



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

**FACULTAD DE PSICOLOGÍA
DIVISIÓN DE ESTUDIOS PROFESIONALES**

**“CONTRIBUCIONES BIOLÓGICAS, PSICOFISIOLÓGICAS
Y CULTURALES A LA PROSODIA DEL LENGUAJE”**

T E S I S

**QUE PARA OBTENER EL TÍTULO DE
LICENCIADO EN PSICOLOGÍA**

**PRESENTA
JONATHAN AZAEL CABALLERO MENESES**

**DIRECTORA DE TESIS:
DRA. JUDITH MARINA MENEZ DIAZ**

**COMITÉ TUTORAL:
DR. FLORENTE LÓPEZ RODRÍGUEZ
MTRA. CONCEPCIÓN MORÁN MARTÍNEZ
DRA. MARÍA DEL CARMEN SUSANA ORTEGA PIERRES
DRA. ALICIA VÉLEZ GARCÍA**

**TESIS APOYADA POR EL PROYECTO:
DGAPA IN304211**



CIUDAD UNIVERSITARIA

NOVIEMBRE, 2011



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

Agradecimientos:

A mis padres.

Sin ustedes no lo habría logrado

A Marina

*Gracias por todo el apoyo,
y por mantenerme siempre dudando*

Al Dr. Florente López

Por el apoyo recibido

A la Mtra. Concepción Morán.

*Por presentarme toda una nueva
perspectiva de pensamiento*

A la Dra. Susana Ortega

Por la amistad y apoyo

A la Dra. Alicia Vélez

*Por el apoyo. Recuerdo con mucho
cariño los primeros semestres*

*A quienes me han influido y
acompañado en este trayecto*

Índice

1. RESUMEN.....	3
2. INTRODUCCION.....	4
3. MARCO REFERENCIAL	12
3.1 Nociones Previas	12
3.1.1 Fonemas y Diferencias Contrastivas.....	12
3.1.2 Diferencias Rítmicas entre Lenguajes.....	13
3.1.3 Mecanismos Generales de la Producción de la Voz	14
3.2 Prosodia Emocional. Abordajes Psicológicos, Acústicos, Computacionales y Psicofisiológicos.....	16
3.2.1 El abordaje neurobiológico.....	17
3.2.2 La perspectiva acústica y aproximaciones computacionales.....	24
3.2.3 Investigación intercultural.....	34
3.2.4 Resumen y discusión.....	44
3.3 Prosodia Lingüística y Abordajes Lingüísticos.....	58
3.3.1 Funciones de la prosodia.....	61
3.3.2 Fisiología de la prosodia lingüística.....	81
3.3.3 Estudios interculturales en prosodia lingüística.....	84
3.3.4 Teorías que integran la perspectiva de influencias biológicas y culturales a la prosodia lingüística.....	92
3.3.5 Resumen y discusión.....	102
3.4. Prosodia Lingüística y Prosodia Emocional.....	110
3.4.1 Contrastes fisiológicos entre prosodia lingüística y emocional y posturas alternativas.....	112
3.4.2 Estudios Acústicos sobre Prosodia Lingüística y Prosodia Emocional	127
3.4.3. Elaboraciones Teóricas.....	134
3.4.4 Resumen y discusión.....	150
3.5 Temas complementarios	158
3.5.1 Acentos y Síndrome del Acento Extranjero.....	158
3.5.2 Música y Prosodia.....	165
3.5.3 Lenguajes Tonaless y Prosodia	175
3.5.4 Prosodia en Lenguajes de Señas	181

3.5.5 Desarrollo de la Prosodia.....	189
3.5.6 Finalización del núcleo empírico	197
3.6 Análisis Conceptual.....	198
4. Conclusión.....	203
5. Referencias	213
6. Anexos	233
6.1 Anexo 1: Tabla de definiciones.....	233

1. RESUMEN

Dado el interés en la influencia de variables biológicas, psicofisiológicas y culturales en la prosodia, así como las dificultades de acuerdo teórico en el campo de la prosodia, específicamente en la división de la misma en las categorías lingüística y emocional, se analizaron dichas fuentes de influencia en la prosodia del lenguaje y se analizó empírica y conceptualmente la noción tanto de prosodia, en general, como de la división dicotómica de la prosodia en las categorías emocional y lingüística. Para lograr este objetivo se revisó de manera amplia la literatura con interés específico en la prosodia lingüística y emocional, así como literatura donde se las compara puntualmente y se buscan efectos de interacción. Se revisaron perspectivas acústicas, perceptuales y neurofisiológicas para aclarar la división. Se revisaron algunos temas relacionados a la prosodia con el potencial de aislar ciertas variables y resolver interrogantes cuyo abordaje podría no ser óptimo, o podría estar sesgado, al revisar sólo los trabajos interesados específicamente en la prosodia. Se encontró apoyo empírico para la división dicotómica de la prosodia y se discutió que podría estar guiado por el uso amplio, *a priori*, del marco teórico que la considera dicotómica. Se exploran modelos alternativos y se sugiere el uso de una división de la prosodia en 5 categorías. Se encontraron algunas diferencias conceptuales en el uso de términos entre distintos campos que podrían generar confusiones al abordar un problema desde diversas disciplinas. Se sugiere el uso de análisis conceptual para la clarificación de términos y el mejor avance de la ciencia en temas que caen en el dominio de diversas disciplinas. Se propusieron dos definiciones para el término “prosodia”, una restringida a lenguajes orales y una con aplicación a lenguajes de señas y potencialmente a otras modalidades, así como una hipótesis integradora, con miras al avance teórico del área: la *“hipótesis de grados de libertad”*. Se proponen algunas metodologías y se sugieren algunas aplicaciones prácticas del estudio de la prosodia en campos diversos como el diagnóstico y tratamiento de alteraciones del lenguaje, enseñanza de lenguas, interfaces computacionales, entre otros.

Palabras clave: Prosodia, Análisis conceptual, Interdisciplina.

2. INTRODUCCION

La prosodia del lenguaje se refiere a distintas características vocales del habla, diferentes de la morfología, sintaxis, pragmática y el significado de lo dicho, aunque en estrecha relación con todas ellas. Dichas características vocales, incluyen la entonación, acento prosódico¹ y características temporales como el ritmo (Thompson & Balkwill, 2006). Dicho de otra manera, se refiere a las características suprasegmentales² (Zakzanis, 1999; Zhang, Li, Lo, & Meng, 2010) que contribuyen a la “melodía” del habla, entre ellas la frecuencia fundamental (tono), duración e intensidad, así como el ritmo, entonación, entre otras. Desde un punto de vista lingüístico, contribuye a señalar distinciones comunicativas léxicas, gramaticales y emocionales (Harwell Rodriguez, 1998), así como énfasis y distinción entre declaraciones, interrogaciones, exclamaciones, etc. (Thompson & Balkwill, 2006).

La comprensión (y producción) de la información prosódica puede darse independientemente de la comprensión del contenido; de hecho, hay lesiones cerebrales (afasias y aprosodias) que pueden afectar principalmente a una de ellas. Esto se expresa en personas que pueden hablar fluidamente y con entonaciones emocionales y de distinción de preguntas, declaraciones y otras funciones, pero sin ningún contenido en el habla, como en la *afasia de Wernicke*; también se observa la capacidad de percibir estas distinciones funcionales, ya que responden algo ante una pregunta entonada como tal, aunque sin relación a lo preguntado. También se han observado casos de personas que pueden producir habla con contenido y entender lo que otros dicen, pero no hacen distinciones prosódicas, o las hacen de manera incorrecta (por ejemplo, en una tarea en la que normalmente participantes sanos enfatizan una palabra, éstos no lo hacen o pueden enfatizar otra) lo que lleva a un habla “plana”; también tienen dificultades para usar la información prosódica del habla de otras personas para desambiguar

¹ Aunque parece llevar en la definición el término a definir, en la literatura, normalmente se refiere como *stress* al mismo concepto, es en español que se le suele denominar *acento prosódico*.

² En lingüística, se denomina *suprasegmentales* a los elementos que afectan a las emisiones a un nivel mayor que el fonema (Gussenhoven, Suprasegmentals, 2001).

enunciados y distinguir funciones e intenciones; estos casos se denominan de *aprosodia* (Buchanan, Lutz, Mirzazade, Specht, Shah, Zilles & Jäncke 2000; Carlson, 2006a).

Esta disociación entre la comprensión de la información prosódica y del contenido se ilustra también en experimentos que exploran la percepción de la información prosódica en lenguas desconocidas para los participantes. En estos experimentos se encuentra que se pueden identificar algunos elementos prosódicos en lenguas desconocidas. El tema ha sido estudiado sobre todo con respecto a la información emocional (Pell, 2006; Pell, Monetta, Paulmann, & Kotz, 2009; Pell & Skorup, 2008; Sauter, Eisner, Ekman, & Scott, 2010; Thompson & Balkwill, 2006) y en menor medida, con respecto a intenciones del habla, con un enfoque más centrado en la pragmática³ (Shochi, Aubergé, & Rilliard, 2006).

La psicología ha estudiado a la prosodia principalmente en su dimensión emocional. Al respecto, algunos autores proponen características más o menos universales de la prosodia emocional (Thompson & Balkwill, 2006), mientras que otros consideran que tiene contribuciones tanto biológicas como culturales (Grandjean & Scherer, 2006; Pell, 2006; Scherer & Bänziger, 2004). De hecho, se ha propuesto que la prosodia emocional tiene un origen filogenético más antiguo a las capacidades léxicas, y que era el principal canal de comunicación, mantenido hasta hoy por su valor evolutivo (Thompson & Balkwill, 2006; Thönnessen, Boers, Dammers, Chen, Norra & Mathiak, 2010).

Por su parte, la lingüística se ha centrado en temas como el desarrollo de la prosodia del lenguaje (DePaolis, Vihman, & Kunnari, 2008) su papel en la desambiguación, énfasis y distinciones funcionales (distintos tipos de enunciados, como declaraciones e interrogaciones, entre otras; Frazier, Carlson, & Clifton Jr, 2006; Hirschberg, 2002; Shapiro & Nagel, 1995; Wilson & Wharton, 2006) y actitudes (como crítica, sarcasmo, incredulidad, dudas, arrogancia etc; Batista

³ Se refiere a las características del habla que comunican significados no literales y muchas veces dependientes de contexto. Entre ellos, intenciones, como la ironía y el sarcasmo. Este tipo de aspectos pueden ser comunicados mediante variaciones prosódicas (Paradis, 1998; Shochi, Aubergé, & Rilliard, 2006).

Antunes, 2007; Shochi, Aubergé, & Rilliard, 2006). Desde esta perspectiva, aunque a veces no sean consideradas “propiamente lingüísticas”, las funciones emocionales de la prosodia (Wilson & Wharton, 2006) también han recibido atención (Cid Uribe & Maluenda Parraguez, 2005), así como las contribuciones biológicas a la prosodia (DePaolis, Vihman, & Kunnari, 2008). La contribución de variables fisiológicas y culturales a la prosodia lingüística y emocional podría contribuir a su diferenciación como entidades independientes: se podría esperar que en la prosodia lingüística, cuya base se puede asumir, *a priori*, aprendida, hubiera un mayor peso de variables culturales; mientras que se podrían esperar más influencias biológicas en la prosodia emocional, si se asume con base evolutiva. La distinción conceptual entre prosodia lingüística y emocional será importante, en especial al revisar los sustratos fisiológicos asociados a la prosodia entendida como emocional o lingüística.

Las distinciones entre las líneas de investigación psicológica y lingüística no son tan tajantes. Efectivamente, entran en contacto continuamente, así como con las neurociencias. Ha habido aproximaciones fisiológicas tanto a la prosodia emocional (Buchanan et al, 2000; Grimshaw, Séguin, & Godfrey, 2009; Péron, Grandjean, Le Jeune, Sauleau, Haegelen, Drapier, Rouaud, Drapier, Vérin, 2010; Ross & Monnot, 2008) como a la lingüística (Mildner, 2004; Shah, Baum, & Dwivedi, 2006). Así como estudios donde se las compara (Grandjean & Scherer, 2006; Van Lancker, Pachana, Cummings, & Sidtis, 2006).

Con todo, existe controversia en la distinción de estos términos, y desacuerdos respecto a que se trate de fenómenos independientes (Seddoh, 2002), así como en lo referente a la organización anatómica y funcional de ambas (Gandour, Dziedzic, Wong, Lowe, Tong, Hsieh, Sattamnuwong & Lurito, 2003).

Esta dificultad para distinguir conceptualmente entre la prosodia emocional y lingüística, así como las observaciones de la influencia de variables biológicas, psicofisiológicas y culturales tanto en la llamada prosodia emocional (Grandjean & Scherer, 2006; Pell, 2006; Scherer & Bänziger, 2004) como en la prosodia lingüística (Cid Uribe & Maluenda Parraguez, 2005; DePaolis, Vihman, & Kunnari,

2008) representan un problema que requiere clarificación, como es reflejado en los trabajos que discuten el tema, y dada la influencia en el curso de la investigación basada en las diferentes posturas existentes hoy en día. Resolver estas dificultades conceptuales y desarrollar un marco unificado apoyaría el curso de la investigación empírica. Asimismo, un análisis conceptual adecuado⁴ permitirá una mejor comprensión de la prosodia. Revisar los sustratos fisiológicos de la prosodia en estas dos acepciones permitirá aclarar si son mejor caracterizados como fenómenos independientes o como un fenómeno único. Además, la revisión de estudios interculturales sirve tanto a los propósitos de aclarar la influencia de la cultura (y por tanto de entornos distintos y del aprendizaje), como la posible independencia de la cultura, y por tanto la identificación de características más universales y presumiblemente biológicas.

Una caracterización adecuada de la prosodia puede, además, tener aplicación clínica al convertirse en una base teórica fuerte en el diseño de terapias para aprosodias y síndrome de acento extranjero (al haberse identificado que la principal alteración asociada es a nivel prosódico) así como aplicaciones educativas y sociales, en la enseñanza de idiomas y entrenamiento de acentos; ya que, nuevamente, la prosodia contribuye de manera importante en la percepción del habla como extranjera y el ser percibido como tal puede tener consecuencias sociales importantes como la exclusión y la discriminación (en un capítulo posterior se desarrollará más ampliamente el tema del síndrome del acento extranjero y se hará una breve mención de las consecuencias sociales del ser percibido como extranjero a partir del habla).

Tomando en cuenta estos antecedentes, el objetivo del presente trabajo de tesis es investigar cómo influyen los factores biológicos y culturales a la prosodia, entendida de manera general como las características suprasegmentales del habla. Y de manera más puntual, investigar cómo influyen diferencialmente estos factores en la prosodia emocional y lingüística.

⁴ El análisis conceptual tiene por objetivo analizar el lenguaje de la ciencia, aclarar significados, las relaciones entre proposiciones teóricas y revisar la consistencia de las teorías (Machado & Silva, 2007)

Otro objetivo es la realización de un análisis conceptual de las nociones de prosodia lingüística y emocional (en el capítulo de “*Prosodia lingüística y prosodia emocional*”), haciendo énfasis en las distinciones fundamentales entre ellas y en la naturaleza funcional de esta distinción; así como del término general “Prosodia” (en el capítulo de “análisis conceptual”).

Se revisarán las bases fisiológicas de la prosodia, en general, es decir, vista como un todo, así como los trabajos que han surgido de las distinciones conceptuales que la dividen en emocional y lingüística. Todo ello para aclarar las características que estas últimas comparten y las que las distinguen en el nivel empírico.

Aunque estos fenómenos (prosodia emocional y lingüística) se den en general de manera concurrente, se propone una forma de analizarlos independientemente, con el objetivo de lograr la mayor claridad en la caracterización empírica y conceptual de los mismos. Para ello se estudiarán tanto las investigaciones con interés puntual en una de ellas, como los que las comparan, tanto a nivel perceptual como fisiológico e intercultural.

En efecto, las similitudes y diferencias entre estos dos tipos de prosodia, ilustradas en estudios tanto exploratorios como experimentales – en cuanto propiedades acústicas y temporales, así como la capacidad de distinguir emociones o funciones interculturalmente (en el nivel perceptual) y la organización anatómica y funcional de los tipos de prosodia (fisiológicamente) –, podrían hablar de las variables que las afectan independientemente y las que las afectan en conjunto. De hecho, Grandjean y Scherer (2006), describen cómo tanto las influencias de la actividad psicofisiológica y de las reglas culturales de expresión influyen tanto a la prosodia emocional como lingüística, aunque ciertamente la actividad psicofisiológica influye más fuertemente a la prosodia emocional mientras que las reglas culturales a la prosodia lingüística.

Ya que el aislar variables es una de las formas en las que se pueden estudiar de manera más específica, surgen las preguntas: ¿cómo aislar los dos tipos de prosodia?, ¿se pueden dar independientemente? ¿no todas las vocalizaciones del

lenguaje llevan implícitos estos dos tipos de prosodia? Se debe recordar que incluso algunos autores critican la distinción argumentando que no es necesario utilizar un calificativo (*lingüística*) a una entidad que se asume caracterizada por poseer dicha cualidad; por lo que *prosodia lingüística* sería un término redundante (Seddoh, 2002).

Una de las formas de determinar la contribución de distintas variables es evaluar todos los aspectos al mismo tiempo, variando sistemáticamente la presencia de los mismos y efectuando análisis estadísticos, de manera análoga al procedimiento empleado por Pell (1999). En su estudio, solicitaron a sus participantes que reprodujeran enunciados, en ellos se variaba el énfasis en una de las palabras de la oración; el modo: pregunta o declaración y el tono emocional: enojado, triste, alegre y neutral.

Otra estrategia es produciendo frases vacías de significado o pseudo-enunciados, como lo realizan Pell, Monetta, Paulmann y Kotz (2009) y Pell, Paulmann, Dara, Alasseri y Kotz (2009) para estudiar prosodia emocional, aunque esto no necesariamente excluye la presencia de factores de prosodia lingüística, ya que se puede esperar que algunas características generales de la prosodia requeridas por el idioma estén conservadas en las frases producidas para representar ciertas emociones; por ejemplo, el patrón de entonación que distingue una pregunta de una declaración.

¿Cómo aislar del habla cualquiera de ellas si se asume que son un componente de la misma? Una propuesta interesante se da en el seno de la investigación sobre música. En ella se ha encontrado que existen áreas cerebrales compartidas en el procesamiento de información musical y lingüística (Levitin & Menon, 2003; Patel, Peretz, Tramo, & Labreque, 1998), y aún más, que manipulaciones similares en las características de tono, duración e intensidad (que en el lenguaje se refieren a la prosodia al ser suprasegmentales) en música y habla, producen consecuencias similares a nivel emocional en los oyentes (Ilie & Thompson, 2006).

Efectivamente, las características de la música que afectan a las emociones (Caballero-Meneses & Menez, 2010; Gomez & Danuser, 2007) son análogas a las características propuestas del lenguaje que contribuyen a la prosodia emocional (Thompson & Balkwill, 2006), a saber, características de ritmo y tono, y sus distintas combinaciones. Quizá esta sea también una aproximación que puede aislar hasta cierto punto la prosodia emocional de la prosodia lingüística, que sería inexistente en la música, por lo que también será abordada en la tesis.

Otra aproximación es el estudio de la prosodia en lenguajes de señas, en los que se mantienen las características de los lenguajes, pero difiere la modalidad por la que se expresan (Newman, Supalla, Hauser, Newport, & Bavelier, 2010). Esta estrategia permitirá aclarar las características que son comunes a la prosodia en lenguajes orales y de señas y abordar un estudio de la prosodia aislada del habla. Por ejemplo, se ha visto que aspectos prosódicos en lenguajes de señas pueden tener funciones similares a la prosodia de lenguaje hablado, como marcas de tópico y focalización, así como señalamiento de fin o continuación de frases (Dachkovsky & Sandler, 2009), estas funciones serán ampliamente abordadas primero en el capítulo sobre prosodia lingüística y después se volverá a hacer mención en el capítulo de temas complementarios. A su vez, ha sido observado que personas que aprenden una lengua de señas extranjera presentan acento extranjero (Hickok, Bellugi, & Klima, 2001); en la literatura sobre síndrome de acento extranjero se pone de relieve que es un fenómeno principalmente prosódico (Blumstein & Kurowski, 2006; González-Álvarez, Parcet-Ibars, Ávila, & Geffner-Sclarsky, 2003), en los lenguajes de señas podría ser también principalmente prosódico el fenómeno.

Mediante todas estas aproximaciones se espera obtener un cuadro amplio, una caracterización conceptual fundamentada de los dos tipos de prosodia, y las contribuciones biológicas, psicofisiológicas y culturales a “la prosodia” y “las prosodias”.

Por supuesto, será hasta el final del presente trabajo cuando se podrá evaluar la utilidad del uso del enfoque propuesto, es decir, el enfoque de la exploración de la literatura perceptual, acústica, fisiológica y transcultural – e inclusive literatura relacionada de manera más distante como la mencionada sobre música y acento – para aclarar el problema de la división de la prosodia; importante dadas sus implicaciones teóricas y para el curso de la investigación empírica.

Se comenzará la exposición con un breve capítulo de nociones básicas y generales que serán empleadas a lo largo del trabajo, se proseguirá con dos capítulos que hacen, el primero, una revisión puntual de la prosodia emocional y, el segundo, de la prosodia lingüística. El siguiente capítulo considerará trabajos que han abordado simultáneamente a la prosodia emocional y lingüística, así como elaboraciones teóricas alternativas. Inmediatamente será presentado un capítulo de temas complementarios: aproximaciones que tocan temas como la música y los lenguajes de señas, y su contribución al estudio de la prosodia, de acuerdo con el enfoque delineado en los párrafos anteriores. Los temas complementarios constituirán el final de la revisión con enfoque más empírico y darán paso al capítulo de análisis conceptual. Finalmente, se dará paso a las conclusiones y se evaluará la utilidad del enfoque propuesto para abordar la problemática estudiada.

3. MARCO REFERENCIAL

3.1 Nociones Previas

Dado que se hará un uso constante de términos y conceptos del campo lingüístico, en este capítulo se ofrece de manera breve la definición más básica de los mismos. También se toca el mecanismo general de producción periférica de la voz, ya que es importante tenerlo en cuenta al leer los estudios que se presentarán en el cuerpo principal del trabajo de tesis. Esta sección puede omitirse en la lectura si se está familiarizado con conceptos como *fonemas*, *diferencias contrastivas* y *diferencias rítmicas entre lenguajes (ritmo basado en moras, sílabas, estrés)*. Asimismo, el mecanismo general de producción de voz no es fundamental para entender el desarrollo del trabajo de tesis, aunque sí es conveniente.

3.1.1 Fonemas y Diferencias Contrastivas.

Se considera una diferencia contrastiva a una variación que, entre la multiplicidad de variaciones posibles, es importante para hacer contrastes de significado, mientras que a las variaciones que no implican contrastes de significado se consideran diferencias no contrastivas. A dos sonidos (del nivel segmental) que comportan una diferencia contrastiva se les considera fonemas, mientras que a dos sonidos que no la comportan (pero pueden ocurrir en el habla) se les considera alófonos. Los fonemas son las diferencias contrastivas mínimas en los lenguajes, es decir, en un conjunto ordenado de ellos, cuando la substitución de uno solo afecta el significado de la palabra, se habla de un fonema; éstos, para diferenciarse de la escritura convencional, se escriben entre diagonales, lo que constituye una anotación fonológica (Hualde, 2001), mientras que se escribe entre corchetes una representación fonética, que captura diferencias acústicas existentes, pero no contrastivas. Por ejemplo, en español la diferencia entre /d/ y /p/ es contrastiva, ya que al sustituirse la /d/ por /p/ en /don/, se obtiene /pon/, con un significado distinto; una diferencia no contrastiva en español está ilustrada entre los fonemas /d/ y /ð/: aunque la palabra “dedo” se pronuncia [deðo], no se considera al sonido [ð] como diferente del sonido [d], a su vez, no hay diferencia de significado si se pronunciara [dedo] o incluso [ðeðo]; de esta manera, [d] y [ð]

son alófonos del fonema /d/ en español. Los fonemas /d/ y /ð/ sí son contrastivos en otras lenguas, por ejemplo, el inglés.

El mismo concepto de diferencias contrastivas se puede aplicar a la prosodia, por ejemplo, entre la multiplicidad de variaciones (ritmo, tonos, etc.) posibles para cualquier frase dada, sólo unas cuantas contrastarán una pregunta de una declaración, por ejemplo, o una sílaba tónica de una sílaba átona. Esto difiere entre lenguas; por ejemplo, en español no importa si se hace una variación de tonos al interior de las sílabas de la palabra “español”, su significado no cambiará; pero este tipo de variaciones afectarían el significado de la palabra en idiomas tonales como el mandarín.

