningún sonido de la L1 (No asimilable - *Non-Assimilable* (NA)).

Hay una serie de experimentos (Best et al. 2001, Best et al. 2003) que confirman las predicciones del modelo PAM. Descubrieron que 1) la experiencia del habla nativa tiene como resultado una mejor percepción de las muestras de habla no nativas que son más similares a las muestras nativas; 2) los oyentes detectan la presencia o ausencia de contrastes fonológicos y automáticamente detectan fonológicamente detalles fonéticos

irrelevantes que pueden ser lingüísticos o no lingüísticos, y, 3) la discriminación se ve afectada por el efecto memoria del contraste SC.

Por otro lado, hay un gran número de estudios que apoyan la teoría de que el relativo grado de dificultad para adquirir los sonidos de la L2 está influido por la L1 del oyente (Best et al. 2003, Best y Strange 1992, Flege, 1981, Flege 1989, Hallé et al. 1999). No obstante, Best et al. (2003) también afirman que cuanto más diferente es el contraste L2 de la L1, más fácil es su discriminación lo cual está en consonancia con la tercera hipótesis planteada en el modelo SLM.

Posteriormente, Best y Tyler (2007) encontraron nuevas implicaciones del modelo PAM. La principal cuestión que plantean es cómo los hallazgos del habla no nativa afectan al aprendizaje auditivo del aspecto fonético y fonológico de la L2. Afirman que el aprendizaje perceptual se regula por principios no nativos del habla. Por lo tanto, toman en cuenta las similitudes entre oyentes experimentados y no experimentados en la L2 contrastando los modelos PAM y SLM. Examinan cómo el aprendizaje de un idioma afecta a la percepción de la información fonética en contra de la información fonológica, el impacto de la percepción monolingüe en contra de la experiencia en diferentes lenguas, y, finalmente, examinan las consecuencias que estos hallazgos podrían tener en la percepción del habla y los cambios que podrían tener en el entorno del lenguaje del oyente.

Sabemos que la percepción de la información fonética depende de la experiencia lingüística del oyente y su historia del desarrollo. La percepción es diferente entre los oyentes expertos e inexpertos porque las influencias experimentales son también diferentes dependiendo de la edad de exposición a la adquisición de la L2. Por lo tanto, argumentan que los oyentes de una L2 o bilingües perciben el habla de la L2 o L1 de la misma forma en la que lo hacen los hablantes monolingües.

El entorno tiene un papel fundamental en la adquisición del lenguaje durante las diferentes edades. La cantidad de exposición y las propiedades fonéticas del input del idioma son importantes en la interacción con el nivel de desarrollo de los oyentes y el estado de la L2.

Las analogías y diferencias fonéticas de los fonemas nativos y no nativos influyen en la percepción no nativa del habla. Además, esta influencia no se limita únicamente a las potenciales peculiaridades fonológicas. La percepción de los contrastes fonéticos no nativos por parte de adultos monolingües se ve afectada por la alteración de ciertas características fonéticas de los hablantes y/o variaciones en las muestras (*tokens*).

La cuestión principal que plantean es cómo podría tener lugar el aprendizaje de las propiedades fonéticas y fonológicas de una L2. En este caso es necesario analizar los entornos o situaciones naturales de comunicación. Estos entornos o situaciones incluyen amplias estructuras gramáticas y fonológicas de la L2. Esta idea resulta más compatible con las teorías que plantea la adquisición de segundas lenguas (*Second Language Acquisition* – SLA en adelante). La teoría central de SLA plantea que una conversación con significado es el contexto más importante en el cual se aprenden las propiedades de una nueva lengua.

La diferencia entre una L2 en un contexto natural de comunicación vs. contextos más limitados emerge para comparar las teorías que plantean la adquisición de segundas lenguas (SLA) y la adquisición de lenguas extranjeras (*Foreign Language Acquisition* – FLA en adelante). En muchos aspectos, la teoría FLA es peor si la comparamos con el contexto natural del aprendizaje del lenguaje, ya que la teoría FLA tiene lugar en un contexto influido por la L1 y no se utiliza demasiado fuera del entorno del aula. Además, normalmente la enseñan profesores los cuales tienen un acento que procede de su L1 y normalmente utilizan una instrucción basada en la enseñanza de la gramática y del vocabulario. Por lo tanto, la teoría FLA abarca un contexto poco apropiado para el aprendizaje de la L2 y los hallazgos perceptuales de los oyentes FLA no deberían mezclarse con los hallazgos encontrados para los oyentes involucrados en el SLA. Por otro lado, los hallazgos del FLA pueden ser útiles ya que los aspectos que abarca la teoría FLA podrían utilizarse en el futuro porque el FLA ofrece una base importante para comparar oyentes de ambas teorías, FLA y SLA.

Best y Tyler (2007) afirman que los oyentes SLA muestran un aprendizaje perceptual de los contrastes de la L2 que al principio pueden resultar difíciles de diferenciar. Sugieren que el aprendizaje perceptual tiene lugar cuando ciertos contrastes de la L2 ocurren, pero el aprendizaje perceptual depende de los contrastes fonológicos de la L1 y de la

relación fonética con la L1. Además, la similitud con la L2 afecta el aprendizaje perceptual. Best y Tyler (2007) demuestran esto comparando aprendices expertos e inexpertos. Los resultados demuestran que los aprendices más experimentados son capaces de una mejor discriminación y categorización de los contrastes no nativos que los aprendices más inexpertos.

So y Best (2010, 2011) afirman que la experiencia lingüística no solo tiene efectos sobre las características segmentales como las vocales y las consonantes, sino que también afecta a las características suprasegmentales, como por ejemplo, los patrones de acentuación o los patrones prosódicos a nivel de frase. Los aspectos fonéticos y fonológicos de la L1 influyen en estas características suprasegmentales. Las características prosódicas de la lengua nativa de los adultos tienen un profundo efecto en su percepción de los contrastes suprasegmentales de las lenguas no nativas. Sin embargo, la experiencia lingüística también afecta a la percepción de los tonos léxicos no nativos. En algunas ocasiones se ha demostrado que la experiencia de la lengua nativa de los oyentes influye sobre la percepción de los tonos no nativos. Se cree que el papel de la experiencia lingüística es fundamental en el aprendizaje de la L2 ya que, por ejemplo, los oyentes con una lengua tonal son capaces de analizar una mejor categorización que aquellos oyentes que no tienen una lengua tonal.

1.4. La percepción de las consonantes inglesas por no nativos: antecedentes experimentales

Como antecedentes directos a la investigación realizada en esta tesis doctoral debemos considerar aquellos realizados sobre la percepción de las consonantes inglesas por no nativos adultos. En este apartado se incluye un resumen de los que se han considerado para la elaboración de los experimentos que la tesis detalla y que se han utilizado para analizar los resultados.

Los estudios realizados en el campo de la percepción del habla se agrupan, generalmente, en los que adoptan un doble enfoque percepción-producción y los que no consideran el aspecto productivo. En el campo específico de la percepción de una L2, la mayor parte de la investigación actual gira en torno al modelo PAM con adaptaciones.

En cuanto a los objetivos de estudio, la adquisición de categorías perceptivas (prototipo cognitivo), la influencia de situaciones adversas (ruido de fondo, resonancia, etc.) y la influencia del entrenamiento auditivo intensivo pueden considerarse como los que han recibido mayor atención.

Este apartado considera, por tanto, estudios que se centran en los siguientes aspectos de la investigación que pueden a la vez servir como antecedentes científicos, a saber:

- Importancia de la transferencia de la L1 hacia la L2.
- Efecto del ruido de fondo sobre la percepción del habla.
- Efecto de la posición silábica en la identificación del sonido.
- Efecto del lugar y modo de articulación de las consonantes en su percepción.
- El entrenamiento auditivo y su efecto en la percepción de sonidos.

Los primeros estudios en el campo de la percepción que utilizaron métodos psicoacústicos se remontan a los años 90. En este período encontramos la formulación de la teoría SLM, que de manera resumida intenta siempre demostrar la importancia que tiene la L1 sobre la creación de categorías fonológicas en una L2 (Skutnabb-Kangas y Toukoma 1976, Cummings et al. 1984, Linde y Lofgren 1988, Flege et al. 1997, Cummins 2000, Piske et al. 2001). Esta importancia se define de manera experimental en la definición del modelo de transferencia fonológica de la L1 a la L2. La hipótesis es, por tanto, siempre la misma: la L1 determina la manera como se adquieren los sonidos en una L2 y los objetivos en el estudio son encontrar las maneras cómo se produce o no dicha transferencia. Dentro de este grupo de estudios, algunos autores eligen para su experimentación sonidos que no existen en la L1 (Best et al. 1988, Best 1995, Best et al. 2001, Best y Tyler 2007, Bradlow et al. 1997, Lively et al. 1993, Logan et al. 1991, Pilus 2005) o seleccionan un conjunto limitado de sonidos en posiciones silábicas diferentes (Lively et al. 1993, Logan et al. 1991, Recasens 2011, Redford y Diehl 1999). Más adelante comienzan a aparecer estudios más completos en los que se hace una revisión en la creación de categorías fonológicas de sistemas completos, como es el caso de Cutler et al. (2004) donde se define el grado de corrección en la identificación de todas las consonantes y vocales inglesas en posición inicial y final de sílaba por hablantes nativos del holandés que estudian inglés como L2.

Como técnica común en la recogida de datos se observa la manipulación de los estímulos auditivos a diferentes niveles de SNR. Presentar estímulos con diferentes grados de ruido de fondo obliga a los participantes en los experimentos a poner a prueba su auténtica capacidad de percepción. Esta técnica comenzó a usarse en los años 50 en estudios que pretendían explicar la percepción del lugar y punto de articulación de las consonantes una vez manipuladas con ruido (Ainsworth y Meyer 1994, Miller y Eimas 1996, Miller et al. 1951, Miller y Nicely 1955, Wang y Bilger 1973). En los últimos tiempos la variable ruido de fondo se ha extendido en los métodos pero también se ha analizado en sí misma para comprender mejor la reacción de los seres humanos en situaciones comunicativas adversas. Los estudios más importantes son los realizados por Cutler et al. (2007), Cutler et al. (2008), García Lecumberri y Cooke (2006) y García Lecumberri et al. (2010). Estos estudios confirman que el ruido de fondo actúa de manera negativa sobre la correcta creación de categorías fonológicas. En experimentación hecha con población clínica (pacientes que reciben implantes cocleares) autores como Hazrati y Loizou (2012) encuentran que factores adversos como la reverberación influyen más negativamente en la correcta identificación de sonidos que el ruido de fondo.

Diferentes lingüistas como Bell y Hooper (1978), Bent et al. (2008), Redford y Diehl (1999), Miller y Eimas (1976) y Jiang, Chen y Alwan, (2006) han investigado sobre la percepción de las consonantes atendiendo a su posición silábica, es decir, atendiendo a la posición de ataque o coda y también a su lugar y modo de articulación. Bell y Hooper (1978) afirmaron que los ataques predominan sobre las codas; las codas son generalmente un subgrupo de los ataques y los grupos consonánticos se ven desfavorecidos a través de las lenguas, pero es más probable que tengan lugar en posición inicial de sílaba que en posición final de sílaba. Estos hechos sugieren que los inventarios silábicos entre las lenguas favorecen los ataques con sílabas formadas por una sola consonante sobre aquellos que carecen de consonantes en el ataque. Por lo tanto, teniendo en cuenta esta afirmación, Redford y Diehl (1999) afirman que es más probable que los ataques predominen en las lenguas que permiten consonantes en posición final de sílaba. Algunos estudios previos y posteriores a Redford y Diehl (1999), como los de Bent et al. (2008), Helfer y Huntley (1991), Miller y Nicely (1955) o Wang y Bilger (1973), indicaron que en muchos casos no hay diferencias en la posición a la hora de identificar las consonantes, pero cuando se produce alguna

diferencia, se tiende a favorecer la percepción de las consonantes iniciales sobre las consonantes finales. Además, Bent et al. (2008) encontraron que cuando se produce un error en posición inicial la inteligibilidad de los sonidos es peor que cuando se produce en posición final.

Más en concreto, Redford y Diehl (1999) afirman que necesitaban encontrar pruebas más completas y que se desarrollasen en contextos naturales para poder corroborar las afirmaciones de los estudios mencionados anteriormente. Para ello, diseñaron un experimento de confusión perceptual para evaluar la identificación de las consonantes iniciales y finales en condiciones que podrían asimilarse a condiciones naturales de donde emergen los sistemas fonológicos. Se presentaron sílabas con la estructura CVC en un contexto natural y en el marco oracional. Estas oraciones ofrecieron un contexto consonántico y vocálico para las consonantes finales. Los segmentos utilizados fueron /p, t, k, f, θ, s, ʃ/ y /i, u, a/ para crear sílabas CVC. Emplearon estos segmentos porque existen en la mayoría de las lenguas, a excepción del segmento /θ/. Sus predicciones se basaban en la hipótesis que establece que la estructura silábica a través de las lenguas refleja diferencias en la percepción entre consonantes en posición inicial de sílaba y en posición final. También predijeron que la percepción de los ataques estaría favorecida sobre la percepción de las codas y que el efecto de esta posición tendría lugar a través de las consonantes, los núcleos vocálicos y los hablantes. Por lo tanto, estas hipótesis asumen que las diferencias en la percepción se asocian con la posición silábica y no con los fonemas, la producción específica del hablante y los contextos y condiciones específicos del hablante. Finalmente, también predijeron que los grupos consonánticos están en desventaja, aunque es más probable que tengan lugar en posición inicial de sílaba que en posición final de sílaba. Los autores intentaban probar tres hipótesis: 1) los ataques se perciben mejor que las codas; 2) existe una ventaja perceptual a través de las consonantes, los núcleos vocálicos y los hablantes, y, 3) las consonantes iniciales de sílaba que preceden a una vocal se identifican mejor que las consonantes finales de sílaba que preceden a una consonante. Los resultados que obtuvieron apoyaron las dos primeras hipótesis pero no la tercera.

Miller y Eimas (1996) realizaron un estudio en el que compararon la percepción del lugar y el modo de articulación de consonantes labio-alveolares y consonantes oclusivas nasales. Después de llevar a cabo tres experimentos, concluyeron que el lugar y el modo

de articulación conllevan un procesamiento similar. Llegaron a esta conclusión afirmando que ambas características hacen el mismo uso de ambos hemisferios cerebrales. Por otro lado, también encontraron que el lugar y el modo de articulación no se procesan de manera independiente ya que encontraron que, aunque, el modo y el lugar de articulación pueden responder a la información en diferentes contextos acústicos, hay una dependencia mutua entre ellos. No obstante, concluyen afirmando que se necesita más investigación para determinar el nivel en el cual sucede la interacción.

Por último, un aspecto importante por la influencia que tiene en el diseño metodológico en el campo de la percepción de lenguas es el efecto del entrenamiento auditivo intensivo en la percepción de sonidos concretos. Por “entrenamiento auditivo intensivo” me refiero a la exposición repetitiva de sonidos presentados de manera aleatoria que estimulan al participante en el experimento de manera que tiene que dar una respuesta de lo que acaba de oír: identificando el estímulo presentado o decidiendo si se trata de un sonido similar o diferente al anteriormente presentado. Este entrenamiento se ha usado para comprender la adquisición de sonidos complejos, como los estudios que comparaban la adquisición del par /l/ - /r/ por parte de hablantes japoneses realizados por Bradlow et al. (1997, 1999), Lively et al. (1993), Logan et al. (1991). Estas experimentaciones medían y analizaban diferentes sistemas de entrenamiento en oyentes nativos japoneses para adquirir las consonantes /l/ - /r/. Los oyentes recibían un entrenamiento de este par de consonantes que consistía identificación de estos sonidos en diferentes contextos fonéticos. Se analizaron este par de sonidos porque no aparecen en la L1 de los oyentes lo cual significa que un aprendiz de una L2 que sí contiene estos sonidos puede tener dificultades a la hora de percibirlos y aprenderlos. Los oyentes inexpertos japoneses demostraron una mala percepción de los sonidos /l/ y /r/ tanto en contextos que habían sido modificados sintéticamente (MacKain et al. 1981, Mochizuki, 1981) como en contextos en los que el habla se presentaba de una manera natural (Goto 1971, Sheldon y Strange 1982). Además, la percepción de estos sonidos no es uniforme pero sí depende del contexto fonético en el que aparece. La identificación de estas consonantes es peor para /r/ y /l/ en posición inicial independientemente de si aparecen solas, en un grupo consonántico o en posición intervocálica. No obstante, se produce una mejor identificación de estos sonidos en posición final tanto en un grupo consonántico como solas (Sheldon y Strange 1982, Mochizuki 1981).

De los resultados aportados por estos estudios pueden establecerse tres generalizaciones fundamentales: 1) por un lado el éxito de utilizar un entrenamiento con una alta variabilidad en los hablantes nativos y en los contextos demuestra que los adultos tienen una neuroplasticidad suficiente para exponerse a modificaciones importantes de sus conocimientos lingüísticos a través del entrenamiento. Sin embargo, los oyentes japoneses que participaron en estos estudios no fueron capaces de adquirir una competencia casi nativa para la identificación de los sonidos /l/ y /r/. Esto indica que están sujetos a alguna limitación impuesta por su L1; 2) la variabilidad ofrecida por los diferentes hablantes ha producido un sólido aprendizaje auditivo acostumbrado a los cambios de información recibida en los sonidos lo que ha causado que los oyentes puedan establecer generalizaciones a largo plazo en la identificación de estos sonidos en el sistema fonético subyacente. Es decir, las mejoras en la identificación de los sonidos /l/ y /r/ se generalizaron en la identificación de nuevos elementos producidos por nuevos hablantes hasta tres meses después del entrenamiento; y 3) el aprendizaje de los sonidos del habla realizado a través de la modalidad auditiva produce modificaciones a largo plazo de la percepción y de la producción de los contrastes estudiados.

Wagner (2013) realizó un estudio multilingüístico donde examinó la percepción de las consonantes fricativas utilizando indicios coarticulatorios. Este autor examinó cómo los oyentes holandeses, italianos, polacos y españoles identificaron consonantes fricativas y oclusivas en posición inicial y final de sílaba. Los resultados demostraron diferencias multilingüísticas en la formación de los indicios coarticulatorios para la identificación de fricativas: los oyentes españoles y polacos extrajeron la información del lugar de articulación de porciones más cortas de sílabas en posición final. No obstante, no encontró diferencias multilingüísticas en la percepción del lugar de articulación en las consonantes oclusivas sugiriendo que una mayor dependencia de los indicios articulatorios no se generaliza a otros tipos de fonemas. Para las consonantes fricativas los resultados demuestran que los oyentes españoles y polacos extraen la información del lugar de articulación de una porción relativamente corta de la fricativación de las consonantes en posición final de sílaba, mientras que los oyentes holandeses e italianos necesitan una porción más grande de fricativación para ganar una cantidad similar de información del lugar de articulación. Además afirma que los oyentes que tienen pares fricativos similares en su L1 dependen más del lugar de articulación de la vocal que precede el ruido fricativo. No obstante, la primera porción de fricativación en posición

final de sílaba llevaba la diferencia más importante para las fricativas. Por lo tanto, podría afirmarse que los oyentes con lenguas que tienen fricativas que son perceptualmente similares han mejorado su percepción de los indicios en el espectro fricativo más que en la dependencia de los indicios coarticulatorias. Sin embargo, según Wagner (2013: 4265)

(...) what speaks against this is that Polish and Spanish listeners were not generally better in identifying fricatives: their performance in the last gate, after hearing the entire frication, was lower than the performance of Dutch and Italian listeners. However, Polish and Spanish listeners gained more information about the place of articulation from a shorter portion of the frication.

Por otro lado, también es importante mencionar la influencia que el ruido enmascarador tiene en la dimensión acústica y cómo esta característica acústica afecta la percepción del lugar de articulación. Jiang, Chen y Alwan (2006) llevaron a cabo un estudio en el que analizaron las consonantes oclusivas /b, d, p, t/ y las consonantes fricativas /f, s, v, z/ en posición inicial y en condiciones de ruido. Los estímulos utilizados consistían en contextos naturales de sílabas en posición inicial combinadas con las vocales /a, i, u/. Jiang, Chen y Alwan (2006: 26) encontraron que

(...) the frequency and relative spectral amplitude were the cues most predictive of place of articulation decisions in quiet decisions, but certain spectral amplitude measurements appeared to be masked at low SNRs, with a contrasting result that formant frequency measurements were better place of articulation cues at low SNRs.

En condiciones óptimas de ruido, todos los sonidos se percibieron correctamente a casi el 100% con frecuencia en formantes o medidas de amplitud espectral relativas. En presencia de ruido, los oyentes percibieron el lugar de articulación labio-alveolar correctamente incluso cuando el nivel de SNR era de -5dB. Sin embargo, en la presencia de -15dB, los oyentes tuvieron grandes problemas para identificar las consonantes y se restringieron las respuestas a una actuación aleatoria. Estos autores demostraron que las fricativas tenían un umbral de SNR menor que las consonantes oclusivas. Por otro lado, atendiendo a las medidas de amplitud espectral relativas, encontraron que los resultados obtenidos de los experimentos indicaban que las medidas de frecuencia de formantes eran más importantes para la percepción del lugar de articulación en un nivel bajo de SNR que las medidas de amplitud espectral relativas.

Afirman que estas medidas son fácilmente influenciadas por el ruido si las comparamos a las medidas de frecuencia de formantes. Asimismo, también afirman que los oyentes utilizan cualquier indicio disponible para percibir los sonidos correctamente, por lo tanto, el efecto del tipo del enmascarador de ruidos debería tenerse también en cuenta. Especulan con la idea de que un habla manipulada o un ruido producido por el balbuceo de diferentes hablantes podrían también corromper las medidas de frecuencia de formantes.

Por consiguiente, las medidas de frecuencia de formantes son indicios más dominantes para el lugar de articulación labial y alveolar que las medidas de amplitud espectral relativa. La percepción del lugar de articulación de las consonantes oclusivas y fricativas en posición inicial depende de la interacción de la sonoridad, el modo de articulación y el contexto vocálico. Por lo tanto, ninguna característica acústica debería ayudar a la percepción del lugar de articulación. Otro hallazgo importante en este estudio es que las diferencias específicas de la lengua se encontraron en posición final de sílaba y no en posición inicial. En las codas, todos los oyentes obtuvieron la misma cantidad de información de la porción de vocal que seguía la constricción. Por consiguiente esto indica que los oyentes se benefician más de la información que viene al final. Esto significa que la información en posición de ataque podría ser más accesible.

Pearman (2007) llevó a cabo un estudio de percepción del habla entre lenguas en el que comparaba inglés y catalán. En su caso, se centró en el reconocimiento de palabras por parte de oyentes de la L2. Los oyentes no nativos utilizaban como indicios acústicos para este reconocimiento procesos de asimilación (lugar, modo o asimilación de nasalidad), elisión y reducción de vocales y consonantes. Al igual que en la percepción segmental de sonidos, en la percepción de palabras, los oyentes no nativos de la L2 también necesitan más tiempo que los oyentes nativos para reconocer palabras, y, además, su reconocimiento es también peor que el de los nativos. Por su parte, los no nativos analizaban la señal de la información apoyándose en la L1 lo cual llevaba a errores en un gran número de ocasiones. Sin embargo, los hablantes nativos se beneficiaban de los efectos de la frecuencia del sonido y de la predicción del mismo para analizar la señal fonética. Por tanto, puede decirse que el proceso de reconocimiento del habla por parte de oyentes nativos y no nativos es diferente.

Además, los datos revelaron que la L1, el contexto y la experiencia tienen un papel importante en el reconocimiento de palabras de una L2 ya que los no nativos tuvieron mejores resultados en el reconocimiento de palabras reducidas por procesos existentes en la L1 en el mismo contexto, que por procesos que se dan en un contexto distinto o por procesos que no se dan en la L1. Por otro lado, es importante resaltar que los estímulos de este estudio están elaborados en un laboratorio en el que contaban con condiciones acústicas óptimas. No se han analizado estos estímulos modificando la señal acústica que perciben los oyentes. Quizás si se hubiera modificado esta señal, los resultados serían diferentes, ya que como hemos dicho anteriormente, cuando se introduce ruido en la señal tanto los oyentes nativos como no nativos tienden a empeorar su percepción. Como hemos mencionado con anterioridad, la modificación de la señal aplicando enmascaramiento es un método que ya es habitual en estudios experimentales dado que hace que se asemeje más a situaciones y contextos reales.

2

Hipótesis y objetivos

2.1. Hipótesis

Tradicionalmente, la percepción de sonidos de L2 se ha atribuido a factores como la influencia de la lengua materna L1 (Flege, 1995, Best et al. 2001, Kuhl, 2004), al ruido de fondo (Nabelek y Donahue, 1984, Cutler et al. 2004, García Lecumberri y Cooke 2006, García Lecumberri et al. 2010), y de manera secundaria, a las características inherentes de los sonidos (Darcy et al. 2007, Werker y Tees, 1984, 2002, Guion et al. 2000, Lalonde y Werker, 1995).

La hipótesis central en esta tesis doctoral está en consonancia con el postulado del modelo PAM el cual establece que hay una serie de rasgos prototípicos, los cuales están compartidos por hablantes de diferentes L1 y, a veces, estos rasgos son idénticos o casi idénticos entre lenguas diferentes. Por ejemplo, recordemos que la teoría de universales lingüísticos (*Markedness Theory*) establece ciertos rasgos fonológicos como no marcados, como por ejemplo la sonoridad, entre otros, y este rasgo es universal a todas las lenguas. Además, en nuestro estudio aceptamos la idea de que los rasgos lingüísticos universales no marcados se aprenden con posterioridad a los rasgos universales marcados.

La hipótesis central de esta tesis también acepta que la creación de prototipos o categorías fonológicas (adquisición de la fonología) se ve favorecida en situaciones adversas. En el campo que nos atañe, esto significaría que el aprendizaje de las unidades fónicas se acelera en presencia, por ejemplo, del ruido de fondo. Consideramos, por tanto, que el enmascaramiento nos sitúa en una situación de percepción de sonidos adversa que obliga al oyente a poner a prueba su capacidad auditiva.

Por otro lado, también asumimos la hipótesis de que el entrenamiento auditivo intensivo produce efectos beneficiosos en la percepción de consonantes inglesas en unidades silábicas en presencia de situaciones adversas de ruido.

Esta tesis pretende demostrar las siguientes hipótesis:

- H0: En oposición a los antecedentes descritos, esta tesis asume que la explicación a la creación de categorías en una L2 se hace a partir de varias L1.
- H1: La presentación de estímulos de manera continuada con entrenamiento auditivo intensivo produce mejores resultados en la percepción de estos estímulos.
- H2: Los estímulos que obtienen mejores resultados en la percepción se corresponden con la presentación continuada de estímulos con mayor nivel de enmascaramiento.
- H3: Se obtienen mejores resultados en la presentación de estímulos en unidades silábicas que en unidades lingüísticas con significado completo, es decir, unidades oracionales.

Para la demostración de la H1 y H3 se han usado métodos experimentales que utilizan oyentes con diferentes L1. Para la demostración de la H2, se han utilizado métodos experimentales que tienen en cuenta oyentes con la misma L1.

Para el análisis estadístico se han tenido en cuenta las siguientes variables: posición silábica (inicial y final o ataque y coda), nivel de SNR, modo de articulación, lugar de

articulación, sonoridad, apertura, aspiración y posición silábica. No se ha incluido la L1 en ningún caso como variable para el análisis de datos. La motivación es la creencia de que los factores que mayormente afectan la creación de prototipos fónicos se encuentran en los rasgos inherentes al sonido y el contexto (enmascaramiento de ruido en nuestro caso).

Para las variables relacionadas con los rasgos articulatorios inherentes a los sonidos hemos seguido el método utilizado por otros estudios (Bent et al. 2008, Bell y Hooper 1978, Helfer y Huntley 1991, Kapoor y Allen 2013, Miller y Nicely 1955, Miller y Eimas 1996, Redford y Diehl 1999, Silbert 2012 y Wagner 2013).

La variabilidad se distribuye de la siguiente manera:

- Variable dependiente: porcentaje correcto de identificación de las consonantes en las posiciones CV y VC.
- Variables independientes: nivel de SNR, modo de articulación, lugar de articulación, sonoridad, apertura, aspiración y posición silábica.

La pregunta que queremos responder es si existen consonantes inglesas que pudieran ser percibidas con mayor o menor dificultad, entendiendo dificultad en relación al proceso de adquisición o creación de categorías fónicas (aprendizaje). También queremos responder a la pregunta de si este proceso de adquisición puede verse acelerado a partir de la exposición de sonidos en contexto adverso (ruido de fondo).

2.2. Objetivos

Los objetivos concretos son analizar la exactitud en la percepción en la identificación de todas las consonantes inglesas en posición inicial y final de sílaba con y sin entrenamiento auditivo intensivo. Más específicamente:

- a) La identificación de las consonantes inglesas en tareas de percepción que se ven favorecidas por un entrenamiento auditivo intensivo y la identificación de las consonantes que obtienen mejores resultados en su identificación.

- b) Conocer la relación entre el índice de precisión en la identificación auditiva de consonantes inglesas y sus rasgos articulatorios: lugar de articulación, modo de articulación y sonoridad.
- c) Conocer cómo afecta el entrenamiento auditivo intensivo cuando se utilizan oraciones como estímulos en las tareas de entrenamiento vs. cuando se utilizan sílabas como estímulos en las tareas de entrenamiento.
- d) Conocer los beneficios del entrenamiento auditivo intensivo con estímulos enmascarados (15dB SNR) en el uso de estímulos en unidades silábicas y oracionales.

Para cumplir con los objetivos (a) y (b) se ha llevado a cabo un primer experimento (Experimento I) consistente en un test de identificación de todas las consonantes inglesas en sílabas con estructura CV y VC (posición inicial y final respectivamente). El test lo han realizado 33 oyentes no nativos con ocho L1s diferentes en el Laboratorio de Fonética de la Universidad de Sevilla.

Para cumplir con los objetivos (c) y (d) se ha llevado un segundo experimento (Experimento II) aplicando un método de entrenamiento auditivo intensivo distribuido en tres grupos: grupo de control, grupo experimental I, grupo experimental II. En el mismo participaron un total de 15 hablantes nativos del español (6 hombres y 9 mujeres) que tomaron diferentes sesiones de entrenamiento auditivo consistente en diferentes tareas de identificación de todas las consonantes inglesas en sílabas con estructura CV y VC y en oraciones. El experimento tuvo lugar en el Laboratorio de Fonética de la Universidad de Sevilla.

En el siguiente capítulo de materiales y método se define con mayor detalle los estímulos utilizados en cada test de percepción realizado en cada experimento así como el procedimiento usado para la recogida de datos.

3

Materiales y métodos

Esta sección incluye la descripción de los materiales utilizados en la experimentación divididos en dos apartados. Estos dos experimentos son los siguientes:

1) Test general de identificación de consonantes inglesas

En el primer experimento, participantes voluntarios con diferentes L1 realizaron tareas de identificación de consonantes inglesas en sílabas (posición inicial y final). Las sílabas fueron grabadas como estímulos naturales para todas las combinaciones consonánticas posibles en la lengua. Los resultados del test se han medido en función del porcentaje de acierto en las consonantes a tenor de la variabilidad descrita en el capítulo anterior:

- Variable dependiente: porcentaje correcto de identificación de las consonantes en las posiciones CV y VC.
- Variables independientes: nivel de SNR, modo de articulación, lugar de articulación, sonoridad, apertura, aspiración y posición silábica.

2) Test de entrenamiento auditivo intensivo: Efectos del contexto adverso (ruido de fondo) y de la combinación de estímulos (sílabas vs. oraciones)

En el segundo experimento, participantes voluntarios nativos de español realizaron diferentes sesiones de entrenamiento auditivo intensivo divididos en 3 grupos. A saber: grupo de control, grupo experimental 1 (entrenamiento con sílabas y ruido de fondo elevado), y grupo experimental 2 (entrenamiento con oraciones y sin ruido de fondo elevado). Las sílabas utilizadas en el primer

experimento sufrieron modificaciones de enmascaramiento por ruido de fondo. A la vez, se creó un nuevo test de identificación de palabras en oraciones. Los resultados de los diferentes modelos de entrenamiento auditivo intensivo se han analizado en función de la variabilidad siguiente:

- Variable dependiente: porcentaje correcto de identificación de las consonantes en las posiciones CV y VC.
- Variables independientes: nivel de SNR, y unidad lingüística entrenada (sílabas vs. oración).

3.1. Experimento I: Test general de identificación de consonantes inglesas

3.1.1. Población

En el experimento participaron 33 oyentes con edades comprendidas entre los 21 y los 39 años (la media de edad es de 24,6). El grupo poblacional estaba compuesto de 8 hombres y 25 mujeres con diferentes L1s (n=8). La mayoría de los participantes eran estudiantes de intercambio en la Universidad de Sevilla en el momento en el que se llevó a cabo el estudio. Sin embargo, había participantes que ya habían concluido sus estudios universitarios. Por lo tanto, esto indica que el nivel educativo entre todos los sujetos del estudio era prácticamente similar. Todos los participantes eran originarios de países europeos: España, Italia, Francia, Bélgica, Alemania, Rusia, República Checa y Grecia. El número total de participantes estaba compuesto por 6 polacos, 9 españoles, 3 alemanes, 5 rusos, 1 checo, 3 italianos, 1 griego, 2 franceses y 3 belgas cuya lengua materna era el francés. Atendiendo al cuestionario demográfico en el cual se pregunta sobre los antecedentes lingüísticos (anexo 1), se observa que todos habían crecido en entornos monolingües, a excepción de dos alemanes y un francés. El sujeto francés era bilingüe en francés y español; uno de los sujetos alemanes era bilingüe en alemán y holandés mientras que el otro sujeto alemán era trilingüe en alemán, polaco y ruso. Este cuestionario demográfico también demostró que todos los participantes usaban regularmente el inglés en su vida diaria.

Con anterioridad al cuestionario demográfico, los sujetos debían firmar un formulario de consentimiento informado (anexo 1) en el cual se exponían los detalles del estudio, se explicaba los beneficios del mismo y los posibles efectos adversos. Además, se les explicaba que todos los datos aportados serían confidenciales y que aceptaban participar voluntariamente y de manera no remunerada.

La población también realizó un cuestionario en el cual se evaluaba el nivel de inglés de los participantes (anexo 1). El cuestionario de inglés se situaba en los niveles B2-C1, nivel considerado como mínimo para que los sujetos voluntarios se consideraran aptos para el estudio.

La última prueba realizada a los voluntarios antes de aceptarlos a formar parte de la experimentación fue una audiometría tonal (anexo 1) que se llevó a cabo en una cabina insonorizada en el Laboratorio de Fonética de la Universidad de Sevilla utilizando un audiómetro *Sibelmed AS5-A*. El test auditivo medía la capacidad auditiva exponiendo a los participantes a diferentes niveles de tonos puros. Estos niveles de tonos estaban compuestos por 250 Hz (a 25 dB), 500 Hz (a 20 dB), 1000 Hz (a 20 dB), 2000 Hz (a 20 dB), 4000 Hz (a 20 dB) and 8000 Hz (a 20 dB). Se consideraba el test apto y, por tanto, se consideraba apto al participante, si todas las respuestas eran correctas o sólo había una respuesta errónea. En nuestro estudio, ningún participante demostró tener problemas auditivos en el test. Además, los cuestionarios de antecedentes demográficos de la población también indicaron que ningún participante había tenido infecciones o dolor en los oídos, pérdida de audición, vértigos o pitidos o una sobreexposición al ruido en su vida cotidiana.

Por lo tanto, nuestros criterios de elegibilidad se centran en tres aspectos concretos: 1) la edad, la población estaba comprendida entre 18 y 45 años; 2) la población debía tener un nivel de inglés comprendido entre el B2 y el C2 según el cuestionario de nivel; y 3) los resultados de la audiometría debían ser claros y no aportar ningún indicio de posible pérdida auditiva.

3.1.2. Materiales

En esta subsección se describen primero los estímulos auditivos incorporados a los tests de percepción y luego los tests de percepción que se diseñaron para el experimento I.

3.1.2.1. Estímulos

Se grabaron sílabas CV y VC (consonante en posición inicial y final) por ocho hablantes del Medio Oeste Americano, concretamente de Indiana, Illinois y Kentucky. El número total de estímulos era de 360, 180 sílabas CV y 180 sílabas de tipo VC. Los estímulos fueron grabados, normalizados y segmentados en una cabina insonorizada y se añadieron tres niveles de SNR (40dB, 15dB y 5dB), con una duración media de 600 ms cada sílaba.

Tabla 1. Clasificación de las consonantes atendiendo a su representación ortográfica y posición silábica

Consonantes en posición inicial		Consonantes en posición final	
Sonidos (representación ortográfica)	Ejemplos	Sonidos (representación ortográfica)	Ejemplos
p-	past, pool	-p	top, soup
pl-	please, plan	-pt	hoped, leaped
pr-	proof, preen	-ps	cups, rips
f-	first, feed	-f	leaf, off
fl-	flat, flee	-th	moth, truth
fr-	fruit, free	-t	eat, pot
th-	third, theme	-ts	treats, shoots
t-	tot, too	-s	peace, boss
tr-	true, tree	-st	east, boost

Consonantes en posición inicial		Consonantes en posición final	
Sonidos (representación ortográfica)	Ejemplos	Sonidos (representación ortográfica)	Ejemplos
s-	sad, city (soft –c)	-sts	feasts, costs
ch-	church, chief	-ch	teach, notch
sh-	shoe, sure	-sh	wash, push
k-	keep, curl	-k	lock, fluke
kl-	class, kleenex	-kt	peeked, looked
kr-	creed, crew	-ks	leaks, socks
kw-	quack, quirk	-b	rob, tube
h-	he, had	-bz	sobs, rubs
b-	be, boot	-v	move, sleeve
bl-	blurt, blue	-vz	halves, proves
br-	brief, brute	-d	need, food
w-	were, we	-dz	noods, moods
m-	me, moon	-z	bees, choose
v-	vast, verse	-j	lodge, stodge
th(v)	these, there	-g	dog, hug
d-	dirt, dude	-m	dream, loom
l-	least, loose	-mz	rooms, chimes
dr-	dream, drew	-n	clean, swan
n-	new, knees	-nt	hint, want
z-	zoo, zee	-nd	pond, leaned
j-	jury, germ (soft –G)	-nts	prince, mints
r-	roost, real	-nz	means, beans
y-	you, year	-ng	bring, song
g-	girl, goon	-ngz	sings, lungs
gl-	glass, glee	-vowel	tee, new
gr-	great, group	-R	stir, cur

Consonantes en posición inicial		Consonantes en posición final	
Sonidos (representación ortográfica)	Ejemplos	Sonidos (representación ortográfica)	Ejemplos
sp-	spur, speak	-L	peel, doll
sw-	swim, sweet		
sm-	smack, smooth		
st-	staff, stew		
sl-	slap, sleep		
str-	street, strap		
sn-	sneak, snooze		
sk-	skirt, school		
skr-	scream, scrap		
vowel-	oops, eat		

La tabla 1 atiende a una representación ortográfica de todos los estímulos que escucharon los oyentes en posición inicial y en posición final con sus diferentes combinaciones con otros sonidos, así como palabras que ejemplifican la posición de los estímulos.

Cada sílaba que aparece en las Tabla 1 se enmascaró a 5dB SNR, a 15dB SNR y a 40dB SNR. Cuanto más bajo es el nivel de SNR más cantidad de ruido incluye la señal.

3.1.2.2. Tests de identificación

Los tests de identificación consistían en un test de identificación de consonantes en posición inicial y un test de identificación de consonantes en posición final.

a) Test de identificación en posición inicial: este test consistía en la presentación de un total de 180 estímulos en posición inicial de sílaba. Los estímulos de este test

corresponden con los que aparecen en la tabla 1 (anexo 2). La tarea de identificación de los sonidos en posición inicial estaba compuesta por 22 consonantes (12 consonantes sonoras y 9 consonantes sordas) combinadas con otras consonantes y distintas vocales, 9 grupos consonánticos encabezados por el sonido /s/ (*s-clusters*) y 4 vocales. El tiempo medio de la presentación del test, medido en todos los participantes del estudio, era de aproximadamente 30 minutos. El número total de muestras en el test era de la siguiente forma: 45 muestras x 4 presentaciones. Las muestras se presentaban de forma aleatoria.

b) Test de identificación en posición final: este test consistía en la presentación de un total de 180 estímulos en posición final de sílaba. Los estímulos de este test corresponden con los que aparecen en la tabla 1 (anexo 2). La tarea de identificación de los sonidos en posición final estaba compuesta por 16 consonantes (8 consonantes sordas, 6 consonantes sonoras, -R y -L) que a su vez se combinaban con diferentes vocales, 9 grupos consonánticos nasales (*nasal clusters*) y 5 vocales. El tiempo medio de la presentación del test, medido en todos los participantes del estudio, era de aproximadamente 30 minutos. El número total de muestras en el test era de la siguiente forma: 36 muestras x 5 presentaciones cada muestra. Las muestras se presentaban de forma aleatoria.

3.1.3. Procedimiento

Los participantes se exponían a la percepción de las consonantes en posición inicial y final en un ambiente tranquilo y silencioso en el Laboratorio de Fonética de la Universidad de Sevilla. Cada oyente participaba en una sesión de identificación, compuesta de 360 estímulos en total. Antes de que la tarea de identificación comenzara, se instruía a los participantes en el entorno de la tarea y en la presentación de estímulos. Los estímulos se presentaron en un ordenador portátil *Toshiba Tecra* utilizando unos auriculares reductores de ruido modelo *Quiet Comfort[®] 15 Bose Acoustic Cancelling[®]*.

La sesión estaba dividida en dos partes que correspondían a la identificación de los sonidos en posición inicial y en posición final. Los sujetos tenían alrededor de diez minutos de descanso entre las partes. Sin embargo, podían pedir un descanso cuando lo consideraran oportuno e incluso podían detener el experimento y abandonar su

participación. Ninguna de estas dos cosas ocurrió en ningún caso. Los participantes se repartieron de manera cuasi-aleatoria en 3 grupos experimentales:

- Grupo Experimental I: 11 voluntarios realizaron los tests de percepción con estímulos enmascarados a 40dB,
- Grupo Experimental II: otros 11 realizaron los tests con estímulos enmascarados a 15dB,
- Grupo Experimental III: y los últimos 11 voluntarios lo hicieron con sílabas enmascaradas a 5dB SNR.

Una vez que comenzaban las tareas de identificación, los participantes tenían que responder a todas y cada una de las pruebas sin tener límite de tiempo. Los estímulos se presentaban sólo una vez, es decir, la producción del estímulo no podía repetirse en ningún momento y los participantes tenían que identificar el sonido que se presentaba clicando en la representación ortográfica de los estímulos en la pantalla del ordenador. Cuando ponían el puntero del ratón sobre los diferentes sonidos, aparecían dos ejemplos en la parte inferior de la pantalla. Además, el test o tarea de identificación de los estímulos incluía también un *feedback* después de cada respuesta. Con este *feedback* se les indicaba si la respuesta dada por los sujetos era correcta o incorrecta. Por tanto, aunque el test contenga únicamente 360 estímulos y se haya diseñado como una tarea de identificación de estímulos, con este *feedback* los sujetos reciben un entrenamiento de los estímulos presentados. Consideramos que la inclusión de este *feedback* es necesaria porque incentiva a la conclusión de la tarea ya que incrementa la interacción del sujeto con la tarea presentada.

3.2 Experimento II: Test de entrenamiento auditivo intensivo. Efectos del contexto adverso (ruido de fondo) y de la combinación de estímulos auditivos (sílabas vs. oraciones).

3.2.1 Población

En este segundo experimento participaron 15 oyentes nativos españoles con edades comprendidas entre los 21 y los 25 años (la media de edad es de 22,5). El grupo poblacional estaba compuesto de 6 hombres y 11 mujeres. Todos los participantes eran estudiantes de la Universidad de Sevilla en el momento en el que se llevó a cabo el estudio.

Al igual que los participantes del primer experimento, esta población también tuvo que rellenar el cuestionario demográfico, firmar el consentimiento informado en el que aceptaron participar en el estudio de manera gratuita y voluntaria, realizaron el test de nivel de lengua inglesa y la audiometría.

Según la información recogida en estos cuestionarios, se observó que todos los participantes contaban con al menos cinco años de aprendizaje del inglés como L2 a nivel universitario a parte de las etapas obligatorias educativas (primaria y secundaria). Además, se exponían regularmente a la L2 escrito y/u oral a través de los medios de comunicación, trabajo o estudios. También todos los participantes habían estudiado fonética y realizado transcripciones fonéticas antes de realizar este experimento. Por otro lado, teniendo en cuenta todo esto y los resultados del test de nivel de inglés, se observó que los participantes tenían un nivel B2.

Por otro lado, los resultados de la audiometría demostraron que ninguno de los voluntarios en el experimento mostraba pérdida auditiva por lo que todos ellos fueron considerados como aptos para seguir en la realización de este estudio.

Teniendo en cuenta los resultados de los diferentes formularios, la población de este segundo experimento se ajusta también a los criterios de elegibilidad establecidos para

la población del primer experimento: 1) edad (población comprendida entre 18 y 45 años); 2) nivel de inglés entre B2 y C1; y 3) los resultados de la audiometría son claros y no aportan ningún indicio de problema auditivo.

3.2.2. Materiales

En esta subsección se describen primero los estímulos auditivos y después, los tests diseñados para el experimento II.

3.2.2.1. Estímulos

Se usaron las mismas sílabas grabadas para el Experimento I (anexo 2). Al igual que en el experimento I, en este segundo experimento todas las muestras se enmascararon con diferentes niveles de SNR. En este caso fueron -5dB, 5dB y 15dB.

Por tanto, el número total de estímulos grabados se especifican a continuación:

- En posición inicial: 8 hablantes x 45 consonantes x 4 vocales x 3 SNRs = 4.320 estímulos.
- En posición final: 8 hablantes x 36 consonantes x 4 vocales x 3 SNRs: 3.456 estímulos.

Respecto a las oraciones, los hablantes grabaron 45 oraciones diferentes enmascaradas a 15dB que se detallan en el anexo 3 (Corpus entrenamiento oracional).

3.2.2.2. Tests de identificación

Para realizar este segundo estudio, se diseñaron dos tests diferentes de identificación. Por un lado, se realizó un test de identificación de sílabas en posición CV y VC y un test de identificación de oraciones. Había dos razones principales para diseñar estos dos tipos de tests. Por un lado, los tests de identificación silábica se usaron como pre-test y post-test y, además, como entrenamiento. Por otro lado, el test de identificación

oracional se usó exclusivamente como entrenamiento. En esta sección se van a describir tanto el test de posición silábica como el test de identificación oracional.

Todos los tests se diseñaron siguiendo el principio de estímulo-respuesta, es decir, consistían en la presentación de un estímulo auditivo (frase o sílaba, según lo apropiado) escuchados a través de unos auriculares BOSE AE2W que se les proporcionaba a los voluntarios en el momento del estudio. Tras la presentación de cada estímulo, los participantes tenían que responder clicando con el ratón en el área correspondiente de la pantalla. Cada clic correspondía a una respuesta relacionada con la identificación auditiva de los participantes. De la misma manera que el Experimento I, este segundo experimento se llevó a cabo en un software especializado para la experimentación de estímulos y respuestas y éstas se cargaban en el sistema para, posteriormente, ser transferidas a un software especial para su análisis posterior. En la hoja de cálculo, cada estímulo presentaba dos valores: la respuesta correcta y la respuesta dada por el participante.

Para este segundo estudio, fue necesario crear cuatro tests de identificación silábica y un test de identificación oracional. La diferencia entre ellos era principalmente el número de estímulos utilizados, la posición de las consonantes y el enmascaramiento de la sílaba. A continuación se presenta una descripción detallada de cada test.

a) Test de identificación de sílabas en posición inicial (CV) 1. Este test consistía en la presentación de un total de 92 estímulos en posición inicial. Las consonantes en el test fueron una selección de las sílabas citadas en la tabla dos anteriormente: vocal, b, ch, d, f, fr, g, h, j, k, l, m, n, p, pr, r, s, sh, st, t, tr, v, w. La media de tiempo para la presentación del test era de 15 minutos, aproximadamente. El número total de muestras en el test es el siguiente: 23 muestras x 4 presentaciones, esto es que cada estímulo se escuchaba 4 veces. Los estímulos se presentaban en un orden aleatorio.

b) Test de identificación de sílabas en posición inicial (CV) 2. Este segundo test consistía en la presentación de un total de 180 estímulos en posición inicial (CV). Estos estímulos correspondían concretamente con los presentados en la tabla 2. La media de tiempo para la presentación del test era de unos 30 minutos, aproximadamente. El

número total de estímulos es el siguiente: 45 estímulos x 4 presentaciones. Los estímulos se presentaban en un orden aleatorio.

c) Test de identificación de sílabas en posición final (VC) 1. Este test consistía en la presentación de un total de 90 estímulos en posición final. las consonantes del test eran una selección de los presentados en la tabla 2: d, k, l, m, n, nd, ng, nt, nts, nz, r, s, st, t, ts, v, vocal, z. La media de tiempo para la presentación de este test era de unos 15 minutos, aproximadamente. El número total de estímulos es el siguiente: 18 estímulos x 5 presentaciones. Los estímulos se presentaban en un orden aleatorio.

d) Test de identificación de sílabas en posición final (VC) 2. Este test consistía en la presentación de un total de 180 estímulos en posición final. Las consonantes de este test coinciden exactamente con las de la tabla 2. La media de tiempo para la presentación de este test era de unos 30 minutos, aproximadamente. El número total de estímulos es el siguiente: 36 estímulos x 5 presentaciones. Los estímulos se presentaban en un orden aleatorio.

Los tests de identificación silábica se diseñaron de tal forma que solo un estímulo auditivo podría ser reproducido en el momento. El participante podía elegir cuándo quería escuchar el siguiente estímulo clicando en la tecla “siguiente” (*Next*). Después de la escucha de cada estímulo, se instruía al participante para clicar en un mapa de etiquetas que aparecían en la pantalla. Este mapa de etiquetas era realmente un set de botones y cada uno contenía una denominación que hacía referencia a todas las consonantes incluidas en el test de percepción silábica. Se crearon cuatro mapas de etiquetas, uno para cada uno de los cuatro test descritos anteriormente (identificación de consonantes en posición inicial 1, identificación de consonantes en posición inicial 2, identificación de consonantes en posición final 1 e identificación de consonantes en posición final 2). El anexo 4 muestra la representación de cada mapa de etiquetas para cada test.

Se garantizó la variabilidad de los test utilizando estímulos que consistían en una combinación de los 8 hablantes y de las 4 combinaciones de las vocales con las consonantes. Por tanto, cada test de percepción consistía en una combinación de los siguientes aspectos:

- Consonantes diferentes/ grupos consonánticos (*consonant clusters*)
- Diferentes vocales combinadas con las consonantes
- Diferentes hablantes para las mismas combinaciones

Al igual que en el experimento I, a pesar de que el participante podía elegir cuándo escuchar la siguiente presentación, los estímulos solo podían escucharse una vez. Una vez que se elegía la respuesta, el sistema proporcionaba un “feedback”. Esto consistía en que una vez que se había realizado la selección en la pantallas, el sistema decía si la respuesta era correcta o no. Si la respuesta era errónea, el sistema proporcionaba la respuesta correcta. Por un lado, la razón para proceder de esta manera era para promover la motivación por parte del oyente y así, minimizar la fatiga y, por otro lado, facilitar un entorno natural de aprendizaje.

e) Test de identificación oracional. Este test consistía en la presentación de 45 oraciones enmascaradas a 15 SNR. El procedimiento de este test varía ligeramente de los anteriores ya que después de la reproducción de cada frase, se mostraban automáticamente una serie de palabras en la pantalla. Esta selección incluye, en orden alfabético, las palabras que construían cada frase y otras que eran fonéticamente parecidas. La captura de la interficie de la figura 27 del anexo 4 muestra la primera oración del entrenamiento.

3.2.3. Procedimiento

En este segundo experimento se ha seguido el siguiente procedimiento: todos los alumnos pertenecían a la Universidad de Sevilla y el estudio se llevó a cabo en el Laboratorio de Fonética. Los participantes eran voluntarios y cada uno de ellos tenía que acudir a un total de tres sesiones para completar el estudio. Cada sesión duraba una hora, aproximadamente. Las sesiones se estructuraron de la siguiente forma:

- Primera sesión: Informe consentido, cuestionario sociodemográfico, audiometría y pre-test.
- Segunda sesión: entrenamiento.
- Tercera sesión: entrenamiento y post-test.

a) Primera sesión. En esta primera sesión, los participantes rellenaron un informe consentido (anexo 1) y un cuestionario sociodemográfico (anexo 1), al igual que los participantes del experimento I. En estos formularios tenían que dar su permiso para la participación en el estudio y contestar a una serie de preguntas sobre su situación sociodemográfica. Para llevar a cabo la audiometría, se utilizó un equipo SIBELMED AS5-A en una cabina insonorizada y unos auriculares para evaluar los umbrales auditivos en las frecuencias 250Hz a 25dB HL y 500, 1000, 2000, 4000 y 8000Hz a 20dB HL (anexo 1)

Una vez que se acabó este primer paso, los participantes comenzaban el pre-test que incluía un test para la posición inicial (Test de identificación de sílabas en posición inicial (CV) 1) y otro para la posición final (Test de identificación de sílabas en posición final (VC) 1). Cada uno presentaba 90 estímulos y ambos estaban enmascarados a 15 SNR.

b) Segunda sesión. La segunda sesión del test tuvo lugar una semana después de la primera, aproximadamente. A cada participante se le asignó un grupo diferente de entrenamiento en un muestreo cuasi-aleatorio. Todos los participantes recibieron tres horas de entrenamiento distribuidos en tres semanas. A continuación se describen los tres grupos diferentes de participantes:

- Grupo de control. El grupo de control recibió un entrenamiento intensivo de identificación de consonantes con estímulos que no habían sido previamente enmascarados a 15 SNR. Además, los tests de entrenamiento sólo incluían estímulos silábicos.
- Grupo experimental I. Este grupo de participantes realizó el pre-test y el post-test de identificación de consonantes y realizó un entrenamiento en condiciones de ruido adversas. Esto quiere decir que recibieron entrenamiento intensivo de consonantes con estímulos que habían sido previamente enmascarados a -5dB SNR. Los tests de entrenamiento sólo incluían estímulos silábicos.
- Grupo experimental II. Este grupo de participantes realizó el pre-test y el post-test de identificación de consonantes y recibieron un entrenamiento mixto sin ruido a 15dB SNR. Esto quiere decir que recibieron entrenamiento intensivo de identificación

de consonantes y oraciones con estímulos que no habían sido enmascarados previamente. La diferencia con el grupo de control es que los tests de entrenamiento incluían estímulos silábicos y oracionales.

c) Tercera sesión. En el tercer día los participantes completaron su última sesión de entrenamiento y después de eso realizaron el post-test. Este post-test incluía uno para la posición inicial (Test de identificación de sílabas en posición inicial (CV) 2 y Test de identificación de sílabas en posición final (VC) 2). Cada uno presentaba 180 estímulos y ambos se enmascararon a 15dB SNR.