3.1.2 Diferencias Rítmicas entre Lenguajes

Se ha propuesto que el ritmo de los lenguajes difiere de acuerdo a la unidad prosódica que lo determina, existiendo lenguajes de ritmo a partir de moras, sílabas o estrés (*mora-*, *syllabe-* y *stress-timed languages* en la literatura en inglés). La idea es que la duración de los componentes base del ritmo (moras, sílabas o sílabas con estrés) de las lenguas tienen aproximadamente la misma duración. Con todo, las evidencias empíricas no han apoyado esta noción, por lo que ha sido considerada una distinción auditiva y subjetiva que, sin embargo, es ampliamente utilizada. Otros trabajos empíricos han encontrado efectivamente diferencias rítmicas en lenguajes de ritmo a partir de sílabas o de estrés no en términos de una duración similar de sus componentes de base, pero sí en términos de variabilidad (mayor variabilidad en las sílabas de lenguas de ritmo a partir de estrés). Lenguajes como el español y el francés son considerados lenguajes de ritmo a partir de sílabas, el inglés y el alemán lenguajes de ritmo a partir de estrés y lenguajes como el japonés de ritmo a partir de moras (Dromey, Silveira, & Sandor., 2005; Patel & Daniele 2003).

Las sílabas están compuestas de un núcleo vocal o “sonorante” (es decir, de actividad de las cuerdas vocales) y de consonantes que lo anteceden (ataque) o que lo siguen (coda); la mora es una unidad más pequeña que la sílaba y mayor que un segmento (o unos de ellos); una sílaba contiene al menos una mora y rara

vez contiene más de dos, la existencia de las moras es ampliamente aceptada en lenguas como el japonés y es controversial en lenguas como el inglés (Shattuck-Hufnagel & Turk, 1996); los sistemas de escritura *kana* en japonés representan, precisamente, moras, compuestas por una consonante y una vocal, por vocales solas o por la consonante /n/ (ん). En cuanto al estrés, recae en sílabas particulares; lenguas como el inglés tienden a reducir las vocales en sílabas que no llevan estrés (aproximarse al sonido schwa (ə) y ser más cortas, entre otras variaciones) y a pronunciar de manera menos articulada sus consonantes asociadas, lo que contribuye a una percepción de “debilitamiento” de dichas sílabas.

3.1.3 Mecanismos Generales de la Producción de la Voz

En el habla se observan dos procesos separados que afectan a la señal acústica: fonación y articulación (García Toro, 1997).

La fonación se refiere a la producción de ondas sonoras en la laringe que posteriormente son modificadas por el proceso de articulación por los distintos órganos articuladores: lengua, labios, velo, paladar, dientes etc. (García Toro, 1997).

La fonación es afectada por dos procesos principales, controlados por la tensión de grupos musculares: el que controla la tensión de las cuerdas vocales, llamado *cricotiorideo* (cuya actividad muchas veces es referida simplemente como *tensión laríngea*) y los que afectan a la presión del aire subglotal: los músculos intercostales y respiratorios (cuya actividad suele ser referida simplemente como *presión subglotal*). El balance entre estos dos procesos determina la *frecuencia fundamental* de la voz o *f0* y la distinción sonoro/sordo de los segmentos, que simplemente indica si hay o no actividad en las cuerdas vocales (Bryant & Fox Tree, 2005; García Toro, 1997; Scott, Clegg, Rudge, & Burgess, 2006).

La vibración de las cuerdas vocales es la clave para la intensidad y frecuencia de la voz, específicamente, depende de variables como la masa, hidratación y estado de turgencia (rigidez celular), forma y tamaño de la laringe, y, como mencionamos

hace un momento, de la actividad muscular. La voz también es afectada por los resonadores: boca, cavidad nasal y senos paranasales, faringe y cavidad torácica (García Toro, 1997)

Finalmente, los articuladores, mediante sus diversas posiciones posibles, filtran la señal acústica, ampliando y reduciendo algunas frecuencias dando lugar a la señal final, con las variaciones espectrales y temporales finas de las que depende la adecuada percepción del habla (García Toro, 1997; Scott et al., 2006).

3.2 Prosodia Emocional. Abordajes Psicológicos, Acústicos, Computacionales y Psicofisiológicos.

La prosodia emocional es definida como la habilidad para expresar emociones a través de diferentes parámetros del habla humana, como el tono, la intensidad y la duración (Besson, Magne, & Schön, 2002; Buchanan et al, 2000; Wildgruber et al, 2005). De este modo, la característica que la distingue de la prosodia definida de manera general, es la inclusión de la función que desempeñan estas variaciones; es decir, el especificar que se refiere a la capacidad de expresar emociones, aunque se mantiene el énfasis en características suprasegmentales.

Sin embargo, también se ha propuesto su definición como “las modificaciones en características suprasegmentales y *segmentales* durante un episodio emocional” (Grandjean, Bänziger, & Scherer, citado en Péron et al, 2010). Algunos autores también consideran la cualidad de la voz en la definición de la prosodia (Grandjean, Bänzinger, & Scherer; 2006; Erickson, 2006; Pell, Monetta, Paulmann, & Kotz, 2009; Pell, Paulmann, Dara, Allasseri, & Kotz, 2009).

Existen diversas perspectivas para el estudio de la prosodia emocional. Entre ellas encontramos la perspectiva neurobiológica, que intenta describir los correlatos y la organización funcional tanto de la producción como de la percepción de la prosodia.

Otra perspectiva se ha centrado en el análisis de las características acústicas de la prosodia. En esta rama se intenta identificar que características del habla determinan la percepción de emociones. También se manipulan muestras de sonido en un intento por clarificar el papel de las variables previamente propuestas. Otros esfuerzos se dirigen a relacionar variables que se sabe afectan a los procesos emotivos y predecir qué efectos tendrían en la prosodia emocional (por ejemplo, el efecto del arousal sobre el tono, intensidad y ritmo del habla, dado que se sabe que afecta al sistema nervioso autónomo y a su vez que éste afecta a la producción vocal (Scherer & Bänziger, 2004). En otra área que desarrolla investigaciones relacionadas, se han creado diferentes sistemas computacionales

de reconocimiento automático de emociones, con diferentes objetivos, como búsqueda de contenidos sonoros basada en categorías emocionales o sistemas que permitan un mejor desempeño en los sistemas de “comunicación natural” con computadoras. Si bien estas aplicaciones no están diseñadas para explicar, se observa cómo muchas de las variables que emplean reflejan los hallazgos teóricos de los otros enfoques, aunque el procedimiento de selección de variables no está determinado por una teoría sino por el desempeño de los programas. Este tipo de aplicaciones también puede indicar variables que no han sido exploradas teórica y empíricamente.

Un enfoque algo distinto ha investigado la percepción de emociones de manera intercultural. Desde esta perspectiva se busca determinar el grado en el que la comunicación de las emociones es independiente del idioma y de la cultura. Si bien la capacidad de reconocer emociones se conserva en cierta medida en este tipo de experimentos, también se observa un mejor desempeño al interior del propio idioma, y a veces un efecto de “gradiente” en el desempeño que depende de la proximidad de los idiomas (por ejemplo, precisión decreciente en la tarea de reconocimiento de emociones en español, portugués, inglés y árabe por un hispanohablante).

Una de las propuestas para la explicación de este fenómeno se refiere a la convencionalización de las formas de expresar emociones y otra a intentos de regulación activa de la expresión de emociones dependientes de enfoques culturales.

En la presente sección se revisarán los principales hallazgos de estas tradiciones, así como los principales debates teóricos y propuestas de solución a los mismos.

3.2.1 El abordaje neurobiológico.

Desde finales de la década de 1870, se empezó a observar que pacientes afásicos, incluso algunos con dificultades severas, podían comunicar efectivamente sus intenciones a partir de la modulación emocional. El habla de

estos pacientes era considerada un habla empobrecida, pero la modulación emocional seguía presente, lo que sugería un papel especial del hemisferio derecho en aspectos emocionales de la comunicación; de hecho, fue a partir de quejas sobre dificultades psicosociales de pacientes después de la recuperación parcial de lesiones cerebrales que se empezaron a estudiar las bases neurobiológicas de la prosodia. Mucho tiempo después, hacia mitades de la década de 1970, iniciaron los primeros estudios formales sobre dificultades en la comunicación del afecto (Ross & Monnot, 2008).

Poco después, se propuso una analogía entre los déficits aprosódicos y las afasias. Es decir, que la organización cerebral del hemisferio derecho con respecto a la producción y percepción de la prosodia, era similar a la del hemisferio izquierdo para el lenguaje (y sus áreas clásicas como Broca y Wernicke; Ross, 1981, citado en Ross & Monnot, 2008). Esta propuesta ha sido recogida e investigada por diversos autores (Buchanan et al, 2000; Wildgruber et al, 2005).

A partir de éstas y posteriores observaciones (en especial en las que se empezó a reportar que las dificultades prosódicas también podían surgir a partir de lesiones en el hemisferio izquierdo), se ha propuesto que el procesamiento de la prosodia es una función principalmente del hemisferio derecho, aunque no exclusivamente. Se postula una organización distribuida en ambos donde el hemisferio derecho juega un papel fundamental (Ross & Monnot, 2008; Pell, 2002; Wildgruber et al, 2005).

¿Por qué lesiones del lado izquierdo podrían traer dificultades prosódicas si se asume que la prosodia es una función lateralizada principalmente al hemisferio derecho? Una de las explicaciones que se ha propuesto tiene que ver con la comunicación interhemisférica. De esta manera, se añadiría a la producción articuladora y verbal (proceso lateralizado hacia el hemisferio izquierdo), las características acústicas y prosódicas relacionadas a las emociones (un proceso más lateralizado al hemisferio derecho), modificando el plan motor para la articulación (Klouda, Robin, Graff-Radford, & Cooper, 1988; Ross & Monnot,

2008); este proceso sería también válido para otros usos pragmáticos del lenguaje, como el sarcasmo y otras actitudes. De manera consistente, se ha observado incapacidad de repetir tonos emocionales en casos de desconexión callosa y se encuentra principalmente daño subcortical en casos de dificultades prosódicas por lesiones en el hemisferio izquierdo (Ross & Monnot, 2008), situación que sugeriría dificultades de tráfico de la información proveniente del hemisferio derecho. Este tipo de observaciones permiten apreciar cómo a pesar de las observaciones de dificultades prosódicas debidas a lesiones del hemisferio izquierdo, la prosodia puede seguir siendo descrita como una función lateralizada principalmente al hemisferio derecho.

Las áreas del cerebro (tanto áreas corticales como subcorticales) que se han relacionado a la prosodia son diversas. Su identificación ha involucrado el uso de diversas metodologías, como los estudios de lesiones, imagenología, mapeo cerebral previo a operaciones neurológicas y procedimientos con resonancia magnética transcraneal.

En el caso de dificultades afásicas y aprosódicas en humanos, suelen ser menos severas cuando involucran exclusivamente regiones subcorticales. En la organización funcional de los procesos de comunicación en otras especies también se observa el involucramiento de regiones subcorticales (Ross & Monnot, 2008).

Los ganglios basales se encuentran entre las estructuras que han recibido mayor atención al respecto. Estos están involucrados principalmente con la conducta motora, aunque algunas regiones también con actividades cognitivas.

Disfunciones en estas áreas (independientemente del hemisferio) han sido asociadas con aprosodias. Y son más graves cuando se presentan de manera conjunta con daño cortical, en especial en el lóbulo temporal (Pell, 2002).

Otra fuente de evidencia con respecto al papel de los ganglios basales en la prosodia proviene del estudio de las dificultades en esta dimensión del lenguaje en personas con Parkinson, una enfermedad caracterizada principalmente por la

degeneración progresiva y más o menos focalizada de los ganglios basales. Estas dificultades se ilustran al comparar la capacidad de identificar el afecto a partir de pistas semánticas (enunciados cuyo significado refiere el estado emocional) contra prosódicas (Pell, 2002). Una de las posibles fuentes de estas dificultades podría estar en deficiencias en el procesamiento de características temporales sutiles de la prosodia, relacionadas, y quizá basadas en observaciones generales de alteraciones en la percepción (*y producción*) temporal en personas con Parkinson y reflejadas también en alteraciones temporales en el habla y la conducta que constituyen parte de las características principales del Parkinson (Breitenstein, Van Lancker, Daum, & Waters, 2001). A partir de un experimento sobre percepción de tiempo en personas con Parkinson, donde se variaron simultáneamente tiempo y tono, se ha sugerido un procesamiento independiente de tiempo y tono, ya que se encontró alterada la percepción temporal pero preservada la percepción tonal en estos pacientes (Harrington et al, citado en Breitenstein et al, 2001).

La interpretación sobre la participación de los ganglios basales basada en el estudio del modelo de Parkinson tiene que tomarse con cuidado, tanto por su posible participación en aspectos prosódicos generales (es decir, que incluyen la discriminación de la prosodia lingüística; Pell, 2002), como por el hecho de que la mayoría de sus proyecciones se dirigen a áreas como la corteza dorsolateral frontal, entorrinal, cíngulo anterior y otras regiones, por lo que parte de las dificultades podrían deberse a una disfunción cortical o a una combinación cortical y subcortical (Ross & Monnot, 2008).

Wildgruber y colaboradores (2005), realizaron un experimento dirigido a explorar las áreas cerebrales involucradas en la percepción de prosodia emocional como un proceso general y áreas involucradas en la percepción de emociones específicas, guiados por observaciones en el campo de percepción de emociones en rostros. Esta analogía de la voz con los rostros ha sido discutida antes, dada la capacidad tanto de la voz, como de los rostros, de comunicar información sobre la

identidad, género, edad aproximada, y estado emocional del emisor, entre otras cosas (Belin, Fecteau, & Bédard, 2004).

En su experimento, Wildgruber et al. (2005), presentaron un conjunto de 100 estímulos auditivos mediante grabaciones en cinta de audio, constituidos por frases declarativas de contenido semántico neutro; asignaron aleatoriamente las frases a una de las categorías: alegría, tristeza, enojo, miedo o asco y fueron grabadas por dos actores profesionales, un hombre y una mujer. El uso de estímulos extraídos a partir de la actuación ha sido ampliamente utilizado en otros estudios (Scherer; 2003) ya sea mediante el uso de actores profesionales (Por ejemplo, Dromey et al., 2005; Scherer, Banse, & Wallbott, 2001) o de personas sin experiencia o con experiencia no profesional (Ej. Erickson, 2006; Pell, Paulmann, Dara, Alasseri, & Kotz; 2009; Pell & Skorup; 2008). La lógica detrás de este tipo de metodología será discutida posteriormente

En una prueba previa, encontraron que las emociones eran reconocidas correctamente en aproximadamente 90% de las ocasiones, excepto para el asco, en el que el reconocimiento promedio fue de 77%⁵.

Compararon la activación cerebral que se dio en la tarea de reconocimiento de emociones (mediante decisión forzada entre las 5 categorías, expresada verbalmente) contra la observada en una tarea fonética, en la que se debía verbalizar la vocal que siguiera al primer fonema /a/ (lo que también resultaba en una decisión forzada de 5 categorías: /a/ /e/ /i/ /o/ y /u/, expresada verbalmente). El uso de esta metodología obedece a que en ambas condiciones (decisión sobre prosodia y decisión sobre vocales) se emplean los mismos estímulos, y el proceso de respuesta es similar (elección forzada, 5 opciones de respuesta). De hecho, reportan que los tiempos de reacción para ambas tareas no difieren. Este hallazgo es interpretado como evidencia de que las tareas tuvieron un nivel de dificultad comparable.

⁵ Esta observación de una mayor dificultad de reconocer el asco también ha sido reportada por otros estudios (Bryant & Barret, 2008; Luengo, Navas, Hernández, & Sánchez, 2005; Pell, Monetta, Paulmann, & Kotz, 2009). Es importante pues resalta que el reconocimiento de emociones básicas no es equivalente entre distintas de ellas. Esto se retomará en las secciones de estudios transculturales.

Encuentran activación relacionada con la percepción de prosodia emocional en corteza frontal derecha (región frontal dorsolateral, también relacionada a discriminación de tonos), en corteza temporal derecha (surco temporal superior, análogo al área de Wernicke en el hemisferio izquierdo); corteza orbitobasal derecha (relacionada con percepción y juicios afectivos en rostros, sabores y entonación emocional).

No encuentran activaciones específicas dependiendo de la emoción, aunque era esperada por hallazgos clínicos previos. Consideran que la ausencia de resultados en este sentido puede deberse al tamaño de la muestra empleada y/o a características específicas de su tarea.

Buchanan y colaboradores (2000) reportan resultados similares. Emplearon estímulos que diferían en la primera vocal (power, bower, tower, dower) y pronunciados cada uno reflejando cuatro emociones distintas: neutra, enojo, alegría y tristeza. Cada participante fue sometido a cuatro condiciones experimentales, en cada una de ellas debía identificar una clase de estímulo particular: en dos de ellas una palabra específica (todas las ocurrencias de power o bower), y en otras 2 una emoción particular, independientemente de la palabra (todas las ocurrencias de entonación triste o alegre).

Relacionado al procesamiento de la prosodia emocional, encuentran una mayor activación en la corteza temporal anterior derecha, así como en el giro frontal inferior (pars opercularis) derecho, el giro cingulado izquierdo (relacionado con la atención y el procesamiento emocional) y el lóbulo parietal inferior derecho.

Resaltan el mayor involucramiento de la corteza temporal derecha al centrarse en la tarea de detección de emociones contra la tarea de detección fonémica, a pesar de tratarse de los mismos estímulos. Esto podría relacionarse a las características temporales de los fonemas contra la prosodia (más breves y más largos, respectivamente). El punto será retomado más tarde en este capítulo.

Consideran como una explicación alternativa el hecho de que el proceso de identificación de prosodia emocional sea una instancia de detección de tonos y

duraciones como procesos básicos y no un proceso independiente. Cabe resaltar que esta interpretación de los resultados concuerda con la conceptualización de la prosodia como constituida principalmente por las variaciones en estas dimensiones. Sin embargo, a partir de dicha interpretación de sus resultados, se esperaría que el procesamiento fuese idéntico al de la prosodia lingüística. En otra sección del presente trabajo se contrastarán los hallazgos relacionados a ambos tipos de prosodias.

En general, se nota convergencia en algunos de los resultados dados por estudios de imagenología, en especial en la participación del lóbulo temporal y frontal derechos. Esto es congruente con lo reportado en diferentes trabajos, en los cuales se han empleado diferentes metodologías (Pell, 2002; Ross & Monnot, 2008; ver también Buchanan et al, 2000; Szymanowski et al. 2007; Wildgruber et al, 2005). Las principales discrepancias vienen del contraste entre los datos provenientes de estudios con metodologías de imagenología al ser comparados con estudios de lesiones (Buchanan et al, 2000; Ross & Monnot, 2008). Esta discusión será retomada más adelante.

La participación del área frontal derecha ha sido relacionada con el procesamiento de emociones (Buchanan et al., 2000) – en especial el área orbitofrontal (Wildgruber et al., 2005) – y memoria de trabajo para tonos (en especial el área dorsolateral; Wildgruber et al, 2005). En cuanto a la participación de las áreas temporales, se observa una mayor activación de la corteza derecha ante las tareas relacionadas con la prosodia (aunque se presenta activación del área tanto en tareas fonéticas y semánticas como prosódicas) Se propone que es por su asociación con el procesamiento auditivo general y de manera más marcada en el procesamiento del tono (Buchanan et al., 2000; Wildgruber et al., 2005). En cuanto a diferencias de género, Schirmer, Kotz y Friederici (2002) exploran si el curso temporal del procesamiento de la prosodia difiere entre hombres y mujeres mediante un procedimiento de *priming* y potenciales evocados. Reportan que en mujeres se encuentra una respuesta dependiente de la prosodia en una etapa

más temprana de procesamiento, mientras que en hombres se encuentra un patrón de respuesta similar más tardío.

En general, se observa una tendencia en el campo al uso de estímulos emocionales específicos (como miedo, enojo, alegría), y a la evaluación de las respuestas cerebrales asociadas. Esto es entendible considerando que este cuerpo de investigación ha sido influido por modelos que proponen un conjunto de emociones básicas y universales (Pell, 2002), aunque esta influencia no se restringe a la perspectiva neurobiológica (ej. Bryant & Barret, 2008; Juslin & Laukka, 2003; Luengo, Navas, Hernández, & Sánchez, 2005; Pell, Monetta, Paulmann, & Kotz, 2009; Pell, Paulmann, Dara, Allasseri, & Kotz, 2009; Russ, Gur, & Bilker, 2008). La manipulación acústica de características de los estímulos no es una metodología que se emplee frecuentemente, aunque los autores, en general, también asumen que son principalmente las características de ritmo y de tono las principales responsables de los efectos observados.

3.2.2 La perspectiva acústica y aproximaciones computacionales

En contraste con el estudio de lo que ocurre en las *personas* en relación a la prosodia (ilustrado claramente en la sección anterior), algunos estudios se interesan en estudiar las características de los *estímulos*, es decir, identificar *qué*, en la prosodia, determina que se perciban emociones en los estímulos, o qué características tiene la voz cuando ha sido producida mientras se experimentaba cierta emoción.

Entre las variables que se han reportado como importantes para la transmisión de información emocional se encuentran la frecuencia fundamental y sus características temporales, así como la velocidad de locución (Luengo, Navas, Hernández, & Sánchez, 2005). Otros autores incluyen también características como la cualidad de la voz (Grandjean, Bänziger, & Scherer; 2006; Erickson, 2006).

El concepto de cualidad de la voz es difícil de analizar (aunque en la literatura sobre acento extranjero que se revisará posteriormente, se ha avanzado en este

aspecto y se han propuesto algunos métodos y conceptos útiles para acercarse a ella, como las configuraciones fonéticas); se refiere a las características del sonido mediante las cuales un escucha puede diferenciar dos sonidos del mismo volumen y tono (Erickson, 2006). La cualidad de la voz puede ser señalada en la onda sonora mediante cambios en la distribución de energía en diferentes bandas de frecuencia producidos por diferentes modos de fonación (Grandjean, Bänzinger, & Scherer; 2006; Erickson, 2006). También puede señalar información como el estrato social del hablante y su origen geográfico y características relacionadas a su físico y salud (Erickson, 2006).

El enojo se ha asociado con un ritmo incrementado en el habla, tensión en la articulación, alta intensidad, un tono promedio alto, alta variabilidad y cambios abruptos de tono; la tristeza con un ritmo lento de habla, pronunciación “arrastrada”, un tono promedio bajo con poca variabilidad, entre otras características (Scott et al., 2006). Si bien algunas de estas características son prosódicas (suprasegmentales), se observa que algunas también pueden reflejarse a nivel segmental (como la articulación “arrastrada”). En la presente sección se describirán estudios que exploran las características prosódicas que transmiten información emocional, desde una perspectiva general, y modelos computacionales dedicados al reconocimiento de emociones, así como las variables que mejores resultados han arrojado.

Desde una perspectiva en la que las emociones son universales, se podría esperar que estos correlatos acústicos con la percepción de emociones se mantuvieran independientemente del idioma en el que se expresaran (Bänzinger & Scherer, 2005; Pell, Paulmann, Dara, Allasseri, & Sonja, 2009), sin embargo, diversos autores suponen que la prosodia emocional puede tener también influencias o restricciones dependientes del lenguaje específico y de la cultura (Pell, 2006; Scherer & Bänzinger, 2004).

Se han empleado métodos lingüísticos para investigar la prosodia emocional. En éstos, se suele estudiar el tono, duración y volumen del habla. En especial, se ha hecho énfasis en la descripción de los *contornos percibidos de tono* (A grandes

rasgos, patrones de tono a nivel de frases o palabras); para la cual se han desarrollado diversos métodos de transcripción y modelos descriptivos. Estos modelos lingüísticos generalmente buscan identificar las funciones de la prosodia (por ejemplo, ver Caffi & Janney, 1994, al respecto de pragmática de la comunicación emocional). A pesar de que desde la perspectiva lingüística no se suele investigar específicamente la prosodia emocional, es desde esta tradición donde provienen las proposiciones más importantes sobre la existencia de patrones específicos de entonación para emociones específicas. Sin embargo, ha habido muchas críticas a esta tradición, por ejemplo, por el uso de estímulos de prueba con contenido semántico que podría sesgar las interpretaciones y ser una variable extraña. Muchos de los estudios empíricos sobre el tema se han realizado de manera independiente a los enfoques lingüísticos. (Bäzinger & Scherer, 2005).

Ladd y colaboradores (1985) consideran que hay variables continuas relacionadas a la emoción, por ejemplo, la frecuencia fundamental (f_0), que correlacionaría con el nivel de *arousal* en un momento dado, y otras categóricas. Estas últimas estarían lingüísticamente delimitadas (es decir, tendrían una palabra asociada) y probablemente también de manera perceptual. Asimismo, enfatizan que variables como la frecuencia fundamental (al reflejar *arousal*) reflejan el estado fisiológico del hablante, mientras que otras variables transmiten información más cognitiva y actitudinal (por ejemplo rechazo o una actitud amistosa).