4

Análisis de datos

4.1. Experimento I: Test general de consonantes inglesas

La distribución de la variabilidad para el análisis de datos del Experimento I fue como sigue:

- Variable dependiente: porcentaje correcto de identificación de las consonantes en las posiciones CV y VC.
- Variables independientes: nivel de SNR, modo de articulación, lugar de articulación, sonoridad, apertura, aspiración y posición silábica.

Para la obtención de resultados se agruparon los porcentajes de acierto en la identificación de las consonantes atendiendo a los diferentes rasgos articulatorios de los sonidos:

- Modo de articulación
- Lugar de articulación
- Sonoridad
- Apertura
- Aspiración

Respecto a los análisis estadísticos se han realizado diferentes pruebas atendiendo a la estadística descriptiva y a la estadística inferencial. Los resultados de los estadísticos descriptivos utilizados nos dan información importante sobre la dificultad en la identificación de consonantes, desde los parámetros arriba expuestos. Los resultados de los estadísticos inferenciales nos apuntan las tendencias que existirían en caso de aumentar la muestra, ampliar las variables o aplicar el modelo de estudio a otros grupos de población. De esta manera nos ayudan a entender futuras líneas de investigación.

Por un lado, se ha realizado un análisis descriptivo de los datos para todas las clasificaciones de consonantes para hallar la media de los resultados que han obtenido peores resultados en la percepción de los sonidos presentados. En este apartado se muestran los resultados en tablas atendiendo a los siguientes parámetros: modo de articulación, lugar de articulación y sonoridad. Además, los resultados de estos estadísticos descriptivos están detallados en base al nivel de enmascaramiento o SNR y posición en la sílaba (ataque o coda).

Además, en las tablas se han resaltado los resultados correspondientes a las medias más bajas en las que se ha cometido un porcentaje de aciertos de consonantes inferior al 75%.

MODO DE ARTICULACIÓN:

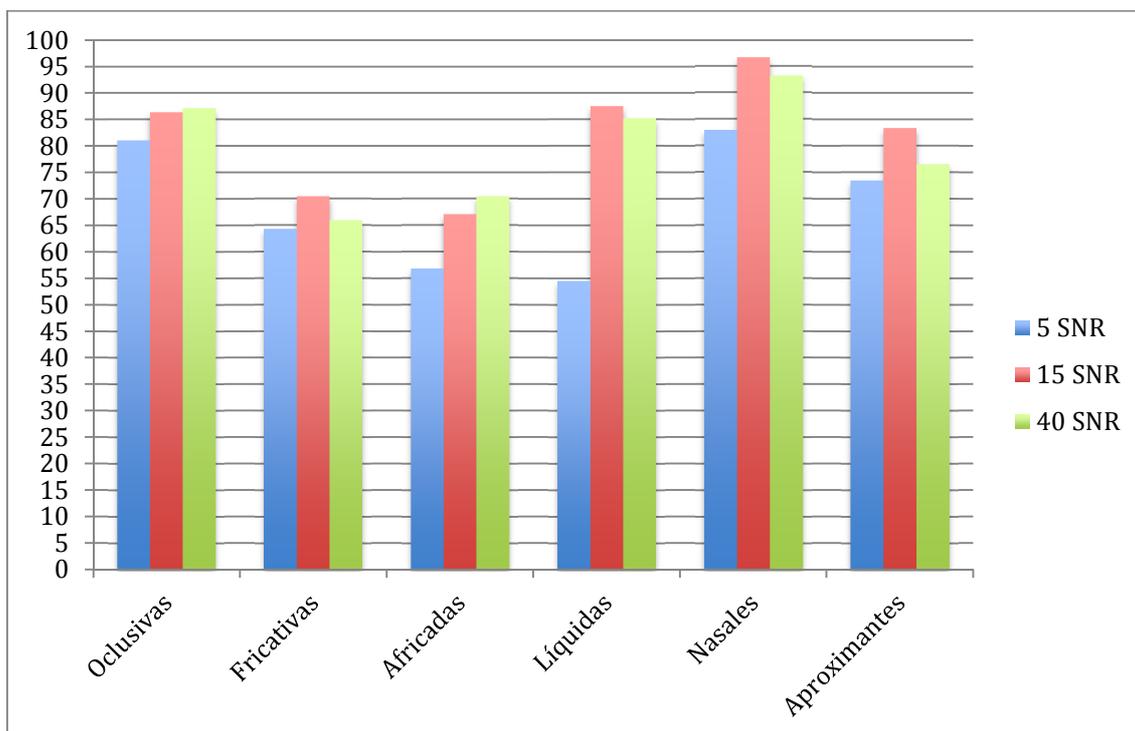


Figura 2. Medias atendiendo al modo de articulación en posición inicial

La figura 2 describe el modo de articulación en posición inicial. Como vemos, las medias más bajas corresponden a las consonantes fricativas y africadas en los tres niveles de SNR. En 5dB SNR, además de las consonantes fricativas y africadas, también encontramos un alto nivel de errores en las consonantes líquidas y aproximantes. En general, sin embargo, apreciamos que los mejores resultados aparecen cuando el enmascaramiento por ruido es de 15dB SNR, siendo ligeramente peores para situaciones de silencio (40dB SNR) o mayor ruido apreciable (5dB SNR).

En cuanto a la media aritmética, las fricativas y las africadas se alejan de los valores generales de la muestra.

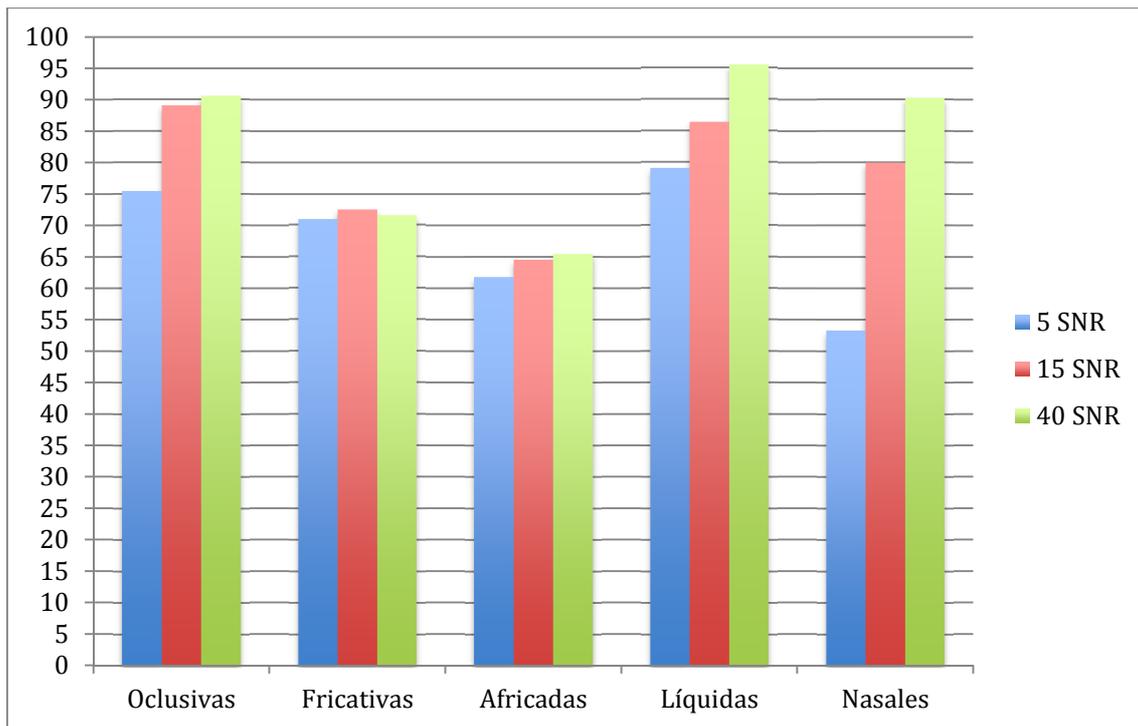


Figura 3. Medias atendiendo al modo de articulación en posición final

La figura 3 muestra los resultados del modo de articulación de las consonantes en posición final. Los sonidos cuyo porcentaje de aciertos es menor son los fricativos y africados en 40dB y 15dB SNR y los sonidos fricativos, africados y nasales en 5dB SNR. Por oposición a la identificación de consonantes en posición inicial, los resultados de los tests de identificación de consonantes en posición final de sílaba muestran que a mayor ruido de fondo, peor identificación de consonante, con especial atención a las nasales.

En cuanto a la media aritmética, las fricativas y las africadas se alejan de los valores generales de la muestra.

LUGAR DE ARTICULACIÓN:

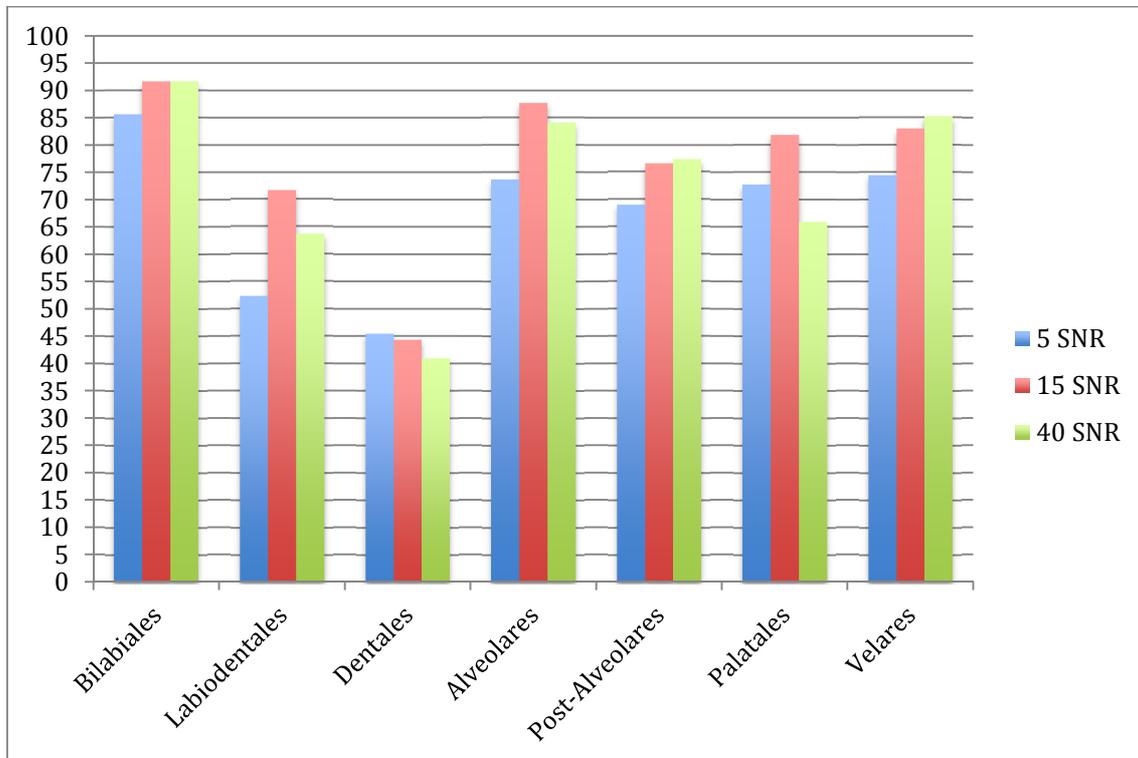


Figura 4. Medias atendiendo al lugar de articulación en posición inicial

La figura 4 describe el lugar de articulación en posición inicial. Observamos que las consonantes con medias más bajas en 40dB SNR son las labiodentales, dentales y palatales. En 15dB, las consonantes labiodentales y dentales son las que tienen las medias más bajas, y, por último, en 5dB SNR, las consonantes labiodentales, dentales, alveolares, post-alveolares, palatales y velares tienen las medias más bajas. La clasificación de los resultados hecha desde la perspectiva del lugar de articulación de las consonantes en posición inicial es bastante parecida a la hecha desde la perspectiva del modo de articulación: el nivel de ruido 15dB es el más estable en los resultados.

En cuanto a la media aritmética, las (labio-)dentales se alejan de los valores generales de la muestra.

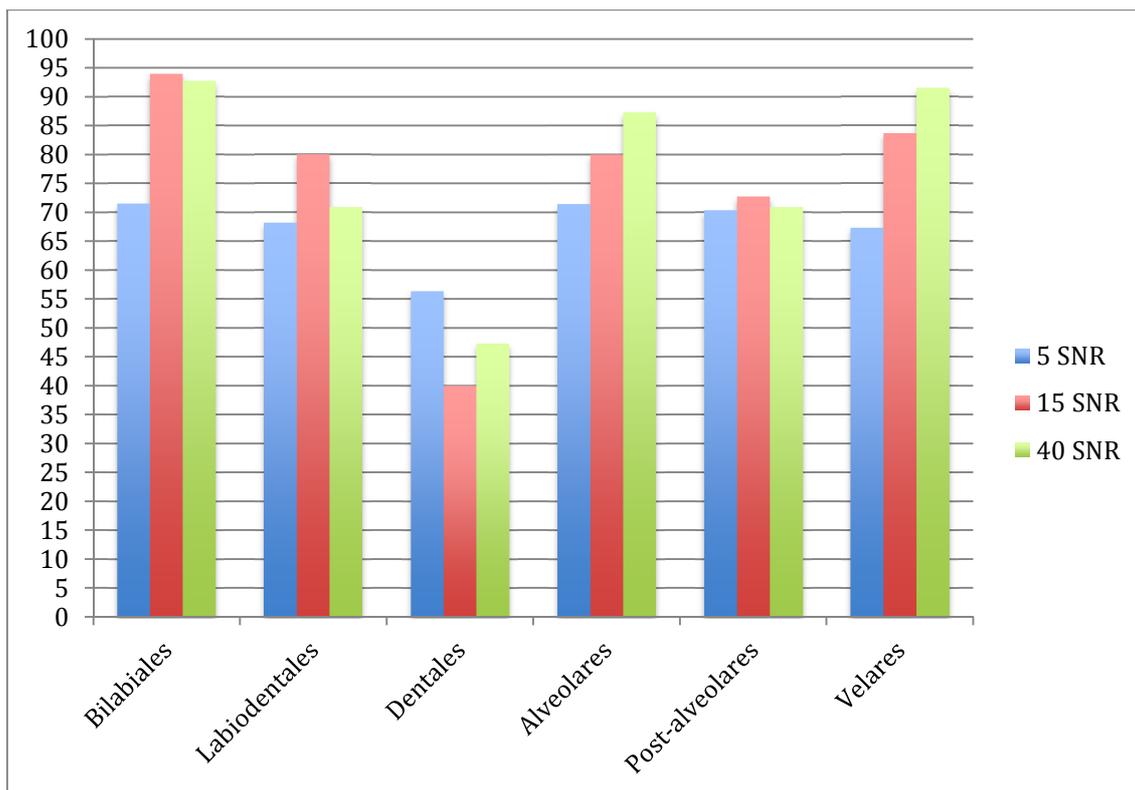


Figura 5. Medias atendiendo al lugar de articulación en posición final

La figura 5 muestra los resultados de los estadísticos descriptivos correspondientes al lugar de articulación en posición final. Podemos observar que las consonantes con menor número de aciertos son las labiodentales, dentales y post-alveolares en 40dB SNR. En 15dB, las consonantes con menos número de aciertos son las dentales y las post-alveolares y en 5dB SNR, las consonantes con menor número de aciertos son las bilabiales, labiodentales, alveolares, post-alveolares, dentales y velares. A excepción de la posición dental, los resultados de identificación de consonantes observados desde la perspectiva del lugar de articulación, posición final, es similar a la del modo: cuanto mayor es el ruido de fondo, peor es el porcentaje de identificación.

En cuanto a la media aritmética, las (labio-)dentales se alejan de los valores generales de la muestra.

SONORIDAD:

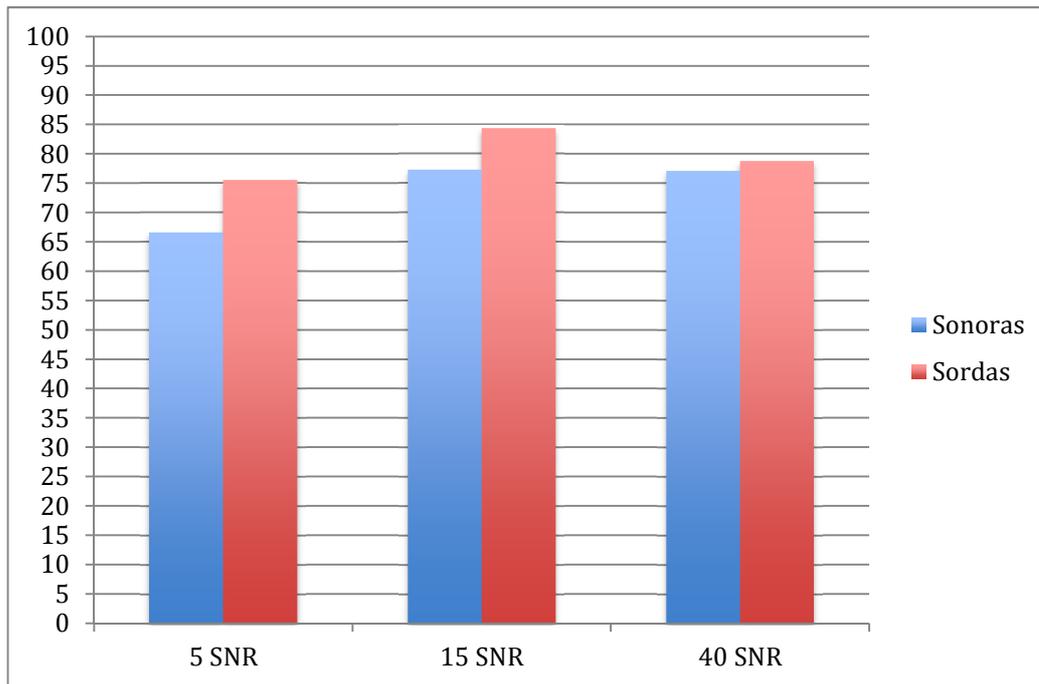


Figura 6. Medias atendiendo a la sonoridad en posición inicial

En la figura 6 observamos los resultados de los estadísticos descriptivos relacionados con la sonoridad en posición inicial. Las medias, aunque altas, son similares en las consonantes sonoras y sordas en 40dB y 15dB SNR. Sin embargo, la sonoridad en 5dB se ve afectada disminuyendo el número de aciertos en ambas categorizaciones (consonantes sonoras y consonantes sordas), correspondiendo la media más baja a las consonantes sonoras.

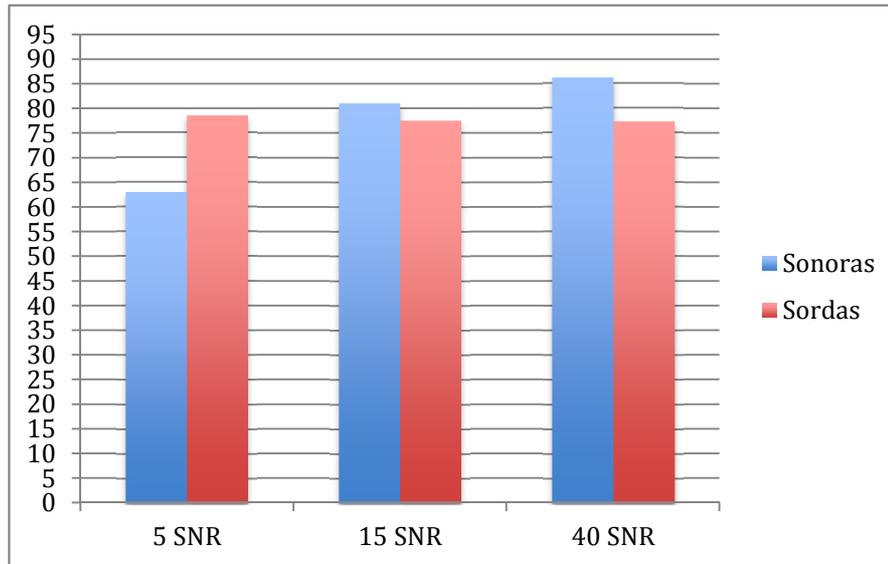


Figura 7. Medias atendiendo a la sonoridad en posición final

La figura 7 refleja los estadísticos descriptivos de la sonoridad en posición final. Al igual que en posición inicial, encontramos medias similares a excepción de las consonantes sonoras en 5dB SNR que es donde aparece el número de aciertos más bajos.

Los aspectos más importantes a tener en cuenta después de la observación de estos resultados podrían ser los siguientes:

- 1) Las consonantes fricativas y africadas obtienen peores resultados en los tests de identificación.
- 2) Las consonantes (labio-)dentales obtienen peores resultados en los tests de identificación.
- 3) El ruido de fondo 15dB SNR presenta mejores tendencias en resultados de identificación ligeramente por encima de valores considerados de “quietud” (40dB, situación de una conversación en tono conversacional) y por encima de valores ya considerados de dificultad perceptiva por ruido de fondo (5dB).
- 4) La sonoridad no parece ser determinante en los resultados de identificación de consonantes, tanto en posición inicial como final.

A tenor de lo resumido en el párrafo anterior, se estimó necesario profundizar el análisis estadístico de las muestras, para conseguir los siguientes objetivos:

- 1) Comprobar que la tendencia a mejor porcentaje de acierto en la identificación de consonantes con 15dB SNR era significativa para toda la muestra.

- 2) Comprobar que, en caso de aumentar la muestra (de población o de estímulos), los resultados serían igualmente significativos.
- 3) Comprobar si la posición en la sílaba es factor determinante para la identificación de las consonantes inglesas por no nativos.

Los tests de estadística inferencial tuvieron en cuenta el tamaño de la muestra y la significatividad (*p-value*). Los tests aplicados fueron ANOVA mixto, ANOVA de una vía para las muestras paramétricas (con distribución normal). Para los datos no paramétricos se usaron los tests H de Kruskal Wallis y U de Mann Whitney (el equivalente no paramétrico para dos muestras independientes al test *t* de Student. Para situar el análisis de resultados nos hemos fijado fundamentalmente en los siguientes estadísticos:

- 1) ***P-value*** (pruebas paramétricas) o H de Kruskal Wallis/ U de Mann Whitney (pruebas no paramétricas) para comprender que los estadísticos aplicados son significativos y que no han ocurrido al azar. **El valor de P hace referencia a la significación cuyo valor se interpreta como significativo si $p < .05$.**
- 2) **Eta_p^2** (tamaño de la muestra): si es grande, consideramos que el número total de muestras tomadas ha sido suficiente para llegar a resultados significativos. Si *Eta* es pequeño, entonces los resultados aquí obtenidos cambiarían significativamente de aumentar la muestra. **El valor de Eta_p^2 se interpreta como pequeño 0.01 - 0.06; medio 0.06 - 0.14; y grande > 0.14 .**

CONSONANTES OCLUSIVAS:

Al realizar ANOVA mixto no se obtuvieron resultados significativos entre la interacción de la posición del sonido en posición inicial o final y el SNR (*p-value*=0,152). No obstante, se realizó un análisis post-hoc de esta misma interacción y observamos que al analizar las comparaciones dos a dos sí existen diferencias significativas en el porcentaje de respuestas correctas en posición final entre 40dB y 5dB SNR (*p-value*=0,025) y entre 15dB y 5dB SNR (*p-value*=0,05).

Al analizar el efecto principal del ruido de fondo se observó $p\text{-value} = 0,025 < 0,05$. Este resultado demuestra que existen diferencias significativas en el porcentaje de respuestas correctas de consonantes oclusivas en función del nivel de enmascaramiento. Además, esto se ve reforzado con el resultado del tamaño del efecto el cual es grande ($\text{Eta}_p^2 > 0,217 > 0,14$).

CONSONANTES FRICATIVAS:

Al analizar el efecto principal de la posición (consonantes en posición inicial y en posición final) y su interacción con ruido de fondo concluimos que no resulta significativo el efecto de la posición en las consonantes fricativas ($p\text{-value} = 0,078$) lo que significa que la posición inicial o final no afecta a la percepción de las consonantes fricativas. Además, tampoco resulta significativo el efecto de interacción ($p\text{-value} = 0,745$), por lo que puede afirmarse que el porcentaje de respuestas correctas en posición inicial y final no es significativamente diferente en las tres categorías de enmascaramiento (5dB, 15dB y 40dB SNR). No obstante, debemos tener en cuenta el tamaño del efecto siendo este grande en la interacción entre la posición y el SNR ($\text{Eta}_p^2 > 0,19 > 0,14$) por lo los resultados obtenidos no dependen del número de datos en la muestra y seguiría la misma tendencia para otros contextos y otras poblaciones.

Por otro lado, los resultados obtenidos al analizar el nivel de ruido observamos que el resultado tampoco es significativo ($p\text{-value} = 0,766$). No obstante, el tamaño del efecto es grande ($\text{Eta}_p^2 > 0,18 > 0,14$).

CONSONANTES BILABIALES:

El análisis del efecto principal de la posición (inicial y final) y su interacción con el ruido de fondo en las consonantes bilabiales demuestra que, por un lado, el efecto de la posición no es significativo ($p\text{-value} = 0,101$). Sin embargo, sí resulta estadísticamente significativa la interacción de la posición inicial y final con el nivel de ruido en sus tres niveles de SNR ($0,006 < 0,05$) teniendo un tamaño del efecto grande ($\text{Eta}_p^2 > 0,291 > 0,14$).

También resulta estadísticamente significativo el resultado del porcentaje de respuestas correctas de consonantes bilabiales en función del nivel de SNR con un tamaño de efecto grande ($\eta_p^2 > 0,498 > 0,14$).

CONSONANTES LABIODENTALES:

En el análisis del efecto de la posición de las consonantes labiodentales obtuvimos un resultado significativo ($0,011 < 0,05$) con un tamaño del efecto grande ($\eta_p^2 > 0,195 > 0,14$). No obstante, el efecto de la interacción entre la posición y el nivel de SNR no es estadísticamente significativo ($0,623$).

Por otro lado, el nivel de SNR no son estadísticamente significativos ($p\text{-value}=0,144$) por lo que la modificación del SNR no tiene un impacto en la percepción de las consonantes labiodentales.