Estos autores realizaron un experimento en el que manipularon la frecuencia fundamental de diversas expresiones y los contornos de entonación, así como las maneras en las que se grabaron las frases: “normal, relajada y amistosa” o “molesta, irritada, enojada” (a esta última manipulación se le llamó “cualidad de la voz”). Presentaron los diversos estímulos para valoraciones perceptuales mediante diferenciales semánticos y analizando qué características acústicas se relacionan con los resultados de las valoraciones. Reportan que un rango amplio de frecuencia fundamental y una cualidad de voz *áspera* en sus expresiones predicen valoraciones de *arousal*, molestia e involucramiento. Consideran que el rango de frecuencia fundamental se relaciona principalmente con el *arousal*

percibido (una dimensión de las emociones referida al estado de alerta, en un continuo que va desde excitado y alerta hasta somnoliento y tranquilo; Caballero-Meneses & Menez, 2010) mientras que la cualidad de la voz transmite el componente de *valencia* (se refiere a una dimensión de las emociones referida a lo agradable o desagradable de las mismas). Consistente con este punto de vista, resaltan que la sorpresa, relativamente neutra en valencia y alta en arousal, es transmitida principalmente mediante la frecuencia fundamental. También encuentran efecto de los contornos tanto en las valoraciones de arousal como en valoraciones de énfasis, involucramiento y cooperación. Específicamente, el uso de contorno con tendencia a subir de tono fue asociado con una percepción más alta de empatía, énfasis y contradicción y con una percepción más baja de cooperación. Reportan resultados negativos para las interacciones entre estas variables (rango de frecuencia fundamental, cualidad de voz y contorno), lo que sugiere contribuciones independientes a la percepción emocional.

Con respecto a los contornos de entonación, Ladd y colaboradores (1985), reportan efectos en variables como empatía, énfasis, involucramiento y cooperación; las cuales no parecen ser dimensiones comúnmente estudiadas al investigar emociones. Este hallazgo de asociación de contornos de entonación con el señalamiento de diferentes actitudes ha sido reportado por otros autores; así como una interacción entre los contornos de entonación y las clases de enunciado (por ejemplo preguntas simples y preguntas de sí/no; Bänziger & Scherer, 2005). Juslin y Laukka (2003), en una revisión extensa de la literatura sobre las características acústicas en la prosodia y las emociones, encuentran que los contornos con tendencia a subir han sido repetidamente relacionados con la expresión de enojo, miedo y alegría; mientras que contornos con tendencia a bajar con tristeza y ternura.

Esta descripción de los contornos de entonación como tendencia a subir o bajar de tono, a pesar de ser relativamente simple (o incluso sobresimplificada y rudimentaria) y constituir un descriptor global, en contraste con una descripción

más específica como las que se emplean en la tradición lingüística⁶, ha sido usada por varios autores (Bänziger & Scherer, 2005). Lo que resalta la relativa independencia de las traducciones lingüística y psicológica de la investigación en prosodia.

Benus, Gravano y Hirschberg (2007) emplean un enfoque distinto. Hacen un análisis de las características prosódicas que diferencian diversos matices de la palabra *whatever*. Este término puede tener diversos usos, desde su empleo como muletilla, un término que denota indiferencia y libertad de elección (como *cualquier*), entre otros. A la vez, puede transmitir diferentes matices, desde la neutralidad hasta diversos grados de actitud negativa. Los autores consideran estas características del término como una gran ventaja al estudiar la prosodia emocional, ya que aunque la asociación de características acústicas ha sido asociada con el arousal, las variables que influyen la valencia no son tan bien entendidas. De esta manera, modificar la prosodia en este término permite hacer comparaciones sobre variables acústicas que afectan a la valencia en un material con contenido semántico y estructura segmental constante, aunque se restrinja al ámbito de la valencia negativa.

Reportan una serie de dos experimentos. En el primero, pidieron a 12 personas que leyeran situaciones en las que se empleaba el término, imaginándose en dichas circunstancias; las diferentes situaciones diferían en el grado de negatividad que implicaba el término (a partir de 3 evaluadores independientes). Analizaron una serie de parámetros más extensa que en otros estudios; entre ellos, la duración del cierre en /t/; máximo, mínimo, y rango de tono para cada sílaba y para la palabra completa, entre otros. Al analizar independientemente las ocurrencias de *whatever*, encuentran que el número de acentos tonales (se refiere como acento tonal a los movimientos de tono que se dan a la par de incrementos en *stress* para las lenguas que presentan esta característica y a los movimientos tonales en cualquier sílaba en aquellas lenguas que no la presentan, Grice y Baumann, 2007; esto puede ser conseguido por diferentes alturas de tono, el

⁶ Por ejemplo, el sistema ToBI (The Ohio State University Department of Linguistics, 1999).

grado de prominencia con respecto a otros acentos tonales puede variar; siendo el acento más importante en una frase prosódica conocido como acento nuclear, Hirschberg, 2002a), y en particular la presencia de éste en la primera sílaba (variaciones de tono que pueden ser ascendentes o descendentes al interior de las sílabas), fue el principal predictor de negatividad de las frases (las frases que se leyeron en los contextos donde *whatever* era más negativa tendieron a tener más acentos tonales). Dichos acentos tonales eran subidas rápidas de tono. También la duración de la frase y de cada sílaba fue mayor dependiendo del grado de negatividad. El rango de tono también fue mayor para las ocurrencias más negativas; tanto en cada sílaba como en general; pero el efecto es más marcado para la primera sílaba.

En el segundo experimento, presentaron un conjunto de estímulos de ocurrencias naturales de *whatever*, tomadas a partir de fuentes como conversaciones telefónicas espontáneas, y por tanto no realizadas con fines experimentales, a tres grupos de participantes; los estímulos habían sido seleccionados tratando de abarcar un significado neutral y 4 grados de negatividad. En un grupo, les pidieron a los participantes que se imaginaran en una situación a la que se respondía con *whatever* a una pregunta neutra (¿Quieres comida china para cenar?). En otro, se presentaron los estímulos junto con su contexto original. En un tercer grupo se presentaron transcripciones de los estímulos, incluyendo el contexto en el que se daban, para explorar la influencia del contexto en la determinación del matiz de la palabra, aún en la ausencia de información auditiva.

Analizan si características acústicas correlacionan con las valoraciones de negatividad. En el primer grupo, encuentran que, análogo al primer experimento, los acentos tonales, en especial al presentarse como un patrón de subida rápida de tono, se relacionan a la percepción de negatividad. Sugieren también que características como un contorno muy plano podría estar asociado a valoraciones más negativas; resaltan que esta asociación entre contornos planos y emociones negativas había sido hipotetizada en otros trabajos (ver Liscombe, Venditti & Hirschberg, citado en Benus et al., 2007).

Encuentran que al presentar los estímulos en contexto (sea de manera sonora o en transcripción), las valoraciones sobre la negatividad son más congruentes. Además, tienden a ser consideradas más negativas al presentarse en contexto, los autores proponen que puede deberse a que la situación que emplearon en la condición en la que suprimieron el contexto original tendía a ser neutra. Existe una correlación entre las valoraciones que incluían contexto (transcripciones y ocurrencias con contexto; $r=0.8$). También se observan correlaciones significativas de la condición libre de contexto con las condiciones de contexto, presentado de manera auditiva (0.595) y en transcripción (0.67), lo que interpretan como evidencia de que las características prosódicas que analizan transmiten en buena medida información sobre la negatividad de la situación en la que se presenta originalmente la ocurrencia; sin embargo, también resalta la importancia del contexto para la aclaración de los matices de palabras como la que se explora. Este estudio, a pesar de que se interesa en las emociones; o, de manera más específica en la valencia, y de encontrar algunas características prosódicas asociadas a la negatividad, realiza un análisis más pragmático de los diferentes usos de una palabra; con todo, ilustra una metodología distinta en el análisis de la prosodia.

Bänziger y Scherer (2005) exploran si existen características prosódicas asociadas a emociones específicas y en general si características acústicas están relacionadas a variables continuas como el arousal; para lo que emplean frases sin sentido, compuestas para combinar características fonéticas de diversos lenguajes europeos y dar la impresión de tratarse de un lenguaje desconocido. Encuentran que el arousal afecta principalmente a la frecuencia fundamental, por lo que medidas agregadas como la frecuencia promedio o el rango de frecuencia, ayudan a distinguir las variaciones más importantes entre categorías emocionales. Reportan también asociaciones entre tendencias de subida de tono asociadas a desesperación y júbilo; y con tendencia de bajada asociadas a tristeza y alegría. Así como una caída de tono al final de la frase más pronunciado para emociones como alegría y júbilo.

Trabajando de manera cercana, ha habido aplicaciones computacionales que se han aproximado a la prosodia emocional. Entre los intereses que se tienen en esta rama se encuentra el desarrollar aplicaciones de reconocimiento automático de emociones, con miras hacia la búsqueda automática de contenidos multimedia basados en el contenido emocional (Seppänen, Väyrynen, & Toivanen, 2003), así como para aplicaciones de atención al cliente (Martínez Sánchez, Montero Martínez, & de la Cerra, 2002), sistemas de detección de mentiras (Saíz Ruiz & Talavera Martín, 1997) y sistemas de comunicación natural con computadoras (Luengo, Navas, Hernáez, & Sánchez, 2005); la síntesis artificial de voz, en especial voz que sea más cercana a la voz natural y con la capacidad de transmitir información emocional (Makarova & Petrushin, 2002; Martínez Sánchez et al., 2002) también ha recibido interés. Este tipo de actividades han recibido el nombre de computación afectiva (Makarova & Petrushin, 2002).

Seppänen, Väyrynen y Toivanen (2003) comparan el desempeño en el reconocimiento de prosodia emocional de un sistema automático y de humanos en un experimento perceptual. El desempeño humano y del sistema es similar. Las frases a las que se aplicó el sistema y las que fueron presentadas en los experimentos perceptuales fueron tomadas de un corpus de frases emocionales en finlandés desarrollado previamente. En un principio se derivaron 43 medidas de parámetros vocales; las cuales incluían medidas de frecuencia fundamental (promedio, rango, tendencia a subir o bajar, entre otras), de intensidad (promedio, rango, máximo, mínimo, entre otras), y otras medidas como características espectrales y proporción de segmentos sonoros contra sordos. Posteriormente, seleccionaron las medidas que ofrecían la mayor precisión con hasta 15 y menos variables. El mejor desempeño del sistema (en una aplicación desarrollada para reconocer prosodia de manera independiente a los datos con los que se entrenó al mismo) fue obtenido con los parámetros Rango de Intensidad, Varianza de intensidad, máximo de subida de f_0 en un segmento sonoro, proporción de segmentos de silencio, rango entre el 5 y 95% de frecuencia fundamental, entre otras relacionadas con la intensidad y la frecuencia fundamental (Toivanen, Väyrynen, & Seppänen, 2004). Cabe resaltar que entre las variables que se

incluyen en el sistema que desarrollaron los autores destacan las medidas de intensidad, f_0 , la variabilidad de ambas variables y la tendencia a subir en el caso de f_0 ; las cuales coinciden con las reportadas en los experimentos acústicos mencionados antes. El método para elegir las no fue basado en la teoría sino en el desempeño a partir de un conjunto más grande de variables. El hecho de que estas variables resulten las mejores para reconocer prosodia emocional de manera automática resalta la convergencia de ambas perspectivas

Luengo, Navas, Hernáez y Sánchez (2005), reportan un trabajo de reconocimiento automático de emociones básicas utilizando un corpus de habla emocional. En un principio, extrajeron 84 características prosódicas para cada frase, y después eliminaron variables hasta encontrar el mejor desempeño con el menor número de variables. Utilizando solamente 6 parámetros logran un desempeño de 92.32% de reconocimiento correcto, similar al desempeño humano determinado en una validación perceptual de su base de datos. Las características que incluyen en su sistema son: “valor medio del *pitch* en escala lineal”; “valor medio de la *energía* en escala logarítmica”, “varianza del *pitch* en escala lineal”, “sesgo del *pitch* en escala logarítmica”, “rango del *pitch* en escala logarítmica”; “rango de la *energía* en escala logarítmica” (Pitch se refiere al tono y energía a volumen, pero se decidió citar los nombres originales de las variables de los autores). Se puede resaltar que estas variables relacionadas al tono y al volumen, coinciden con lo reportado en los experimentos sobre análisis acústicos reportados anteriormente. Aunque los autores no emplean variables de duración (que podrían reflejar características de ritmo y velocidad de locución) discuten la posibilidad de mejorar los resultados al incluir este tipo de variables.

También empleando un enfoque de clasificación de emociones basado en dimensiones (valencia y arousal), Devillers y Vasilescu (2004) encuentran que los parámetros prosódicos que diferencian emociones de valencia positiva de aquellas con valencia negativa son el rango de f_0 ; el máximo incremento de f_0 entre dos segmentos sonoros y el máximo de f_0 . Más tarde, en un trabajo donde emplean grabaciones de interacciones de un servicio de emergencias, donde las

emociones presumiblemente son genuinas (cabe recordar que muchos de los trabajos revisados emplean estímulos de voces actuadas); encuentran que variables relacionadas al tono (medidas derivadas y varias medidas de dispersión); medidas de volumen (varias medidas de dispersión); ritmo del habla entre otras medidas son, entre un conjunto amplio de variables, los mejores clasificadores para reconocer automáticamente emociones (Devillers & Vidrascu, 2006).

Finalmente, Pérez Espinosa y Reyes García (2010), reporta los resultados de un sistema de reconocimiento de emociones a partir de los parámetros de arousal, valencia y dominancia. A partir del primer cálculo de dichos parámetros, realizan un segundo proceso para categorizar los resultados obtenidos en una emoción discreta (enojo, felicidad, aburrimiento, miedo, tristeza, disgusto y neutral). Encuentra como variables importantes para calcular la valencia, parámetros como contorno de entonación, cualidad de voz y algunas características espectrales (como la frecuencia); para el arousal y la dominancia, encuentran características espectrales, volumen, contorno de entonación y cualidad de la voz, sin embargo, afirman que para el cálculo de la dominancia se requiere información sobre la distribución de energía en diferentes rangos de frecuencia de la voz; es decir, información sobre la cualidad de la voz.

Este trabajo ilustra cómo a partir del cálculo de características continuas (arousal, valencia y dominancia) de las emociones, también se puede llegar a un reconocimiento categórico.

Aunque esta literatura no se interesa por explicar la prosodia, ha hecho algunas propuestas. Por ejemplo, al discutir la dificultad de crear sistemas de reconocimiento aplicables a una variedad de hablantes distintos, sostienen que la forma en que las características vocales señalan el afecto podría ser, hasta cierto punto, dependientes de cada hablante (Seppänen, Väyrynen & Toivanen, 2003). Aunque esto parece poco probable, considerando la capacidad de detectar prosodia por parte de los humanos en voces que no se habían oído antes como los estímulos experimentales utilizados en los experimentos revisados en esta

sección. Si la propuesta de Seppänen y colaboradores (2003) fuese cierta, se requeriría de tener experiencia con cada persona antes de poder inferir sus estados emocionales.

3.2.3 Investigación intercultural.

La principal utilidad de las metodologías interculturales es que permiten evaluar el efecto de la cultura en las variables que se estudian. Al trabajar con esta perspectiva, se asume que las diferencias observadas en un fenómeno dependen del entorno cultural; del aprendizaje y las condiciones específicas de los mismos. Asimismo, la falta de diferencias en un fenómeno a pesar de diferencias culturales suele asociarse a una base biológica; en especial si se trabaja con culturas muy distantes.

Partiendo de esta lógica, si la prosodia emocional tiene bases biológicas y características específicas, se puede esperar que en estudios interculturales los participantes puedan reconocer correctamente la prosodia emocional de hablantes de una cultura lejana, y, efectivamente, en la literatura se encuentra que personas de culturas diferentes tienen la capacidad de reconocer prosodia emocional producida por hablantes de culturas distintas a niveles superiores al azar (Dromey, et al., 2005).

Esto ha sido interpretado como evidencia de que el proceso de reconocimiento de expresiones emocionales es compartido por los humanos y es una consecuencia neuropsicológica de experimentar emociones básicas (Pell, Monetta, Paulmann, & Kotz, 2009). Asimismo, se ha considerado que al estar experimentando una emoción, una consecuencia neurofisiológica es que el habla sea afectada (Pell, Paulmann, Dara, Alasseri, & Kotz, 2009), lo que completa el cuadro de la influencia de las emociones tanto en la codificación como en la decodificación prosódica.

Dada la relevancia biológica y la importancia para la conducta adaptativa de estas expresiones emocionales para los demás miembros de la especie, se ha supuesto

que comparten características invariables, por lo que pueden ser reconocidas de manera independiente del aprendizaje y la cultura (Pell, Paulmann, Dara, Alasseri, & Kotz, 2009). Incluso se ha propuesto que si las expresiones emocionales no tuvieran características universales, su ambigüedad podría tener un alto costo evolutivo, por ello, se tendería a seleccionar mecanismos universales para la producción y percepción de expresiones emocionales (Bryant & Barret, 2008).

Sin embargo, la literatura sobre investigación transcultural muestra un panorama más complejo.

A pesar de la existencia de características universales de las emociones, también es ampliamente reconocido que existen variaciones culturales en la expresión de las mismas; en especial en cuanto a las situaciones en las que es apropiado expresarlas y en la forma en la que puede hacerse; esto ha sido denominado *reglas de exposición* (Bryant & Barret, 2008; Pell, Paulmann, Dara, Alasseri, & Kotz, 2009).

De hecho, inclusive autores que trabajan desde la perspectiva más biológica, en la fisiología de la prosodia emocional, han enfatizado la influencia de la cultura en la misma; critican como una visión “demasiado restringida” el considerar a las emociones como básicas y trabajar a partir de esa premisa solamente, ya que puede obscurecer el importante papel del hemisferio derecho en la conducta social y los juicios sociales y pragmáticos sobre el lenguaje (Pell, 2002). Pell (2002) resalta la necesidad de considerar que incluso la prosodia asociada a las llamadas emociones básicas suele presentarse en contextos de interacción interpersonal y que la forma precisa que toma la expresión prosódica de dichas emociones está considerablemente dictada de forma social. Por ello, poder entender el significado de la prosodia emocional requiere de la adquisición de conocimiento pragmático compartido entre el hablante y el oyente.

En la prosodia – al ser una característica del lenguaje y considerando lo que hemos venido describiendo acerca de su capacidad de expresar emociones – se puede esperar encontrar una interacción de los factores biológicos y culturales que

afectan a las emociones, y entre las contribuciones culturales, tanto de las formas de expresarlas como de las características del lenguaje.

Características asociadas a eventos “naturales” (el experimentar emociones) y características aprendidas propias del hablar un lenguaje (patrón prosódico general, sílabas tónicas, distinciones léxicas de la prosodia, etc.) estarían representadas a la vez en las propiedades acústicas del habla en condiciones emocionales (Pell, Monetta, Paulmann, & Kotz, 2009). Se podría considerar que las expresiones vocales (entendidas como sonidos producidos oralmente, no necesariamente referidos al lenguaje) de la emoción estén limitadas por la necesidad de mantener las características segmentales y suprasegmentales de un lenguaje, así como por el aprendizaje relacionado a la comunicación lingüística (Pell, Monetta, Paulmann, & Kotz, 2009; Sauter et al., 2010). Se podría considerar también que las expresiones vocales más “puras” de la emoción son aquellas en las que no “interfiere” el lenguaje (como gritos y risas; Sauter et al., 2010).

A partir de esto se pueden avanzar dos hipótesis:

1. – La prosodia emocional debe hacer uso de las características no fundamentales para la comprensión de un mensaje, es decir, la señal acústica se puede modificar mientras no afecte a las diferencias contrastivas en un lenguaje (tanto segmentales como suprasegmentales). En adelante nos referiremos a ésta como *hipótesis de grados de libertad*. Esta hipótesis será desarrollada ampliamente en el curso del presente trabajo y se probará como la contribución más importante al final del mismo.

2. – Si bien el estudio de expresiones vocales emocionales como gritos y risas, por definición no puede formar parte del estudio de la prosodia emocional, podría brindarnos algunas pistas sobre el tipo de características acústicas que se encuentran en una situación emocional sin lenguaje. Quizá parte de estas características se podrían esperar como parte de la señal acústica de prosodia emocional relacionada con una emoción similar; siempre y cuando no interfiera con las características contrastivas.

En concreto, la literatura sobre investigación transcultural en prosodia emocional resalta que aunque puede ser reconocida a niveles superiores al azar (lo que apoya la hipótesis de la universalidad), también se suele encontrar un mejor desempeño al interior de la propia cultura (lo que apoya la influencia cultural), es decir, un mejor desempeño al evaluar estímulos provenientes del mismo idioma y cuyas fuentes presumiblemente comparten otras variables culturales (Dromey et al., 2005; Pell, Monetta, Paulmann, & Kotz, 2009; Pell, Paulmann, Dara, Alasseri, & Kotz, 2009; Pell & Skorup, 2008).

Sin embargo, también se ha reportado en algunos estudios un mejor desempeño por parte de un grupo ajeno a la cultura de la que se obtuvieron los estímulos muestra (Erickson, 2006; McCluskey & Albas, citado en Dromey et al., 2005), para lo que se ha propuesto un efecto de neutralidad de lenguaje (que sea más difícil inferir significado emocional a partir de la prosodia en una palabra neutra del propio lenguaje, lo que llevaría a alguien de una cultura ajena y que no nota la neutralidad emocional semántica a desempeñarse mejor), e influencia de la importancia cultural de poner atención en esas características (lo que llevaría a un miembro de una cultura que enfatiza la atención a la prosodia emocional a desempeñarse mejor en un idioma extranjero que un nativo de ese idioma cuya cultura no enfatiza esta capacidad).

Otro resultado relacionado a la ventaja al interior de los grupos en el reconocimiento de prosodia emocional es que el efecto no es simétrico para distintas emociones. Pell, Monetta, Paulmann y Kotz (2009), en un estudio que involucró el reconocimiento de prosodia emocional en frases sin sentido producidas por hablantes de Inglés, Alemán, Árabe y Español por parte de hispanohablantes, encuentran que independientemente del idioma de origen de los estímulos, el desempeño fue mejor para enojo y tristeza y más baja para asco. El enojo y la tristeza, seguidos de miedo y expresiones neutrales también tuvieron buenos niveles de identificación en el trabajo de Pell, Paulmann, Dara, Alasseri, y Kotz (2009). Scherer y colaboradores (2001) también reportan buenos niveles de reconocimiento para enojo y bajos niveles para alegría. En un trabajo que

involucró el reconocimiento de expresiones vocales no relacionadas al lenguaje (risas, gemidos, entre otras), Sauter y su equipo (2010) también encuentran un menor desempeño en el caso de la alegría.

Por otra parte, también se ha propuesto un efecto de *gradiente de proximidad* de los lenguajes: entre menos emparentados (lingüísticamente hablando) sean dos lenguajes, menor sería el desempeño al reconocer prosodia emocional entre ellos (Scherer et al., 2001). En un estudio, Scherer y colaboradores (2001), encuentran que el orden de desempeño entre hablantes de distintas lenguas era similar al orden de similaridad lingüística con el idioma de los hablantes que codificaron originalmente los estímulos: los actores hablaban alemán y el mejor desempeño en el reconocimiento de la prosodia emocional fue el de hablantes de alemán, seguido por hablantes de lenguas germánicas (Inglés y Holandés), posteriormente por hablantes de lenguas romances (Italiano, Francés y Español; las lenguas romances y germánicas pertenecen a la familia de lenguajes indo-europeos), y finalmente por hablantes de malayo (fuera de la familia de lenguajes indo-europeos)⁷.

Sin embargo la hipótesis de la influencia de la proximidad lingüística no parece haber recibido mucho apoyo, como muestran los estudios de Pell, Paulmann, Dara, Alasseri y Kotz (2009) y Pell, Monetta, Paulmann y Kotz (2009). Con todo, en el trabajo de Pell, Monetta, Paulmann y Kotz (2009) encuentran que la condición con el lenguaje más distante fue *subjetivamente* la más difícil, sin que esto repercutiera en el desempeño.

Pell, Monetta, Paulmann y Kotz (2009) consideran que definir la proximidad de lenguajes a partir de características entonacionales o rítmicas en lugar de emplear el enfoque de familias de lenguajes más tradicional podría ayudar a dirigir la investigación en este campo.

⁷ Cabe mencionar que los estímulos eran frases carentes de significado, no frases en alemán, y que habían incluido sílabas de todos los lenguajes de la familia indo-europea del estudio.

Ciertamente la hipótesis de la cercanía lingüística requiere más investigación. Sin embargo, dado que los estudios requeridos para responder una pregunta como esta requieren la inclusión de idiomas que varíen en su distancia lingüística, es entendible que sean escasos, pues plantean dificultades logísticas importantes.

Hasta el momento, sólo se tiene conocimiento de dos artículos en los que se hace una comparación extensa de los desempeños en reconocimiento de prosodia emocional en varias culturas.

El primero de ellos fue realizado por Scherer, Banse y Wallbott en 2001. Dicho estudio fue conducido en nueve países, siete europeos, Estados Unidos e Indonesia. Se presentó a los participantes un conjunto de estímulos codificados por actores profesionales de habla alemana con la intención de expresar enojo, miedo, alegría, tristeza y expresión neutral. Los estímulos eran frases sin sentido compuestas de sílabas de lenguajes europeos. Entre sus resultados encuentran que las emociones son reconocidas correctamente a niveles superiores al azar; así como que el mejor desempeño fue por parte de los hablantes de alemán (ventaja dentro del grupo, ya descrita); encuentran un efecto de género, en el que las mujeres se desempeñaron mejor (este tipo de efectos ya había sido reportado en la sección de fisiología del presente capítulo por Schirmer et al., 2002, y se repite en varios de los artículos que se reportarán en esta sección); también reportan diferencias en el desempeño para distintas emociones, siendo mejor para enojo y relativamente bajo para alegría, y efectos del país (que fue la forma en la que operacionalizaron a la cultura). Se encuentra que el patrón de confusión para los diferentes grupos es similar, lo que sugiere un proceso común. Es en este artículo en el que se sugiere el efecto de *gradiente de proximidad* ya descrito.

El segundo fue realizado por Pell, Paulmann, Dara, Alasseri y Kotz (2009). En este estudio, los autores prepararon estímulos que consistían en frases sin sentido con tonos de enojo, asco, miedo, tristeza, alegría y sorpresa agradable. Fueron codificados por hablantes nativos de inglés, alemán, hindi y árabe con experiencia no profesional en actuación o habla pública. Las frases sin sentido fueron construidas por un hablante nativo de cada lengua de modo que no violaran las

reglas fonéticas ni las sílabas posibles de la lengua, pero carecieran de significado. Presentaron estímulos previamente validados para mostrar la emoción que se pretendía expresar a grupos de hablantes nativos de todas las lenguas incluidas en el estudio. Exploraron la capacidad de reconocer perceptualmente los estímulos producidos por hablantes del mismo idioma o por hablantes de otros idiomas y caracterizaron acústicamente los estímulos. Encuentran que la prosodia emocional puede ser reconocida correctamente a niveles superiores al azar al interior de cada lenguaje. El desempeño en la tarea varió dependiendo de la emoción que fue codificada originalmente y del idioma de los escuchas: En hablantes de inglés, hindi y árabe, hubo mejor desempeño al identificar tristeza, enojo y miedo; además, en hablantes de árabe también hubo un buen desempeño al identificar expresiones neutrales. En hablantes de alemán, el mejor desempeño fue el de expresiones neutrales y de enojo, seguidas de asco, tristeza y miedo. Al tomar en cuenta a todos los participantes, el enojo y la tristeza fueron las emociones más fáciles de reconocer, seguidas de cerca por miedo y expresiones neutrales; la alegría y el asco fueron las emociones más difíciles de reconocer.

Los autores resaltan que la capacidad de reconocer miedo y enojo depende del grado de intensidad que representen. Por otra parte, sugieren que una expresión neutral, más que una carencia de intención, podría reflejar una estrategia adoptada por los hablantes para resaltar la intención proposicional de un mensaje (es decir, señalar fuertemente que se trata de una expresión neutral), como puede ser el caso, también, de alguien que desea parecer “objetivo”. De manera consecuente, las expresiones neutrales tuvieron un reconocimiento relativamente bueno (inclusive fue el mayor en el caso de escuchas de habla alemana). También llaman la atención hacia la influencia de diferencias individuales y de personalidad para comprender conducta emocional expresiva, así como al expresar emociones, sin embargo, no realizan un análisis específico al respecto.

Por otra parte, encuentran que es posible identificar automáticamente, mediante métodos computacionales, las emociones expresadas a partir de parámetros de promedio y variabilidad de f_0 , y ritmo del habla (sílabas por segundo); sin embargo

el desempeño es cualitativamente diferente al de los humanos: a partir de estas variables se obtuvo un buen resultado al reconocer sorpresa y asco, aunque los jueces humanos tuvieron dificultades al hacerlo, lo que sugiere que estas emociones tienen un patrón acústico particular en estas dimensiones, aunque no basta para que puedan ser reconocidas adecuadamente por participantes humanos; al mismo tiempo, las dos variables de f_0 y ritmo del habla no tuvieron buen desempeño al reconocer enojo, emoción que puede ser reconocida por humanos de manera consistente. Los autores consideran estos resultados como evidencia indirecta de la necesidad de incluir otras variables acústicas como la variación de volumen y la cualidad de la voz (diferencias en la distribución de energía del espectro). En este trabajo no se encontró evidencia que apoyara la hipótesis del *gradiente de proximidad*.

En otros trabajos que emplean metodologías similares, presentando estímulos de prosodia emocional producidos en un idioma y por miembros de una cultura y presentándolos a miembros de otra cultura y muchas veces también hablantes de una lengua distinta, se resaltan efectos de ventaja en el propio idioma para políglotas (Dromey et al., 2005), efectos de interacción entre la duración del estímulo y lenguaje propio-ajeno (Pell & Skorup, 2008), correlación entre edad y desempeño (Dromey et al., 2005; ligera pero significativa $r = -.249$; $p = .003$), efectos de género, donde las mujeres se desempeñan mejor (Erickson, 2006; Dromey, Silveira, & Sandor, 2005) y efectos de personalidad y diferencias individuales (Pell, Monetta, Paulmann, & Kotz, 2009).

Una de las posibles críticas a este cuerpo de investigación es que las culturas que estudian, si bien son diferentes, están en contacto continuamente, por ejemplo por la influencia de medios masivos de comunicación. Además, el uso de muestras universitarias para hacer los estudios reduce las diferencias culturales.

Al respecto, dos trabajos han analizado culturas considerablemente aisladas y con poca o nula convivencia con miembros de otras culturas.

Sauter y colaboradores (2010), estudian la capacidad de hablantes de inglés europeo y de hablantes de himba, miembros de una tribu aislada del norte de Namibia, de reconocer expresiones vocales de la emoción (risas, gritos, gemidos) producidas por miembros de ambas culturas. Encuentran que ambos grupos son capaces de reconocer las expresiones vocales de miembros del otro grupo. También encuentran un efecto de ventaja al interior del grupo. Resaltan que en el caso de las emociones positivas su reconocimiento intercultural en el grupo himba fue menor, es decir, los himbas tuvieron más problemas para reconocer los estímulos provenientes de hablantes de inglés. Lo interpretan en términos evolutivos, ya que si la función de las emociones positivas es aumentar la cohesión al interior del grupo, pero no con miembros ajenos a él, las emociones positivas podrían estar más sujetas a variaciones culturales (como características del lenguaje y estereotipos culturales; ver Pell, Monetta, Paulmann, & Kotz, 2009). Encuentran un menor desempeño en general de la población himba, esto podría reflejar simplemente falta de familiaridad de los mismos con los procedimientos psicológicos utilizados para medir el reconocimiento de los estímulos. Sin embargo este trabajo no estudia la prosodia como tal.

Bryant y Barret (2008) comparan la capacidad de reconocer prosodia emocional en hablantes de inglés de Estados Unidos y hablantes de shuar y español de Ecuador. La comunidad Shuar era relativamente aislada, hablaban español por un programa de educación bilingüe de su país. Presentaron frases codificadas originalmente por hablantes de inglés y encontraron que los hablantes de shuar pueden reconocerla a niveles superiores al azar. Más aún, su desempeño es similar al de hablantes de inglés que oyeron los mismos estímulos manipulados digitalmente para eliminar la información semántica. Resaltan que entre los hablantes de shuar respondían “miedo” al estar oyendo una frase que pretendía expresar alegría, pero el error contrario no se presentaba. Los autores consideran que esto podría deberse a las *reglas de exposición* (variaciones culturales en la expresión de las emociones, en especial en cuanto a las situaciones en las que es apropiado expresarlas y la forma en la que puede hacerse) que manejaban los participantes Shuar. Cabe señalar que las instrucciones fueron dadas en español

a los miembros de la cultura shuar, por lo que se puede inferir un buen dominio de dicho idioma por parte de esta comunidad; antes ya se habían encontrado efectos de hablar más de un idioma para este tipo de tareas (Dromey et al., 2005).

Los trabajos interculturales son relativamente escasos, aún más los que examinan culturas aisladas. Esto es entendible considerando los requisitos metodológicos y las limitaciones logísticas para realizar este tipo de estudios.

Definitivamente, el conjunto de estudios interculturales es fundamental para responder la pregunta eje de este trabajo de tesis: ¿cuáles son las influencias biológicas y culturales a la prosodia? Y, en los presentes, específicamente en la prosodia emocional.

Entre los hallazgos de esta literatura se puede resaltar que:

- Se observa una capacidad de reconocer emociones interculturalmente, lo que refleja la contribución biológica. Los resultados de este cuerpo de investigación convergen con las hipótesis evolutivas que guiaron a varios de los otros trabajos (particularmente a los estudios sobre neurobiología del presente capítulo).
- Se observan variaciones culturales, principalmente en las *reglas de exposición*, efectos de aprender idiomas y posibles efectos del énfasis que se da en algunas culturas a poner atención a la prosodia.
- Las características acústicas en las que se basa la transmisión de la prosodia convergen con lo que había sido descrito en la sección de enfoques acústicos y computacionales. Más aún, estas características se revelan en diversos idiomas estudiados (Pell, Paulmann, Dara, Alasseri, & Kotz, 2009).
- La capacidad de reconocer emociones tanto de la misma cultura como de otras varía dependiendo de la emoción, lo que sugiere una interpretación en términos evolutivos. Emociones negativas como el miedo, el enojo y la tristeza podrían ser comunicadas efectivamente dada su significancia o urgencia biológica (señalan agresión, pérdida y peligro físico), por que es

importante poder comunicarlas a distancia y claramente (Pell, Paulmann, Dara, Alasseri, & Kotz, 2009; Scherer, 2003), y esto favorecería su claridad comunicativa en el caso de la voz, al permitir comunicación a distancia. En el caso del asco, dado que podría prevenir a otros de comer comida podrida, sería más útil su reconocimiento facial, además de que implica acciones como regurgitar y bloquear olores desagradables (Scherer, 2003), lo que no facilita la comunicación vocal simultánea.

- A pesar de divergencias en resultados, las convergencias entre las diferentes tradiciones de investigación revisadas permiten tener un panorama consistente referido a la prosodia emocional.

En la siguiente sección, la última del capítulo, se retomarán algunos de los temas que se han dejado inconclusos a lo largo del mismo, considerando sus implicaciones teóricas y resaltando algunos puntos metodológicos de consideración para la interpretación de los resultados.

3.2.4 Resumen y discusión.

En esta sección presentaremos de manera sintetizada algunas consideraciones teóricas que surgen a partir de lo revisado en el capítulo así como consideraciones metodológicas necesarias para la interpretación de los resultados.

- Variables de personalidad y diferencias individuales.

A pesar de que se reconoce la importancia de variables de personalidad y diferencias individuales en la capacidad de codificar y decodificar prosodia emocional (Pell, Monetta, Paulmann, & Kotz, 2009; Pell, Paulmann, Dara, Alasseri, & Kotz, 2009), los estudios no suelen incluir estas variables como parte del análisis.

Comparaciones entre el tipo de variables a las que ponen atención personas con buena aptitud para decodificar y las que no, así como contrastes entre las características acústicas de quienes son aptos para codificarla contra los que no, podrían revelar algunas sutilezas en el campo así como informar a la investigación

aplicada (por ejemplo para personas que pueden requerir la habilidad de detectar pequeñas variaciones, como psicólogos, fonetistas, miembros de la policía etc.).

- Uso de variables subjetivas y porcentajes de reconocimiento.

Es fundamental el empleo de estudios con variables subjetivas. Los estudios perceptuales son tan importantes como aquellos que emplean análisis acústicos.

Más aún, además de las variables acústicas (*tono, ritmo, volumen*), complementar en estos estudios con variables subjetivas sería muy fructífero (*tono percibido, ritmo percibido, volumen percibido*). ¿Por qué? Las variables acústicas no equivalen necesariamente a sus equivalentes percibidos. La señal acústica es transformada al interior del sistema auditivo (Carlson, 2006b; Scherer, 2003). Por otra parte, variables perceptuales correlacionan más fuertemente con algunas características de la señal que con la señal en su conjunto: por ejemplo, el volumen percibido correlaciona más fuertemente con la distribución de energía en algunas bandas de frecuencia que con el volumen general (Scherer, 2003).

Esto puede estar relacionado a resultados como los de Pell, Paulmann, Dara, Alasseri y Kotz (2009), donde se muestran diferencias entre los rendimientos humanos y automatizados para reconocer prosodia emocional, a saber, los humanos detectaron fácilmente el enojo mientras que el sistema automático tuvo dificultades al hacerlo; el patrón opuesto de resultados fue observado al reconocer asco. Esto puede reflejar diferencias cualitativas en la forma de operar de humanos y sistemas automáticos, y considerando las transformaciones acústicas al interior del sistema auditivo humano que refiere Scherer (2003), también diferencias “*cuantitativas*”⁸ en la señal.

El reconocimiento de prosodia, tanto de humanos como de sistemas automatizados dista mucho de ser perfecta. Con todo, el rendimiento de ambos está alrededor del 60% (Scherer, 2003; Seppänen et al, 2003; ver los estudios de la sección acústico y computacional del presente trabajo, la variabilidad de estos

⁸ Aunque para aclarar una cuestión como esta sería necesario poder medir acústicamente una señal al interior del sistema auditivo.

es mayor, puede llegar a más de 90% y depende de la metodología usada para la implementación del sistema automático).

La forma específica en el que ambos sistemas obtienen su rendimiento puede diferir. De esta manera resalta la importancia del uso de diversas metodologías al abordar un tema como este.

- Uso de palabras sin sentido, palabras con contenido semántico emocionalmente neutro y palabras con contenido emocional.

Varios estudios en el área emplean estímulos constituidos a partir de frases o palabras sin sentido. Esto obedece a un esfuerzo metodológico por aislar los efectos de la prosodia en el lenguaje mediante la eliminación de la información semántica, es decir, mantener una neutralidad semántica con respecto a la información emocional. Otra forma de buscar la neutralidad semántica es el uso de palabras o frases no relacionadas a un significado emocional, también referidas como palabras o frases neutras (por ej. Erickson, 2006, emplea la palabra japonesa /banana/ con el mismo significado que en español). Un efecto muy similar se obtiene al usar estímulos provenientes de idiomas desconocidos para los participantes.

También se ha recomendado el uso de sílabas reduplicadas (ej. *ma mamama ma mama*), en especial para estudios acústicos como una forma de reducir las variaciones de f_0 asociadas a las vocales, ya que dada la posición del tracto vocal necesaria para producir las diferentes vocales existe una tendencia de variación de f_0 ; en términos generales, las vocales altas tienden a estar asociadas a una f_0 más elevada (Harwell Rodriguez, 1998).

Sin embargo, el uso de frases sin sentido no permite evaluar posibles efectos de interacción con el contenido semántico de las frases, para lo que algunos autores comparan los resultados obtenidos a partir de estímulos sin significado contra estímulos con significado y donde pueden manipular congruencia o incongruencia de la prosodia y el significado. Se puede esperar un reconocimiento más sencillo en condiciones congruentes contra condiciones incongruentes.

Un abordaje final que se propone en el presente trabajo y que aún no ha sido empleado es una aproximación a la prosodia lingüística a partir de la cual se pueden hacer las comparaciones. La forma de implementarlo requeriría el uso de frases repetidas en distintos tonos emocionales e inclusive un tono neutro. La idea básica de la metodología propuesta consiste en la determinación de características acústicas compartidas por todas las ocurrencias de una frase particular, a pesar de los diferentes tonos emocionales o tono neutro. Dicho conjunto de características compartidas podría ser considerado como la señal requerida para la codificación del mensaje; dicho de otra manera, podría ser considerado como la señal acústica en la que se han eliminado las contribuciones emocionales. A partir de la señal agregada, la comparación de la prosodia emocional con la prosodia “libre de información emocional” arrojaría las características específicas de la prosodia emocional sobre la que se trabajase. Este tipo de enfoque tiene la ventaja de que el agregado de emociones puede ser realizado sólo entre las frases de un participante – lo que permite ver las variaciones del mismo sujeto al producir diferentes entonaciones emocionales – o entre grupos de participantes – lo que llevaría a una determinación independiente del hablante de las características acústicas requeridas para codificar un mensaje –, aproximándose más a la caracterización de la prosodia libre de emoción.

- Convergencia en los parámetros acústicos referidos a la emoción.

Se ve en diversos estudios que las variables de promedio, rango y variabilidad de la frecuencia fundamental (f_0), así como el volumen y el ritmo del habla son variables importantes para diferenciar entre emociones, tanto en corpus de emociones actuadas como de emociones genuinas en varios idiomas (ver sección de estudios acústicos y computacionales). Sin embargo también existe evidencia de que las variables que revelan estudios acústicos y las que emplean los humanos pueden diferir (Pell, Paulmann, Dara, Alasseri, & Kotz, 2009).

- Uso de modelos continuos contra modelos discretos de la emoción.

El uso de diferentes enfoques teóricos para caracterizar las emociones tiene un impacto en los trabajos. El uso de modelos discretos de la emoción (emociones categóricas y normalmente un enfoque de emociones básicas), suele perder la noción de intensidad de las emociones (pero véase Bänziger & Scherer, 2005). El enfoque de emociones en modelos continuos (divididas en dimensiones continuas, como *arousal* y *valencia*, también a veces *dominancia*), permite identificar diferenciaciones más sutiles de las emociones, pero suele perder la capacidad para probar hipótesis evolutivas o patrones distintivos de emociones categóricas, a cambio, gana la capacidad de encontrar variables asociadas a las dimensiones que emplea. El equipo mexicano de Pérez Espinosa y Reyes García (2010) muestra un enfoque original y sugerente en el que partir de un modelo continuo, integra las variables en un modelo categórico, usando simultáneamente los dos modelos. El uso simultáneo de modelos podría ser útil en la aclaración de preguntas en el área.

Se ha sugerido que la capacidad de reconocer emociones es dependiente de la intensidad con la que se expresan (Pell, Paulmann, Dara, Alasseri, & Kotz, 2009), pero al manejarlas exclusivamente como categóricas es difícil responder a esta interrogante. Una gradación en los parámetros asociados a la emoción mediante los enfoques de manipulación digital podría ser útil para explorar esta cuestión: basándose en los datos acústicos asociados a la expresión de la emoción podrían construirse estímulos que variaran en la magnitud de estas características acústicas, explorando si esta manipulación influye en los juicios sobre el *grado* de la emoción expresada, o si lo que se obtiene es un umbral a partir del cual se logran discriminaciones de emoción (que no mejora considerablemente por encima de dicho umbral), siendo nula o muy difícil por debajo del mismo.

- Uso de términos y sutilezas semánticas en las traducciones.

A pesar de que se hacen esfuerzos por hacer las traducciones de la forma más equivalente posible (ej. Scherer et al., 2001), o de evitar estas dificultades mediante el uso de metodologías que no requieren de etiquetas verbales (Bryant & Barret, 2008; Sauter et al., 2010), el hecho de la equivalencia de términos entre

idiomas puede ser difícil de controlar. A manera de ilustración, en el trabajo de Pell, Monetta, Paulmann y Kotz (2009) en el que trabajaron con hispanohablantes, se empleó el término “*repugnancia*” para referirse a “*disgust*”. No es descabellado suponer como hablante nativo de español que se respondería diferente si el término hubiera sido “*asco*” en lugar de “*repugnancia*”.

En el trabajo de Pell, Paulmann, Dara, Alasseri y Kotz (2009), encuentran que los hablantes de alemán, al codificar estudios que pretendían expresar *miedo* (*fear*, *Angst*), fueron los únicos en producir frases con ritmo de habla lento y un rango de *f0* muy limitado, características contrarias a las frases producidas por hablantes de inglés, hindi y árabe. Los autores consideran que esto puede haber ocurrido si los participantes de la condición en alemán conceptualizaron un miedo angustiado, en contraste con un miedo que refleja *pánico*. El término que los autores emplearon para referirse al miedo (*Angst*) podría, en su semántica, estar más asociado a un miedo angustiado que al pánico. De hecho en el trabajo de Bänziger y Scherer (2005), conducido con estímulos en alemán, se hace una distinción específica entre los dos tipos de miedo, se refiere al miedo angustiado (*anxious fear*) como *Angst* y al pánico (*panic fear*) como *Panische Furcht*.

Ambos grupos de investigación consideran que otra fuente de confusiones en el área puede darse al interior de un término, por ejemplo al referirse al enojo, sin hacer una distinción entre caliente o frío (*hot anger* y *cold anger* que en español podrían traducirse tentativamente como furia y enojo, respectivamente). Si no se hiciera la distinción específica, ambas emociones caerían en la categoría de enojo aunque difieren en intensidad, es decir en *arousal*. Inclusive a pesar de hacer una mención explícita de esta distinción y describir una situación asociada se podrían dar dificultades en la codificación de uno u otro (Pell, Paulmann, Dara, Alasseri, & Kotz, 2009). Por supuesto, este punto se relaciona al uso de modelos teóricos distintos de la emoción; por ejemplo Bänziger y Scherer (2005) emplean un enfoque de familias de emociones, derivado de las teorías de la atribución o *appraisal*, a favor de cuyo uso argumenta Scherer (2003).

Con todo, los experimentos en el área de investigación intercultural en prosodia emocional son relativamente escasos y se suele observar un celo metodológico evidente. Incluir una condición de contrastes entre posibles términos agregaría dificultades a la conducción de dichos estudios. Sin embargo, este tipo de contrastes también podría brindar antecedentes y una base para proponer que el uso de términos puede ser una variable importante, de modo que se pudiera tomar en cuenta para futuros estudios.

Los límites semánticos diferentes entre distintas lenguas no son un tema nuevo, es conocido desde hace mucho en la literatura lingüística (ver por ejemplo Roberson & Hanley, 2010; Winawer, Witthoft, Frank, Wu, Wade, & Boroditsky, 2007). No sería sorprendente que el fenómeno de la existencia de diferentes límites semánticos en distintos idiomas también se diera para términos emocionales.

- Metodologías de análisis de lesiones y metodologías de imagenología.

Ambas metodologías permiten aclarar qué áreas son importantes para funciones específicas.

En el caso de las metodologías de imagenología, se observan las áreas que contribuyen a un proceso, por lo que pueden incluirse áreas con funciones relacionadas que están participando en el momento de la tarea; asimismo, requieren de condiciones de contraste, los resultados pueden variar dependiendo de la tarea de contraste empleada.

Las metodologías de análisis de lesiones muestran aquellas áreas críticas para una función, sin embargo, a veces pueden resultar en confusiones. Variables que influyen en estos estudios y son difícilmente controladas (normalmente por métodos estadísticos y clínicos solamente) son el tamaño y localización exacta de la lesión, el tiempo desde la lesión, la edad, el sexo, entre otras.

Las características de ambas metodologías llevan a algunas incongruencias en los resultados (Ross & Monnot, 2008). Ross y Monnot (2008), favorecen la

perspectiva clínica y ofrecen algunos argumentos a favor de su postura. También mencionan la capacidad de metodologías como la resonancia magnética transcraneal y el mapeo de funciones cerebrales previo a cirugías ablativas para ayudar a aclarar algunas de las interrogantes que surgen al estudiar la organización de funciones específicas en el cerebro.

- Procesamiento de eventos temporales de distinta duración en el cerebro y su relación con los hemisferios

Se ha sugerido que los hemisferios cerebrales difieren en cuanto a la duración temporal de los estímulos sonoros que procesan de manera óptima, de modo que el hemisferio izquierdo sería más apto para el procesamiento de estímulos cortos y rápidamente cambiantes, como la información fonémica, mientras que el derecho sería más apto para el procesamiento de estímulos de duración más larga, como la prosodia (Ross & Monnot, 2008; Wildgruber et al., 2005).

En la sección de neurobiología reportamos el estudio de Buchanan y su equipo (2000), donde se reporta un mayor involucramiento de la corteza temporal derecha en una tarea de detección de prosodia emocional contra una tarea con los mismos estímulos pero de discriminación fonémica. La base de esto podría ser el proceso que se plantea en este punto, las características temporales de los fonemas contra la prosodia (más breves y más largos respectivamente).

Asimismo, Pell (2002) y Breitenstein et al. (2001), discuten acerca del hallazgo de que personas con Parkinson tienen dificultades para comprender prosodia emocional. Dado que en esta población también son conocidos problemas en percepción de tiempo, es posible que a las deficiencias de percepción de prosodia emocional las subyazgan las dificultades de percepción temporal. Viéndolo desde este punto, la percepción de prosodia sería un proceso compuesto por tareas de discriminación temporal y tonal. Recordemos que en el trabajo del equipo de Breitenstein, se reporta un resultado previo en el que se muestra percepción temporal alterada pero percepción tonal conservada en pacientes con Parkinson. Más aún, una función compartida entre prosodia y *timing* podría ser la capacidad

de identificar patrones temporales estructurados o periódicos (Menez, 2011; comunicación personal).

De hecho, debemos recordar que la interacción entre hemisferios para la *producción* de prosodia emocional ha sido propuesta en el sentido de que el hemisferio derecho modifica planes motores del izquierdo para hacer diferencias en duración y/o tono asociadas a la prosodia (emocional y lingüística; Klouda, et al., 1988; Ross & Monnot, 2008; sección de “*el abordaje neurobiológico*” en el presente capítulo).

Otros autores también han sugerido la noción de un procesamiento de tono y de tiempo al procesar prosodia emocional (Buchanan et al., 2000). Desde esta perspectiva, el hemisferio izquierdo estaría involucrado principalmente en detección de duración mientras que el derecho en detección de tono.

- Interpretación evolutiva.