CONSONANTES ALVEOLARES:

Los resultados obtenidos del análisis del efecto principal de la posición (inicial y final) y su interacción con el SNR demuestran que no son significativos ($p\text{-value}=0,096$). Tampoco es significativo el efecto de la posición de las consonantes alveolares ($p\text{-value}=0,261$).

CONSONANTES VELARES:

En las consonantes velares, los factores intra-sujetos son las consonantes alveolares en posición inicial y final y los factores inter-sujetos son los tres niveles de SNR (40dB, 15dB y 5dB SNR).

El análisis del efecto de la posición en las consonantes velares no resultó ser significativo ($p\text{-value}=0,976$). No obstante, si nos fijamos en el efecto de la interacción de la posición y el SNR, sí comprobamos que, en este caso, el resultado es estadísticamente significativo ($p\text{-value}=0,038$) y, además, el tamaño del efecto es grande ($\eta_p^2 > 0,195 > 0,14$).

Además, también se encontraron diferencias significativas en el porcentaje de respuestas correctas de consonantes velares en función del nivel de SNR, con un tamaño de efecto grande ($\text{Eta}_p^2 > 0,416 > 0,14$).

CONSONANTES SONORAS:

Al analizar el efecto de la posición, comprobamos que este no resulta estadísticamente significativo ($p\text{-value}=0,093$). No obstante, el efecto de la interacción entre el SNR y la posición sí resulta significativo ($p\text{-value}=0,021$), teniendo, además, un tamaño del efecto grande ($\text{Eta}_p^2 > 0,227 > 0,14$).

También existen diferencias significativas en el porcentaje de respuestas correctas de consonantes velares en función del nivel de SNR ($p\text{-value}=0,001$), con un tamaño de efecto grande ($\text{Eta}_p^2 > 0,368 > 0,14$).

CONSONANTES APROXIMANTES POSICIÓN INICIAL:

Tras realizar ANOVA de una vía y en relación al contraste de las consonantes aproximantes en posición inicial, se analizaron las diferencias en función del nivel de SNR (5dB, 15dB y 40dB). A nivel descriptivo, los datos obtenidos fueron para 40dB tenían un porcentaje de aciertos de 76,51%, para 15dB encontramos un porcentaje de aciertos de 83,33% y para 5dB un porcentaje de aciertos de 73,48%. Se observa cómo la mejor ejecución se realizó en 15dB siendo en 5dB dónde se obtuvo una peor ejecución.

Al obtener los resultados de la ANOVA de una vía, nos permitió conocer si existían diferencias estadísticamente significativas entre los tres grupos de SNR. Los resultados de la prueba indican que no se encontraron diferencias estadísticamente significativas entre los tres grupos ($p\text{-value}=0,317$). Aunque el tamaño del efecto es grande ($d > 1,194 > 0,80$).

CONSONANTES PALATALES POSICIÓN INICIAL

En relación al contraste de las consonantes palatales en posición inicial, se analizaron las diferencias en función del nivel de SNR (5dB, 15dB y 40dB). A nivel descriptivo, los datos obtenidos fueron para 40 SNR tenían un porcentaje de aciertos de 65,91%, para 15dB encontramos un porcentaje de aciertos de 82% y para 5dB un porcentaje de aciertos de 72,72%. Por tanto, se puede observar como la mejor ejecución se realizó, de nuevo, en 15dB, con un porcentaje medio de aciertos de 81,82%. Además, el nivel de 40dB fue donde se observó una peor ejecución.

Al obtener los resultados de la ANOVA de una vía, nos permitió conocer si existían diferencias estadísticamente significativas entre los tres grupos de SNR. Los resultados de la prueba indican que no se encontraron diferencias estadísticamente significativas entre los tres grupos ($p\text{-value}=0,38$). Aunque el tamaño del efecto es grande ($d > 1,00 > 0,80$)

Como se comentaba anteriormente, hubo una serie de variables (africadas, nasales y líquidas) que no siguieron una distribución normal. Se ejecutaron diferentes transformaciones (Logarítmicas y Box-Cox) con el objetivo de normalizar las variables, no obstante, el problema de normalidad no pudo ser subsanado. Es por ello, que se tuvo que aplicar una estrategia no paramétrica para el análisis de los resultados. Es importante que tengamos en cuenta, que esta estrategia es menos potente y perdemos información, ya que no nos permite analizar el efecto de interacción entre SNR y la posición de la consonante. Se han aplicado estos análisis por separado en las consonantes africadas, nasales, aspiradas y líquidas en posición inicial y final.

Por tanto, se aplicó la H de Kruskal Wallis, entendido como el homólogo no paramétrico de la ANOVA de una vía, que nos permite analizar las diferencias entre más de dos grupos, en este caso, entre los tres niveles de SNR (5dB, 15dB y 40dB). Posteriormente, se aplicó la U de Mann-Whitney para hacer comparaciones por pares con el objetivo de detectar entre qué grupos de SNR se hallaba la diferencia.

En los resultados de H de Kruskal Wallis y U de Mann Whitney de las consonantes mencionadas anteriormente se han tenido en cuenta los valores de P o *P values* y del

tamaño del efecto o r . El valor de P hace referencia a la significación asintótica cuyo valor se interpreta como significativo si $p < .05$. El tamaño del efecto o r se interpreta como pequeño $< .30$, medio $.30-.50$ y grande $> .50$.

CONSONANTES AFRICADAS, NASALES, LÍQUIDAS EN POSICIÓN INICIAL Y FINAL:

Al realizar el análisis estadístico comprobamos que se encontraron diferencias estadísticamente significativas en las consonantes nasales en posición inicial ($p\text{-value}=0,032$) y en posición final ($p\text{-value}=0,001$). Las consonantes líquidas en posición inicial ($p\text{-value}=0,002$) y en posición final ($p\text{-value}=0,001$).

Realizamos estadísticos de contraste de las consonantes analizadas en 40dB y 15dB SNR. Comparando los niveles de 40dB y 15dB sólo se encontraron diferencias estadísticamente significativas entre las consonantes nasales en posición final ($p\text{-value}=0,021$), siendo el tamaño del efecto mediano ($r = - 0,49$).

Realizamos también estadísticos de contraste de las consonantes analizadas en 40dB y 5dB. Se encontraron diferencias estadísticamente significativas entre las consonantes líquidas en posición inicial ($p\text{-value}=0,004$) con tamaño del efecto grande ($r = -.62$), consonantes líquidas en posición final ($p\text{-value}=0,003$) con tamaño del efecto grande ($r = -.61$), y consonantes nasales en posición final con tamaño del efecto grande ($r = -.80$).

Finalmente, realizamos estadísticos de contraste de las consonantes analizadas en 15dB y 5dB SNR. Se encontraron diferencias estadísticamente significativas entre las consonantes líquidas en posición inicial ($p\text{-value}=0,002$) con tamaño del efecto grande ($r = -.65$), consonantes nasales en posición inicial ($p\text{-value}=0,014$) con tamaño del efecto grande ($r = -.52$), y consonantes nasales en posición final ($p\text{-value}=0,001$) con tamaño del efecto grande ($r = -.74$).

En resumen:

- 1) El porcentaje de identificación de las consonantes en posición final es, en general, mejor que en posición inicial.
- 2) Las consonantes fricativas y africadas tienen peores resultados en tests de identificación en su conjunto.
- 3) La sonoridad no presenta grandes dificultades en su identificación.
- 4) El comportamiento entre grupos y para todas las consonantes es significativo fundamentalmente para el modo de articulación, en especial: fricativas, africadas, pero también nasales y líquidas. Esto quiere decir que la amplitud media (siendo amplitud baja las oclusivas y amplitud alta las vocales) es la que presenta peor significatividad entre las muestras, es decir, el comportamiento de los participantes ha sido más variado.
- 5) El efecto del ruido sobre las muestras es significativo especialmente en 15dB SNR. Esto quiere decir que cuando hay poco o mucho ruido los resultados son menos predecibles, por lo que el contexto estable de percepción es el medio (algo de ruido de fondo).

4.2. Experimento II: Entrenamiento auditivo intensivo. Efectos del contexto adverso (ruido de fondo) y de la combinación de estímulos auditivos (sílabas vs. oraciones).

En esta sección se muestran los resultados del experimento en entrenamiento. Los objetivos se perfilaron una vez analizados los datos del Experimento I, y su justificación fue básicamente la que sigue:

- 1) Comprobar el efecto del entrenamiento sobre las consonantes medidas como unidades en sí mismas sin tener en cuenta aspectos relativos a su articulación, en posición inicial y final. Para lo cual:
 - a. Comprobar el efecto del ruido de fondo sobre dicho entrenamiento.
 - b. Comprobar el efecto de la unidad lingüística entrenada (sílabas vs. palabras).

A tenor de los resultados obtenidos en el Experimento I, donde el nivel de ruido de fondo 15dB SNR resultó el significativo a través de todos los participantes y para todas las unidades, el enmascaramiento de control se realizó a dicho nivel de SNR. De hecho, los objetivos antes mencionados se analizaron con cuidado tal y como se detalla en la descripción del método siendo los grupos de población los siguientes:

- Grupo de control. Entrenamiento de sílabas (CV y VC) a 15dB.
- Grupo experimental I. Entrenamiento de sílabas (CV y VC) a -5dB.
- Grupo experimental II. Entrenamiento de palabras en oraciones a 15dB.

La variabilidad en el análisis se consideró tal y como se explica en el capítulo anterior, a saber:

- Variable dependiente: porcentaje correcto de identificación de las consonantes en las posiciones CV y VC.
- VARIABLES INDEPENDIENTES: nivel de SNR, y unidad lingüística entrenada (sílabas vs. oración).

Para el estudio estadístico, de todas las unidades silábicas incluidas en la experimentación, nos hemos centrado en 12 consonantes teniendo en cuenta su importancia desde el punto de vista contrastivo L1-L2. (No olvidemos que los participantes en este experimento, a diferencia de los que lo hicieron en el primero, son nativos del español).

Todos los resultados se midieron de acuerdo al porcentaje de aciertos. Por tanto, cada valor indica el grado correcto de identificación de los sonidos elegidos en el estudio. Para el análisis estadístico, se considera el valor de p o “P-value” estadísticamente significativo si $p < 0,05$ mientras que el valor de p se considera altamente significativo si $p < 0,01$. Respecto a las estadísticas de correlación, se considera el coeficiente de correlación de Pearson r^2 muy correlacionado si r^2 está cercano a 1.

Para organizar y aclarar los resultados y discusión de este experimento, en las siguientes subsecciones se comparan los resultados finales de cada grupo de entrenamiento. Esta organización ayudará a clarificar las hipótesis (entrenamiento intensivo en condiciones adversas o con enmascaramiento producen mejores resultados de acuerdo al pre-test y al

post-test; el entrenamiento intensivo con variabilidad en el input producen mejores resultados de acuerdo con el pre-test y el post-test). Además, también se intentará proporcionar una respuesta más profunda al planteamiento inicial que plantea demostrar que el entrenamiento auditivo intensivo mejora la adquisición fonológica de los estudiantes de una L2.

Los resultados muestran, para los 3 grupos de entrenamiento, el porcentaje de acierto en la identificación de consonantes oclusivas y fricativas/africadas, en posición inicial y final.

4.2.1 Grupo de control: entrenamiento de sílabas a 15dB SNR

Los participantes del grupo de control presentaron mejores resultados en la posición silábica inicial. La diferencia en el porcentaje de acierto es, sin embargo, significativa sólo en el caso de las consonantes oclusivas. Las consonantes fricativas/africadas tuvieron poco o nulo beneficio después de las sesiones de entrenamiento.

Los resultados muestran:

- 1) La identificación de consonantes oclusivas tras el entrenamiento mejora 15 puntos en las consonantes iniciales, pero empeora 8 puntos en el caso de las finales.
- 2) La identificación de consonantes fricativas/africadas no parece presentar beneficios del entrenamiento auditivo intensivo, no obstante, es muy inferior si lo comparamos con los resultados de identificación de las oclusivas.

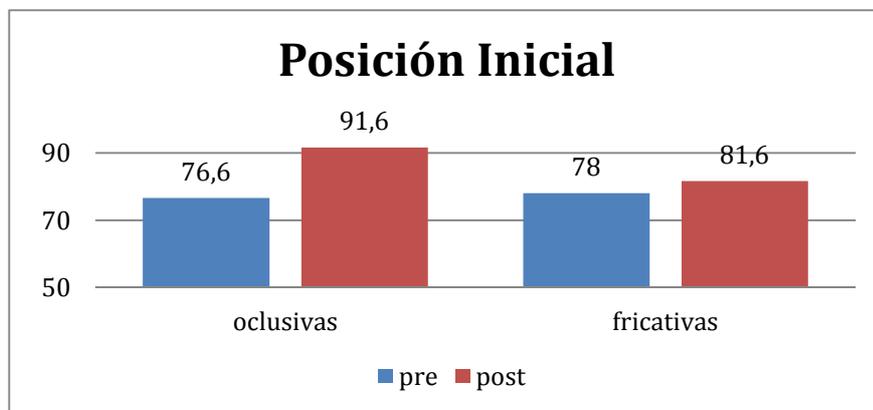


Figura 8: Porcentaje de aciertos consonante inicial (Grupo Control)

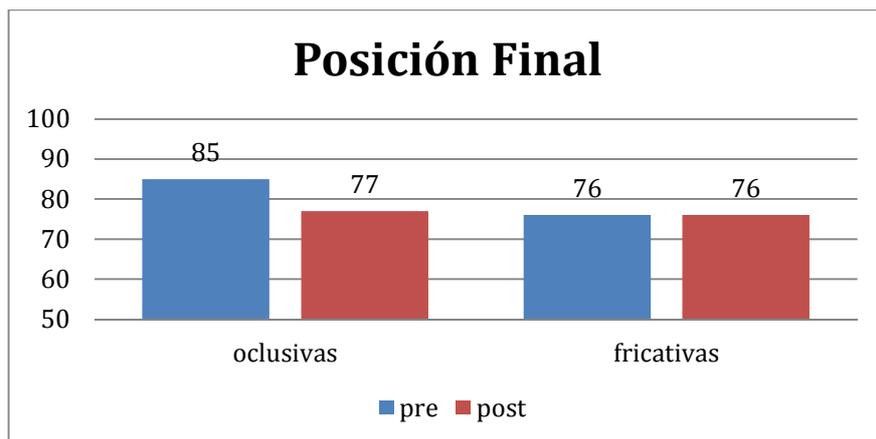


Figura 9: Porcentaje de aciertos consonante final (Grupo Control)

Los efectos bajo el entrenamiento en buenas condiciones de ruido demuestran que los resultados detallados anteriormente para las consonantes en posición inicial son estadísticamente significativos ($p=0,05$ obtenidos por un t-test pareado a una cara¹). Esto quiere decir que estos resultados parecen indicar que incluso si aplicamos un SNR alto (silencioso) a los estímulos en el entrenamiento auditivo, los resultados mejoran.

Al contrario de lo que sucede en la posición silábica inicial, parece ser que el entrenamiento con buenas condiciones de SNR o enmascaramiento no tiene un efecto significativo. Esto puede afirmarse después de aplicar un t-test pareado a una cara sobre el porcentaje de aciertos del pre-test y el post-test de las sílabas en posición final ($p = 0,15$).

Finalmente, se han contrastado los resultados de las sílabas en posición inicial y final y se ha encontrado que la correlación de Pearson producto-momento o “Pearson Product Moment Correlation”². El valor de r^2 es 0,9008. Esto muestra un efecto positivo en la correlación entre lo que ocurre en la posición inicial y lo que ocurre en la posición final.

¹ Un test pareado compara dos ejemplos en casos en los que cada valor en un ejemplo tenga un compañero natural en el otro. Un t-test pareado se centra en las diferencias entre los valores pareados en dos ejemplos, tiene en cuenta la variación de los valores dentro de cada ejemplo y produce una única cifra que se conoce como valor de t o *t-value*.

²La correlación de Pearson producto-momento o “Pearson's product moment correlation”, r , representa la fuerza de la relación lineal entre dos variables métricas, tales como la extraversión y la inteligencia. Los valores se escalan de -1 a +1. Es necesario un diagrama de puntos para interpretar r . La proporción de la variabilidad considerada en la relación lineal viene proporcionada por r^2 . La correlación entre grupos de datos se mide cómo de bien están relacionados (Nagelkerke, N. J. D. 1991: 691).

Esto da significatividad al hecho de que la pérdida de acierto en las oclusivas en posición final está correlacionada con el porcentaje de subida en posición inicial. Se produce aprendizaje porque hay cambio de patrones.

4.2.2. Grupo Experimental I: entrenamiento con sílabas a -5dB SNR

Los participantes en el grupo experimental I mostraron una tendencia completamente estable entre las dos posiciones (inicial vs. final) y los dos grupos de consonantes (oclusivas vs. fricativas/africadas). La media en posición inicial fue 85,75% en el pre-test y 87% en el post-test. La media aritmética en posición inicial fue 83,3% en el pre-test y 84% en el post-test.

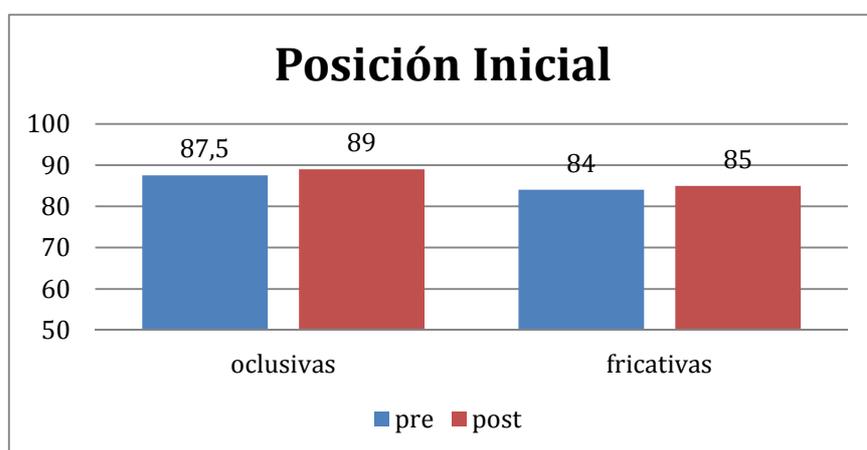


Figura 10: Porcentaje aciertos consonante inicial (Grupo experimental I)

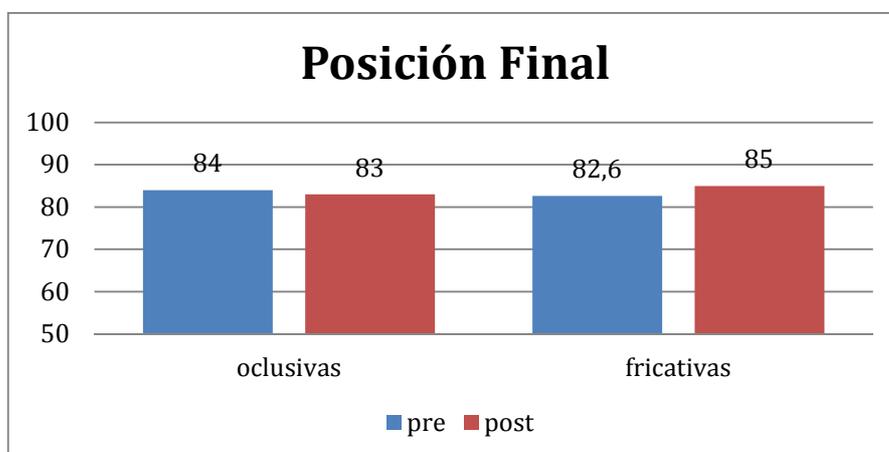


Figura 11: Porcentaje aciertos consonante final (Grupo experimental I)

Los efectos del entrenamiento en condiciones de ruido discutidos en los resultados anteriores en posición inicial son, sin embargo, no significativos estadísticamente ($p = 0,32$ en el t-test pareado a una cara). Esto quiere decir que estos resultados parece indicar que si se aplica SNR a los estímulos en el entrenamiento auditivo, la percepción de las consonantes iniciales inglesas se verá afectada tal y como se demuestra en las tareas de identificación.

Parece que el entrenamiento en posición final con condiciones de ruidos no tiene efecto significativo. Esto puede aplicarse después de aplicar un t-test pareado a una cara del porcentaje de aciertos del pre-test y el post-test ($p = 0,44$).

Por último, se han contrastado los resultados de los test en posición inicial y final y se ha aplicado la correlación de Pearson. El valor de r^2 es 0,9065 el cual muestra una alta correlación positiva entre la posición inicial y final. Esto significa que resultados altos en la variable X produce resultados altos en la variable Y y viceversa.

4.2.3. Grupo experimental II: entrenamiento con palabras a 15dB SNR

El grupo Experimental II recibió un entrenamiento basado en oraciones. Los resultados del test pre-post en identificación de consonantes sigue, sin embargo, una tendencia similar al mostrado por el grupo de control, que entrenó sin ruido de fondo pero sólo con la presentación continuada de sílabas. Las figuras 12 y 13, respectivamente para la identificación en posición inicial y final, muestran una mejoría significativa para las oclusivas de 6 puntos, que se convierte en pérdida en posición final (-5 puntos de beneficio). Los beneficios del entrenamiento en el caso de fricativas-africadas no es evidente (2 puntos en posición inicial y 6 puntos en posición final, pero sin llegar a porcentajes altos de en torno a 90%).

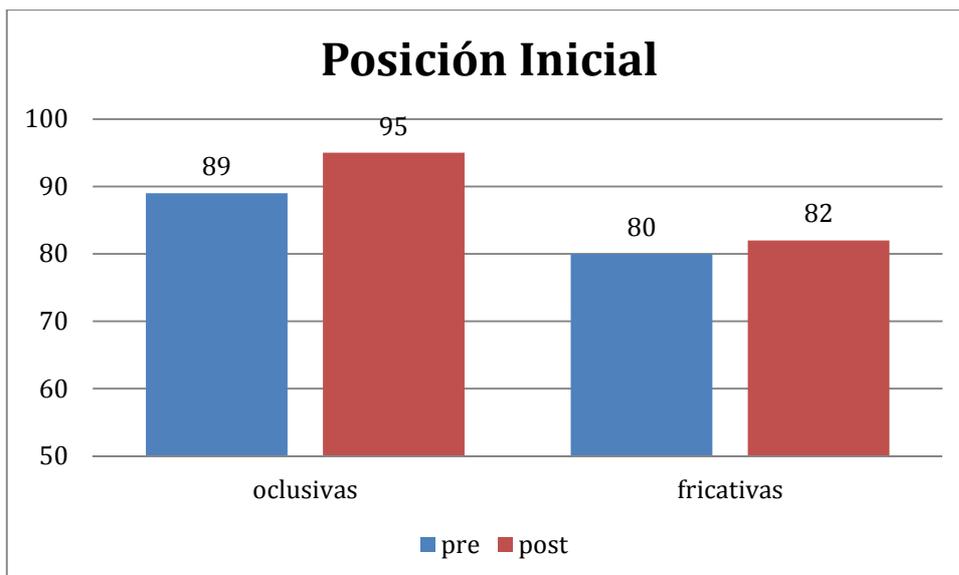


Figura 12: Porcentaje acierto consonante inicial (Grupo experimental II)

Después de aplicar un t-test pareado a una cara sobre el porcentaje de aciertos del pre-test y el post-test en posición inicial se observa que $p = 0,19$. Por tanto, el análisis de los resultados de todas las consonantes no es estadísticamente significativo.

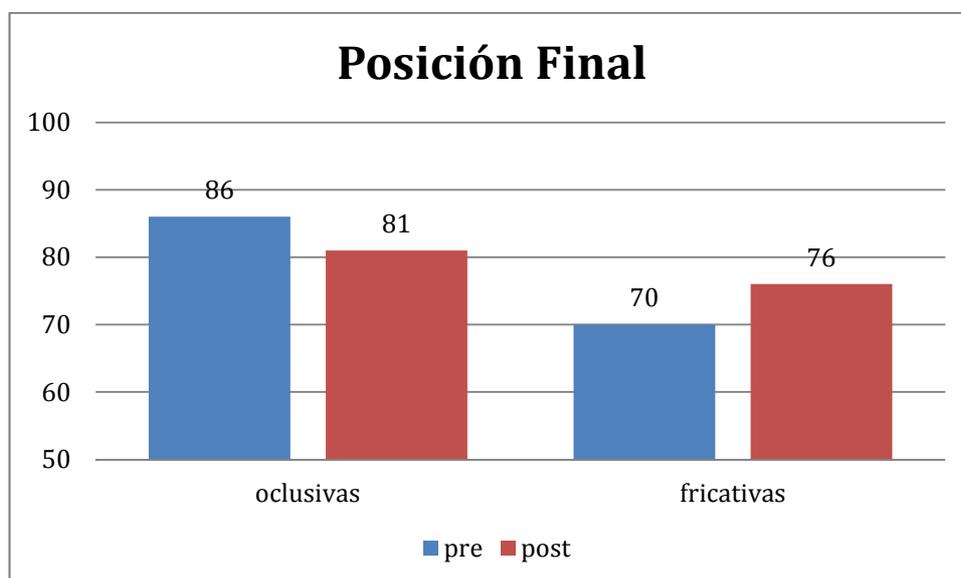


Figura 13: Porcentaje acierto consonante final (Grupo experimental II)

Después de aplicar un t-test pareado a una cara en el porcentaje de aciertos entre el pre-test y el post-test en posición final se observa que $p = 0,44$ y, por tanto, el análisis de los resultados de todas las consonantes no es estadísticamente significativo.

Por último, se han contrastado los resultados de los test en posición inicial y final y se ha aplicado la correlación de Pearson. El valor de r^2 es 0,9598 el cual muestra una alta correlación positiva entre la posición inicial y final. Esto significa que resultados altos en la variable X produce resultados altos en la variable Y y viceversa.

A continuación voy a realizar un resumen de los resultados del Experimento II, comparando el grupo de control con cada uno de los grupos experimentales. En la figura 12 la comparación en porcentajes del grupo de control (sílabas entrenadas a 15dB SNR) con el grupo Experimental I (sílabas entrenadas a -5dB SNR).