El enfoque evolutivo ha guiado a esta literatura en más de una manera. Una de las interpretaciones más valiosas está en el campo de la capacidad diferencial de reconocer emociones, válida tanto para emociones producidas en el propio idioma como en uno distinto. Siguiendo los razonamientos de Scherer (2003), Pell, Paulmann, Dara, Allasseri y Kotz (2009) y Sauter y su equipo (2010), las presiones evolutivas para comunicar algunas emociones negativas a distancia, así como la mejor aptitud del rostro contra la voz para expresar repugnancia, y la función de cohesión social de emociones positivas al interior de un grupo pero no al exterior del mismo, podrían ofrecer una explicación (aunque dicho argumento no está exento de críticas: que la alegría sea fácilmente reconocida interculturalmente en el rostro va en contra de su argumento sobre la función de cohesión social de las emociones positivas al interior, y no al exterior del grupo; si esto último fuese así, debería haber variaciones interculturales en la expresión de la alegría en el rostro). Con todo, hipótesis alternativas y más estudios empíricos deberían ser conducidos

para resolver estas interrogantes. De cualquier forma, este tipo de razonamientos y perspectiva ofrecen ideas a contrastar e informan a la investigación.

- Reconocimiento intercultural y ventaja al interior de los grupos.

La investigación intercultural destaca estos dos puntos en mucha de la investigación realizada. Esto realza la influencia de variables biológicas y apoya hipótesis evolutivas (reconocimiento intercultural) y también enfatiza las diferencias debidas a la cultura (ventaja al interior de los grupos). El conocer el lenguaje permite distinguir qué en el patrón sonoro es requisito del mensaje (contrastes semánticos y probablemente sutilezas de significado compartidas, como funciones pragmáticas relacionadas a intenciones: ej. Sarcasmo, admiración, sufrimiento etc. que podrían ser transmitidas por las mismas características sonoras pero significar una u otra cosa dependiendo de la cultura; Erickson, 2006) y qué no lo es (lo que podría señalar distinciones emocionales). Este tipo de discriminación no sería posible para alguien a quien el idioma le es ajeno. También, si uno asume que las personas con conocimiento sólo de su lengua suponen que todas las lenguas funcionan de manera similar y se basan en el conocimiento de la suya al hacer inferencias en otra, podría ofrecer una base teórica para la hipótesis del *gradiente de proximidad*. Esta hipótesis, con todo, no ha recibido mucho apoyo en su variante definida a partir de tipología de las lenguas, de hecho, son más los resultados negativos, sin embargo, podría tenerlo en su variante definida a partir de características rítmicas y entonacionales.

- Longitud de los estímulos

Aunque algunos estudios emplean frases como estímulos, otros utilizan palabras. Esto no impide que se muestren los efectos de reconocimiento de emociones, y se observa que incluso a partir de relativamente poca información auditiva se pueden reconocer. Esto llega al extremo en el estudio de Sauter y colaboradores (2010), en el que ni siquiera se usan palabras y los estímulos son en extremo cortos (expresiones vocales de la emoción sin lenguaje, disponibles en internet). Sin embargo, sí se han mostrado efectos de la longitud de los estímulos (Pell &

Skorup, 2008). Cabe preguntarse si independientemente de la longitud de los estímulos, las características acústicas referidas a descriptores globales (promedios y rangos de variables) y aquéllas relativas al tiempo (como *contornos de tono*) sean similares. ¿Si se normalizaran los *contornos* relativos a la duración del estímulo se encontraría la misma forma general?

- Efectos de género

Tanto en la literatura intercultural (Dromey et al., 2005; Erickson, 2006) como en la literatura de neurobiología (Schirmer et al., 2002), se encuentran efectos de género relacionados al desempeño en los primeros y al curso temporal de actividad eléctrica en el segundo. La superioridad de las mujeres en tareas emocionales también se ha observado a partir de expresiones faciales, tareas léxicas y otras (Dromey et al., 2005). Schirmer y colaboradores (2002) proponen que en el caso de las mujeres, la información emocional es integrada con otras tareas, mientras que en los hombres podrían funcionar de manera más independiente, y que una organización del lenguaje en el cerebro diferente entre géneros podría influir. También consideran que podría existir una diferencia en las formas preferidas de expresión: las mujeres prefiriendo expresar emociones mediante la prosodia, mientras que los hombres mediante el contenido de lo dicho. No obstante, estos hallazgos también podrían estar influidos por variables culturales y biológicas.

- Uso de emociones actuadas contra emociones genuinas

¿El uso de emociones genuinas o emociones actuadas hace alguna diferencia? Aparentemente, no. Los estudios revisados convergen a pesar del hecho de que algunos emplean emociones presumiblemente genuinas y otros, emociones actuadas, en especial del área acústica y computacional. Aún más, algunos emplean actores sin experiencia mientras que otros emplean actores profesionales (este tema es discutido ampliamente por (Scherer, 2003).

En un primer momento se podría pensar que la mejor metodología posible sería emplear emociones genuinas, sin embargo esto es difícil en la práctica, dada la

dificultad de conseguirlas con la calidad que se requiere para estudios de laboratorio (ruido de fondo, longitud de estímulos, neutralidad semántica, etc.). Las dificultades para grabar estímulos con emociones genuinas pronunciando frases carentes de significado serían mucho mayores. También se presentan problemas en identificar cuál emoción corresponde a la expresión obtenida espontáneamente (Scherer, 2003). Otra posibilidad es elicitación de emociones en un ambiente controlado, aunque este enfoque tampoco está exento de dificultades.

La segunda mejor metodología parecería el uso de actores profesionales, ya que estos podrían, por la naturaleza de su profesión, tener una mejor idea de cómo expresar las emociones (Scherer, 2003, recomienda el uso de técnicas de actuación en las que se induce el estado emocional, y la inclusión de descripciones específicas de la emoción y el antecedente que lo provoca). Sin embargo, tener muy bien estudiada la forma de mostrarlas podría reflejar sesgos culturales y no características de las emociones, lo que sería indeseable en los estudios que desean saber las contribuciones biológicas a la prosodia emocional. Emplear estímulos codificados por actores no profesionales podría aminorar el estudio intensivo de las emociones, pero no resuelve los problemas de la influencia de la cultura.

Aún en las ocurrencias más genuinas, no se pueden descartar los efectos de la cultura, tal como demuestran los argumentos que aducen la influencia de las *reglas de exposición*. También se ha discutido la ocurrencia casi exclusiva de prosodia emocional en situaciones interpersonales, en donde el conocimiento pragmático compartido entre los hablantes juega un papel (Pell, 2002). También debe considerarse que existen en todo momento restricciones sociales, situacionales y tendencias inconscientes a auto-presentarse de cierta manera (Scherer, 2003).

Scherer (2003) argumenta que al usar emociones actuadas se podrían exagerar algunas pistas perceptuales relativamente obvias y se pierdan otras más sutiles que se dan en situaciones reales. Considera prioritaria la realización de un estudio

que explore específicamente las diferencias entre emociones actuadas y genuinas.

Finalmente, los requerimientos prosódicos para un mensaje así como los requerimientos culturales para señalar ciertas cosas, afectan a la señal sonora. Por ejemplo, Scherer y Bänziger (2004) ejemplifican el caso de una cultura donde se requiere sonreír para señalar cortesía, en este caso, el mismo hecho de sonreír modifica la forma del tracto vocal y tiene consecuencias en la señal sonora. Al respecto podemos acomodar este ejemplo a una de las tesis fundamentales de este trabajo (la *hipótesis de grados de libertad*): que la prosodia emocional hace uso de las características sonoras no contrastivas de los mensajes, por lo que a partir de un ejemplo como este se puede considerar que se aprende a adjudicar las características sonoras a la fuente relevante, prosodia emocional o requerimiento lingüístico o cultural del mensaje. A la larga se consolidaría en aprendizaje perceptual, permitiendo una evaluación rápida e inconsciente sobre la fuente a la cual adjudicarla. Estas diferencias serían muy variables dependiendo de la cultura y harían posible que (dependiendo del grado de solapamiento entre las señales acústicas de una fuente u otra en diversas culturas), la distinción de situaciones particulares variara en dificultad de una cultura a otra.

El punto más importante sobre esta pequeña discusión es notar la convergencia de resultados de distintas metodologías.

- Influencia de variables biológicas, psicofisiológicas y culturales a la prosodia emocional

Queda claro que existen influencias biológicas y culturales a la prosodia emocional. Entre las biológicas, los mismos cambios fisiológicos que subyacen a las emociones (Scherer & Bänziger, 2004). Quizá la existencia de factores biológicos que subyacen a la prosodia emocional es algo fácilmente esperado, en especial dada la interpretación evolutiva de la comunicación de emociones; sin embargo, en el presente capítulo abordamos en más detalle la organización específica en el cerebro de este fenómeno.

La influencia de factores culturales (que conllevan aprendizaje) puede ser algo menos esperada, y se encuentra principalmente en cuanto a las *reglas de exposición*, así como las características del lenguaje y los efectos de hablar un idioma particular (o varios). Reflejada principalmente por los efectos de ventaja al interior del grupo.

Con un énfasis en la prosodia emocional, Scherer y Bänziger (2004) incluyen ambos tipos de influencia en su perspectiva, llamando factores de *empuje (push)* a los primeros y factores de *jale (pull)* a los segundos. Específicamente definen a los factores de empuje como: “las externalizaciones biológicamente determinadas de procesos que ocurren naturalmente en el organismo, particularmente procesamiento de la información y preparación de la conducta”, y a los factores de *jale* como: “las normas o *modelos* determinados socioculturalmente sobre las características de la señal requeridas por los códigos compartidos socialmente para la comunicación de estados internos e intenciones conductuales”.

En el siguiente capítulo se abordará la prosodia lingüística. Es esperable encontrar contribuciones culturales importantes y relativamente menos contribuciones biológicas, pero se revisarán ambas perspectivas.

3.3 Prosodia Lingüística y Abordajes Lingüísticos

Así como en la sección anterior se presentó principalmente trabajos desde la perspectiva psicológica y fisiológica, o informados principalmente por las teorías y conceptos de estos campos (de hecho, a propósito de abordajes lingüísticos sólo se presentó el trabajo de Benus et al., 2007, e incidentalmente el trabajo de Caffi & Janney, 1994, sobre comunicación emocional), en la presente sección se presentarán principalmente estudios de abordajes lingüísticos, ya que estos son los que más se han ocupado de la prosodia lingüística.

Es importante tener en mente que la conceptualización de prosodia lingüística y emocional no es la misma desde la perspectiva lingüística: ¿qué sentido tendría hablar de prosodia lingüística en el campo lingüístico? ¿No sería tanto como hablar de “emociones psicológicas” para un psicólogo? Aparentemente el determinante “psicológicas” simplemente sobra al estar dentro del campo de la psicología.

En el campo lingüístico se suele hablar de “funciones del lenguaje”, y dentro de estas funciones se encuentra la expresión de emociones. Extendiendo esta perspectiva a la prosodia se puede hacer un boceto de la división de la prosodia en el campo de la lingüística en el que la prosodia puede dividirse en funciones, entre las cuales se encuentra la de expresar emociones. Esto contrasta con la división que se presentó en las secciones anteriores (emocional vs lingüística) pues la división de la prosodia en el campo lingüístico incluye varias ramas (las funciones) que se encuentran al mismo nivel jerárquico que la prosodia emocional.

Esta diferencia entre los dos campos en la conceptualización de la prosodia tiene sentido al considerar que la cantidad y especificidad de la división de un concepto como este debe ser más amplia para un campo cuyo objeto de estudio es el lenguaje: simplemente denominar lingüística y emocional a las funciones de la prosodia no provee de una distinción clara y adecuada para los estudiosos de la lengua.

Otro contraste a resaltar es la conceptualización diferente de algunos términos que son centrales en psicología pero no en la lingüística, por ejemplo, Caffi y Janney (1994) escriben respecto del uso del concepto “afecto” en el campo lingüístico que:

“En lingüística, por otra parte, el término “afecto” es usado de manera simple como un sinónimo amplio para “sentimiento”, y es considerado como incluyendo no sólo las nociones psicológicas tradicionales de emoción, estado de ánimo y actitud, sino también nociones de carácter y personalidad⁹, y nociones relacionadas a fenómenos interaccionales lingüísticos como el enmascaramiento, cobertura (hedging), undercutting, y otros”.

Este tipo de diferencias en el uso de términos “psicológicos” como *afecto*, se observa también en trabajos como el de Gussenhoven (2002).

Es evidente que no se comparte la cantidad de diferencias sutiles que se emplean en el campo de la psicología, donde son requeridas por la naturaleza del trabajo de los psicólogos, dado su objeto de estudio y el grado de profundidad con el que se requiere tratar ciertos temas.

No quiero decir con esto que el marco de la psicología sea superior al de la lingüística; de hecho, la investigación lingüística a partir de sus conceptualizaciones ha propuesto muchas ideas importantes y fructíferas, como las que se abordarán en el presente capítulo. Simplemente son campos diferentes que requieren niveles de especificación diferentes para distintos términos.

Este tipo de consideraciones conceptuales serán elaboradas de manera más amplia en el capítulo de “Prosodia lingüística y prosodia emocional” y posteriormente en el dedicado al análisis conceptual, pero resultan pertinentes para introducir este capítulo.

Definitivamente los campos de la lingüística, la psicología y la psicobiología, aunque convergen, divergen, y también difieren en sus enfoques e intereses de estudio incluso si se dirigen en un momento dado a un mismo fenómeno; por ejemplo, la prosodia.

⁹ Aquí también no queda claro si el uso de términos como carácter y personalidad se pueden asumir equivalentes a los del campo psicológico.

Este capítulo enriquecerá el trabajo de tesis mediante la incorporación del cuerpo teórico de una disciplina hermana. Se tratarán de plantear claramente los conceptos teóricos y de hacer comparaciones con los de la psicología cuando sea necesario, sin embargo, una gran cantidad de conceptos son relativamente independientes – aunque semejantes – a los que se manejaban en la sección anterior, y en algunos momentos “ensamblan” mejor entre ellos que en el conjunto formado por los conceptos de ambas disciplinas.

Indudablemente esta aproximación enriquece la comprensión de los fenómenos: algunos temas difícilmente vendrían a la mente siguiendo un enfoque psicológico, pero son centrales desde un enfoque lingüístico y viceversa. También se puede apreciar que algunos problemas fundamentales desde una perspectiva son epifenómenos o incluso ruido en la otra: Mientras que en los estudios reportados en el capítulo anterior, guiados principalmente por la perspectiva biológica, se intenta aproximarse lo más posible a los sonidos más naturales (más auténticos y consecuencia directa de estados emocionales) de la prosodia emocional, en trabajos lingüísticos se considera a veces un fenómeno paralingüístico (una manera muy lingüística de decir epifenómenos); pero a la vez, para ellos es importante la competencia de las personas para reproducir voluntaria y estratégicamente este tipo de sonidos, independientemente del estado fisiológico real (por ejemplo, Caffi & Janney, 1994 ; Gussenhoven, 2002).

Como en el capítulo anterior, se tomarán en cuenta estudios que permitan explorar las interrogantes básicas del presente trabajo de tesis: influencias biológicas y culturales, en este caso a la prosodia lingüística. Sin embargo, las secciones de las que se compone este capítulo no siguen el mismo esquema que las del anterior.

Comenzaremos presentando el enfoque general de funciones de la prosodia, deteniéndonos sólo en unas pocas que han sido más consistentemente estudiadas, incorporaremos los trabajos sobre fisiología, los cuales generalmente analizan una o pocas de estas funciones. Posteriormente presentaremos los estudios interculturales. A dicha sección la seguirá la presentación de las teorías

más importantes y cómo permiten integrar la perspectiva de influencias biológicas y culturales al estudio de la prosodia lingüística. Finalmente se resaltarán los hallazgos relacionados a influencias biológicas y culturales y se dará paso a la discusión y consideraciones finales del capítulo.

3.3.1 Funciones de la prosodia.

La idea de dividir al lenguaje en funciones tiene una larga historia. El funcionalismo de Praga es una corriente lingüística que lo puede ilustrar bien, y dentro de ella, la división de las seis funciones del lenguaje (referencial, fática, poética, apelativa, metalingüística y emotiva) propuesta por Jakobson (Caffi & Janney, 1994). Cabe destacar que esta perspectiva se refiere a las funciones del lenguaje en su totalidad, no de la prosodia.

Muchos de los estudiosos de la prosodia también han propuesto sus funciones, que son muy diversas, y no se refieren a una división que compita con las funciones del lenguaje de Jakobson; más bien se refieren a una división más sutil que cabe dentro de las funciones del lenguaje de dicho autor.

Entre las principales funciones de la prosodia se encuentran el foco y el énfasis (Fitzpatrick, 2000; Grice & Baumann, 2007; Gussenhoven, 2002; Gutiérrez Bravo, 2008; Hirschberg, 2002a; Meng, Tseng, Kondo, Harrison, & Viselgia, 2009; Wilson & Wharton, 2006); el fraseo y desambiguación sintáctica (Frazier et al., 2006; Grice & Baumann, 2007; Hirschberg, 2002a; Kouassi, 2008; Meng et al, 2009; Shah et al., 2006; Wilson & Wharton, 2006); marcar el tópico (Bryant & Fox Tree, 2005; Fitzpatrick, 2000; Gutiérrez Bravo, 2008; Hirschberg, 2002a; Meng et al, 2009); la distinción entre información dada e información nueva (Bryant & Fox Tree, 2005; Fitzpatrick, 2000; Grice & Baumann, 2007; Meng et al, 2009); la distinción de los actos de habla (Fitzpatrick, 2000; Grice & Baumann, 2007; Meng et al, 2009); indicar toma de turnos en una conversación y si el enunciado ha finalizado o va a continuar (Grice & Baumann, 2007; Hirschberg, 2002a; Meng et al, 2009); funciones emocionales/emotivas (algunas veces consideradas paralingüísticas; Fitzpatrick, 2000; Grice & Baumann, 2007; Gussenhoven, 2002; Kouassi, 2008; Meng et al, 2009). En menor medida se ha resaltado la capacidad

de la prosodia para cumplir funciones pragmáticas, interpersonales y sociales¹⁰ (Caffi & Janney, 1994; Campbell, 2004; Gussenhoven, 2002; Hirschberg, 2002a), la transmisión de actitudes (Meng et al, 2009) y apoyo y diferenciación sutil del contenido semántico (Gussenhoven, 2002; Hirschberg, 2002a; Kouassi, 2008; Seddoh, 2002; Wilson & Wharton, 2006). En algunas ocasiones se han propuesto también funciones como la transmisión de certidumbre (Fitzpatrick, 2000; Gussenhoven, 2002; Hirschberg, 2002a), cortesía (Fitzpatrick, 2000; Gussenhoven, 2002; Hirschberg, 2002a), interés (Fitzpatrick, 2000), relación predicado-argumento (Fitzpatrick, 2000), estado de ánimo (Fitzpatrick, 2000); estructura de eventos (Fitzpatrick, 2000), estilo (Fitzpatrick, 2000), actos de habla directos contra indirectos (Fitzpatrick, 2000), factores/funciones del discurso (Fitzpatrick, 2000; Meng et al, 2009), negación (Gussenhoven, 2002), alejamiento (Gussenhoven, 2002), servicialidad (Gussenhoven, 2002), transmitir autoridad o dar la impresión de credibilidad en lo que se dice (Bryant & Fox Tree, 2005; también Pell, Paulmann, Dara, Allasseri, & Kotz, 2009 adelantan un comentario al respecto), señalar sumisión (Gussenhoven, 2002), simpatía (Gussenhoven, 2002), vulnerabilidad (Gussenhoven, 2002), confianza, (Gussenhoven, 2002), protección (Gussenhoven, 2002), agresión (Gussenhoven, 2002), masculinidad y feminidad (Gussenhoven, 2002), señalamiento de implicaturas conversacionales (Hirschberg, 2002b; Wilson & Wharton, 2006), entre otras.

No se supone que estas sean funciones exclusivas de la prosodia, sino funciones en las que participa, en mayor o menor medida y en las que interactúa con otros niveles del lenguaje. Por ejemplo, en el caso del foco (énfasis prosódico que implica un énfasis en la información; será tratado en detalle en el apartado 3.3.1.2), se asume que debe existir información semántica con la que la prosodia interactúe para que se pueda dar la función (Bryant & Fox Tree, 2005).

Algunas funciones requieren de pasos inferenciales por parte del oyente para que se puedan dar, de otra manera se mantendría solamente el significado literal (Caffi

¹⁰ Si bien todas las funciones antes enumeradas deberían satisfacer estas (Menez, 2011, comunicación personal), las mencionamos aparte al tener una relación más íntima con estas categorías, como será ilustrado en la sección de "otras funciones pragmáticas y sociales de la prosodia" del presente capítulo.

& Janney, 1994). Por ejemplo, se requiere tomar en cuenta el contexto para poder entender la ironía, donde se puede intentar transmitir lo opuesto a lo que de hecho se dice (Bryant & Fox Tree, 2005). Estas funciones difieren en el grado de automaticidad y de pasos inferenciales que se requieren para que se puedan dar: así, se puede asumir que distinguir una pregunta de una declaración (por su contorno melódico) es una tarea relativamente automática, mientras que detectar ironía requiere estar inmerso en el contexto.

Wilson y Wharton (2006) consideran que si algo requiere de inferencia, no puede considerarse como una función de la prosodia, ya que si bien podría señalarlo, no implicaría que hacerlo sea su función. Por ejemplo, si bien el estado físico de una persona (por ejemplo, estar cansado) podría influir en su habla, incluyendo características prosódicas, de modo que un interlocutor pueda notar el estado físico de la persona sólo por el habla, no sería la función de la prosodia transmitir esta información; mientras que si la prosodia emocional está determinada por códigos innatos, se puede considerar que señalar el estado emocional *sí* es una función de la prosodia. En el caso de la prosodia lingüística se puede hablar de funciones adquiridas, producto del aprendizaje y la conformación de códigos lingüísticos, por ejemplo la distinción de preguntas y declaraciones por el contorno melódico; sin embargo, otros autores también han propuesto la influencia de códigos biológicos en la prosodia lingüística, por lo que no sería completamente arbitraria, y se ha identificado la existencia de tendencias lingüísticas universales; por ejemplo, en la mayoría de las lenguas las preguntas son señaladas por subida de tono al final de la frase (ver Gussenhoven, 2002).

Dada la interacción entre varios niveles del lenguaje (fonología, sintaxis, semántica, etc.) y las diferencias entre idiomas, la misma función (por ejemplo, foco) podría ser señalada por medios prosódicos o sintácticos en diferentes lenguas o en diferentes ocasiones en la misma lengua.

Sólo abordaremos algunas de estas funciones. No nos centraremos en las caracterizaciones acústicas (aunque en casos pertinentes se mencionarán a

manera de ejemplo) ya que éstas pueden diferir entre idiomas, pero mencionaremos tendencias universales identificadas.

Al mencionar ejemplos de características acústicas es importante recordar que una gran cantidad de la información está basada en las tendencias del idioma inglés, por lo que no necesariamente son válidas para otros idiomas. El contraste específico entre características acústicas de varios idiomas al respecto de diferentes funciones, si bien podría ser un esfuerzo importante en la determinación de características universales e influencias culturales, va más allá de los alcances del presente trabajo.

La información fisiológica será incluida en los apartados de las funciones; mientras que las consideraciones generales sobre fisiología de la prosodia lingüística (por ejemplo las hipótesis generales de organización de prosodia lingüística en el cerebro) serán abordadas en la sección posterior: fisiología de la prosodia lingüística.

3.3.1.1 Fraseo y desambiguación sintáctica

La atención en la prosodia como parte importante para el procesamiento del lenguaje tanto en lingüistas como en desarrolladores de tecnología relacionada con el habla ha ido en aumento; se recurre cada vez más a la prosodia para desambiguar instancias de ambigüedad sintáctica y semántica (Hirschberg, 2002a). Uno de los mecanismos en los que la prosodia contribuye a la desambiguación es el fraseo. Hay varios factores que intervienen en el proceso del fraseo, principalmente los parámetros temporales.

La prosodia divide el habla en “porciones”. A estas “porciones” se las ha denominado *grupos de aliento*, *grupos de sentido*, *unidades tonales*, *frases fonológicas*, *frases entonacionales*, entre otras. Quizá el indicador más importante de los límites entre estas porciones son las pausas: Entre más larga sea la pausa, el límite es percibido con mayor claridad; aunque también es posible percibir límites aunque no existan pausas, esto se puede dar por un cambio repentino en el tono, siempre y cuando no sea atribuible a otra función como la enfatización. Estos cambios repentinos de tono pueden corresponderse, con límites sintácticos,

pero no necesariamente deben hacerlo, aunque es mejor considerarlos relativamente independientes (Grice & Baumann, 2007) precisamente por que su correspondencia no es obligada y por que en diversos idiomas se observan grados diferentes de coocurrencia de límites sintácticos y prosódicos (Grice & Baumann, 2007). Este proceso de dividir el habla en porciones es denominado fraseo.

Mediante el fraseo se pueden resolver instancias de ambigüedad como las siguientes.

- *¿De qué color son los carros de Ana y de Nancy?*

- *Rojo y negro y azul.*

En un ejemplo como este, sin signos de puntuación ni información prosódica, no resulta claro de qué color es cada automóvil. Si señalamos las pausas posibles de esta oración (|) se puede apreciar la capacidad de desambiguación sintáctica del fraseo.

- *Rojo | y negro y azul.*

VS

- *Rojo y negro | y azul.*

Una pausa en un lugar específico de la oración puede dar lugar a la resolución de la ambigüedad, y se aprecia como tiene un profundo efecto sobre su significado. El mismo tipo de fenómeno se puede observar en alemán (Grice & Baumann, 2007) e inglés (Kouassi, 2008). De hecho, se ha propuesto la existencia de algoritmos universales para el fraseo con pocas restricciones dependientes de lenguaje (Fitzpatrick, 2000).

Con todo, Grice y Baumann (2007) remarcan que no hay que esperar que los hablantes hicieran siempre este tipo de distinciones prosódicas, ya que la ambigüedad puede desaparecer en un contexto claro a pesar de la falta de distinción prosódica.

Se ha identificado que las palabras que están antes de finalizar la frase prosódica incrementan en duración (Shah et al., 2006) dado que hay una ralentización de los articuladores (Grice & Baumann, 2007). Este efecto es más pronunciado entre más larga sea la frase. El fenómeno se ha encontrado en varias lenguas y se asume de base fisiológica, aunque el grado en que se da obedece, en parte, a la lengua (Grice & Baumann, 2007). A su vez, las primeras sílabas de la frase posterior a un límite de frases prosódicas pueden ser pronunciadas de manera más veloz; así, un cambio abrupto en ritmo y tono, conforman una pista perceptual importante para percibir límites entre frases (Grice & Baumann, 2007).

Esta función interactúa de manera importante con la función de “Foco” para marcar distinciones semánticas sutiles. También está relacionada con la función de señalamiento de toma de turnos en la conversación e indicar si la frase que alguien está en proceso de decir ha concluido o falta algo más por ser dicho. Ambas serán abordadas posteriormente.

Shah y colaboradores (2006) realizaron un estudio comparando la capacidad de personas con daño en hemisferio izquierdo, derecho y participantes control (sin daño cerebral) para producir las características acústicas esperadas en diferentes tareas de fraseo; también analizaron la capacidad de un grupo independiente de hablantes sanos para percibir los límites en las frases producidas por los 3 grupos mencionados. Encuentran – de manera consistente con la literatura en el área – patrones similares a los normales para las distinciones prosódicas realizadas tanto por individuos con daño en hemisferio izquierdo como en el derecho. Sin embargo, encuentran que los individuos con daño en el hemisferio izquierdo muestran un uso menos consistente de parámetros temporales para señalar límites de frases prosódicas (y con esta dificultad se implica una para señalar distinciones sintácticas). Las pausas al interior de frases que no requerían una distinción de límite de oración fueron más largas de lo normal, lo que interpretan como evidencia de que los individuos con daño en hemisferio izquierdo podrían estar produciendo pausas en lugares inapropiados. No encuentran en ningún grupo patrones de F0 asociados a límites entre frases. En contraste, en el estudio

perceptual (en el que un cuarto grupo juzgó las producciones de los tres primeros) fueron las producciones de personas con daño en hemisferio derecho, y no en izquierdo, aquellas en las que fue más difícil reconocer los límites de frases.

Los autores remarcan que, al no encontrar patrones claros en F0, podría bastar con características temporales para señalar los límites de frases prosódicas. Como se remarcó en este apartado, variables como pausas y su duración, así como el ritmo (ambas variables temporales) parecen ser las más importantes para la función de fraseo y su contribución a la desambiguación sintáctica y semántica.

Frazier y colaboradores (2006) enfatizan que todos los lenguajes hacen uso del fraseo, aunque en diferentes formas. El procesamiento de aspectos relacionados al fraseo, como los límites entre las frases es procesado de manera temprana; los mismos autores reportan un efecto inmediato en potenciales relacionados a eventos. Resulta particularmente interesante la evidencia que citan sobre procesamiento prosódico durante lectura en silencio: mediante potenciales evocados, se encuentra un efecto de las comas en estímulos escritos, análogo al efecto de límites de frases en estímulos auditivos; también reportan correlación entre el movimiento de los ojos y la prosodia: en las frases escritas, el tiempo de movimiento de los ojos y ligeras pausas corresponde con los sitios donde hay límites de frase obligatorios en estímulos auditivos.

3.3.1.2 Funciones relacionadas a la estructura informativa: Foco, Tópico y distinción de Información dada/Información nueva.

Además de las categorías sintácticas (ej. frase adverbial, oración subordinada, etc.) existen categorías pragmáticas referidas a la *estructura informativa* de la oración (Gutiérrez Bravo, 2008). Entre estas categorías encontramos el *foco*, el *tópico* y la distinción entre *información dada/información nueva*. La prosodia juega un importante papel para señalar estas funciones en el lenguaje oral.

La principal propiedad formal de los focos es de hecho, prosódica, como se ilustrará en breve. Se ha propuesto la existencia de dos tipos de foco, el foco contrastivo y el foco informativo; el primero resalta a una entidad específica entre

un conjunto de alternativas y el segundo corresponde a una información nueva que contrasta con información ya compartida entre los interlocutores. Se puede entender de manera general que algunas palabras están focalizadas, o se hace un énfasis prosódico en ellas, que conlleva un énfasis informacional. Otros autores han descrito de manera general la función de *énfasis* que puede cumplir la prosodia, pero al ser un concepto muy relacionado con el de *foco*, se describirán al mismo tiempo en la presente exposición.

En la interacción oral, las palabras focalizadas hacen un contraste con la información que podría llamarse “de fondo”; qué información es focalizada y cuál es considerada de fondo depende de la estructura previa del discurso y de las intenciones del hablante (Grice & Baumann, 2007). Un ejemplo del uso de la prosodia en la focalización es el siguiente (de Grice & Baumann, 2007).

- *Liebst Du Maria oder Anna? (¿Amas a María o a Ana?)*

- *Ich liebe MARIA. (Amo a María)*

La forma de marcar el foco difiere entre lenguas; por ejemplo, en algunas lenguas no se pueden focalizar prosódicamente verbos, y otras tienen un sistema que dependiendo del “*peso semántico*” de las palabras, son o no focalizables (Grice & Baumann, 2007); incluso algunas lenguas poseen marcas morfológicas específicas de foco (Gutiérrez Bravo, 2008). Asimismo, en algunas lenguas, la distribución de los focos en la oración es bastante libre, mientras que en otras, los elementos focalizados deben ocupar una posición específica en la cláusula, como en el húngaro y muchas lenguas mayas (Gutiérrez Bravo, 2008). A su vez, puede ser señalado de diferentes maneras al interior de una lengua, por ejemplo, en inglés puede ser señalado mediante prominencia entonacional, variación de fraseo y entonación, o mediante modificaciones en el orden de las palabras y el uso de construcciones sintácticas particulares (Hirschberg, 2002a).

Con todo, la distribución de los focos en las lenguas parece depender más de la flexibilidad prosódica que de la flexibilidad sintáctica (Gutiérrez Bravo, 2008); es decir, parece ser una función condicionada principalmente por el sistema

prosódico de las lenguas. Los siguientes ejemplos (de Gutiérrez Bravo, 2008) ilustran como en inglés se puede expresar el foco mediante medios prosódicos, mientras que en español se requiere un cambio en el orden de las palabras:

A.-*TRACY* voted for Peter VS *La que* votó por Pedro fue Tracy.

B.-Tracy *DANCED* that night. VS *Lo que* hizo Tracy esa noche fue bailar.

En el ejemplo B, la traducción “Tracy bailó esa noche” no captura el significado original del inglés, en el que se expresa que lo que hizo fue bailar, y *no otra cosa*; en contraste, la traducción “Lo que hizo Tracy esa noche fue bailar” es más cercana a este significado. Un razonamiento análogo se puede aplicar al ejemplo A.

Sin embargo, esto no implica que en español, el foco no se exprese por un medio prosódico. El sistema prosódico del español requiere que el acento prosódico nuclear de la frase esté en el extremo derecho, por lo que al expresar una frase con un elemento focalizado se modifica el orden de las palabras para poder mantener este requerimiento, así, la modificación sintáctica es producto de un requerimiento prosódico. En muchos lenguajes de distintas familias se observa una correlación entre el foco y la entonación, los elementos focalizados se señalan mediante énfasis prosódico (Gutiérrez Bravo, 2008).

Es necesario señalar que el énfasis prosódico que señala el foco no necesariamente debe incluir toda la información a ser focalizada. El fenómeno de proyección de foco permite que aunque el énfasis prosódico esté en una palabra, la interpretación de la información focalizada pueda extenderse a la frase de la cual forma parte (Gutiérrez Bravo, 2008). La preferencia por la palabra que recibe el acento (llamada exponente de foco) es dependiente de lenguaje (Grice & Baumann, 2007).

Shah y colaboradores (2006) reportan que las habilidades para marcar prosódicamente el foco se conservan sin alteraciones (al menos acústicamente) en personas con daño cerebral en el hemisferio izquierdo o derecho; sin embargo, al ser escuchados por hablantes control, la localización de las palabras focalizadas

fue más difícil de identificar en las producciones de personas con daño cerebral en cualquiera de los dos hemisferios.

La segunda de estas funciones, el *tópico* ha sido dividido en tópico oracional y tópico discursivo. El primero, a grandes rasgos, se refiere a una frase sintáctica que en lugar de aparecer en su posición canónica aparece desplazada hacia el principio de la oración. El segundo tipo de tópico, el tópico discursivo, se refiere a una entidad que ha sido previamente introducida en el discurso, es decir, el tema o tópico común de un conjunto de oraciones que juntas forman una unidad discursiva mayor. La propiedad más importante de los tópicos oracionales es que se desplazan hacia el inicio de la oración; pero en algunas lenguas también existen marcas de tópico (Gutiérrez Bravo, 2008), por ejemplo la partícula は (wa) del japonés. Los tópicos oracionales también muestran propiedades entonacionales particulares que los distinguen de los focos. Para muchas lenguas se ha observado que los tópicos forman una unidad entonacional por sí misma, separada de la unidad entonacional que corresponde al resto de la oración (este fenómeno se da en los tópicos desplazados al inicio de las oraciones, Gutiérrez Bravo, 2008), lo que permite identificar la relación de la función tópico con la de fraseo. El siguiente ejemplo (de Gutiérrez Bravo, 2008) ilustra este fenómeno:

Al paciente, la enfermera lo salvó en el último momento.

En esta oración, el tópico es “*al paciente*”, se observa que está al inicio de la oración (en contraste con la forma: “*la enfermera salvó al paciente en el último momento*”), y, como señala la coma, la frase se realiza de manera separada del resto de la oración. Aunque en español no siempre se observa esta independencia prosódica de los tópicos, hay lenguas, como el italiano, en las que sí (Gutiérrez Bravo, 2008).

En cuanto al tópico discursivo (aquél que señala el tema general del que se está hablando), se ha encontrado que inicios de un nuevo tópico son señalados por subidas de tono al inicio de las frases mientras que continuaciones de tópico son señalados por bajadas de tono (Gussenhoven, 2002; Grice & Baumann, 2007), de

manera relacionada, subidas de tono al final de las frases indican que se continúa en el mismo tópico y bajadas que se finaliza. Esto se da en varias lenguas y según la perspectiva de Gussenhoven (2002) podría ser una tendencia universal. Hirschberg (2002b) reporta que el rango de tono incrementa al iniciar nuevos tópicos en una variedad de lenguas.

La distinción de *información dada/información nueva* determina qué tan compartida o qué tan nueva es la información entre los participantes de la conversación en un momento dado (Grice & Baumann, 2007). En español, por ejemplo, se emplean los artículos indefinidos para introducir información nueva en el discurso, mientras que los artículos definidos indican que se comparte el referente. Nótese el contraste entre las siguientes dos oraciones (Fernández Meneses, comunicación personal).

A.- *Había una vez un rey en un gran castillo. El rey tenía un hijo...*

B. - **Había una vez el rey en un gran castillo. El rey tenía un hijo...*

Mientras que el ejemplo A es completamente gramatical, el ejemplo B no lo es, por el uso del artículo “el” en el primer enunciado (en lingüística se indica por convención mediante [*] las oraciones agramaticales).

Un grado más alto de información dada es expresado mediante el uso de pronombres como en el ejemplo siguiente.

C.- *Había una vez un rey en un gran castillo. Él tenía un hijo...*

En C se ha omitido el sustantivo “rey”, pero se sigue entendiendo que es de él de quién se habla; así se puede ver la relación de la información dada y nueva con el tópico descrita anteriormente. En otras lenguas también se emplean artículos definidos e indefinidos, así como pronombres para expresar el grado de novedad de la información (Grice & Baumann, 2007).

La distinción entre información dada y nueva también puede indicarse mediante la prosodia. Un énfasis prosódico suele indicar la información nueva. Los ejemplos siguientes (de Grice & Baumann, 2007) lo ilustran:

A.- *Thomas hat Hunger. Also isst er einen APFEL.* (Thomas tiene hambre así que come una manzana).

B.- (En un diálogo)

- *Hast Du gesagt, dass Thomas mit einem Apfel jongliert?* (¿Dijiste que Thomas está haciendo malabares con una manzana?)

- *Nein, er ISST einen Apfel* (No, está comiendo una manzana).

En el ejemplo A, la mención de la manzana es información nueva, y esto es marcado mediante un énfasis prosódico, mientras que en el ejemplo B, como ya ha sido mencionada, no recibe énfasis prosódico. La tendencia a marcar la novedad de la información de esta manera difiere dependiendo del lenguaje. Mientras que algunos lenguajes omiten el énfasis prosódico para la información dada, otros no lo hacen y otros emplean diferentes tipos de acento prosódico (por ejemplo un acento de tono alto para la información nueva, un acento que decae para información no completamente dada pero no del todo nueva, etc. Grice & Baumann, 2007). Es importante señalar que se requiere de una referencia o línea base prosódica para poder detectar un énfasis prosódico (Bryant & Fox Tree, 2005), de otra manera no se podría diferenciar (por ejemplo en una palabra aislada de un discurso) si dicha información está enfatizada prosódicamente y si es información nueva o dada; de hecho, en el concepto mismo de esta distinción se asume que es información que se comparte en el discurso, por lo que se requiere tener referencia al contexto.

Sobre las funciones de focalización y distinción de información dada/información nueva, Grice y Baumann (2007, pag 10) resaltan que aunque tienen cierta relación tanto en la forma de ser señaladas como conceptualmente, son dimensiones independientes.

3.3.1.3 Distinción léxica

La prosodia permite hacer distinciones léxicas de palabras; esto es, que una característica suprasegmental marque la diferencia entre dos palabras con significado diferente. Las características prosódicas que permiten hacer esto difieren entre lenguajes. Ejemplos de esto son el tono en lenguas como el chino

(/ma/ puede significar madre si es pronunciado con un tono alto o caballo si es pronunciado con un tono que baja y posteriormente sube, u otras cosas dependiendo del tono); el *stress* en lenguas como el inglés y el español (ejemplos *access*(acceso) vs *access*(acceder); *papa* vs *papá*), y la duración, por ejemplo en el japonés /koko/ (ここ) significa “aquí”, pero /ko:ko:/ (こうこう) significa preparatoria (los dos puntos indican que la vocal es larga). Esta función es muy conocida, por lo que sólo abordaremos un estudio al respecto de este tipo de contrastes, aplicado a la duración: el trabajo de Mildner (2004), que es particularmente interesante.

A grandes rasgos, la metodología que emplea es muy similar a la metodología de bisección temporal. A partir de dos duraciones de la palabra croata /pas/ que en su duración larga significa *cinturón* o *cintura* y en su duración corta significa *perro*, construyeron diferentes estímulos con duraciones intermedias; combinaron esta manipulación con una metodología de escucha dicótica. Los resultados fueron analizados en términos del porcentaje de respuestas que indicaban que “perro” era el significado de lo oído. Encuentran que el porcentaje de respuestas “perro” decrementa con el incremento de duración, resultado que no es sorprendente. Aunque no realizan análisis como punto de bisección y otros que se emplean en los experimentos de percepción de tiempo, resulta interesante preguntarse si se comportarían de la misma manera que en dichos experimentos, ya que es una instancia de la lógica general de los experimentos en dicha área con alta validez ecológica. Con todo, el objetivo de los autores, más que evaluar los efectos de la duración en la percepción es el de evaluar la lateralización de la función; concluyen que la función no está fuertemente lateralizada y que bien podría reflejar procesamiento subcortical o procesamiento diferenciado en los dos hemisferios de distintos atributos de la señal.

3.3.1.4 Actos de habla

Quizá una de las funciones más conocidas de la prosodia es la de la indicación de actos de habla; esto es señalar diferencias entre tipos de enunciados, donde lo más conocido es la distinción entre declaraciones y preguntas.

Sin embargo, estos no son los únicos tipos de enunciados. La prosodia distingue, por ejemplo, si un enunciado tiene la intención de solicitar información o un favor (pregunta/petición), dar una orden (imperativo), y otros. Grice y Baumann (2007) identifican entre los actos de habla a las declaraciones, peticiones, promesas y disculpas. Al respecto de las peticiones, es particularmente importante el uso de la prosodia en japonés; Kamiyama y Yamamoto (2007) identifican usos específicos de la prosodia que incluyen parámetros de cualidad de voz, sin los cuales el habla de extranjeros puede ser interpretada de manera descortés y tener consecuencias sociales; esto será retomado en el siguiente apartado.

La prosodia es particularmente importante para distinguir preguntas de declaraciones si no hay una sintaxis (como en inglés) o partícula específica (como en japonés) para identificar a los enunciados interrogativos (Grice & Baumann, 2007); el español es un ejemplo de este tipo de idiomas.

La prosodia puede presentar diferencias al interior de una categoría, por ejemplo, en inglés difiere la prosodia de preguntas de la *serie wh*¹¹ de las preguntas polares (donde la respuesta es “sí” o “no”). A la vez, la prosodia puede hacer que un enunciado que sea semántica y sintácticamente interpretado como perteneciente a cierta categoría sea asignado a otra categoría de actos de habla; por ejemplo, en alemán, una pregunta de *serie wh* puede ser interpretada como petición si se pronuncia con una caída de tono al final (Grice & Baumann, 2007).

Aunque normalmente se resalta la contribución del tono a la distinción de actos de habla (de hecho Grice & Baumann hablan de la influencia de la *entonación*, y no de la *prosodia*, en los actos de habla), es importante resaltar que también los parámetros temporales influyen fuertemente en esta función.

Gussenhoven (2002) señala que ha habido reportes de idiomas en los que un pico tonal más tardío se emplea para señalar preguntas, aunque también indica que esto ha sido interpretado como una influencia del *timing* sobre el tono, en la que

¹¹ Suelen ser llamadas en la literatura “*wh- questions*” y se refieren a las preguntas que incluyen una de las palabras: *where, why, when*, etc.

picos tonales más tardíos serían interpretados como más altos. Esto es, posibles efectos de interacción entre percepción de tono y percepción de tiempo; específicamente, un efecto en el que el tiempo específico en el que se detecta un cambio de tono influye en la magnitud percibida de dicho cambio.

Shah y colaboradores (2006) reportan distinciones prosódicas adecuadas entre enunciados declarativos e interrogativos en personas con daño en el hemisferio derecho.

3.3.1.5 Otras funciones pragmáticas y sociales de la prosodia: Toma de turnos, Actitudes y Señalamiento de relaciones interpersonales.

Finalmente, presentaremos algunas funciones de la prosodia relacionadas a la interacción social y a usos pragmáticos del lenguaje.

La prosodia permite señalar la toma de turnos y si las intervenciones han sido finalizadas o hay algo más por decir, lo que permite interacciones orales adecuadas y ordenadas; también permite señalar *intenciones* que incluyen diferencias sutiles de significado, por ejemplo, señalar ironía; esto también suele requerir referencias al contexto. Asimismo, se ha visto que la prosodia puede señalar relaciones interpersonales y que difiere dependiendo de la cercanía emocional entre los interlocutores, se aprecia cómo esta función se conecta con la función emocional. Presentaremos brevemente estas funciones y finalmente haremos una mención a las llamadas funciones emocionales o emotivas de la prosodia; estas últimas no serán elaboradas pues ya se ha dedicado un capítulo entero a su tratamiento.

Las funciones de toma de turnos y de índice de continuación de frase, o final de la misma, están muy relacionadas: indicar fin de frase es una pista para cambio de turnos e indicar que la frase está por continuar informa que el turno continúa. Al final de frases fonológicas independientes pero que forman parte de una sola frase sintáctica (por ejemplo listas del tipo “*se venden perros, gatos, ratones y ranas*”) se identifica una subida de tono al final de cada frase intermedia y una bajada de tono al final de la frase sintáctica (Grice & Baumann, 2007); estos hallazgos se refieren al alemán, pero desde la perspectiva de Gussenhoven (2002), podrían ser

válidos para muchas lenguas. Shah y colaboradores (2006) resaltan que, en inglés, se indica mediante “*tonos de continuación*” que la frase aún no ha sido finalizada, estos implican una caída y luego una subida de tono (una forma general similar a la reportada para el alemán). Chen (2007) indica que las realizaciones prosódicas en chino pueden variar dependiendo de la posición en la toma de turnos. Por ejemplo, un marcador de pregunta como los estudiados en su trabajo, va de un tono bajo a un tono alto si está al final de turno pero presenta el patrón contrario (alto a bajo) si se localiza a la mitad del turno y es precedido por otras frases; otros patrones prosódicos le conferirían al marcador una categoría diferente al de marcador de pregunta. El autor también identifica que si se dan pausas en los turnos se pueden presentar *muletillas de relleno* (*conversational fillers* en el original); frases sin relevancia en la conversación y de significado prácticamente vacío pero que siguen indicando que el turno no ha sido concluido.

Los usos pragmáticos y sociales de la prosodia incluyen varias funciones, como las actitudes, señalamiento de relaciones interpersonales, entre otros. Entre las funciones relacionadas a las actitudes identificamos muchas, como la ironía (Bryant & Fox Tree, 2005), certidumbre (Fitzpatrick, 2000; Gussenhoven, 2002; Hirschberg, 2002a), cortesía (Fitzpatrick, 2000; Gussenhoven, 2002; Hirschberg, 2002a), interés (Fitzpatrick, 2000), entre otras.

Algunas de ellas, como la ironía, requieren de referencia específica al contexto. Por ejemplo la frase siguiente: “¿*Lo estamos interrumpiendo, señor Jiménez?*” no puede ser identificada como irónica sin un contexto adecuado, por ejemplo, si esta frase es dicha por un profesor a un alumno durmiendo en clase es interpretada de manera distinta a si es preguntada por unos empleados a su jefe al entrar en su oficina. Aunque se ha propuesto que la prosodia puede indicar actitudes como la ironía por medios prosódicos, el contexto es fundamental, así como la interacción con el significado de las palabras. Bryant y Fox Tree (2005) concluyen que la idea de una entonación que señala ironía es sólo una ilusión perceptual. Como variable prosódica el ritmo de habla podría estar más relacionado con una interpretación de ironía y también los contrastes contra una línea base prosódica; consideran que si

bien otros parámetros acústicos como la nasalidad (variable perteneciente a cualidad de la voz, dada su capacidad de distinguir a un sonido de otro con el mismo tono y volumen, como fue mencionado en el apartado 3.2.2) podrían contribuir a señalarla, la influencia de otras fuentes de información, diferentes al canal acústico (variables contextuales principalmente) es tan importante que es difícil proponer características específicas de prosodia para señalar ironía.

Otros usos, como la cortesía, están mejor caracterizados. En japonés difiere la prosodia empleada para una petición que para una pregunta, aunque ambas tienen la forma sintáctica de una pregunta (Kamiyama & Yamamoto, 2007). El uso de prosodia estándar de pregunta al intentar hacer una petición puede llevar a que la frase sea considerada descortés e incluso ofensiva. En japonés las preguntas implican una subida de tono al final de la frase (normalmente en la partícula de pregunta “ka” (か)), pero al ser dichas con intención de petición la subida es muy ligera o prácticamente inexistente y también hay diferencias en el volumen. En diferentes idiomas, los tonos altos están relacionados con interpretaciones de cortesía (Gussenhoven, 2002), aunque no necesariamente es cierto para el japonés (Ito, citado en Kamiyama & Yamamoto, 2007), lo que permite resaltar diferencias culturales. Los autores señalan un ejemplo aún más interesante, en donde cierto tipo de frases donde además de los parámetros de tono y volumen, se emplea cualidad de la voz, frases con la intención “*kyoshuku*”, que intentan transmitir “una mezcla de sufrimiento, vergüenza y pena”. Como en otros artículos que señalan variables de cualidad de la voz, aquí no es especificado acústicamente y se asume que debe manifestarse en diferencias observables en diferentes bandas espectrales. La prosodia se usa para facilitar interacciones sociales, por ejemplo, dado que a partir de diferencias de utilización de variables prosódicas se pueden producir malinterpretaciones que, en el caso de peticiones, pueden llevar a la falta de cooperación del interlocutor.