Para realizar dicha comparativa me fijo en las variables siguientes:

- a) Presencia de armónicos vs. ausencia de armónicos (para entender oclusivas vs. fricativas/africadas)
- b) Posición inicial vs. posición final
- c) Entrenamiento con ruido de fondo
- d) Entrenamiento con sílabas vs. oraciones

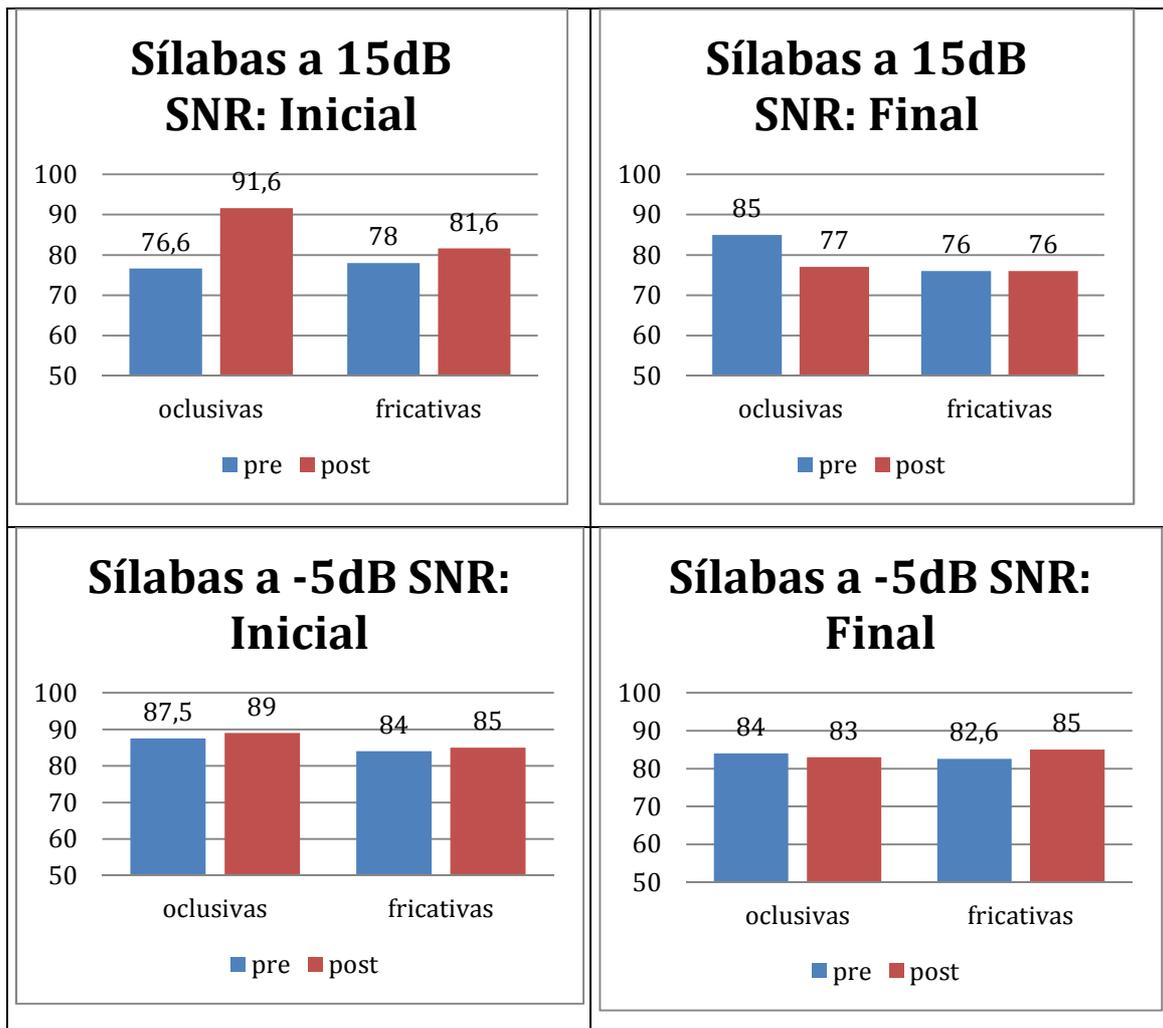


Figura 14: Comparativa Grupo de Control vs. Grupo Experimental I

- a) Variable estructura de armónicos (oclusivas=estructura armónica vs. fricativas/africadas=no estructura armónica): Según se aprecia enfrentando los resultados, el ruido de fondo (recordemos que -5dB SNR es un ruido especialmente elevado que tapa la señal), marca la complejidad en la identificación de consonantes inglesas atendiendo a su modo de articulación. Las consonantes oclusivas se caracterizan por presentar bajos niveles de amplitud, pero armonía en frecuencias, mientras que las consonantes fricativas muestran mayores niveles de amplitud, pero falta total de armonía frecuencial. Cuando se aplica ruido de fondo elevado, los resultados de identificación se igualan. Sin ruido de fondo, las oclusivas se identifican mucho mejor que las fricativas/africadas.
- b) Variable posición en la sílaba: En todos los casos (exceptuando las fricativas/africadas del grupo Experimental I), la posición inicial presenta mayores beneficios tras el entrenamiento que la posición final.

c) Variable estímulos con ruido de fondo: El grupo Experimental I presenta beneficios en el entrenamiento de fricativas/africadas cuando los estímulos están enmascarados a -5dB SNR.

La figura 20 muestra la comparativa en los resultados pre-post tras entrenamiento enfrentando el Grupo de Control (entrenamiento con sílabas a 15dB SNR) con el Grupo Experimental II (entrenamiento con oraciones a 15dB SNR).

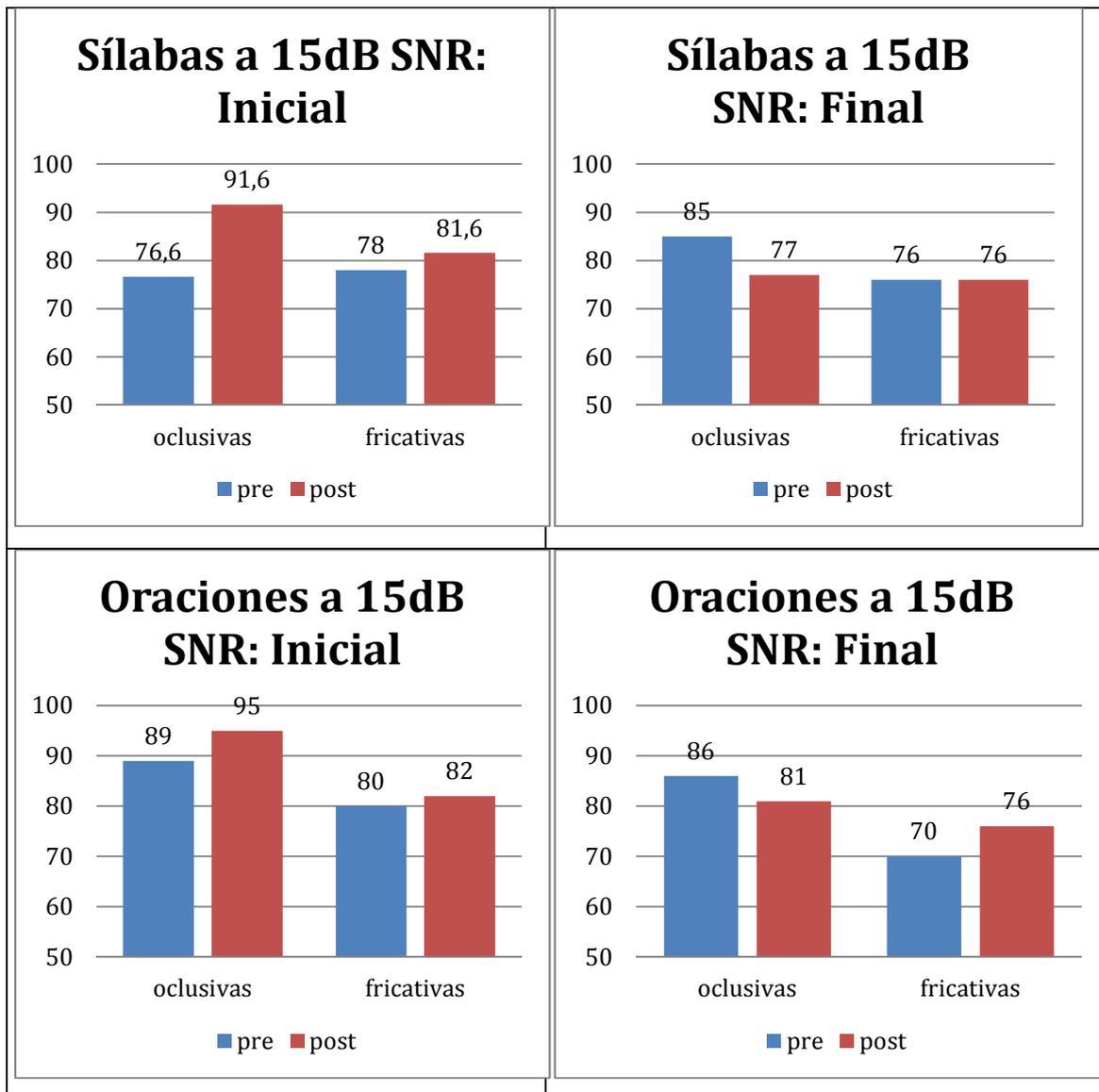


Figura 15: Comparativa Grupo de Control vs. Grupo Experimental II

- a) Variable estructura de armónicos (oclusivas=estructura armónica vs. fricativas/africadas=no estructura armónica): El entrenamiento con ruido de fondo a 15dB SNR confirma que las consonantes oclusivas se benefician mucho más del entrenamiento intensivo que las fricativas/africadas.
- b) Variable posición en la sílaba: En todos los casos (exceptuando las fricativas/africadas del grupo Experimental II), la posición inicial presenta mayores beneficios tras el entrenamiento que la posición final.
- c) Variable estímulos sílaba vs. oración: El grupo Experimental II se beneficia ligeramente más de un entrenamiento que incluye presentación intensiva a oraciones si lo comparamos con el grupo de control, que fue entrenado exclusivamente con estímulos silábicos.

Por último, la tabla siguiente muestra un resumen de los estadísticos t y r^2 , resultados de aplicar un test t pareado y correlación Pearson, respectivamente en el contraste posición inicial vs. posición final de sílaba. El estadístico t nos informa sobre si los resultados son significativos (es decir, si los mismos se deben al azar o incluso si cambiarían de aumentar la muestra) mientras que la correlación Pearson nos informa sobre si los resultados del entrenamiento se correlacionan entre sí (al mejorar en una posición se empeora en la otra y viceversa).

Tabla t-test y Pearson Correlation para los 3 grupos.

	r^2	t -test	
		Inicial	Final
Grupo de Control	0,9008	0,05	0,15
Grupo Experimental I	0,9065	0,32	0,44
Grupo Experimental II	0,9598	0,19	0,44

5

Discusión

La investigación en el campo de la percepción de L2 ha demostrado que el principal factor que incide en la categorización fonológica de los nuevos sonidos es la relación que existe con la L1 (Flege, 1995, Best et al. 2001, Kuhl en Iverson 1995, Kuhl 2004). En este sentido, se ha considerado normalmente el factor transferencia (positivo o negativo) como elemento fundamental para explicar y predecir cómo se adquieren los sonidos de la L2. Otros estudios han intentado abordar el problema desde otras perspectivas. Una de las líneas de investigación más recurrente es aquella que intenta comprender cómo afecta el ruido de fondo a la señal. Así, autores como Nabelek y Donahue (1984), Cutler et al. (2004), García Lecumberri y Cooke (2006), García Lecumberri et al. (2010), entre otros, describen cuáles son los porcentajes de acierto en tareas que implican creación de categorías fonológicas cuando los estímulos utilizados en la investigación aparecen enmascarados con diferentes niveles de SNR. Por último, en los últimos años son cada vez más frecuentes aquellos estudios que intentan llegar a conclusiones generalizables en relación a las características intrínsecas de los sonidos para entender cómo se produce la categorización fonológica. Entre otros destacamos los siguientes: Darcy et al. (2007), Werker y Tees (1984, 2002), Guion et al. (2000), Lalonde y Werker (1995).

En uno y otro caso, el estudio de la percepción de lenguas desde una disciplina psicoacústica se ha centrado en el análisis de los resultados a partir fundamentalmente

del porcentaje de acierto. Aquellos que tienen en consideración la clasificación de los sonidos, como hemos hecho en esta tesis, analizan los resultados fundamentalmente a partir de sus rasgos articulatorios, aunque en la mayor parte de los casos se limitan a considerar cada sonido como una categoría única en sí misma. Es decir, si el objetivo del estudio son las consonantes, los resultados suelen medirse en función del nivel de acierto de cada consonante. Son pocas las investigaciones que han intentado categorizar las variables del análisis de resultados en función de las características inherentes de las unidades fónicas. Existen estudios que, de hecho, hablan de la coarticulación vocal-consonante en posición final (Martin y Bunnell 1982, Öhman 1966, Redford y Diehl 1999) y miden los movimientos formánticos, pero esta variable acústica debe medirse atendiendo a múltiples niveles como frecuencia, tiempo o amplitud (Kapoor y Allen 2012), lo cual sería materialmente inmanejable en estudios con tantas unidades como el nuestro, ya que la variabilidad acústica de los materiales sería muy difícil abordar debido a la gran cantidad de datos y variables con las que habría que trabajar.

Es por ello que la mayor parte de los estudios que abordan el problema desde el análisis acústico se centran en pocos sonidos. El motivo principal es, como decimos, la dificultad metodológica que encontraríamos al tener que analizar un aspecto tan sumamente variable de la lengua: los rasgos acústicos son difícilmente generalizables, pues sobre ellos incide el hablante, la situación de habla, la rapidez, y multitud de otros factores que harían poco práctico su análisis estadístico. Este ha sido fundamentalmente el motivo por el que este estudio no incorpora la variabilidad acústica. A pesar de que existen algunos estudios que abordan la complejidad de la percepción de una L2 desde el enfoque acústico, para este proyecto de tesis no hemos considerado posible esta variabilidad. Nuestro interés ha sido tratar todos los contextos CVC en inglés.

Continuando con nuestra argumentación, observamos que la mayor parte de los estudios han centrado su atención en un par de sonidos difíciles que no existen en la L1 como es el caso recurrente del par /l/ y /r/ (Best y Tyler 2007, Bradlow et al. 1997, Bradlow et al. 1999, Goto 1971, Lively et al. 1993, Logan et al. 1991, MacKain et al. 1981, Mochizuki 1981, Sheldon y Strange 1982) y el impacto del SNR en los estímulos (Cutler et al. 2004, 2007, 2008). Si en nuestro estudio nos hubiésemos centrado en pocos sonidos difíciles, habríamos acabado diseñando un estudio sesgado desde su origen y no nos habría permitido llegar a conclusiones generales sobre la percepción del inglés.

Tenemos también que aceptar postulados más que validados como el que establece el modelo PAM en el cual se asume la existencia de rasgos comunes en los fonemas de todas las lenguas que hacen que se unan entre ellas y en este punto es donde se centra el oyente para formar sus categorías (Best 1993, Best 1994a, Best 1994b, Best 1995, Best et al. 2001, Best et al. 1988, Best y Tyler 2007). Si es cierto que existen rasgos comunes que luego se categorizan en unidades ligeramente diferentes, quizás es más productivo conocer el grado de dificultad en la percepción de los sonidos analizados en función de sus rasgos, para poder llegar a conclusiones más aplicables sobre la predicción en la creación de nuevas categorías fonológicas.

La hipótesis central en esta tesis doctoral parte del postulado de PAM el cual establece que existen en los sonidos de todas las lenguas una serie de rasgos que en cierta manera son universales. Es decir, puesto que se ha demostrado que el ser humano nace con la capacidad de distinguir todos los sonidos de la lengua, y que es a la edad aproximada de 6 meses cuando esta capacidad empieza a disminuir pues el niño se especializa en un único sistema, es también posible afirmar que, de conocer estos rasgos comunes, y a partir de los datos de la L1 y L2, se podrán predecir los patrones de adquisición de categorías fonológicas entre diferentes lenguas por oyentes nativos y no nativos.

Por otro lado, la formulación de la hipótesis en este estudio parte de los resultados encontrados en la investigación con diferentes niveles de SNR en la señal que apuntan a que los estímulos que se usen en una investigación de estas características deben incluir diferentes niveles de enmascaramiento. Se trata de un aspecto metodológico, pero importante, pues los resultados válidos y significativos de cualquier estudio de estas características deben reproducir, en la medida de lo posible, las condiciones que se dan en una situación comunicativa real. Cabe aquí decir que en nuestro estudio la población analizada es de normo-oyentes, por lo que el ruido no puede considerarse una variable en sí mismo. Es decir, en personas con deficiencias auditivas, el nivel de ruido puede afectar la correcta identificación y/o discriminación de sonidos. Es por ello que la investigación en el campo de personas con disminución auditiva, o con implantes cocleares, presenta una línea de interés específica en la que se intenta medir cuáles son los niveles de ruido óptimos para una y otra situación (el caso de los implantados es especialmente significativo, pues el dispositivo que se implanta puede presentar limitaciones a la hora de procesar ruido de fondo). En el caso de estudios como el que

aborda esta tesis, el nivel de SNR no tiene por qué considerarse como una variable independiente en sí misma, como objetivo imprescindible de estudio. Sin embargo, los resultados medidos en forma de porcentaje de acierto en la identificación de sonidos, como es nuestro estudio, sí tienen que incorporar la presentación de los estímulos a diferentes niveles de enmascaramiento. El objetivo metodológico no es otro sino el de encontrar los niveles idóneos, para evitar los denominados “efectos de nivel techo y nivel suelo”. Es decir, la experimentación debe diseñarse de tal manera que las conclusiones a las que se lleguen han de analizarse a partir de aquellos resultados que no fueron ni demasiado difíciles ni demasiado fáciles para los participantes. En línea con los principales estudios en el campo de la percepción de sonidos de la L2 por normo-oyentes (Nabelek y Donahue, 1984; Cutler et al., 2004; García Lecumberri, 2006, 2010), hemos elegido los niveles 5dB, 15dB y 40dB de SNR para el enmascaramiento de sonidos. Nuestras conclusiones vienen derivadas de los resultados de acierto (variable independiente) de los tests presentados al nivel de enmascaramiento que, estadísticamente, es significativo. En nuestro caso, 15dB. Esto se detalla más a continuación en esta sección.

Desde el punto de vista metodológico, esta tesis también es consecuente con los resultados hallados en los antecedentes, como es el más importante que se refiere a la importancia que tiene la L1 en la categorización de los sonidos de una L2. Es por ello que nuestra población alcanza hasta ocho L1s diferentes. No obstante, no hemos incorporado esta variable al análisis de los datos, pues hemos centrado nuestra atención en características propias del sonido o rasgos articulatorios, a saber:

- Lugar de articulación
- Modo de articulación
- Sonoridad
- Apertura
- Aspiración

Nuestro objetivo primordial se ha centrado en conocer cómo estas características (consideradas variables dependientes) interactúan o se correlacionan con otras variables, más en concreto:

- Posición: ataque y coda silábicos
- Nivel de SNR: 5dB, 15dB y 40dB

En definitiva, los resultados de nuestro estudio deben entenderse como un esfuerzo por saber si, al menos para el inglés, existen rasgos articulatorios que deban entenderse como más o menos difícil para su categorización por no nativos, y si la posición también interactúa para determinar este grado de dificultad.

Para articular toda la variabilidad de los datos en relación con la hipótesis de partida, hemos organizado su análisis en los siguientes objetivos:

- a) Conocer la relación entre el índice de precisión de identificación en la identificación de consonantes inglesas y el grado de SNR que se aplica a la señal.
- b) Conocer la relación entre el índice de precisión de identificación en la identificación de consonantes inglesas y su posición silábica (inicial y final).
- c) Conocer la relación entre el índice de precisión de identificación en la percepción de consonantes inglesas y las características intrínsecas de estos sonidos que se miden en términos de articulación (lugar de articulación y modo de articulación), sonoridad, amplitud (apertura de los resonadores y entendida en inglés como *sonority*) y aspiración.
- d) Conocer cómo afecta el entrenamiento auditivo intensivo realizado con estímulos enmascarados (15dB SNR) en el índice de precisión de identificación en la percepción de consonantes inglesas medidas en términos de la posición que ocupa en la sílaba (inicial o final).
- e) Conocer cómo afecta el entrenamiento auditivo intensivo realizado con estímulos enmascarados (15dB SNR) en el índice de precisión de identificación en la percepción de consonantes inglesas medidas en términos de articulación (lugar de articulación y modo de articulación), sonoridad, amplitud (apertura de los resonadores y entendida en inglés como *sonority*) y aspiración.
- f) Conocer cómo afecta el entrenamiento auditivo intensivo cuando se utilizan oraciones como estímulos en las tareas de entrenamiento vs. cuando se utilizan sílabas como estímulos en las tareas de entrenamiento.

Teniendo en cuenta nuestra hipótesis y objetivos hemos analizado nuestros datos atendiendo al nivel de SNR, a su posición silábica y a los rasgos articulatorios de los sonidos.

5.1. Cómo afecta la percepción del SNR en la muestra

Atendiendo a la forma en la que hemos tabulado los datos hemos analizado el SNR como una variable dependiente. En nuestro estudio hemos intentado averiguar cuál es el nivel de SNR más estable ya que un SNR alto produce resultados de techo y un SNR bajo produce resultado de suelo. Es necesario evitar estos resultados por cuestiones metodológicas ya que no serían relevantes. Hemos elegido 40 SNR, 15 SNR y 5 SNR tal y como hacen otros estudios (Friesen et al. 2001, García Lecumberri y Cooke 2006, Helfner y Huntley 1991, Helfer y Wilber 1990, van Winjgaarden 2001, van Winjgaarden et al. 2002) .

En las figuras 16, 17, 18, 19, 20, 21 y 22 se describe el comportamiento del SNR para las consonantes analizadas en cada figura. En las figuras observamos que el eje Y muestra los porcentajes de aciertos y el eje X la posición (inicial y final).

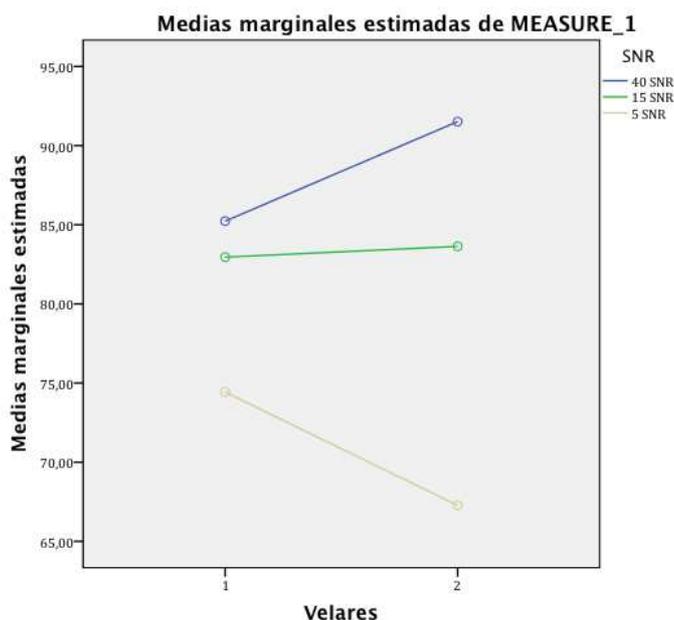


Figura 16. Comportamiento del SNR en las consonantes velares

La figura 16 muestra que 15dB SNR es el nivel más constante analizando los tests de identificación de manera pareada atendiendo a la posición (inicial vs. final). Por norma general, 40dB SNR es el nivel que tiene los resultados más altos y 5dB SNR es el que tiene los resultados de aciertos más bajos alejándose de 15dB y 40dB SNR (tal y como podemos comprobar en el resto de figuras del anexo 4).

Si nos fijamos en las figuras 17-22, observamos que el nivel de 40dB produce una peor identificación que el de 15dB SNR. Esto puede deberse al hecho de que, en condiciones óptimas de ruido, los oyentes no hacen realmente el proceso que harían en una situación normal porque las personas normo-oyentes estamos acostumbradas a escuchar en condiciones de ruido. Esto nos lleva a afirmar que el SNR tiene que estar presente en los estudios relacionados con la percepción de estímulos para que los resultados sean más consistentes. Este nivel de ruido de fondo parece acertado afirmar sería el de 15dB. Con este nivel, la experimentación en percepción de consonantes extranjeras se vería doblemente beneficiada. Por una parte, evitaríamos reproducir una situación de laboratorio en condiciones alejadas de la situación comunicativa normal y, por otra parte, mantendríamos la atención del participante en la tarea obligándole a reflejar sus auténticas capacidades perceptivas.

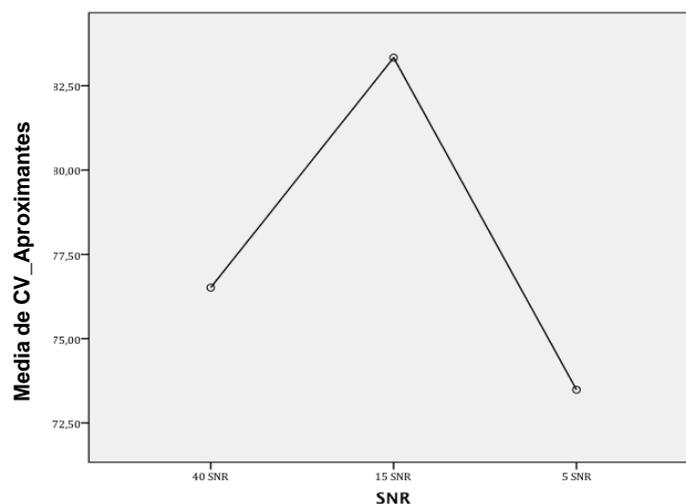


Figura 17. Comportamiento del SNR en las consonantes aproximantes en posición inicial

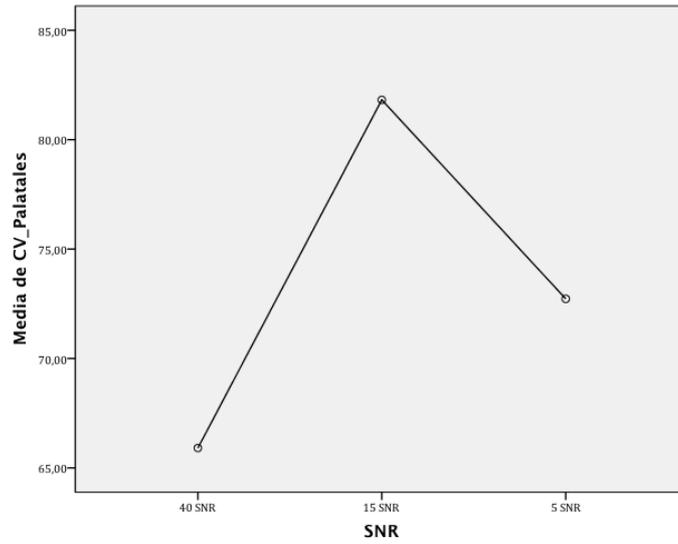


Figura 18. Comportamiento del SNR en las consonantes palatales en posición inicial

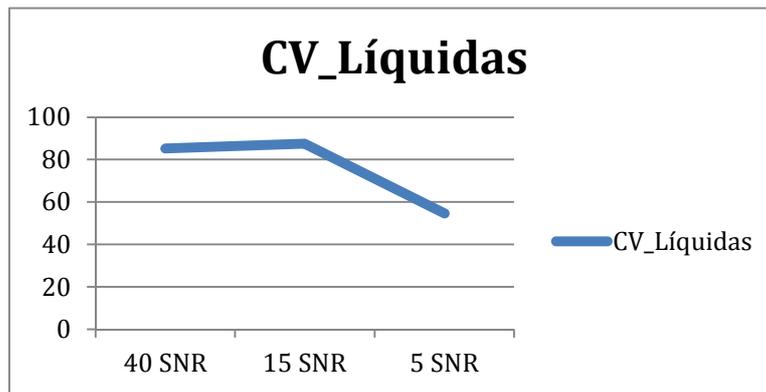


Figura 19. Comportamiento de del SNR en las consonantes líquidas en posición inicial

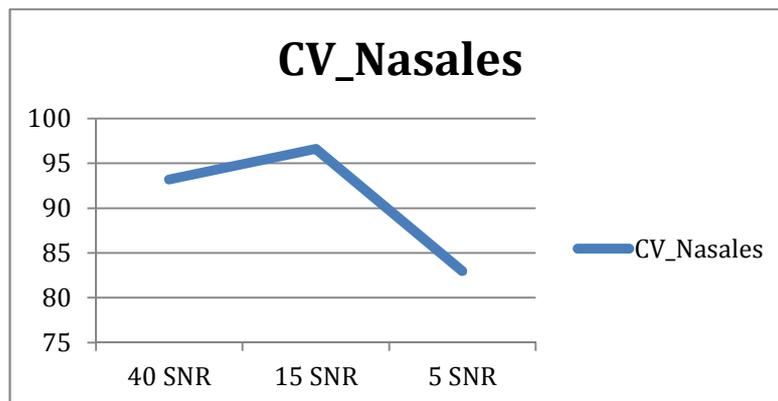


Figura 20. Comportamiento de del SNR en las consonantes nasales en posición inicial

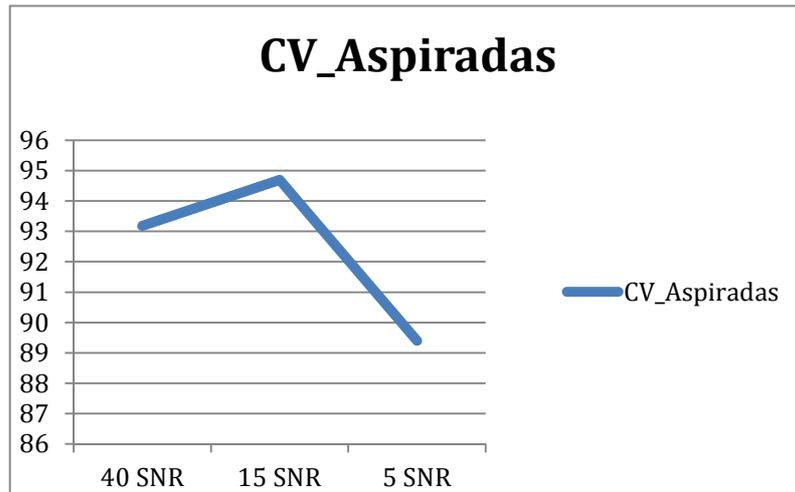


Figura 21. Comportamiento de del SNR en las consonantes aspiradas en posición inicial

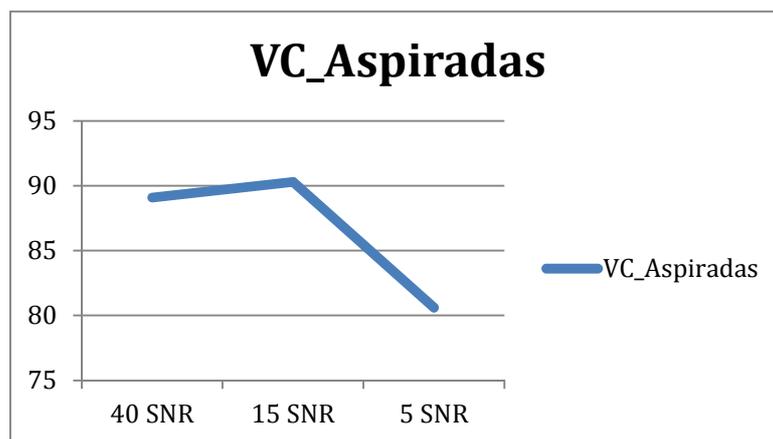


Figura 22. Comportamiento del SNR en las consonantes aspiradas en posición final

La identificación del habla de una lengua en condiciones de ruido es difícil para nativos y no nativos (García Lecumberri & Cooke, 2006). Sin embargo, el nivel de competencia lingüística en esa L2 determinará la capacidad de identificar los sonidos de la lengua, tal como afirman Masuda y Arai (2010). Es posible que hablantes con un nivel de competencia muy alto, o “bilingües”, puedan ser capaz de acercarse a los oyentes nativos en la percepción en condiciones de ruido (Mayo et al, 1997). Estudios realizados por Masuda y Arai (2012) demostraron que la identificación de las consonantes inglesas /r/ y /l/ por oyentes japoneses era tan buena como la que podían hacer los nativos, incluso en condiciones adversas. Su estudio enmascaró las sílabas a 0dB SNR. El tipo de ruido de fondo es el conocido como “babble noise”, que es el característico ruido de fondo de un grupo de personas hablando, como el que ocurre en cafeterías y otros

lugares públicos cerrados. En su estudio, los oyentes japoneses encontraron dificultades en la identificación de estas dos consonantes claves que, en situaciones no adversas, no habían mostrado.

En mi estudio, el efecto del entrenamiento no favorece la adquisición de las categorías cuando el contexto es adverso (entrenamiento intensivo con sílabas a -5dB SNR). De hecho, ningún grupo experimental (Experimento II) se benefició claramente del entrenamiento auditivo intensivo con ruido. Lo que sí se observó es que la confusión de consonantes debida a su posición en la sílaba se redujo, como se recoge de los valores de correlación del estadístico de Pearson (r^2 0,9 para los 3 grupos de entrenamiento). En la siguiente sección resumo y discuto los resultados obtenidos relativos a los beneficios del entrenamiento intensivo en relación a la posición de la consonante entrenada en la sílaba.

5.2. Cómo afecta la percepción de la posición silábica en la muestra

Nuestro análisis de datos apunta que el efecto de la posición no es estadísticamente significativo ya que sólo ha tenido impacto en las consonantes labiodentales y no aspiradas en las que la percepción se ha visto perjudicada por la posición silábica de los estímulos por lo que nuestros resultados contradicen las afirmaciones de Bell y Hooper (1978), Redford y Diehl (1999), Wagner (2013) las cuales establecen que las consonantes en posición inicial son más fáciles de identificar que las consonantes en posición final. Sin embargo, nuestros resultados son más consecuentes con aquellos estudios que afirman que en la mayoría de los casos no hay diferencias en la percepción de los estímulos en relación con la posición silábica de los mismos (Bent et al. 2008, Helfer y Huntley 1991, Miller y Nicely 1955, Wang y Bilger 1973) .

Asimismo, nuestros resultados contradicen a otros estudios que afirman que se produce una peor percepción de los sonidos en posición inicial que en posición final ya sea en grupos consonánticos o en solitario y en posición intervocálica (Lively et al. 1993, Logan et. al 1991). Estos estudios analizan la percepción de los sonidos /l/ y /r/ por lo que afirmamos que sus datos están sesgados ya que la elección de dos únicos sonidos no

aporta grandes generalizaciones sobre la posición silábica. Es necesario elegir una muestra de sonidos más grande para llegar a generalizaciones más estables. No podemos decir lo mismo de los resultados encontrados en los estudios de Bent et al. (2008) y Helfer y Huntley (1991) en los cuales se analizan una gran variabilidad de consonantes, vocales y grupos consonánticos. Al tener una mayor variabilidad en la muestra, sí son aceptables sus conclusiones las cuales, como se ha dicho anteriormente, afirman que no hay diferencias en la identificación de los estímulos en relación con la posición silábica de los mismos. Además, cuando en los casos en los que se produzca una peor identificación en posición inicial se producirían más errores de inteligibilidad.

Hemos encontrado que el efecto de la interacción entre SNR y posición silábica es significativo en las consonantes oclusivas, bilabiales, velares, sonoras, bajas, altas, no aspiradas, líquidas y nasales en posición final. Además, el SNR también afecta la percepción de las consonantes nasales, altas y líquidas en posición final. Como vemos el SNR afecta más a la posición final que a la inicial. Asimismo, respecto a las consonantes altas las cuales se ven afectadas por el SNR en posición inicial y final, la significación es más alta en posición final ($.000 < .005$) que en posición inicial ($.027 < .005$). Por tanto, al modificar la señal con SNR, estos resultados están más en consonancia con las afirmaciones de Redford y Diehl (1999) que establecen que las consonantes en posición final son más difíciles de identificar que las consonantes en posición inicial. Para entender estos datos tendríamos que asumir que el ruido de fondo funciona como factor determinante para forzar la adquisición de las categorías, pues añade un elemento adverso.

5.3. Cómo afectan los rasgos articulatorios en la muestra

Atendiendo al modo de articulación (tablas 2 y 3 anexo 5), si nos centramos en 15dB SNR, apreciamos que las consonantes cuya percepción está más afectada y por lo tanto se ha producido una peor identificación son las fricativas y africadas en posición inicial y final. Además, si nos fijamos en 40dB y 5dB SNR también se ve afectada la identificación de estos sonidos por lo que podemos afirmar que respecto al modo de articulación, las consonantes fricativas y africadas son las que más problemas de

identificación producen a nivel general (ver figuras 2 y 3). Este es el caso para el Experimento I, donde los participantes tenían diferentes L1, y para el Experimento II, con participantes españoles únicamente. En relación al lugar de articulación, en nuestro caso son las consonantes (labio-)dentales las que presentan valores más bajos en los porcentajes (ver figuras 4 y 5).

Varios estudios se han centrado en la identificación de las consonantes fricativas y africadas (Jiang, Chen y Alwan 2006, Dorman et al. 1980, Johnson y Babel 2007, Nittrouer 2002, Nittrouer y Miller 1997, Redford y Diehl 1999, Tsao et al. 2006, Zeng y Turner 1990). Estos estudios afirman que para que se produzca una correcta identificación de estos sonidos los oyentes recurren a indicios acústicos como la amplitud, las transiciones formánticas o la calidad de la vocal. Estos indicios acústicos dependen en gran medida del contexto recibido y de la calidad de la señal. Al estar la señal enmascarada produce que estos indicios articulatorios se difuminen produciendo problemas de identificación de estos sonidos. Teniendo en cuenta esto, es posible que se produjeran errores de identificación en las consonantes fricativas y africadas debido a que los oyentes no identificaron los indicios acústicos mencionados anteriormente. Desde el punto de vista acústico, la articulación de las consonantes fricativas (y de las africadas por el mismo motivo) conlleva la acción de dos articuladores que se aproximan y se mantienen cercanos el tiempo suficiente para producir que el aire que escapa a través de ellos lo haga de forma turbulenta creando un sonido aperiódico. Por lo tanto, las fricativas y africadas se caracterizan por poseer un componente de ruido. Añadir ruido de fondo a todo un conjunto de estímulos (fricativos o no), producirá mayor confusión al oyente no entrenado, que busca en el ruido el rasgo articulatorio que define la consonante en cuestión. Esta motivación podría explicar por qué las fricativas son más difíciles de identificar, en condiciones adversas de ruido. Eckman e Iverson (1993, 1994) encontraron que atendiendo a los rasgos universales del lenguaje, los sonidos más marcados son las consonantes africadas, seguidas por las consonantes fricativas y por las oclusivas. Atendiendo a nuestra clasificación de consonantes para el rasgo de apertura, encontramos que estas consonantes se clasifican como consonantes medias (fricativas y africadas) y bajas (oclusivas). Por tanto, los resultados encontrados en nuestro estudio están en consonancia con los universales lingüísticos por lo que podemos afirmar que las consonantes medias son sonidos universalmente difíciles de identificar.

Atendiendo a la sonoridad (tablas 4 y 5 anexo 5) y tomando como referencia el nivel 15dB SNR, vemos que la sonoridad de las consonantes no produce ningún problema de identificación y lo mismo ocurre en 40dB. En 5 SNR sí que se producen problemas de identificación de las consonantes sonoras en posición inicial y final. No obstante, centrándonos en 15dB SNR afirmamos rotundamente que la sonoridad no es una característica universal de los sonidos tal y como podríamos pensar ya que no produce ningún efecto negativo en la identificación tal y como Stockwell y Bowen (1970) afirman. Además, nuestros resultados difieren de las afirmaciones de algunos estudios que afirman que hay una tendencia natural en la percepción de los sonidos sonoros en posición inicial (Cebrián 2000, Flege 1995, Flege et al. 1992, Mora y Fullana 2007, Wang y Wu 2001) que tiende a ensordecir los sonidos sonoros en posición final. En nuestro estudio la sonoridad no ha tenido un impacto significativo en la identificación de los sonidos de nuestro estudio por lo que no es un rasgo articulatorio que sea determinante en la percepción. Cebrián (2000), Flege (1995), Flege et al. (1992), Mora y Fullana (2007) y Wang y Wu (2001) afirman que los hablantes experimentados del inglés como L2 son capaces de distinguir la sonoridad al mismo nivel que lo hacen los hablantes nativos. Por tanto, podríamos afirmar que el hecho de que no hayamos obtenido resultados significativamente estadísticos puede ser porque nuestros participantes se encuentran en estadios tardíos del aprendizaje y todos ellos tienen un nivel alto de la lengua por lo que son capaces de hacer la correcta identificación de los rasgos sonoros y sordos.

También se estudió el efecto de la aspiración en relación a aquellas consonantes presentes en nuestros tests. En entrenamiento intensivo ≥ 15 dB SNR, no se observan errores de identificación de las consonantes aspiradas. Los errores sólo aparecen en el caso de ruido adverso (5dB SNR). No parece definitorio afirmar que la aspiración pueda considerarse como un rasgo cuantificable a la hora de determinar la dificultad en la adquisición de categorías fonológicas de las consonantes inglesas. No obstante, para autores como Jenkins (2000) en la definición de su teoría del inglés como *Lingua Franca Core* (LFC), la aspiración es una característica fonética que debería ser prioritaria en la enseñanza del inglés como lengua internacional (EIL). De mis experimentaciones no he podido concluir que esta afirmación sea cierta.

5.4. Cómo afecta el entrenamiento auditivo intensivo en relación a los rasgos articulatorios en la muestra, la posición en la sílaba y el ruido de fondo

Los resultados del entrenamiento realizado durante el Experimento II demuestran un cierto beneficio del entrenamiento para ciertos casos, como en el entrenamiento de las consonantes fricativas a -5dB SNR, tanto en el grupo de control como en el grupo Experimental I (entrenamiento con sílabas) (ver figura 14). Además, también observamos una mejora en los resultados tras el entrenamiento en la posición inicial silábica (exceptuando las consonantes fricativas y africadas del grupo Experimental II). Por otro lado, al comparar los resultados obtenidos en la percepción de los estímulos en oraciones observamos que estos también obtienen una ligera mejora tras el entrenamiento sobre los obtenidos en los estímulos silábicos. Por tanto, podemos decir que un entrenamiento auditivo intensivo, produce, en líneas generales mejoras en la percepción de sonidos de una L2.

El estudio del impacto del entrenamiento auditivo intensivo ha sido profundamente experimentado y estudiado a lo largo de los últimos años. Diversos estudios realizados y aplicados al reconocimiento de sonidos inherentemente difíciles como el sonido /l/ y /r/ inglesas por parte de aprendices japoneses del inglés como L2 llevados a cabo por Logan et al. 1991, Lively et al. (1993), Lively et al. (1994), Bradlow et al. (1997), Bradlow et al. (1999) demuestran que el entrenamiento intensivo produce, en general, beneficios en el reconocimiento de estos sonidos. Logan et al. (1991) y Lively et al. (1993) encontraron que se produce una mejora tras un entrenamiento intensivo en la percepción de sonidos intrínsecamente difíciles para un grupo de aprendices de inglés como L2 dado que estos sonidos (/r/ y /l/) carecen de equivalente en la lengua materna de los mismos (japoneses). Encontraron que se produce una mejora en los resultados obtenidos en los sonidos en posición inicial de sílaba, en posición intervocálica y en posición inicial en un grupo consonántico (*cluster*), lo cual contrasta con los resultados encontrados por otros investigadores como por ejemplo Gillete (1980), Goto (1971), Mochizuki (1981), Sheldon y Strange (1982) o Strange y Dittmann (1984), quienes observaron que los contrastes producidos en diferentes entornos fonéticos,

especialmente en inicial de sílaba y grupos consonánticos iniciales obtienen peores resultados en la identificación de los mismos. Para identificar los motivos por los cuales ciertos sonidos tienen peores resultados en la percepción dependiendo de su posición en la palabra, se han realizado diferentes análisis acústicos teniendo en cuenta las características temporales y espectrales, ya que estas pueden estar relacionadas con los errores en la percepción. Por ejemplo, Logan et al. (1991), Mochizuki (1981) y Sheldon y Strange (1982) demostraron que se produce una mejor percepción de los sonidos /r/ y /l/ en los contextos fonéticos en los que la duración de los mismos es más larga mientras que en los contextos fonéticos con una duración más corta se produce una peor identificación de estos sonidos.

Sin embargo, es posible que la duración no sea el único indicio acústico que indica una dificultad perceptiva de ciertos sonidos en diferentes posiciones silábicas. Henly y Sheldon (1986) contrastaron la percepción de estos sonidos (/r/ y /l/) por aprendices de inglés japoneses y chinos, teniendo estos últimos más dificultad en la identificación de estos sonidos en posición final y en posición inicial en un grupo consonántico. Por tanto, Henly y Sheldon (1986) afirman que esto se produce porque la lengua nativa de los oyentes tiene un papel importante en la adquisición de nuevos sonidos en diferentes contextos fonéticos. La interferencia de la lengua materna ha sido algo muy probado y demostrado a lo largo de la historia, como ya hemos dicho anteriormente, y esta interferencia produce graves hándicaps cuando se aprende una L2. A pesar, de ello, se puede contrarrestar su efecto en la percepción y adquisición de nuevos sonidos con un adecuado entrenamiento auditivo, tal y como estamos demostrando.

Al igual que los participantes del Experimento II, los participantes del estudio de Logan et al. (1991) demostraron una mejora en la percepción de los sonidos con una tarea de entrenamiento de laboratorio consistente en el contraste de pares mínimos (*minimal pairs*). Al igual que en nuestro estudio, los participantes obtuvieron peores resultados dependiendo del lugar silábico en el que se producía los sonidos los cuales eran objeto de estudio. Sin embargo, estos autores encontraron que no sólo la posición de los sonidos afectaba a su percepción, sino que también la variabilidad de los oyentes, es decir, cuanta más variabilidad, peor identificación se producía. Este aspecto no ha podido ser demostrado en nuestra investigación. En este punto, es importante recordar que nos hemos centrado en la posición más que en la variabilidad ya que es una variable

que no hemos tenido en cuenta en nuestro experimento, por lo que nuestras conclusiones no pueden ser consideradas como definitivas.

Posteriormente, utilizando los resultados encontrados por Logan et al. 1991, Lively et al. (1994) demostraron que los participantes del estudio obtuvieron una mejora en los sonidos /r/ y /l/ después de tres semanas de entrenamiento intensivo. Además, tres y seis meses después de la conclusión del experimento volvieron a realizar post-tests en los mismos participantes y descubrieron que los resultados obtenidos en estos post-tests eran sorprendentemente buenos. Esto demuestra que un entrenamiento intensivo produce cambios y mejoras en la percepción fonética a largo plazo. Lively et al. (1994) y Bradlow et al. (1999) afirman que los cambios en la percepción están producidos por cambios en una atención selectiva de los indicios de los contrastes fonéticos. Lo que es más interesante de este aspecto es que parece ser que las modificaciones en la atención de estos indicios se retienen a lo largo del tiempo a pesar de que los oyentes no estén expuestos a los sonidos que han sido objeto del estudio. Por tanto, a pesar de que no lo hemos comprobado, una futura línea de investigación para nuestro Experimento II sería comprobar si en las consonantes oclusivas y fricativas, se siguen obteniendo buenos resultados o manteniendo los obtenidos después de un tiempo sin entrenamiento auditivo de los mismos. De esta forma, podríamos afirmar más rotundamente que es necesario un entrenamiento de sonidos que son inherentemente difíciles para un aprendiz de una L2 ya que uno de los objetivos principales de los experimentos de entrenamiento auditivos entre diferentes lenguas es desarrollar nuevas categorías para los contrastes fónicos no nativos y que estas categorías sean permanentes y sólidas. Además, estas categorías son necesarias para estímulos que normalmente muestran una alta variabilidad, dado que estos sonidos aparecen en multitud de posiciones (inicial, final, intervocálica, etc.) y producidas por diferentes hablantes o diferentes situaciones que se ven afectadas por diferentes elementos contextuales como, por ejemplo, un alto nivel de ruido.

No obstante, ya hemos dicho anteriormente que en un estudio de nivel acústico debe aparecer la variable SNR para obtener resultados más cercanos a la realidad ya que nuestro entorno está repleto de diferentes elementos que interfieren en la señal acústica de los sonidos y producen su diferente percepción. Ya vimos en el Experimento I que a 40dB SNR, el cual es un nivel óptimo de ruido por la inexistencia de elementos que

interfieren en la señal, y a 5dB SNR, el cual es un nivel que tiene bastante ruido por lo que empeora gravemente la señal acústica, se produce peor percepción de los sonidos. Analizando los resultados obtenidos observamos que es en 15dB SNR donde se produce una mejor percepción, lo cual nos llevó a afirmar que es necesaria la presencia del SNR para analizar los resultados. Por tanto, en el Experimento II, 15dB SNR es el nivel de ruido de referencia. Teniendo en cuenta esto, los resultados de Logan et al. (1991), Lively et al. (1993), Lively et al. (1994), Bradlow et al. (1997), Bradlow et al. (1999) podrían estar sesgados, al ser estudios de entrenamiento acústico en los cuales no se modifica la señal, por lo que los sonidos objetos de estudio se producen en un laboratorio con una calidad óptima de ruido lo que se aleja de un entorno real en el que se producirían esos sonidos y, por tanto, los resultados obtenidos podrían ser diferentes al introducir nivel de SNR. Aunque los resultados obtenidos en el Experimento II pueden compararse a los resultados de los estudios mencionados anteriormente ya que se produce una mejora en la percepción de los sonidos en posición inicial después de un entrenamiento intensivo, no pueden compararse plenamente por la ausencia de nivel de SNR.

Actualmente, es cada vez más común encontrar estudios de entrenamiento auditivo, no sólo en el campo de la L2, sino también en el campo de la rehabilitación auditiva y de los implantes cocleares dado la importancia de entrenar a estas personas para poder llevar una vida lo más plena y normal posible. En 2010, la Organización Mundial de la Salud (OMS) ya informó que cada vez es mayor el número de personas con pérdida de capacidad auditiva debido el envejecimiento de la población y por el ritmo de vida y vaticinó el incremento de la pérdida de audición debido a que el número de población anciana es cada vez mayor. Cualquier estudio realizado con personas normo-oyentes, no obstante, podría considerarse como referente en las conclusiones obtenidas de estudios similares realizados con personas con discapacidad auditiva.

El ruido de fondo es una variable determinante en la experimentación con personas con discapacidad auditiva. A pesar de que la principal forma de rehabilitación de adultos con discapacidad auditiva es la utilización de implantes, muchos adultos deciden no usarlos a pesar de que experimentan dificultades en la percepción del habla. Según Burke y Humes (2008) estas dificultades aumentan con la presencia de ruido de fondo. No obstante, una solución a estas dificultades en las que la percepción se produce de

forma ineficiente por la presencia de SNR, es entrenar a los oyentes para que utilicen de forma más eficiente la existencia del SNR. Por tanto, teniendo en cuenta la afirmación de Brouns et al. (2010), el entrenamiento auditivo se ha convertido en nuestros días en una herramienta clave para mejorar la habilidad de los individuos con problemas de audición para percibir el habla correctamente.

A pesar del reciente y renovado interés en el entrenamiento auditivo como un método para mejorar las habilidades de percepción del habla, especialmente en condiciones adversas de ruido, autores como Anderson y Kraus (2013), Burke y Humes (2007, 2008), Kilgard et al. (2007) o Sullivan et. al (2012) no se ponen de acuerdo en los métodos que se deben utilizar en la rehabilitación auditiva en niños y adultos. Si bien es cierto que todos coinciden en la necesidad de un entrenamiento con ruido de fondo, no llegan a un consenso sobre los procedimientos que deben seguir o los niveles de SNR que deben emplearse. Todos han llegado a la conclusión de que tras un período de entrenamiento se producen mejoras en la percepción del habla. Además, tal y como ya afirmaron Bradlow et al. (1999) y Burke y Humes (2008), también encontraron que después de un entrenamiento intensivo se mantuvieron los resultados por un período largo de tiempo. Estos últimos autores llevaron a cabo una tarea de entrenamiento de 150 palabras lexicalmente difíciles definidas de acuerdo al Modelo de Activación de Vecinos (*Neighborhood Activation Model – NAM*) de Luce y Pisoni (1998). Se entiende por palabras lexicalmente difíciles aquellas que tienen lugar con poca frecuencia y tienen muchos sonidos o palabras similares (*neighboring*). Los autores partieron de la hipótesis de que si los participantes eran capaces de mejorar el reconocimiento de palabras lexicalmente difíciles en condiciones adversas de ruido con un entrenamiento auditivo, serían capaces de mejorar la identificación de palabras lexicalmente fáciles que no reciben entrenamiento. Después del entrenamiento, observaron que se produjo una gran mejora en el reconocimiento de las palabras difíciles, no obteniendo la misma mejora en las palabras fáciles las cuales no habían sido entrenadas. Por tanto, una vez más podemos destacar la necesidad y la importancia de un entrenamiento auditivo para el reconocimiento, no sólo de sonidos o palabras difíciles, sino también de aquellos que resultan más fáciles a priori.

Por otro lado, también se produjo una mejora en la percepción de las palabras lexicalmente difíciles, no solo de forma aislada sino también en un contexto oracional,

es decir, se produjo también una mejora en la identificación de estas palabras en oraciones que las contenían. Por tanto, el entrenamiento auditivo no solo tiene implicaciones y mejoras a nivel de palabras aisladas, sino que también tienen un impacto importante en la identificación de estas palabras dentro de la cadena hablada.

En este caso también se analizó el impacto del SNR en el entrenamiento y reconocimiento de las palabras que eran objeto de estudio y, de nuevo, a pesar de las condiciones adversas, los oyentes obtuvieron una mejoría en la identificación de los mismos en los post-tests. Esto puede deberse a la presentación de estos sonidos en contextos más naturales y más cercanos a la vida cotidiana. Por otro lado, afirman que si es posible que el entrenamiento auditivo tenga los resultados esperados en palabras que son lexicalmente difíciles debido a su poco uso en la vida diaria de los oyentes, es lógico que si se entrenan con palabras fáciles, es decir, más comunes, obtenga mejores resultados.

Schuman et al. (2015) diseñaron también una tarea de entrenamiento auditivo consistente en una discriminación de fonemas por ordenador para la rehabilitación auditiva de pacientes con implantes cocleares. Al igual que nuestro experimento II, este consiste en un pre-test y un post-test para medir los beneficios del entrenamiento auditivo. El experimento de estos autores se prolongó durante tres semanas en las que los participantes fueron entrenados en la discriminación de fonemas dos veces en semana a través de un ordenador. No sólo se midió el impacto del entrenamiento a nivel segmental sino también los sonidos a nivel suprasegmental, es decir, en el reconocimiento de los mismos a nivel oracional en diferentes condiciones adversas de ruido. Después del experimento se observaron mejoras en el reconocimiento, no sólo en la tarea de discriminación de fonemas sino también en la tarea de reconocimiento de oraciones en las que se había modificado la señal acústica con diferentes niveles de ruido. Además, se comprobó que las mejoras en la percepción de los fonemas y de las oraciones se mantuvo durante un largo período de tiempo, tal como ya sucediera en los experimentos llevados a cabo por Bradlow et al. (1999), Lively et al. (1994), Logan et al. (1991) y Burke y Humes (2008). Vemos de nuevo la importancia de la experimentación con entrenamiento auditivo intensivo y el impacto que este puede tener en la mejora, no solo de las habilidades auditivas de los sujetos, sino también en la

mejora de la calidad de vida de los mismos al poder producirse una mayor relación e interacción con el entorno que les rodea.