Otro ejemplo de uso de la prosodia para funciones de interacción social es reportado en el trabajo de Chen (2007). El uso de pausas entre partículas (las mismas partículas para preguntas que fueron mencionadas en la sección de actos

de habla), puede generar interpretaciones de “*suavización del lenguaje*” y transmitir cortesía y solidaridad. Se aprecia la influencia de variables temporales de la prosodia (pausas) en significados sociales de la misma. Otros patrones prosódicos en las mismas partículas pueden señalar desacuerdo, acuerdo, negación, búsqueda de información o confirmación, énfasis, incitación del interés, muletillas de relleno, entre otras. Esto permite apreciar la centralidad de la prosodia (de hecho, junto con el contexto, es la única variable que influye pues las partículas siempre son las mismas) para especificar funciones sociales como solidaridad y desacuerdo, toma de turnos (mencionada en la sección anterior), actos de habla (reportado en la sección del mismo nombre del presente apartado), etc. Los autores resaltan que se requieren estratégicas prosódicas específicas en diferentes situaciones sociales. En conjunto con el trabajo de Kamiyama y Yamamoto (2007), estos estudios permiten enfatizar la importancia de la prosodia en las interacciones interpersonales, de manera general.

Campbell (2004) encontró diferencias en la prosodia dependiendo de diversos factores paralingüísticos. Analiza principalmente la cualidad de la voz y encuentra relación de los cambios en esta variable con cambios en las variables “interlocutor”, “acto de habla” y “cortesía”. En cuanto a la influencia del interlocutor, encuentra que tanto f_0 como la cualidad de la voz difieren dependiendo de si se habla con niños, familia, amigos, sí mismo y extraños; f_0 y cualidad de la voz se comportan como variables independientes. A la vez, la cualidad de la voz difiere dependiendo del miembro de la familia, reflejando aparentemente el grado de “cuidado” en el habla. Cabe mencionar que el estudio fue realizado a partir del análisis de los datos de una hablante japonesa grabada durante dos años; el orden específico de “cuidado” propuesto por los autores parece reflejar tendencias culturales.

El autor utiliza un enfoque informado por la perspectiva evolutiva. Considera que la voz podría haber sustituido en algunas ocasiones a la interacción cara a cara y por lo tanto es apta para comunicar estados relacionados al bienestar, relaciones sociales entre las personas (mayor o menor involucramiento), entre otras. Esto

refleja la integración a su visión del nivel emocional, de hecho hace mención específica a la prosodia emocional y como interactúa con la prosodia lingüística.

Recordemos que en la literatura lingüística, que suele dividir a la prosodia en funciones, también se ha tomado a la prosodia emocional como función; la función emocional del lenguaje está presente en la división de Jakobson descrita al principio de esta sección. La función emocional o emotiva de la prosodia no siempre coincide con la idea general de la prosodia emocional, ya que vista como función se suele entender como la capacidad de la prosodia para *transmitir* información emocional, y esto no asume que de hecho este ocurriendo un evento emocional (Caffi & Janney, 1994), por lo que usos estratégicos de la prosodia con fines de hacer creer al interlocutor que se experimenta un estado emocional también caben en esta perspectiva. Cuando incorpora este tipo de instancias, suele ser denominada *emotiva* en vez de emocional. Otros autores consideran a la función emocional/emotiva como algo paralingüístico, y que además tiene una representación continua y en gradiente en vez de categórica (Grice & Baumann, 2007). Cuando se identifica que la prosodia varía de manera categórica en vez de continua se dice que es un uso gramaticalizado (Gussenhoven, 2002); por ejemplo, la distinción de preguntas contra declaraciones en español y en otras lenguas se marca mediante la subida de tono al final del enunciado: si sube es pregunta, si no, no lo es, sin posiciones intermedias ni preguntas parciales. Con todo, hay variaciones en el grado en el que el tono sube, o el momento en el que lo hace, etc, estas son llamadas variaciones fonéticas (variaciones en el sonido que no hacen diferencia en el significado). En el caso de usos emocionales continuos, se pueden marcar grados, por ejemplo de enojo, y esto no suele tener impacto sobre variables como el tipo de enunciado que se expresa (declarativos, interrogativos, imperativos, etc.)

3.3.1.6 Sobre la perspectiva de funciones

Quizá una de las mejores formas de resumir la perspectiva de las funciones del lenguaje sea siguiendo la conceptualización de Kouassi (2008). Él considera que las unidades suprasegmentales o prosódicas constituyen un conjunto de hechos que se superponen a las estructuras segmentales para producir juntos la

significación. Caffi y Janney (1994), al mencionar la necesidad de tomar pasos inferenciales por parte del interlocutor, enfatizan que esto permite expresar *sobresentido*¹² en los enunciados.

Ha habido intentos de aplicación de este conocimiento, principalmente en el ámbito de los avances computacionales tanto para la síntesis como para el reconocimiento automático de voz (Campbell, 2004; Hirschberg, 2002a).

Hirschberg (2002a) menciona el caso de una evaluación de sistemas de síntesis de voz, donde con solo cambiar los parámetros prosódicos de un sistema para ajustarse a los producidos por humanos llevaba a que el sistema que había sido peor evaluado pasara a ser el sistema mejor evaluado. La misma autora comenta que la incorporación del conocimiento de características prosódicas conocidas, por ejemplo las asociadas a foco, permite una interacción más natural y efectiva con los sistemas y resalta que investigar los parámetros prosódicos que transmiten cosas como la cortesía y la certidumbre puede tener aplicaciones futuras.

Con todo, aún no están del todo bien caracterizadas todas las características de la prosodia para señalar diversas funciones, y algunas distinciones sutiles entre categorías pueden generar confusión. Grice y Baumann (2007) ilustran el ejemplo de una conocida política que, siguiendo los consejos de consultores sobre habla, bajaba su tono al hablar con el objetivo de sonar más confiable; sin embargo, su bajada de tono solía ser muy pronunciada y era confundida con una pista de finalización de turno, por lo que era interrumpida frecuentemente durante sus discursos.

Cabe destacar la interacción entre el nivel prosódico y el nivel segmental. Por ejemplo, el modificar la prosodia para señalar foco y fraseo tiene efectos en la duración de los segmentos, en qué tan marcados se pronuncian, etc (Grice & Baumann, 2007; Gutiérrez Bravo, 2008). De manera ilustrativa, si bien no hay diferencia fonológica (al menos en español) entre dos /p/ producidas con mayor o menor duración o con mayor o menor esfuerzo, sí hay diferencias fonéticas. Esto ilustra la interacción de los niveles segmental y suprasegmental; fenómeno que el

¹² *Surplus de sens* en el original.

cuerpo de investigación sobre acento extranjero ha abordado con más detenimiento y será abordado brevemente en el capítulo de temas complementarios.

3.3.2 Fisiología de la prosodia lingüística.

Hay diversas perspectivas acerca de la forma en la que está organizada la prosodia lingüística en el cerebro.

Entre las principales encontramos la hipótesis de lateralización derecha de la prosodia, en la que se propone que tanto las funciones emocionales como lingüísticas están lateralizadas al hemisferio derecho (Arciuli & Slowiaczek, 2007; Carlson, 2006a; Dromey et al., 2005; Mildner, 2004). Otra hipótesis, la de la lateralización funcional, propone que el involucramiento de los hemisferios varía dependiendo de la significancia de los estímulos prosódicos y la tarea; la prosodia lingüística sería procesada principalmente en el hemisferio izquierdo, mientras que la emocional sería procesada en el derecho (Arciuli & Slowiaczek, 2007; Dromey et al., 2005; Mildner, 2004; Shah et al., 2006). La hipótesis de dependencia de pistas perceptuales (*cue-dependent hypothesis*) propone que el involucramiento de los hemisferios varía dependiendo de atributos de la señal (Mildner, 2004; Shah et al., 2006): los parámetros temporales de la prosodia estarían más lateralizados al hemisferio izquierdo, mientras que los parámetros espectrales estarían lateralizados más hacia el hemisferio derecho (Arciuli & Slowiaczek, 2007; Gandour, Tong, Wong, Talavage, Dziedzic, Xu, Li & Lowe, 2004; Shah et al., 2006). Otra perspectiva, que bien podría ser llamada hipótesis de dominio temporal, considera que los estímulos de duración corta serían procesados principalmente en el hemisferio izquierdo (principalmente información segmental), mientras que aquellos de duración larga (suprasegmental, y por lo tanto, prosodia) serían procesados principalmente en el hemisferio derecho (Arciuli & Slowiaczek, 2007; Shah et al., 2006); esta teoría introduce la dificultad de especificar los límites exactos entre información de duración corta y larga; como será abordado en la sección de acento extranjero, los límites entre el dominio segmental y el suprasegmental no siempre son claros. Incluso se ha propuesto que la prosodia

no está lateralizada (Arciuli & Slowiaczek, 2007), siendo probablemente una función subcortical (Mildner, 2004).

De manera relacionada se ha propuesto que, en general, aspectos comunicativos del lenguaje como un todo (aspectos lingüísticos, paralingüísticos, extralingüísticos, pragmáticos, dependientes de contexto, prosodia, discurso) están más lateralizados al hemisferio derecho; mientras que las habilidades independientes de contexto (aspectos formales del lenguaje, gramática en general, fonología, sintaxis, morfología, etc.) están más lateralizadas hacia el hemisferio izquierdo (Fonseca, Wagner, Rinaldi, & Parente, 2007; Paradis, 1998).

Como se aprecia, algunas perspectivas hacen referencia al contraste con la prosodia emocional, otras, al contraste dependiendo de la duración de los estímulos y otras consideran la distribución de la prosodia en general (tanto emocional como lingüística). Esto no está guiado necesariamente por consideraciones teóricas sino por trabajo empírico que intenta descubrir las zonas en las que está representada esta función. El debate sobre cual es la hipótesis que mejor explica la lateralización de la prosodia está lejos de aclararse, pues diversos trabajos apoyan diversas hipótesis (Arciuli & Slowiaczek, 2007; Gandour et al., 2004; Mildner, 2006; Shah et al., 2006).

Al hablar de prosodia se habla de características suprasegmentales, pero estas varían en su duración. La unidad más pequeña en la que se puede brindar información prosódica es la sílaba (Gandour et al., 2004); por ejemplo en español e inglés existe el *stress*, en japonés vocales cortas y largas y en chino hay diferencias de tono para distinguir palabras. Sin embargo, en el término de prosodia y características suprasegmentales también caben variaciones de entonación a nivel de enunciados (como las que hemos visto para distinguir preguntas de declaraciones). Esto, aunado al hecho de que hay contribuciones diferentes a la prosodia: temporales, tonales, volumen, cualidad de la voz, etc., hace surgir la pregunta de si efectivamente es una buena conceptualización considerar a la prosodia como *todo* lo suprasegmental, o si diferenciaciones más

sutiles de las variables que la influyen, o de su dominio temporal (por ejemplo, sílabas, palabras, enunciados) permitirían un mejor estudio.

Siguiendo esta idea es importante remarcar que factores específicos del lenguaje influyen en la lateralización. El trabajo de Gandour y colaboradores (2004) y Gandour, Dzemidzic, Wong, Lowe, Tong, Hsieh, Sathamnuwong y Lurito (2003), ha puesto de relieve algunas de las dificultades de las teorías antes descritas, ya que mediante el estudio de lenguajes tonales han encontrado disociación de procesamiento de tono a nivel léxico (en el que la variación tonal se da al nivel de sílaba) contra tono a nivel oracional (por ejemplo para la distinción de actos de habla). El autor resalta que la organización cerebral de los lenguajes varía dependiendo de las características prosódicas específicas de los lenguajes y que los hallazgos de organización funcional dependen de variables de la tarea, de la modalidad en la que se presenta y del lenguaje. En el capítulo de temas complementarios se dedicará una sección a los hallazgos sobre lenguajes tonales y su contribución al estudio de la prosodia.

En contraste con la caracterización neurofuncional de la prosodia emocional que presentamos en el capítulo anterior, se observa que la organización de la prosodia lingüística presenta más debate. Lo más importante es resaltar que varias hipótesis consideran al hemisferio izquierdo como parte importante del procesamiento de prosodia lingüística¹³. Éste es un punto de partida para el contraste entre las dos prosodias, que será abordado en detalle en el siguiente capítulo.

La perspectiva de influencia de los factores específicos del lenguaje de Gandour (2004) y Gandour, Dzemidzic, Wong, Lowe, Tong, Hsieh, Sathamnuwong y Lurito, (2003) es particularmente importante, pues refleja influencia del aprendizaje, aunque no necesariamente de la cultura.

¹³ De los estudios fisiológicos que revisamos, de hecho ninguno apoya la perspectiva de la lateralización exclusiva hacia el hemisferio derecho.

En la siguiente sección abordaremos los trabajos interculturales en prosodia lingüística.

3.3.3 Estudios interculturales en prosodia lingüística.

Al abordar los estudios culturales en el capítulo sobre prosodia emocional resaltamos que son escasos. Pues bien, en el área de prosodia lingüística son aún más escasos. Esto no es algo sorprendente, ya que el cuerpo teórico del campo lingüístico no suele incorporar las perspectivas evolutivas, que llaman a la realización de estudios interculturales. Sin embargo, sí hay algunas teorías que incluyen nociones de tendencias universales, como la teoría de los tres códigos biológicos de Gussenhoven (2002) y la perspectiva de tipología entonacional de Fitzpatrick, (2000) que serán abordadas en la siguiente sección, pero aun así no son teorías con una perspectiva evolutiva clara pues incluyen muchos otros conceptos que las vuelven más complejas y no necesariamente llaman a los estudios interculturales. En contraste, en el área lingüística se asume la influencia del aprendizaje y de especificidades de los lenguajes; consecuentemente hay estudios extensivos sobre sistemas prosódicos de lenguas particulares y se han desarrollado sistemas de transcripción muy finos para capturar las características más importantes de diversos sistemas prosódicos, como el sistema británico (Grice & Baumann, 2007), el sistema de Índices de tonos y paradas (ToBI por sus siglas en inglés; The Ohio State University Department of Linguistics, 1999), el sistema INTSINT (International Transcription System for Intonation), que pretende ser un sistema adecuado para describir la entonación de diversas lenguas de manera análoga al Alfabeto Fonético Internacional (Fitzpatrick, 2000) y sistemas no muy conocidos pero que pueden capturar incluso características de cualidad de la voz como el desarrollado con fines didácticos por Kamiyama y Yamamoto (2007).

En esta sección abordaremos los estudios interculturales sobre prosodia lingüística, varios de los cuales han sido realizados en el marco de investigación para la enseñanza de lenguas y otros en el marco de investigación en fisiología.

Quizá el trabajo que mejor encaja en la perspectiva de estudios de prosodia lingüística intercultural es el de Shochi, Aubergé y Rilliard (2006). Los autores analizan la percepción de hablantes de francés con tres niveles de aprendizaje de japonés (sin conocimiento del idioma, principiantes e intermedios) sobre actitudes de japoneses expresadas prosódicamente.

Destacan que las actitudes y las emociones son a veces confundidas en la literatura, lo que en el caso de la prosodia conlleva dificultades. Ejemplificando con la sorpresa, consideran que se puede asumir como emoción si es parte de un proceso involuntario, pero como actitud si es expresada voluntariamente en el proceso de comunicación. Esto es consistente con las conceptualizaciones de “funciones emotivas/emocionales” que presentamos en la sección de funciones de la prosodia.

Desde su perspectiva, las actitudes son construidas socialmente por y para la lengua, por lo que pueden existir específicamente en una lengua y no existir o ser interpretados de manera diferente por hablantes de otra. Analizan doce actitudes japonesas que han sido estudiadas en la literatura: duda-incredulidad, evidencia, exclamación de sorpresa, autoridad, irritación, arrogancia-descortesía, sinceridad-cortesía, admiración, *kyoshuku*¹⁴, cortesía-simple, declaración, y pregunta-simple. Algunas de estas actitudes son consideradas particulares de la cultura japonesa, en especial las relacionadas a la cortesía y el *kyosuku* (cuyo contrario es “arrogancia-descortesía”). Los autores desarrollaron un corpus tratando de mantener estímulos semánticamente neutros donde sólo la prosodia permitiera identificar la actitud. La validación perceptual hecha por hablantes nativos de japonés demostró buenos índices de identificación en la mayoría de las actitudes analizadas y algunos patrones de confusión: admiración con cortesía simple; autoridad con evidencia; evidencia con arrogancia-descortesía; sinceridad-cortesía con *kyoshuku*, y sinceridad-cortesía con cortesía-simple. Los autores interpretan la

¹⁴ Esta actitud fue mencionada en la sección de funciones de la prosodia y definida en su momento como una característica transmitida mediante cualidad de la voz que transmite “una mezcla de sufrimiento, vergüenza y pena”. Normalmente se da en el contexto de peticiones o al tener un punto de vista contrario frente a alguien considerado socialmente superior.

mayoría de estos patrones de confusión en términos de la proximidad semántica de los términos en la cultura japonesa.

Encuentran que los participantes sin conocimiento de japonés pueden reconocer adecuadamente admiración, autoridad e irritación; las declaraciones son distinguidas de las preguntas sólo por el grupo de hablantes intermedios; y que los niveles extremos de cortesía japonesa son confundidos con descortesía por todos los grupos de hablantes japoneses. Los resultados se describen de manera un poco más específica en los siguientes párrafos.

Los participantes sin conocimiento de japonés confunden el “*kyoshuku*”, que implica un grado muy alto de cortesía con “irritación”, “arrogancia-descortesía” y “autoridad”; los autores consideran que esto se da en parte porque la actitud de *kyoshuku* no se da en francés y por que la cualidad de voz asociada al *kyoshuku* se utiliza en francés para expresar emociones negativas.

Los hablantes principiantes de japonés identifican adecuadamente “sinceridad-cortesía” y “duda-incredulidad”. Confunden, como el grupo de principiantes, preguntas de declaraciones, y también confunden “sinceridad-cortesía” con “declaración” y “arrogancia-descortesía” con “evidencia”. También confunden “duda-incredulidad” con “exclamación de sorpresa”. Aún confunden *kyoshuku* con “arrogancia-descortesía”, “irritación” y “autoridad”.

Para el nivel intermedio la distinción entre preguntas y declaraciones es posible. *Kyoshuku* aún es confundido con “irritación”, “arrogancia-descortesía” y “autoridad”. A la vez, “arrogancia-descortesía” es confundido con “declaración” y “evidencia”.

Kamiyama y Yamamoto (2007) investigan la capacidad de hablantes nativos de francés para producir patrones prosódicos adecuados en peticiones en japonés. A diferencia de una hablante nativa, los aprendices de japonés mantenían una prosodia con subida de tono al final, consistente con la producción de preguntas. Aunque las peticiones en japonés se pueden presentar en forma de pregunta, la prosodia varía en esos casos.

Los autores resaltan que para aprendices del idioma japonés que han asociado una subida de tono a las preguntas puede ser difícil aprender a asociar otro tipo de contorno tonal al mismo tipo de construcciones sintácticas. A su vez, consideran que patrones prosódicos del idioma de origen pueden interferir con las producciones en otros idiomas. Resaltan que producir una petición cortés sin una subida de tono puede parecer poco natural para los aprendices, considerando que en varias lenguas las subidas de tono están asociadas a la cortesía.

Reportan resultados preliminares de un programa de apoyo al aprendizaje de las características prosódicas. Desarrollaron un sistema de transcripción de las características prosódicas que permite calcular parámetros de tono, volumen e incluso cualidad de la voz (por ejemplo para indicar *kyoshuku*) y lo emplearon en conjunto con explicaciones de los parámetros prosódicos a los cuales se debe prestar atención. En ejercicios de repetición de las producciones del instructor se apoyaron del sistema de transcripción desarrollado. Las producciones de los estudiantes sonaron menos descorteses según juicios de los instructores. A su vez, los estudiantes comentaron de manera informal que el sistema de apoyo fue muy útil y que permitió aumentar la conciencia en la importancia de la prosodia, incluso en parámetros específicos de ella. Al ser sólo resultados preliminares no presentan análisis estadísticos, pero expresan la importancia de continuar las investigaciones en el área y sus posibilidades de aplicación a la enseñanza de idiomas.

Ward y Bayyari (2010) realizan un experimento intercultural sobre las pistas prosódicas que indican la expectativa de una *respuesta de involucramiento*¹⁵. Las *respuestas de involucramiento* se presentan en muchas lenguas; los autores citan el ejemplo del árabe, inglés y español: en todas estas lenguas, los contextos en los que se dan, sus funciones pragmáticas y semánticas así como su frecuencia (aproximadamente cada 12 segundos) es similar. Estas respuestas son importantes para la comunicación exitosa y para el desarrollo de *rappport*, a pesar

¹⁵ Originalmente *Back channels*, se refieren a respuestas cortas que se dan cuando el interlocutor tiene el turno en la conversación, por ejemplo *sí, ajá, ok, bien, etc.* de manera común muestran que se sigue escuchando.

de que no son muy sobresalientes, ya que se dan en muchas ocasiones sin ser conscientes de ello.

El punto central de su estudio es que a pesar de que la producción o no de *respuestas de involucramiento* depende en última instancia del escucha, hay algunos momentos en el diálogo donde son particularmente bien recibidas, y que estos momentos pueden ser indicados mediante pistas prosódicas por parte del hablante, pero las pistas prosódicas con las que se marcan estos momentos pueden diferir entre lenguajes.

Los autores evalúan la diferencia en la percepción de pistas prosódicas que marcan la expectativa de *respuestas de involucramiento* en árabe por parte de hablantes de árabe, hablantes de inglés y hablantes de inglés con un entrenamiento previo de 25 minutos sobre cómo actuar apropiadamente como interlocutor al oír el idioma árabe, haciendo énfasis en las *respuestas de involucramiento*. Los estímulos empleados carecían de información semántica, y sólo difería el final de las “frases” en cuanto a características prosódicas. Las características prosódicas que diferían tenían el objetivo de indicar continuación (por ejemplo en una lista donde se espera que algo más siga en la oración), fin de turno, o expectativa de *respuesta de involucramiento*.

Los autores reportan que la pista prosódica de expectativa de *respuesta de involucramiento* es percibida como tal por hablantes de árabe, pero no por hablantes de inglés. Los hablantes de inglés muestran confusión de esta pista prosódica con la indicación de afecto negativo, pero los hablantes de inglés con entrenamiento previo no.

Los autores enfatizan que esto demuestra que la interpretación de estas pistas es dependiente de la cultura. Resaltan la importancia que esto puede tener en contextos de conversación intercultural (personas de diferentes culturas conversando en un idioma común), ya que podrían existir malentendidos por el uso de códigos prosódicos no compartidos; por ejemplo, al interpretar afecto negativo donde no lo hay, y a la vez, no proveer de una *respuesta de*

involucramiento esperada por el interlocutor. Es importante el hallazgo de esta diferencia intercultural, ya que se ha propuesto que la mecánica de interacción en los diálogos muestra características universales o al menos características similares en muchos idiomas, incluyendo toma de turnos, saludos, entre otras. Los autores consideran que su trabajo tiene aplicaciones educativas (por ejemplo para entrenamiento específico), considerando que hablantes de lenguas extranjeras mantienen los patrones de toma de turnos de la lengua materna; así como en ámbitos sociales, por ejemplo para evitar malentendidos interculturales como la atribución consistente de afecto negativo a hablantes de árabe.

Los estudios de Meng et al. (2009) y Zhang et al. (2010) se interesan ambos en la prosodia intercultural en inglés y lenguas asiáticas. Ambos enfatizan la necesidad del entrenamiento prosódico específico en el aprendizaje de otras lenguas.

Meng y colaboradores (2009) resaltan que los aprendices de segundas lenguas son influidos por patrones perceptuales y articulatorios de sus lenguas maternas; proceso conocido como *transferencia*. Estos procesos de *transferencia* pueden tener consecuencias en la inteligibilidad de las producciones de extranjeros. La *transferencia* se puede dar a niveles segmentales y suprasegmentales, pero es en estos últimos donde se presentan efectos más fuertes en la inteligibilidad.

Consideran que aunque el entrenamiento prosódico específico en lenguas extranjeras es importante, es difícil, debido a la variabilidad prosódica al interior de las lenguas y la falta de consenso de la significancia comunicativa de algunas características prosódicas. Este artículo es un proyecto para la recolección de un corpus, por lo que no ofrece resultados.

Zhang y colaboradores (2010) evalúan la capacidad de hablantes nativos de chino en proceso de aprender inglés de percibir y entender adecuadamente características prosódicas del inglés: *stress* a nivel de palabra, *stress* a nivel de oración (foco), entonación y fraseo, y desambiguación prosódica. El énfasis de su estudio se centra en el análisis de las características prosódicas asociadas a énfasis (principalmente foco) y fraseo. Evalúan la capacidad de los estudiantes de inglés para identificar las características prosódicas asociadas a las funciones de

énfasis y fraseo a nivel de conocimiento y perceptual. La prueba de conocimiento no incluyó estímulos auditivos, intentaba medir si los estudiantes sabían *a priori* sobre las características prosódicas asociadas a las funciones; la prueba de percepción fue similar a la primera pero incluía estímulos auditivos sobre los que se hacían las preguntas. Los autores encuentran que el desempeño aumenta al hacer las pruebas perceptuales, en comparación con las pruebas de conocimiento; sin embargo no siempre llegan a un nivel de ejecución muy alto. Los aprendices tienen buen conocimiento de algunos patrones prosódicos (por ejemplo, que la mayoría de las preguntas conllevan una subida de tono al final), pero no de otros (por ejemplo que las preguntas de la serie wh- llevan una bajada de tono al final y que la continuación de turno se marca con una subida de tono), sin embargo el desempeño mejora en la prueba perceptual. Los autores resaltan la necesidad de incluir entrenamientos prosódicos específicos en la enseñanza de lenguas y consideran que su estudio puede ayudar a desarrollar tecnología de apoyo didáctico para estos temas.