A diferencia de otros autores, para Ferguson et al. (2014) el entrenamiento no debe trabajarse únicamente a nivel auditivo sino también debe trabajarse la memoria de trabajo ya que está muy relacionada con la comprensión del lenguaje tal y como ya establecieron Rönnberg et al. (2008). Moore et al. (2005) y Zhang et al. (2012) afirman que es necesario entrenar directamente la cognición para así mejorar a la vez la habilidad auditiva. Kronenberg et al. (2011) demostraron la veracidad de esta afirmación al llevar a cabo un estudio de entrenamiento de la memoria de trabajo. Este estudio produjo mejoras tanto en las habilidades relacionadas con el lenguaje como en las habilidades relacionadas con la memoria en niños con implantes cocleares. Por tanto, según Ferguson et al. (2014) las evidencias aportadas por el estudio de Kronenberg et al. (2011) dan a entender que para que se produzca una mejora de la percepción del habla, es necesario e incluso más importante el desarrollo de las habilidades cognitivas y memorísticas que el desarrollo de las habilidades sensitivas. De esta forma, la cognición puede tener un papel más importante en el desarrollo de la audición o de las habilidades perceptivas que un entrenamiento que sea únicamente auditivo y que no tenga en cuenta otras habilidades. No obstante, estos resultados no son del todo concluyentes al carecer de una investigación más profunda sobre la efectividad de los estímulos de entrenamiento (auditivo versus cognitivo o ambos) para mejorar la percepción del habla por parte de personas con problemas auditivos. Teniendo en cuenta los resultados aportados por este grupo de autores, nuestro experimento podría ampliarse de tal forma que se mida si la memoria también juega un papel importante en adquisición de sonidos en una L2, al igual que, aparentemente, lo tiene en el campo de la rehabilitación auditiva.

Hemos visto cómo diferentes autores han llevado a cabo estudios con relación al entrenamiento auditivo y a su efecto en la percepción de sonidos inherentemente difíciles, o más complejos, resultando en una mejora de la identificación de los mismos por lo que es necesario destacar la importancia de este tipo de entrenamiento para una mayor mejora en la percepción de estos sonidos. En nuestro Experimento II también mostramos resultados en relación a esto ya que tras un entrenamiento se producen mejoras en la percepción de los sonidos oclusivos y fricativos en posición inicial,

especialmente de las consonantes oclusivas. Además, hemos visto que el entrenamiento intensivo auditivo no es sólo importante en la percepción y adquisición correcta de sonidos de una L2, sino que es cada vez más importante en el campo de la rehabilitación auditiva. No obstante, a diferencia de como se hace en el campo de adquisición de segundas lenguas, no se entrenan los sonidos de forma aislada sino en palabras y oraciones.

En nuestro experimento también demostramos la importancia del enmascaramiento de la señal para que los resultados no estén sesgados y se alejen de un entorno de laboratorio con condiciones óptimas produciendo un acercamiento a situaciones más cotidianas y reales. No obstante, no hemos encontrado estudios en los que se analicen los resultados en función de los rasgos articulatorios, lo que podría ser útil para poder acotar aquellos sonidos que sean más difíciles de percibir e identificar y, así, poder trabajar y diseñar tareas que resuelvan problemas de identificación en varios sonidos a la vez, en lugar de sonidos aislados.

A pesar de que al igual que en los estudios de Bradlow et al. (1999), en nuestro Experimento II se produce una mejora de la percepción de las consonantes en posición inicial en el grupo de control (figura 8), vemos que en el grupo de experimental I se producen resultados estables en posición inicial y final (figuras 10 y 11) y en el grupo experimental II se produce una mejoría en las consonantes oclusivas en posición inicial. Por tanto, estos resultados no nos permiten afirmar que la posición silábica es un rasgo que influya en el aprendizaje y percepción de los sonidos. Estos resultados están en consonancia con aquellos encontrados en el primer experimento en el que la posición silábica sólo tuvo impacto en las consonantes labiodentales y no aspiradas. Por tanto, este segundo experimento vuelve a contradecir las afirmaciones de Bell y Hooper (1978), Redford y Diehl (1999) y Wagner (2013) las cuales establecen que las consonantes en posición inicial son más fáciles de identificar que las consonantes en posición final. No obstante, los resultados de este segundo experimento, al igual que los resultados del primer experimento, son más consecuentes con las afirmaciones de Bent et al. (2008), Helfer y Huntley (1991), Miller y Nicely (1955) y Wang y Bilger (1973) que establecen que en la mayoría de los casos no hay diferencias en la percepción de los estímulos en relación con la posición silábica de los mismos. Por tanto, puede afirmarse que la posición silábica no es una característica a tener en cuenta en un estudio de

identificación de sonidos ya que la identificación de sonidos concretos consonánticos no se ve afectada por la posición en una sílaba o en una palabra.

Después de contrastar y discutir nuestros resultados con los resultados analizados por otros autores en estudios con entrenamiento intensivo podemos afirmar:

- El entrenamiento auditivo intensivo es necesario, no sólo en la mejora de la percepción de sonidos de una L2, sino que también es necesario en pacientes con problemas auditivos, ya sean implantados cocleares o que tengan pérdidas de audición moderadas. Por tanto, el entrenamiento debe incluirse en cualquier estudio de índole acústica en la que se mida la percepción o identificación de sonidos o palabras.
- La percepción e identificación de sonidos no se ve afectada por la posición de los mismos en diferentes contextos silábicos (posición inicial o final).
- La presencia de SNR es también necesaria en cualquier estudio acústico sobre percepción de sonidos ya que, como hemos visto, los resultados obtenidos están menos sesgados y se acercan más a la realidad cuando el SNR está presente en la señal. Esto es así porque los seres humanos vivimos rodeados de ruidos que afectan continuamente a la percepción de sonidos.

6

Conclusión

Si bien es cierto que Cutler et al. (2004, 2008) y García Lecumberri y Cooke (2006) analizaron la identificación de todas las consonantes inglesas combinadas con todas las vocales utilizando diferentes niveles de SNR, en nuestro estudio se ha realizado el análisis de datos atendiendo a una clasificación categorial inexistente hasta ahora por lo que no existen antecedentes comparables. Hemos organizado los datos atendiendo a características articulatorias de todas las consonantes inglesas en posición inicial y final en combinación con las vocales inglesas y, además, se ha modificado el nivel de SNR. Se ha analizado la interacción del SNR con las categorías inherentes las cuales son modo de articulación, lugar de articulación, sonoridad, aspiración y apertura (*sonority*) para establecer los sonidos los cuales son comunes a todas las lenguas tal y como afirma Best (1995) y, además, hemos afirmado que estos sonidos tienen características universales ya que se ha producido errores de identificación por parte de aprendices del inglés como L2 con diferentes L1s.

Por tanto, es cierto que tal y como establecen Flege (1995), Best (1995, 1993, 1994), Kuhl e Iverson (1995) la L1 está presente en la percepción de los sonidos de la L2 pero nosotros hemos encontrado que la adquisición de los fonemas ingleses no se basa solo en la búsqueda de sonidos similares en la L1, sino que también hay características universales que afectan a la percepción de estos sonidos.

Es importante recordar que en nuestro estudio no hemos utilizado las L1s de los sujetos que componían nuestra población como variable independiente porque los rasgos articulatorios analizados habrían perdido peso en nuestro análisis y nuestro estudio sería

un estudio más en el que se compara la L1 con la L2 y no habríamos aportado nada significativo.

Hemos encontrado que el SNR es necesario en cualquier estudio acústico para que los datos sean más fiables lo que produce que este hallazgo sea una cuestión metodológica importante. Esto se ha corroborado con el hecho de que en nuestro estudio los datos hallados en 15dB SNR son más estables y fiables que en 40dB SNR ya que en diferentes casos se han encontrado mejores identificaciones de los estímulos en 15dB SNR. Esto se produce porque estamos acostumbrados a escuchar en condiciones de ruido por lo que la ausencia del mismo nos lleva a errores en la identificación de los estímulos.

Nuestra investigación partía de la hipótesis de que existen rasgos articulatorios que son más difíciles para una población homogénea de oyentes no nativos del inglés. Estos rasgos articulatorios (lugar de articulación, modo de articulación, sonoridad, aspiración y apertura) tan sólo han resultado ser significativos en dos tipos específicos: labiodentales/dentales y fricativas. Los rasgos articulatorios en sí mismos no presentan una graduación clara. Quizás de haber elegido un método de recogida de datos distinto habríamos conseguido otros resultados y, por ende, habríamos llegado a otras conclusiones. En futuras investigaciones deberíamos quizás desarrollar un método de entrenamiento pre/post test para corroborar las conclusiones a las que aquí hemos llegado. Consideramos, no obstante, que nuestro estudio aporta de manera significativa al campo de la percepción de segundas lenguas, pues hemos abordado la percepción de las consonantes inglesas en posición inicial y final, y en combinación con todas las vocales. Nuestro estudio también corrobora, para este corpus tan exhaustivo, datos tan importantes como que la posición silábica no afecta a un mayor grado de identificación de las consonantes inglesas por no nativos, ni elementos como la aspiración, sonoridad o apertura, que tradicionalmente se han considerado afectados por la posición en la sílaba. En este sentido, nuestros resultados son tajantes.

Otra posible modificación al método de recogida de datos habría sido el incorporar la variable dependiente de la producción. Puesto que el ciclo percepción-producción ha sido más que demostrado en el campo, encontrar qué sonidos se producen con mayor acierto que otros, valorados por evaluadores nativos entrenados, podría confirmar o

modificar nuestras conclusiones. También utilizar una población con diferentes niveles de competencia en la L2 podría ayudar a refinar las conclusiones.

Este último aspecto no deja de tener importancia. La población elegida para nuestro estudio estaba compuesta por sujetos en estadios tardíos del aprendizaje de la L2 por lo que estaban acostumbrados a utilizar la misma y a desenvolverse en ella. Por lo tanto, se debería analizar la identificación de las consonantes inglesas por sujetos que estén inmersos en estadios más tempranos del aprendizaje del inglés. Además, incluso deberíamos analizar la adquisición de estas categorías en niños para establecer comparaciones con la adquisición de las mismas por parte de adultos y analizar su adquisición atendiendo a las teorías de Kuhl (2004), Kuhl e Iverson (1995) las cuales afirman que los niños adquieren mejor estas categorías y que es la experiencia la que empeora esta adquisición. Siguiendo con la línea de Kuhl, que en sus estudios enfatiza el hecho de que mayor input determina mejor identificación, un método de recogida de datos que dividiera entre población experimental y población de control podría mejorar las conclusiones.

En resumen, este estudio puede ser de utilidad a otros campos, en especial el de enseñanza del inglés como lengua extranjera (EIL) y el inglés como lengua franca (ELF). Por un lado, nos permite diseñar estrategias metodológicas adecuadas para evitar errores de identificación de estos sonidos en contextos multilingües y, además, tiene también un impacto importante en el LFC diseñado por Jenkins (2000). Para Jenkins, el efecto de la aspiración es muy importante para la inteligibilidad en contextos en los que estén presentes varios hablantes no nativos del inglés. La incorrecta identificación de esta categoría podría producir errores irreversibles en la comunicación. No obstante, nosotros hemos demostrado que la aspiración no es importante para la inteligibilidad, lo que podría suponer una reestructuración de los postulados del LFC defendido por Jenkins. También, hemos encontrado que la sonoridad no es importante para la correcta identificación de los sonidos tal y como establecen Stockwell y Bowen (1970).

Por último, es muy importante resaltar de nuevo el hecho de que hemos utilizado sujetos con diferentes L1s, lo cual no se había hecho anteriormente. Se han realizado estudios croslingüísticos inglés-español (García Lecumberri y Cooke 2006, García Lecumberri et al. 2010), inglés-japonés (Bohn y Best 2010, Bradlow et al. 1997, Bradlow et al. 1999,

Lively et al. 1993, Logan et al. 1991) o inglés-zulú (Best 1995, Best y Tyler 2007), entre otros, pero nunca hasta ahora se había realizado un estudio que abarcara la identificación de los sonidos inglés por hablantes del inglés como L2 con diferentes L1s. Lo innovador de nuestro estudio es que no importa tener o no diferentes L1s si éstas no se analizan como variable independiente. En nuestro estudio existen diferentes oyentes con distintas L1s pero estas L1s no se han analizado como variables independientes, sin embargo, los resultados de mi estudio son significativos. Además, podríamos aventurarnos a afirmar que tal y como indican los resultados de la ANOVA, si aumentáramos el número de participantes los resultados de nuestros datos seguirían siendo estadísticamente significativos. Este hecho nos ayuda a mantenernos en nuestro supuesto para seguir investigando en la misma línea y establecer mayores generalizaciones sobre qué sonidos del inglés determinan un mayor o menor nivel de identificación correcta para una población de aprendices de esa L2 para una población con diferentes L1s.

Finalmente, los datos aportados por este estudio puede ayudarnos a entender mejor cómo se identifican las categorías fonéticas y puede ayudarnos a establecer qué sonidos son comunes en diferentes lenguas para poder diseñar estrategias metodológicas que ayuden a los aprendices del inglés como L2 a identificar mejor estos sonidos evitando errores de inteligibilidad de los mismos. Las nuevas líneas de investigación en el campo son cada vez más ambiciosas, con poblaciones útiles como la de personas con implantes cocleares. En cualquier caso, esperamos que este estudio exhaustivo de la percepción de todas las consonantes inglesas en posición inicial y final de sílaba pueda servir como recopilatorio general en el campo, tanto por las aportaciones metodológicas que aporta en la fase de recogida de datos, como en las conclusiones finales a las que llega. De las más importantes, nos quedamos con la dificultad de percibir consonantes fricativas y labiodentales/dentales a niveles de SNR estables (15dB) para la posición inicial y final de forma independiente. Por último, se trata del único estudio que utiliza oyentes no nativos de 7 L1s diferentes. Creemos por tanto en la eficacia de nuestro método y consideramos que el paso adelante sería incorporar los sonidos en contextos más amplios de palabra, oración así como en tests que evalúen otros aspectos lingüísticos, en especial el léxico-semántico, por lo que afirmamos que será en esa línea donde se sitúen nuestras futuras investigaciones.

Como conclusión final, me gustaría resumir los hallazgos más importantes encontrados tras el análisis de los resultados para los 2 experimentos. En el Experimento I, se realizó un test general de identificación de consonantes inglesas en sílabas, en posición inicial y final, para analizar los porcentajes de acierto en la identificación por una población de diferentes L1 con un nivel B2 de la lengua. Las variables dependientes del estudio se midieron en función de rasgos articulatorios, posición en sílaba pero también degradación por ruido de fondo: nivel de SNR, modo de articulación, lugar de articulación, sonoridad, apertura, aspiración y posición silábica. Este resumen incluiría lo siguiente, tal como se aprecia en la sección 3 de esta tesis doctoral:

- 1) El porcentaje de identificación de las consonantes en posición final es, en general, mejor que en posición inicial.
- 2) Las consonantes fricativas y africadas tienen peores resultados en tests de identificación en su conjunto.
- 3) La sonoridad no presenta grandes dificultades en su identificación.
- 4) El comportamiento entre grupos y para todas las consonantes es significativo fundamentalmente para el modo de articulación, en especial: fricativas, africadas, pero también nasales y líquidas. Esto quiere decir que la amplitud media (siendo amplitud baja las oclusivas y amplitud alta las vocales) es la que presenta peor significatividad entre las muestras, es decir, el comportamiento de los participantes ha sido más variado.
- 5) El efecto del ruido sobre las muestras es significativo especialmente en 15dB SNR. Esto quiere decir que cuando hay poco o mucho ruido los resultados son menos predecibles, por lo que el contexto estable de percepción es el medio (algo de ruido de fondo).

Comparando con antecedentes, la conclusión 1) y la 2) son las más significativas, pues refrendan resultados obtenidos en estudios previos. El nivel de ruido de fondo también es interesante desde el punto de vista metodológico. De hecho, nos sirvió para enmascarar los estímulos del Experimento II y utilizarlo como medida de referencia (grupo de control). Más en concreto, el Experimento II buscaba analizar los efectos del entrenamiento intensivo comparando el entrenamiento con sílabas vs. oraciones, así como los posibles beneficios del ruido de fondo en la adquisición de categorías fonológicas. Los resultados resumidos podrían enumerarse así:

- a) Variable estructura de armónicos (oclusivas=estructura armónica vs. fricativas/africadas=no estructura armónica): El entrenamiento con ruido de fondo a 15dB SNR confirma que las consonantes oclusivas se benefician mucho más del entrenamiento intensivo que las fricativas/africadas.
- b) Variable posición en la sílaba: En todos los casos (exceptuando las fricativas/africadas del grupo Experimental II), la posición inicial presenta mayores beneficios tras el entrenamiento que la posición final.
- c) Variable estímulos sílaba vs. oración: El grupo Experimental II se beneficia ligeramente más de un entrenamiento que incluye presentación intensiva a oraciones si lo comparamos con el grupo de control, que fue entrenado exclusivamente con estímulos silábicos.

Por tanto afirmamos que hay una mejoría producida por el entrenamiento auditivo intensivo lo cual ya había sido apoyado por autores como Bradlow et al. (1999), Lively et al. (1994), Logan et al. (1991), Burke y Humes (2008) y Schuman et al. (2015), por tanto nuestros resultados confirman también sus resultados. Por otro lado, no hemos comprado si los resultados de mejora obtenidos tras el entrenamiento se mantuvieron pasado un tiempo para comprobar de esta forma si el entrenamiento fue verdaderamente efectivo, ya que como afirman Ferguson et al. (2014), se deben realizar las pruebas de percepción para comprobar que los resultados no han variado o han variado muy ligeramente con el paso del tiempo.

Para concluir, los resultados más destacables que se pueden utilizar en futuros estudios acústicos que se pueden aplicar tanto al campo del aprendizaje de L2 como al campo de rehabilitación auditiva son los que están relacionados con la necesidad de utilizar SNR ya que se producen resultados más estables y, además, los que demuestran que el entrenamiento auditivo produce claros beneficios tanto en aprendices de lenguas como en pacientes con problemas auditivos.

Referencias bibliográficas

- Ainsworth, W.A. & Meyer, G.F. (1994). Recognition of plosive syllables in noise: Comparison of an auditory model with human performance. *Journal of the Acoustical Society of America* 96, 687-694.
- Anderson, S., & Kraus, N. (2013). Auditory Training: Evidence for Neural Plasticity in Older Adults. *Perspectives on Hearing & Hearing Disorders: Research & Diagnost*, 17, 37-57.
- Bell, A. & Hooper, J.B. (1978). *Syllables and Segments*. Amsterdam: North-Holland.
- Bent, T., Bradlow, A.R. & Smith, B.L. (2008). Segmental errors in different word positions and their effects on intelligibility of non-native speech. En O.-S. Bohn y M.J. Murray J. (Eds). *Language Experience in Second Language Speech Learning: in honor of James Flege*. Amsterdam: John Benjamins Publishing Company.
- Best, C.T. (1993). Emergence of language-specific constrains in perception of non-native speech: A window on early phonological development. En *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*, ed. B. De Boysson-Bardies, S. De Schonen, P. Jusczyk, P. McNeilage, y J. Morton. Amsterdam: Kluwer Academic Publishers.
- Best, C.T. (1994a). The emergence of native-language phonological influences in infants: A perceptual assimilation model. En *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*, ed. J. Goodman and H.C. Nusbaum. Cambridge MA: MIT Press.
- Best, C.T. (1994b). Learning to perceive the sound pattern of English. En C. Rovee-Collier and L. Lipsitt (Ed.) *Advances in Infancy Research*. Hillsdale NJ: ABlex.

- Best, C.T. (1995). A direct realist view of cross-language speech perception. En W. Strange (Ed.), *Speech Perception and linguistic experience: Issues in cross-language speech research* (pp. 172-206). Timonium, MD: York Press.
- Best, C.T., McRoberts, G.W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America* 109, 775-794.
- Best, C.T., McRoberts, G.W., & Sithole, N. (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* 4, 345-360.
- Best, C.T., & Tyler, M.D. (2007). Non-native and second-language speech perception: Commonalities and complementarities. En M.J. Munro & O.-S. Bohn (Eds.). *Second language speech learning: The role of language experience in speech perception and production* (pp. 13-34). Amsterdam: John Benjamins.
- Best, C., Hallé, P. Bohn, O. & Faber, A. (2003). Cross-language perception of non-native vowels: Phonological and phonetic effects of listeners' native languages. En M. J. Solé, D. Recasens y J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences*, 2889-2892. Barcelona: Causal Productions.
- Best, C. & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics*, 20, 305-330.
- Bradlow, A.R., Pisoni, D.B., Yamada, R.A. & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/ IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America*, 101, 2299-2310.
- Bradlow, A.R., Yamada, R.A., Pisoni, D.B. & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Perception and Psychophysics*, 61, 977-985.

- Brouns, K., Refaie, A. E., & Pryce, H. (2010). Auditory Training and Adult Rehabilitation: A Critical Review of the Evidence. *Global Journal of Health Science*, 3, 49-63.
- Burk, M.H., & Humes, L. E. (2007). Effects of Training on Speech-Recognition Performance in Noise using Lexically Hard Words. *Journal of Language and Hearing Research*, 50, 25-40
- Burk, M. H., & Humes, L. E. (2008). Effects of long-term training on aided speech-recognition performance in noise in older adults. *Journal of Speech, Language and Hearing Research*, 51, 759-771.
- Cebrián, J. (2000). Transferability and productivity of L1 rules in Catalan-English interlanguage. *Studies in Second Language Acquisition*, 22, 1-26.
- Cummins, J. (2000). Interdependence of first and second language proficiency in bilingual children. En E. Bialystok (Ed.). *Language Processing in Bilingual Children*. Cambridge: Cambridge University Press.
- Cutler, A., Cooke, M., García Lecumberri, M.L. & Pasveer, D. (2007). L2 consonant identification in noise: Cross-language comparisons. *Proceedings of Interspeech*, 1585-1588.
- Cutler, A., García Lecumberri, M.L., & Cooke, M. (2008). Consonant identification in noise by native and non-native listeners: Effects of local context. *Journal of the Acoustical Society of America* 124, 1264-1268.
- Cutler, A., Weber, A., Smits, R. & Cooper, N. (2004). Patterns of English phoneme confusions by native and non-native listeners. *Journal of the Acoustical Society of America* 116, 3668-3678.
- Darcy, I., Peperkamp, S. & Dupoux, E. (2007). Bilinguals play the rules: perceptual compensation for assimilation in late L2 learners. En Jennifer Cole y José I. Hualde (eds). *Laboratory Phonology*. Berlin: Mouton de Gruyter.

- Dorman, M.F., Raphael, L.J., & Isenberg, D. (1980). Acoustic cues for a fricative-affricate contrast in word-final position. *Journal of Phonetics*, 8, 397-405.
- Eckman, F. R., & Iverson, G. (1993). Sonority and markedness among onset clusters in the interlanguage of ESL learners. *Second Language Research*, 9, 234-252.
- Eckman, F. R., & Iverson, G. (1994). Pronunciation difficulties in ESL: Coda consonants in English interlanguage. En M. Yavas (Ed.), *First and Second Language Phonology*. San Diego, CA: Singular.
- Ferguson, M.A., Henshaw, H, Clark, D.P.A., & Moore, D.R. (2014). Benefits of phoneme discrimination training in a randomized controlled trial of 50- to 74- year olds with mild hearing loss. *Ear and Hearing*, 35, 110-121.
- Fernald, A. & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, 20, 104-113.
- Flege, J.E. (1981). The Phonological Basis of Foreign Accent: A hypothesis. *TESOL Quarterly*, 15, 443-453.
- Flege, J.E. (1989). Chinese subjects' perception of the word final English /t-/d/ contrast: Before and after training. *Journal of the Acoustical Society of America*, 86, 1684-1697.
- Flege, J.E. (1995). Second language speech learning: Theory, findings and problems. En Winifred Strange (ed.) *Speech Perception and Linguistic Experience Issues in Cross-Language Research*. Timonium, MD: York Press.
- Flege, J.E. (1999). Age of learning and second language speech. En D. Birdsong (Ed.), *Second Language Acquisition and the Critical Period Hypothesis* (101-131). Mahwah, NJ: Lawrence Erlbaum Associates.

- Flege, J.E. (2002). Interactions between the native language and second language phonetics systems. En P. Burmeister, T. Piske y A. Rohde (Eds.), *An Integrated view of Language Development: Papers in Honor of Henning Wode*. Trier: Wissenschaftlicher Verlag.
- Flege, J.E., Frieda, E.M., & Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics*, 25, 169 – 186.
- Flege, J. E., Munro, M. J., & Skelton, L. (1992). Production of the word-final English /t/-/d/ contrast by native speakers of English, Mandarin and Spanish. *Journal of the Acoustical Society of America*, 92, 128-143.
- Friesen, L.M., Shannon, R.V., & Baskent, D., y Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. *Journal Acoustic Society of America*, 110, 1150-1163.
- García Lecumberri, M.L., & Cooke, M. (2006). Effect of masker type on native and non-native consonant perception in noise. *Journal Acoustic Society of America*, 119, 2445-2454.
- García Lecumberri, M.L., & Cooke, M., Cutler, A. (2010). Non-native speech perception in adverse conditions. *Speech Communication* 52, 864-886.
- Gillete, S. (1980). Contextual variation in the perception of L and R by Japanese and Korean speakers. *Minnesota Papers in Linguistic and Philosophy of Language*, 6: 59-72.
- Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds ‘L’ and ‘R’. *Neuropsychologia*, 9: 317-323.

- Guion, S.G., Flege, J.E., Akahane-Yamada, R. & Pruitt, J.C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal Acoustical Society of America*, 107 (5), 2711-2724.
- Hallé, P.A., Best, C.T. & Levitt, A. (1999). Phonetic versus phonological influences of French listeners' perception of American English approximants. *Journal of Phonetics* 27, 281-306.
- Hazrati, O. & Loizou, P. C. (2012). The combined effects of reverberation and noise on speech intelligibility by cochlear implant listeners. *International Journal of Audiology* 51: 437-443.
- Helfer, K.S., & Huntley, R.A. (1991). Aging and consonant errors in reverberation and noise. *Journal Acoustical Society of America*, 90, 1786-1796.
- Helfer, K.S., & Wilber, L.A. (1990). Hearing Loss, Aging, and Speech Perception in Reverberation and Noise. *Journal of Speech and Hearing Research*, 33, 149-155.
- Henly, E., & Sheldon, A. (1986). Duration and context effects on the perception of English /r/ and /l/: A comparison of Cantonese and Japanese speakers. *Language Learning*, 36: 505-521.
- Jenkins, J. (2000). *The Phonology of English as an International Language*. Oxford: Oxford University Press.
- Jiang, J., Chen, M. & Alwan, A. (2006). On the perception of voicing in syllable-initial plosives in noise. *Journal of the Acoustical Society of America*, 119 (2), 1092-1105.
- Johnson, K. & Babel, M. (2007). Perception of fricatives by Dutch and English speakers. *UC Berkeley Phonology Lab Annual Report*, 302-325
- Kapoor, A. & Allen, J.B. (2012). Perceptual effects of plosive feature modification. *Journal of the Acoustical Society of America*, 131, 478-491.

- Kilgard, M.P, Vázquez, J.L., Engineer, N.D., & Pandya, P.K. (2007). Experience dependent plasticity alters cortical synchronization. *Hearing Research*, 229, 171-179.
- Kronenberg, W.G., Pisoni, D.B., Henning, S.C. (2011). Working memory training for children with cochlear implants. *Journal of Speech, Language and Hearing Research*, 54, 1182-1196.
- Kuhl, P. K. (1992). Infants' perception and representation of speech: development of a new theory. En J. Ohala, B. Derwing, M. Hodge y G. Wiebe (Eds.). *Proceedings of the International Conference on Spoken Language Processing*, 449-456.
- Kuhl, P. K. (1993a). Infant speech perception: A window on psycholinguistic development. *International Journal of Psycholinguistics*, 9, 33-56.
- Kuhl, P. K. (1993b). Innate predispositions and the effects of experience in speech perception: The Native Language Magnet Theory. En B. de Boysson-Bardies, S. de Schonen, P- Jusczyk, P. MacNeilage, y J. Morton (Eds.), *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*, 259-274. The Netherlands: Kluwer Academic Publishers.
- Kuhl, P.K. (1994). Learning and representation in speech and language. *Current Opinion in Neurobiology*, 4, 812-822.
- Kuhl, P. K. (2000). A new view of language acquisition. *Proceeding of the National Academy of Science*, 97 (22), 11850-11857.
- Kuhl, P. K. (2004). Early language Acquisition: Cracking the speech code. *Neuroscience* 5, 831-843.
- Kuhl, P.K. & Iverson, P. (1995). Linguistic experience and the “perceptual magnet effect”. En W. Strange (Ed.) *Speech perception and linguistic experience: Issues in cross language research*, 121-154, Timonium, MD: York Press.

- Lalonde, C. E. & Werker J.F. (1995). Cognitive Influences on cross-language speech perception in infancy. *Infant Behavior and development*, 18, 459-475.
- Linde, S.G., & Lofgren, H. (1988). The relationship between medium of instruction and school achievement for Finnish-speaking students in Sweden. *Language, Culture and Cognition*, 1, 131-146.
- Lively, S.E., Logan, J.S. & Pisoni, D.B. (1993). Training Japanese listeners to identify English /r/ y /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America*, 94, 1242-1255.
- Logan, J.S., Lively, S.E. & Pisoni, D.B. (1991). Training Japanese listeners to identify English /r/ y /l/: A first report. *Journal of the Acoustical Society of America*, 89, 884-886.
- Luce, P. & Pisoni, D.B. (1998). Recognizing Spoken Words: The Neighborhood Activation Model. *Journal of the American Audiology Society* 19, 1-36.
- MacKain, K., Best, C.T. & Strange, W. (1981). Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics*, 2, 369-390.
- Martin, J.G. & Bunnell, H.T. (1982). Perception of anticipatory coarticulation effects in vowel-stop consonant-vowel sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 473-488.
- Masuda, H., & Arai, T. (2010). Processing of consonant clusters by Japanese native speakers: Influence of English learning backgrounds. *Acoustical Science and Technology*, 3, 320-327.

- Masuda, H., & Arai, T. (2012). Perception of /r/ and /l/ in quiet and multi-speaker babble noise by Japanese and English native listeners. *Proceedings of the Spring Meeting of Acoustical Society of Japan 2012*, 477-480
- Miller, J.L., & Eimas, P.D. (1976). Studies on the selective tuning of feature detectors for speech. *Journal of Phonetics*, 4, 119-127.
- Miller, J.L. & Eimas, P.D. (1996). Internal structure of voicing categories in early infancy. *Perception and Psycholinguistics*, 58, 1157-1167.
- Miller, G.A. & Nicely, P.E. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the American Audiology Society*, 27, 338-352.
- Miller, G.A., Heise, G.A. & Lichten W. (1951). The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology*, 41. 329-335.
- Mochizuki, M. (1981). The identification of /r/ and /l/ in natural and synthesized speech. *Journal of Phonetics*, 9, 283-303.
- Moore, D.R., Rosenberg, J.F., & Coleman, J.S. (2005). Discrimination training of phonemic contrast enhances phonological processing in mainstream school children. *Brain Language*, 94, 72-85.
- Mora, J.C., & Fullana, N. (2007). Production and perception of voicing contrasts in English word-final obstruents: Assessing the effects of experience and starting age. *New Sounds 2007: Proceedings of the Fifth International Symposium on Acquisition of Second Language Speech*, 207-221.
- Nabelek, A.K., & Donahue, A.M. (1984). Perception in reverberation by native and non-native listeners. *Journal of the Acoustical Society of America*, 75, 632-634.
- Neufeld, G. (1979). Towards a theory of language learning ability. *Language Learning* 29: 227-241.

- Nittrouer, S. (2002). Learning to perceive speech: How fricative perception changes, and how it says the same. *Journal of the Acoustical Society of America*, 112, 711-719.
- Nittrouer, S., & Miller, M.E. (1997). Developmental weighting shifts for noise components of fricative-vowel syllables. *Journal of the Acoustical Society of America*, 112, 572-580.
- Öhman, S.E.G. (1966). Coarticulation in VCV utterances: Spectrographic measurements. *Journal of the Acoustical Society of America*, 39, 151-168
- Pearman, A. (2007) Native and npn-native perception of casual speech: English and Catalan. *Tesis doctoral*. Universidad Autónoma de Barcelona.
- Pilus, Z. (2005). Perception of voicing in English word-final obstruents by Malay speakers of English: Examining the Perceptual Assimilation Model. *Journal International Islamic University* 1, 1-12.
- Piske, T., MacKay, I.R.A. & Flege, J.E. (2001). Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics*, 29, 191-215.
- Recasens, D. (2011). A cross-language acoustic study of initial and final allophones of /l/. *Speech Communication* 54 (2012), 368-363.
- Redford, M.A. & Diehl, R.L. (1999). The relative perceptual distinctiveness of initial and final consonants in CVC syllables. *Journal of the Acoustical Society of America*, 106, 1555-1565.
- Rönnerberg, J., Rudner, M., Foo, C., & Lunner, T. (2008). Cognition counts: A working memory system for ease of language understanding (ELU). *International Journal of Audiology*, 47, 99-105.

- Schumann, A., Serman, M., Gefeller, O., & Hoppe, U. (2015). Computer-based auditory phoneme discrimination training improves speech recognition in noise in experienced adult cochlear implant listeners. *International Journal of Audiology*, 54, 190-198.
- Sheldon, A. & Strange, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics*, 3, 243-261.
- Silbert, N. (2012). Syllable structure and integration of voicing and manner of articulation information in labial consonant identification. *Journal of the Acoustical Society of America*, 131, 4076-4086.
- Skutnabb-Kangas, T. & Toukomaa, P. (1976). *Teaching migrant children's mother tongue and learning the language of the host country in the context of the sociocultural situation of the migrant family*. Informe para la Unesco. Tampere: University of Tampere.
- Snow, C. & Hoefnagel-Höhle, M. (1979). Individual differences in second-language ability: A factor-analytic study. *Language and Speech* 22: 151-162.
- So, C.K. & Best, C.T. (2010). Categorizing Mandarin tones into listeners' native prosodic categories: The role of phonetic properties. *Poznan Studies in Contemporary Linguistics* 47 (1): 133-145.
- So, C.K. & Best, C.T. (2011). Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. *Language and Speech* 53(2): 273-293.
- Stockwell, R.P. & Bowen, J.D. (1970). *The Sounds of English and Spanish*. Chicago: The University of Chicago Press.

- Strange, W. & Dittmann, S. (1984). Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. *Perception and Psychophysics*, 36, 131-145.
- Sullivan, J.R., Thibodeau, L.M., & Assmann, P.F. (2012). Speech recognition in noise by children with Hearing Loss as a function of Signal-to-Noise Ratio. *Journal of Educational Audiology*, 18, 24-30.
- Tsao, F.-M., Liu, H.-L. & Kuhl, P.K. (2006). Perception of native and non-native affricate-fricative contrasts: Cross-language tests on adults and infants. *Journal of the Acoustical Society of America*, 120, 2285-2294.
- Van Wijngaarden, S. J. (2001). Intelligibility of native and non-native of Dutch speech. *Speech Communication*, 35, 103-113.
- Van Wijngaarden, S. J., Steeneken, H.J.M. & Hougast T. (2002). Quantifying the intelligibility of speech in noise for non-native listeners. *Journal of the Acoustical Society of America*, 111, 1906-1916.
- Wagner, A. (2013). Cross-language similarities and differences in the uptake of place information. *Journal of the Acoustical Society of America*, 133 (6), 4256-4267.
- Wang, M.D., & Bilger, R.C. (1973). Consonant confusions in noise: a study of perceptual features. *Journal of the American Audiology Society*, 54, 1248-1266.
- Wang, H.S., & Wu, J.C. (2001). The effect of vowel duration on the perception of postvocalic voiced/ voiceless consonants. *Concentric: Studies in English Literature and Linguistics*, 35-52.
- Werker, J.F. & Tees, R.C. (1984). Speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7, 49-63.
- Werker, J.F. & Tees, R.C. (2002). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 25, 121-133.

Zeng, F.-G & Turner, C.W. (1990). Recognition of voiceless fricatives by normal and hearing-impaired subjects. *Journal of Speech, Language and Hearing Research*, 33, 440-449.

Zhang, Y.X., Dorman, M.F., Fu, Q.J., & Spahr, A.J. (2012). Auditory training in patients with unilateral cochlear implant and contralateral acoustic stimulation. *Ear and Hearing*, 33, 70-79.

ANEXO I

Documentos:

- **Formulario de consentimiento informado**
(Informed Consent Form)
- **Cuestionario sociodemográfico**
(Language Background Questionnaire)
- **Prueba de nivel de inglés**
(English Proficiency Test)
- **Audiometría tonal**
(Hearing screening speech perception Lab)

Study # _____

**UNIVERSITY OF SEVILLE
INFORMED CONSENT FORM**

CONSONANTS CONFUSION BY NON-NATIVE SPEAKERS OF ENGLISH

You are invited to participate in a research because your mother tongue is different from English and you are not bilingual in any language.

INFORMATION

To help us in this task, you will be asked to fill out two questionnaires at the beginning of the study. After that we may ask you to participate in a listening experiment. You will be invited to listen to different English sounds by a computer program and will judge which sounds were played in the recordings. These listening sessions will not take longer than an hour and you may be asked to take two of them on different days.

Information gathered from your sessions will be grouped with information gathered from other participants to provide observations regarding pronunciation and perception of English as a second language.

RISKS OR DISCOMFORTS

No available scientific evidence suggests that there are risks that could result from these sessions. In order to lessen fatigue due to the long duration of the task, you will take brief breaks if you consider it necessary.

BENEFITS

Participation in this study may help improve your listening and pronunciation skills of spoken English.

CONFIDENTIALITY

The data obtained in this study will be treated as confidential information. The recordings of productions will be kept for the duration of the study and will eventually be archived in a secure place as documentation of the research in case there are any questions about data analysis. Your name will not be used in reports or publications. Your recordings may be used for the purpose of exemplifying speech characteristics in academic conference settings as well as results obtained from the analysis of the data.

PAYMENT

You will not receive any payment for taking part in the present study.

PARTICIPATION

Participation in this study is voluntary. You may decline to participate without penalty or loss of benefits to which you are otherwise entitled. If you decide to participate, you may discontinue participation at any time without penalty, or loss of benefits to which you are otherwise entitled. If you withdraw from the study before data collection is completed, your data will be used if they are complete enough to be of scientific value.

Participant's initials: _____

CONTACT

If you have any questions at any time about the study or the procedures, or if you experience any adverse effect, you may contact the principle researcher, Dr. Dario Barrera Pardo at dario@us.es

CONSENT

I have read the above information and have had the choice to ask questions and/or have had questions answered to my satisfaction. I acknowledge receiving a copy of this form and I agree to take part in the study.

Participant's printed name: _____

Participant's signature: _____ Date: _____

Printed Name of the person obtaining the consent: _____

Signature of the person obtaining the consent: _____ Date: _____

LANGUAGE BACKGROUND QUESTIONNAIRE

STUDY TITLE: Perception of English consonants by non-native speakers

Participation Date: _____

Please, complete the following information (write/circle as needed)

1. Name: _____
2. Present Age: _____ Yrs. _____ Mo.
3. Gender (circle one): Male Female
4. Address: _____
5. Telephone number: (H) _____ (W) _____
6. Birthplace (town/province/country): _____
7. Father's Birthplace (town/province/country): _____
8. Languages your father speaks fluently: _____
9. Mother's Birthplace (town/province/country): _____
10. Languages your mother speaks fluently: _____
11. As a child, what languages were spoken at home? (By parents, guardians, relatives, etc.)

12. As a child, what languages were spoken outside your home? (At school, in public places, etc.) _____
13. Do you have an accent in your language? YES _____ NO _____
14. If you have answered YES to the previous question, which regional area/dialect is your accent from? _____
15. Places where you have lived **in your home country** for longer than 3 months:

City/State/Province	From (month/year)	To (month/year)
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____

LANGUAGE BACKGROUND QUESTIONNAIRE (Continued)

16. Places where you have lived outside your home country for longer than 3 months:

City/State/Province	From (month/year)	To (month/year)
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____

17. Apart from Spanish, do you speak other languages fluently? YES NO

18. If YES, what other languages are they? _____

19. Have you ever studied Phonetics (the scientific study of speech sounds) at the college level? YES NO

20. If YES, have you ever done phonetic transcription? YES NO

21. Have you noticed difficulty with your hearing recently? YES NO

22. Do you have a history of ear infection or any pain in your ears? YES NO

23. Is there a family history of hearing loss? YES NO

24. Do you have any history of exposure to noise in recreational activities, at work, or in the military? YES NO

25. Do you experience dizziness or ringing (tinnitus) in ears? YES NO

26. Did you have any speech therapy when you were younger? YES NO

ENGLISH PROFICIENCY TEST

ID_CODE: _____ DATE: _____

Directions: The questions here test your knowledge of English grammar. Each question consists of a short written conversation, part of which has been omitted.

Four words or phrases, labeled 1, 2, 3 and 4, are given below the conversation. Choose the word or phrase that will correctly complete the conversation. Underline the answer you think is correct.

1. What year did you __ university?

graduate

graduate from

graduating

graduating from

2. Chicago is a large city, __?

aren't it

doesn't it

won't it

isn't it

3. Do you enjoy __?

to swim

swimming

swim

to swimming

4. Do you have __ to do today? We could have a long lunch if not.

many work

much work

many works

much works

5. When will the meeting ___?

hold on

hold place

take on

take place

6. Why don't you ___ us?

go to the house party with

go together the house party with

go the house party with

together the house party with

7. They didn't ___ John's plan?

agree with

agree to

agree

agree about

8. Tomorrow is Paul's birthday. Let's ___ it.

celebrate

praise

honor

congratulate

9. It's snowing. Would you like to ___ on Saturday or Sunday?

skiing

go to ski

go skiing

go ski

10. Do you feel like __ now?

swimming

to swim

swim

to go swimming

11. "What happened to them last night? They look depressed"

"I don't think __ happened."

nothing

everything

something

anything

12. "Those students will perform the annual school play."

"Yes, it is __ for next week."

due

scheduled

time-tabled

put on

**Hearing Screening
Speech Perception Lab**

Participant #: _____

Date: _____

Tester: _____

PURE-TONE SCREENING

Right

Left

_____	250Hz (at 25dB)	_____
_____	500Hz (at 20dB)	_____
_____	1000Hz (at 20dB)	_____
_____	2000Hz (at 20dB)	_____
_____	4000Hz (at 20dB)	_____
_____	8000Hz (at 20dB)	_____

PASS / FAIL

RECOMMENDATION

_____ Refer for full evaluation to Hearing Clinic

ANEXO II

Corpus estímulos silábicos

Tabla 1. Clasificación de las consonantes atendiendo a su representación ortográfica y posición silábica

Consonantes en posición inicial		Consonantes en posición final	
Sonidos (representación ortográfica)	Ejemplos	Sonidos (representación ortográfica)	Ejemplos
p-	past, pool	-p	top, soup
pl-	please, plan	-pt	hoped, leaped
pr-	proof, preen	-ps	cups, rips
f-	first, feed	-f	leaf, off
fl-	flat, flee	-th	moth, truth
fr-	fruit, free	-t	eat, pot
th-	third, theme	-ts	treats, shoots
t-	tot, too	-s	peace, boss
tr-	true, tree	-st	east, boost
s-	sad, city (soft -c)	-sts	feasts, costs
ch-	church, chief	-ch	teach, notch
sh-	shoe, sure	-sh	wash, push
k-	keep, curl	-k	lock, fluke
kl-	class, kleenex	-kt	peeked, looked
kr-	creed, crew	-ks	leaks, socks
kw-	quack, quirk	-b	rob, tube
h-	he, had	-bz	sobs, rubs
b-	be, boot	-v	move, sleeve
bl-	blurt, blue	-vz	halves, proves
br-	brief, brute	-d	need, food
w-	were, we	-dz	noods, moods
m-	me, moon	-z	bees, choose
v-	vast, verse	-j	lodge, stodge
th(v)	these, there	-g	dog, hug
d-	dirt, dude	-m	dream, loom

Consonantes en posición inicial		Consonantes en posición final	
Sonidos (representación ortográfica)	Ejemplos	Sonidos (representación ortográfica)	Ejemplos
l-	least, loose	-mz	rooms, chimes
dr-	dream, drew	-n	clean, swan
n-	new, knees	-nt	hint, want
z-	zoo, zee	-nd	pond, leaned
j-	jury, germ (soft – G)	-nts	prince, mints
r-	roost, real	-nz	means, beans
y-	you, year	-ng	bring, song
g-	girl, goon	-ngz	sings, lungs
gl-	glass, glee	-vowel	tee, new
gr-	great, group	-R	stir, cur
sp-	spur, speak	-L	peel, doll
sw-	swim, sweet		
sm-	smack, smooth		
st-	staff, stew		
sl-	slap, sleep		
str-	street, strap		
sn-	sneak, snooze		
sk-	skirt, school		
skr-	scream, scrap		
vowel-	oops, eat		

ANEXO III

Corpus de entrenamiento oracional

1. Don't look at the sun.
2. There are bass in Lake Monroe.
3. I just need to sit down.
4. You should get some rest.
5. Go ahead with the plan.
6. They tried hard to impress them.
7. I am late for the show.
8. It is always nice to hear your voice.
9. That is a pretty song.
10. Try to be downtown on time.
11. Can you afford it?
12. Go into the next room and wait.
13. The tomatoes are not ripe.
14. Buckle your seat belt.
15. We have five minutes left.
16. It was great to see you yesterday.
17. That sound is getting annoying.
18. There was a fire drill today.
19. The meat is too tough to cut.
20. Thanks for the flowers.
21. The symphony concert is free.
22. When should we expect you?
23. Should I buy us some snacks?
24. It has been a great year.
25. My son isn't going to join us.
26. The elevator doors close automatically.
27. Please turn left at the light.
28. Keep that wound covered.
29. Set up an emergency family meeting place.
30. Thank you for the nice gift.
31. Could you pick me up some cheese?
32. How much is the newspaper?
33. The market is located on Morton Street.
34. We are staying in a hotel.
35. I had a hard time getting up.
36. He calls every night.
37. Sorry, I forgot your name.
38. How do you like your steak?
39. He brought in the water.
40. Gas is much cheaper in Indianapolis.
41. They are following new procedures.
42. He just joined the club.
43. I was sick all last week.
44. Why don't you come over tonight?
45. My doctor is very thorough.

ANEXO IV

Figuras de los mapas de los tests de entrenamiento

p-	f-		t-	s-	ch-	sh-	k-	h-	Voiceless
pr-	fr-		tr-						
b-	v-		d-		j-		g-		Voiced
			l-				r-		
w-									
m-			n-						S Cluster
			st-						
Vowel-									
						START			
									SNR Adapt OFF
									SNR

Figura 23. Mapa de etiquetas para el test de identificación de ataques (CV) 1

			-t	-s			-k	Voiceless
				-st				
			-ts					
								Voiced
	-v		-d	-z				
								Nasal Clusters
-m			-n				-ng	
			-nt	-nd				
			-nts	-nz				
-Vowel			-R				-L	

Figura 24. Mapa de etiquetas para el test de identificación de codas (VC) 1

p-	f-	th-	t-	s-	ch-	sh-	k-	h-	Voiceless
pl-	fl-						kl-		
pr-	fr-		tr-				kr-		
							kw-		
b-	v-	th(v)-	d-	z-	j-		g-		Voiced
bl-			l-				gl-		
br-			dr-			r-	gr-		
w-						y-			
m-			n-						
									S Cluster
sp-			st-				sk-		
			sl-						
			str-				skr-		
sw-									
sm-			sn-						
Vowel-									
						START			
									SNR Adapt OFF
									SNR

Figura 25. Mapa de etiquetas para el test de identificación de ataques (CV) 2

-p	-f	-th	-t	-s	-ch	-sh	-k	Voiceless
-pt				-st			-kt	
-ps			-ts	-sts			-ks	
-b	-v		-d	-z	-j		-g	Voiced
-bz	-vz		-dz					
-m			-n				-ng	Nasal Clusters
			-nt	-nd				
-mz			-nts	-nz			-ngz	
-Vowel			-R				-L	

Figura 26. Mapa de etiquetas para el test de identificación de codas (VC) 2

	(1)	(2)	(3)	(4)	(5)
<div style="border: 1px solid red; padding: 5px; display: inline-block;"> TEMPORAL PENALTY 1 </div> <p>Click the mouse on the words you heard.</p>	AN	AND	AS	AT	DON'T
	DOSED	DOTE	DOVE	LOOK	LOOKED
	DOSED	DOVE	LOOK	LOOKED	LOOKS
	DOSED	DOVE	LOOK	LOOKED	LOOKS
	DOSED	DOVE	LOOK	LOOKED	LOOKS
	DOSED	DOVE	LOOK	LOOKED	LOOKS
	DOSED	DOVE	LOOK	LOOKED	LOOKS
	DOSED	DOVE	LOOK	LOOKED	LOOKS
	DOSED	DOVE	LOOK	LOOKED	LOOKS
	DOSED	DOVE	LOOK	LOOKED	LOOKS
	SOME	SONS	SUCH	SUN	THE
	THEM	THEN	THUS		

Figura 27. Mapa de etiquetas para el test de identificación de oraciones

ANEXO V

**Tablas de estadísticos descriptivos: modo
de articulación y sonoridad en posición
inicial y final**

Tabla 2. Medias atendiendo al modo de articulación en posición inicial

SNR	N	Mínimo	Máximo	Media	Desv. típ.	
40 SNR	CV_Oclusivas	11	75,00	95,83	87,1212	6,57609
	CV_Fricativas	11	53,57	78,57	65,9091	9,89275
	CV_Africadas	11	37,50	100,00	70,4545	16,07864
	CV_Líquidas	11	50,00	100,00	85,2273	14,59686
	CV_Nasales	11	75,00	100,00	93,1818	8,59440
	CV_Aproximantes	11	41,67	100,00	76,5152	16,16674
	N válido (según lista)	11				
15 SNR	CV_Oclusivas	11	62,50	100,00	86,3636	11,65042
	CV_Fricativas	11	60,71	89,29	70,4545	11,07614
	CV_Africadas	11	12,50	100,00	67,0455	24,54125
	CV_Líquidas	11	62,50	100,00	87,5000	12,50000
	CV_Nasales	11	75,00	100,00	96,5909	8,08337
	CV_Aproximantes	11	58,33	100,00	83,3333	13,94433
	N válido (según lista)	11				
5 SNR	CV_Oclusivas	11	62,50	91,67	81,0606	9,56767
	CV_Fricativas	11	46,43	92,86	64,2857	14,28571
	CV_Africadas	11	12,50	87,50	56,8182	21,18640
	CV_Líquidas	11	25,00	100,00	54,5455	23,89608
	CV_Nasales	11	50,00	100,00	82,9545	16,07864
	CV_Aproximantes	11	50,00	91,67	73,4848	15,73133
	N válido (según lista)	11				

Tabla 3. Medias atendiendo al modo de articulación en posición final

SNR	N	Mínimo	Máximo	Media	Desv. típ.	
40 SNR	VC_Oclusivas	11	73,33	100,00	90,6061	10,52174
	VC_Fricativas	11	46,67	100,00	71,5152	16,01136
	VC_Africadas	11	40,00	100,00	65,4545	18,63525
	VC_Líquidas	11	80,00	100,00	95,4545	6,87552
	VC_Nasales	11	73,33	100,00	90,3030	9,12318
	N válido (según lista)	11				
15 SNR	VC_Oclusivas	11	66,67	100,00	89,0909	10,96367
	VC_Fricativas	11	46,67	86,67	72,4242	11,06455
	VC_Africadas	11	40,00	100,00	64,5455	16,94912
	VC_Líquidas	11	20,00	100,00	86,3636	22,48232
	VC_Nasales	11	66,67	93,33	80,0000	9,42809
	N válido (según lista)	11				
5 SNR	VC_Oclusivas	11	36,67	90,00	75,4545	15,65409
	VC_Fricativas	11	26,67	93,33	70,9091	22,06510
	VC_Africadas	11	40,00	100,00	61,8182	19,40009
	VC_Líquidas	11	50,00	100,00	79,0909	15,13575
	VC_Nasales	11	40,00	86,67	53,3333	12,99573
	N válido (según lista)	11				

Tabla 4. Medias atendiendo a la sonoridad en posición inicial

SNR	N	Mínimo	Máximo	Media	Desv. típ.
CV_Sonoras	11	65,38	82,69	77,0979	5,33317
40 SNR CV_Sordas	11	69,44	88,89	78,7879	6,72116
N válido (según lista)	11				
CV_Sonoras	11	59,62	96,15	77,2727	9,87414
15 SNR CV_Sordas	11	66,67	94,44	84,3434	8,89520
N válido (según lista)	11				
CV_Sonoras	11	48,08	88,46	66,6084	13,41400
5 SNR CV_Sordas	11	52,78	88,89	75,5051	11,30512
N válido (según lista)	11				

Tabla 5. Medias atendiendo a la sonoridad en posición final

SNR	N	Mínimo	Máximo	Media	Desv. típ.
VC_Sonoras	11	78,18	98,18	86,1157	8,18274
40 SNR VC_Sordas	11	57,50	100,00	77,2727	12,72078
N válido (según lista)	11				
VC_Sonoras	11	45,45	98,18	80,9917	13,56621
15 SNR VC_Sordas	11	67,50	87,50	77,5000	7,74597
N válido (según lista)	11				
VC_Sonoras	11	34,55	90,91	62,9752	14,68732
5 SNR VC_Sordas	11	42,50	95,00	78,4091	14,88593
N válido (según lista)	11				