Los trabajos de Gandour y colaboradores (2004) y Gandour, Dzemidzic, Wong, Lowe, Tong, Hsieh, Sattthamnuwong y Lurito (2003) también incluyen comparaciones interculturales. Estos trabajos reflejan un interés por el estudio de la organización cerebral de la prosodia lingüística, específicamente en cómo las características prosódicas propias de los idiomas influyen en los patrones neurofisiológicos. Para estudiar esta interrogante hacen comparaciones de hablantes de inglés contra hablantes de una lengua tonal: el chino.

Los hallazgos principales de sus estudios muestran un efecto cultural: algunas características prosódicas están asociadas a patrones de activación diferentes entre hablantes de una y otra lengua; a saber, el tono léxico (que sólo es una diferencia contrastiva en chino) muestra un patrón de activación que incluye actividad en el hemisferio izquierdo, mientras que en hablantes de inglés muestra un patrón de activación lateralizado al hemisferio derecho. Desde la perspectiva de los investigadores, los patrones neurobiológicos muestran ajustes a los “requerimientos prosódicos de las lenguas” (Gandour, Dzemidzic, Wong, Lowe,

Tong, Hsieh, Satthamnuwong y Lurito, 2003). Los autores resaltan que el patrón de activación en hablantes de inglés (predominantemente hacia el hemisferio derecho) refleja un procesamiento principalmente acústico, mientras que el procesamiento con involucramiento del hemisferio izquierdo en hablantes de chino refleja acceso a información lingüística; apoyan su interpretación resaltando que el procesamiento con involucramiento del hemisferio izquierdo no se da en hablantes de chino al ser expuestos a una lengua tonal diferente: tailandés (Gandour, Dziedzic, Wong, Lowe, Tong, Hsieh, Satthamnuwong y Lurito, 2003). Proponen que este tipo de patrones de respuesta podría reflejar que los sistemas de procesamiento perceptual más profundo (ej. lingüístico) se impondrían a sistemas de procesamiento perceptual más temprano (por ejemplo, percepción de tono). En otras palabras, la lateralización de los patrones de activación es influida por la experiencia con el lenguaje, que moldea las representaciones internas (prosódicas) asociadas a ciertas características acústicas externas (Gandour et al., 2004).

Para finalizar, se pueden presentar otros hallazgos como el hecho de que hablantes de un idioma pueden percibir pausas que no se encuentran en la señal acústica y que a hablantes de idiomas de patrones de *stress* muy predecibles, como el francés, donde normalmente está en la última sílaba (pero véase DePaolis et al, 2008), les es difícil percibir patrones más complejos de *stress* como los del inglés (Frazier et al., 2006).

Al respecto de la relativa escasez de trabajos interculturales, Fitzpatrick (2000) enfatiza que la descripción del sistema entonacional de una lengua dada es, por sí mismo, una tarea ardua; por lo que hacer comparaciones de los sistemas entonacionales de varios idiomas presenta grandes dificultades.

Los trabajos presentados en esta sección permiten resaltar (1) cómo la significación de algunos patrones prosódicos pueden ser reconocidos en lenguas ajenas; (2) que algunas características de prosodia lingüística pueden ser interpretados como reflejo de actividad emocional por personas que no comparten el código lingüístico; (3) que existen patrones prosódicos que se presentan en

algunas lenguas, pero a usentes en otras, y (4) que el idioma y su influencia en la representación de las características prosódicas relevantes influye en la actividad cerebral asociada a la percepción de la prosodia. Todo esto refleja la influencia de la cultura en la prosodia emocional, tema cuyo análisis es uno de los dos objetivos principales de este capítulo.

En la siguiente sección presentaremos a grandes rasgos las teorías que integran la perspectiva de influencias biológicas y culturales a la prosodia lingüística.

3.3.4 Teorías que integran la perspectiva de influencias biológicas y culturales a la prosodia lingüística.

Kouassi (2008) menciona la existencia de las hipótesis universalista y gramatical a la prosodia, y más específicamente a la entonación. Donde una propone que la entonación es derivada de códigos biológicos y la otra propone funciones de la prosodia; principalmente las asociadas a los actos de habla, influencia en la interpretación de los mensajes y transmisión de información asociada al diálogo y las actitudes.

Sin embargo son la perspectiva de la tipología entonacional de Fitzpatrick (2000) y la teoría de los tres códigos biológicos de Gussenhoven (2002) las que mejor permiten incorporar tanto las influencias biológicas como las culturales a la prosodia. A continuación se mencionan de manera sintética, pero serán desarrolladas en los siguientes apartados.

La perspectiva de la tipología entonacional de Fitzpatrick (2000) extiende a la prosodia la perspectiva de la tipología de los lenguajes, muy desarrollada en el campo de la lingüística, aunque con pocos trabajos en el caso de la prosodia, siguiendo a la autora. Es una perspectiva que se opone a una visión universalista pura.

La perspectiva de los tres códigos biológico de Gussenhoven (2002) es algo más compleja y propone una base biológicamente determinada que al desarrollar usos gramaticales se sujeta a los procesos de cambio lingüístico y a desviaciones del

código biológico original. Esta teoría fue desarrollada posteriormente por Hirschberg (2002b) al incorporar la perspectiva de implicaturas conversacionales de Grice (citado en Hirschberg, 2002b), una aproximación teórica también muy influyente dentro del campo lingüístico.

Presentaremos a continuación las características esenciales de estas dos teorías y de qué manera integran las contribuciones biológicas y culturales.

3.3.4.1 Teoría lingüística tipológica de la entonación.

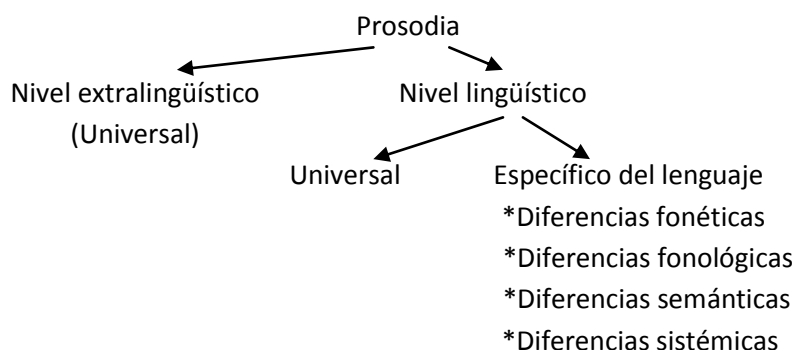
La teoría lingüística tipológica de la entonación se opone a una visión universalista de la prosodia, principalmente en el ámbito de la entonación. Fitzpatrick (2000) presenta a las teorías universalistas como aquellas que proponen que la entonación es extralingüística, esencialmente emocional, y, por tanto, universal. Resalta que estas teorías se basan en mecanismos físicos como los grupos de respiración y la presión subglotal, ya que existe una tendencia a presentar un tono más alto al principio de las frases, donde se dispone de más aire y una tendencia del tono a caer hacia el final, cuando el aire y la presión subglotal han disminuído; fenómeno conocido como *declinación* (DePaolis et al., 2008; Gussenhoven, 2002). A partir de esta teoría se han propuesto algunas “tendencias universales”:

- Asociación de frases completas, fin de frases y declarativos con tonos bajos o que bajan.
- Asociación de no completitud, indicación de que no se ha terminado y preguntas con tonos altos o que suben.
- Asociación de información nueva o sobresaliente con picos de tono locales y acentos tonales.
- Declinación en frases entonacionales
- Reinicialización de tono al inicio de frases entonacionales.

La autora considera que este tipo de enunciados son vagos y no logran capturar toda la variación a la que se refieren. También considera que hay muchos contraejemplos para estas “tendencias universales”. El mayor inconveniente de esta perspectiva es que, desde el punto de vista de la autora, no es falseable.

En contraste, la perspectiva tipológica considera que la entonación está estructurada lingüísticamente – en lugar de poseer exclusivamente características biológicas y universales – y que puede formarse una taxonomía de acuerdo a diversos niveles de la gramática.

La tipología de la entonación realiza primero una distinción entre los niveles extralingüístico (que Fitzpatrick asume de características universales) y lingüístico. A su vez, subdivide el nivel lingüístico en características universales y específicas de lenguajes. Dentro de las características específicas del lenguaje resalta las diferencias 1) fonéticas, 2) fonológicas, 3) semánticas y 4) sistémicas. Esto se ilustra en el siguiente diagrama.



Entre las funciones prosódicas que se consideran en el nivel extralingüístico universal se encuentran las emociones, el interés, la certidumbre, la cortesía, entre otras. Entre las funciones prosódicas universales en el nivel lingüístico se consideran tendencias no generalizables, como la tendencia en oraciones declarativas a bajar el tono; tendencia de las preguntas a tener un tono alto o que sube al final; y que la información nueva o más sobresaliente es marcada con acentos tonales; a diferencia de la perspectiva universalista, la tipología entonacional hace énfasis en que estas características no se presentan siempre.

Entre las características prosódicas de nivel lingüístico y específicas del lenguaje se encuentran diferentes posibilidades de realización de una misma *melodía lingüística*¹⁶ (nivel fonético); diferencias fonológicas en las características de los

¹⁶ Se refiere a los contornos asociados a funciones.

tonos (nivel fonológico); diferencias en el significado de *melodías fonológicas* entre diferentes lenguajes, por ejemplo, un contorno asociado a preguntas de la serie wh- en un idioma podría estar asociado a declarativos en otro (nivel semántico); diferencias en el número y función de las *melodías fonológicas*, por ejemplo, un idioma puede tener diferentes contornos para diferentes tipos de preguntas, mientras que otro emplea el mismo contorno para todos esos tipos (diferencias sistémicas).

Fitzpatrick (2000) enfatiza que el análisis de las características entonacionales requiere el análisis simultáneo de otros niveles del idioma, como la estructura de la información (descrita en la sección de funciones del presente capítulo), ya que la prosodia interactúa con diferentes niveles del lenguaje.

La idea principal de esta perspectiva es dividir las características prosódicas de las lenguas de acuerdo a la tipología propuesta.

Una de las principales objeciones que se podrían hacer sobre esta teoría tiene que ver precisamente con qué tanto valor teórico tiene, ya que presenta una tipología, no una teoría muy elaborada. Quizá lo realmente importante es que el desarrollo de la tipología se basa en la noción teórica de que la entonación está estructurada lingüísticamente, en contraste con la visión completamente universalista. Con todo, la taxonomía propuesta incluye un nivel universal paralingüístico y un nivel universal lingüístico, además de varios niveles dependientes de lenguaje.

La perspectiva de tipología entonacional presenta esencialmente un programa de investigación, ya que se argumenta un conocimiento pobre de los sistemas entonacionales de la mayoría de las lenguas incluso sin tomar en cuenta ni siquiera la *variación dialectal* (al interior de las mismas). La autora considera que estudios de lenguajes individuales y de comparación de lenguas pueden avanzar la clasificación de las mismas de acuerdo a esta taxonomía y que el uso de este sistema permitiría la incorporación del conocimiento sobre el origen y parentesco de lenguas. También permite la incorporación de otras teorías al hacer análisis

específicos de lenguas, como la *Teoría Jerárquica de la Prosodia* que se puede emplear al abordar fraseo.

La autora resalta que la gran cantidad de trabajo requerido sólo para describir el sistema entonacional de un idioma hace que trabajos comparativos de muchas lenguas sean virtualmente imposibles, al menos para el estado del arte del momento en el que publicó su trabajo (en el año 2000).

3.3.4.2 Teoría de los Tres Códigos Biológicos de Gussenhoven.

En contraste con la perspectiva de la sección anterior, Gussenhoven (2002) ofrece una teoría más elaborada; basada en la existencia de tres códigos de naturaleza biológica: código de frecuencia, código de esfuerzo y código de producción. Estos códigos y sus características pueden ser usados voluntariamente por los hablantes y llegar a ser gramaticalizados, pudiendo, así, someterse a las mismas reglas de cambio lingüístico de otros niveles del lenguaje, proceso por el cual pueden cambiar y reflejar patrones poco comunes. Más tarde Hirschberg (2002b) integra nociones de la perspectiva de implicaturas conversacionales. Con esta elaboración, el contexto cobra mayor importancia y permite explicar por qué en algunas ocasiones los códigos biológicos no cumplen las funciones propuestas sin que esto deba ser considerado como evidencia en contra de la teoría.

Describiremos a continuación las características principales de la teoría de los tres códigos biológicos y posteriormente la elaboración de implicaturas conversacionales.

Gussenhoven (2002) considera que su propuesta permite integrar la existencia de tendencias universales y específicas del lenguaje a la prosodia. Desde su perspectiva, las tendencias universales se expresan en la implementación fonotáctica mientras que las tendencias específicas del lenguaje son expresadas en la morfología y fonología entonacional. Propone que los significados universales están basados en características que reflejan condiciones biológicas que influyen el proceso de producción lingüística, como la F0. En el caso de los significados específicos de lenguaje, considera que son potencialmente arbitrarios,

aunque muchas veces reflejan significados similares a los paralingüísticos, es decir, significados similares a los propuestos por los códigos biológicos. Los tres códigos biológicos que propone el autor están asociados a variaciones acústicas y éstas a tendencias de interpretación. Las interpretaciones pueden estar asociadas a significados afectivos¹⁷ (reflejando atributos del hablante), o informacionales (reflejando atributos del mensaje). Todas las interpretaciones deberían ser accesibles a todos los humanos, pero los diferentes significados derivados de los códigos podrían ser incompatibles en algunas ocasiones y hablantes de diferentes lenguas podrían hacer diferentes elecciones entre los significados posibles. Además, en los lenguajes donde un significado posible de los códigos esté gramaticalizado pueden sesgarse las interpretaciones de sus hablantes, haciéndolos especialmente aptos para percibir los significados gramaticalizados sobre los otros posibles significados.

El código de frecuencia se basa en la asociación de laringes más pequeñas con tonos más agudos, dado que tienen cuerdas vocales más pequeñas y ligeras que pueden vibrar más rápidamente (este código fue propuesto originalmente en el trabajo de Ohala, citado en Gussenhoven, 2002). Se considera que la asociación del tamaño de la laringe y la velocidad de vibración de las cuerdas vocales se utiliza para expresar relaciones de poder. La asociación de vocalizaciones con tonos más graves por parte de animales más agresivos y de vocalizaciones con tonos más agudos por animales sumisos. Dicho patrón, entre otras asociaciones biológicas, sería reflejado por este código, y los humanos tendrían una tendencia evolutiva para captar estas diferencias. Entre los significados afectivos relacionados al código se encuentran las dimensiones de sumisión-dominancia; masculinidad-feminidad; así como la asociación de amabilidad y simpatía, cortesía, vulnerabilidad con tonos altos y de agresividad y confianza con tonos bajos. Entre los significados informacionales asociados al código se encuentran la certeza (poca certeza asociada a tonos altos y alta certeza a tonos bajos), lo que permitiría derivar la asociación de preguntas a tonos altos y de declaraciones a

¹⁷ Debe recordarse la discusión sobre el uso de este término en el campo lingüístico presentada al inicio del capítulo.

tonos bajos. Entre los usos gramaticalizados del código de frecuencia se encuentra la asociación de tonos que suben al final de las frases con preguntas, característica presente en más del 70% de los lenguajes del mundo; aunque hay lenguajes que hacen esta distinción al principio y no al final de las frases; también se ha observado de manera excepcional, lenguajes que marcan preguntas con caídas de tono y de manera aún menos común, asociación de subida de tono con declaraciones.

El código de esfuerzo se basa en que incrementos en el esfuerzo empleado en la producción del lenguaje, llevan a una mayor precisión articulatoria y a la vez a cambios más bruscos en el tono. Entre los significados afectivos del código de esfuerzo se incluyen la sorpresa y amabilidad, así como disposición servicial, de manera parecida al habla dirigida a los infantes, si se considera como un intento de ayuda para la comprensión de los mismos. También se puede asociar a niveles de transmisión de autoridad y credibilidad en lo que se dice, así como a una actitud de simpatía. Entre los significados informacionales asociados a este código se encuentran el énfasis, si se asume que el hablante emplea mayor esfuerzo dado que cree que el contenido de su mensaje es particularmente importante; también se puede asociar un rango tonal comprimido con intentos de ocultar información y negatividad. Los usos gramaticalizados del código de esfuerzo incluirían la expresión del foco, tanto contrastivo como informativo (descritos en la sección de funciones del presente capítulo) y servicialidad, al menos en holandés. En algunas lenguas en las que el foco es expresado por medios morfosintácticos podría no haber distinción en la magnitud de los movimientos de tono.

El código de producción sería un poco diferente de los otros dos códigos en el sentido de que sólo tendría significados informacionales. Este código se basa en el hecho de que al inicio de las frases la presión de aire subglotal es más alta que hacia el final de las frases, lo que conllevaría a un decremento gradual en intensidad y un incremento leve, también gradual en F0; fenómeno conocido como *declinación*. Esto permitiría asociar tonos altos con inicios de frase y tonos bajos con los finales de frases. Los significados informacionales de este código incluirían

la asociación de tonos altos al inicio de frases con inicio de nuevos tópicos y tonos bajos con la continuación de tópicos; la asociación de tonos altos al final de las frases con continuación del turno y de tonos bajos con fin de la frase y de turnos de conversación. Los usos gramaticalizados de este código reflejarían el mismo patrón de los significados informacionales, al menos en algunas lenguas.

El autor cita ejemplos de estudios que apoyan tanto los significados afectivos como informacionales de códigos, cuya revisión detallada rebasa los alcances del presente trabajo. También resalta que los códigos pueden interactuar entre ellos dando lugar a tonos diferentes con la misma forma general; por ejemplo una bajada de tono puede señalar una oración declarativa pero también fin de turno, pero los tonos que señalan declaraciones simples, fin de turno y declaraciones que se hacen a la vez que se finaliza el turno pueden variar.

En la teoría de Gussenhoven, si un significado de los códigos se gramaticaliza, se vuelve objeto de procesos de variación lingüística, lo que puede llevar a formas poco comunes en las lenguas (por ejemplo a señalar preguntas con tonos altos al principio y no al final de las oraciones), pero el proceso de variación lingüística nunca puede llevar a revertir por completo el significado del código biológico en el que se basa el uso gramaticalizado. De esto se deriva que en los casos en los que se observan tendencias poco naturales en las lenguas, estas deben ser necesariamente producto de un proceso de cambio lingüístico que opera sobre un uso gramaticalizado (que implica un significado sistematizado) y por lo tanto discretas (en oposición a continuas, ya que en un uso gramaticalizado se puede distinguir, por ejemplo, una pregunta de una declaración, pero no diferentes grados al interior de la categoría de preguntas).

La teoría de Gussenhoven no asume que las condiciones biológicas y fisiológicas en las que se basan los códigos deban estar presentes para que se puedan dar los usos de los códigos; se asume que el uso de los códigos es controlado por los hablantes y no un efecto automático. Se asume que los hablantes (y oyentes) de los lenguajes tienen un conocimiento implícito de los códigos y sus asociaciones, por lo que *hacen uso* de este conocimiento para poder aprovechar los códigos

para efectos comunicativos. A partir de esto, se proponen variables sustitutas a la variación de la frecuencia fundamental; estas variables sustitutas constituyen otro tipo de variaciones que pueden reflejar indirectamente los códigos. Por ejemplo, un pico tonal tardío (una variable temporal) podría sustituir un pico tonal pronunciado (una variable tonal). La descripción detallada de las variables sustitutas no será abordada en esta presentación breve de la teoría.

En contraste con la teoría de Fitzpatrick (2000), en esta teoría se nota una orientación más teórica y también un mayor grado de elaboración. La teoría de Gussenhoven (2002) integra desde sus cimientos la importancia de las influencias biológicas (los tres códigos) y culturales (principalmente los cambios lingüísticos que operan sobre los códigos gramaticalizados), y ofrece un mecanismo de integración entre ellos.

En la descripción de las características fundamentales de la teoría de Gussenhoven se nota que aunque integra una perspectiva biológica y universalista, no es tan similar al marco teórico que se utiliza en trabajos evolutivos que guían al trabajo intercultural en el área de prosodia emocional; un marco teórico como este no sugiere de inmediato la realización del mismo tipo de estudios interculturales que fueron revisados en el capítulo anterior, como fue resaltado en la sección de estudios interculturales del presente capítulo.

Posteriormente, Hirschberg (2002b) integra la noción de implicaturas conversacionales¹⁸ al trabajo de Gussenhoven. La autora considera que, dado que hay muchos contraejemplos de la perspectiva de la existencia de asociaciones específicas entre características prosódicas con significados particulares y que en el marco de Gussenhoven queda vaga la noción sobre cómo la experiencia de los aspectos físicos del habla se relaciona con el significado que se asume para las realizaciones de los hablantes, se requiere de un marco diferente para pensar en este fenómeno. Considera que integrar la dependencia de contexto a la asociación de prosodia y significado permite conceptualizarla de una manera más adecuada;

¹⁸ Noción que viene de la perspectiva de la Conversación Cooperativa de Grice (citado en Hirschberg, 2002b); muy importante en el campo de la lingüística.

la principal propuesta de la autora es que integrar la perspectiva de implicaturas conversacionales puede resolver estas dificultades.

Desde la perspectiva de implicaturas conversacionales se asume que tanto el hablante como el oyente comparten conocimiento sobre las características esperadas en la conducta comunicativa: las máximas conversacionales. Cuando el hablante hace caso omiso a una de estas conductas esperadas, se considera que lo hace con una intención comunicativa, lo que lleva al oyente a buscar un significado más allá del significado literal del mensaje; en estos casos se habla de implicaturas conversacionales. También las inferencias que pueden seguir a una frase, pero, en sentido estricto no son dichas, pueden ser consideradas instancias de implicaturas conversacionales (ej. “*Algunas personas salieron temprano*” implica que “no todas las personas salieron temprano”) y también están basadas en la suposición de que se conocen las *máximas conversacionales* y se ajusta a ellas. Las implicaturas siempre son dependientes de contexto.

Hirschberg (2002b) considera que los códigos biológicos de Gussenhoven pueden constituir la base de máximas conversacionales prosódicas. Por ejemplo, en el caso del código de frecuencia se podría proponer la *máxima de tono*, donde se espera que los tonos altos transmitan menor certidumbre que los tonos bajos; sin embargo en algunas ocasiones se espera que se viole esta máxima. Así, si bien en una pregunta estándar (*¿Lo interrumpimos, señor Ramírez?*) se puede asumir incertidumbre; en una pregunta formulada en un contexto determinado (por ejemplo en la que un profesor pregunta “*¿Lo interrumpimos, señor Ramírez?*” a un alumno distraído) no refleja incertidumbre, sino una implicatura conversacional (una petición de volver a poner atención, a pesar de que no es dicho de manera directa).

Desde la perspectiva de la autora, integrar esta visión dependiente de contexto permite comprender, por ejemplo, por qué si bien los tonos altos al principio de oraciones pueden señalar un nuevo tópico no siempre sucede así; razonamientos análogos podrían ser propuestos para otros significados propuestos por Gussenhoven.

La autora enfatiza que las máximas que ella propone sólo son ilustrativas de la utilidad del enfoque, no exhaustivas.

Una elaboración posterior que integra la teoría de la relevancia al estudio de la prosodia lingüística puede encontrarse en Wilson y Wharton (2006); será retomado en el próximo capítulo, dado que estos autores elaboran sobre la prosodia lingüística y emocional.

Finalmente, vale la pena retomar el comentario de Campbell (2004) sobre los modelos de la prosodia. Él considera que es importante integrar las funciones sociales a los modelos sobre prosodia (a lo que Hirschberg, 2002b, se adelanta al integrar la perspectiva de implicaturas conversacionales) y también la cualidad de la voz. Efectivamente, las teorías presentadas aquí parecen estar interesadas principalmente en características tonales, dejando un poco de lado otros parámetros prosódicos.

3.3.5 Resumen y discusión.

A lo largo del capítulo hemos resaltado algunas de las influencias biológicas y culturales en la prosodia lingüística. Si retomamos una perspectiva evolutiva podemos considerar como influencias biológicas a las características que han sido propuestas como “universales” por diversos autores. Las principales influencias biológicas y culturales que hemos visto a lo largo del capítulo son las siguientes:

Para las influencias biológicas encontramos principalmente el fenómeno de la *declinación* (DePaolis et al., 2008; Gussenhoven, 2002) así como la ralentización de los articuladores al final de las frases, que se refleja en la expansión de los segmentos al final de las frases y es más marcado entre más larga sea la frase; esto se da en muchos idiomas y se asume de base fisiológica (Grice & Baumann, 2007).

Ha habido diversas tendencias universales, que se pueden asumir de base biológica en un marco evolutivo: si son universales podrían ser heredadas. Sin embargo también reflejan susceptibilidad a ser moldeadas por el aprendizaje, por lo que no son universales categóricos. Entre estas encontramos la asociación de

tonos altos con preguntas, donde más del 70% de las lenguas presentan subida de tono al final de las frases asociada a preguntas (Fitzpatrick, 2000; Gussenhoven, 2002) y la asociación de tonos altos con cortesía (Kamiyama & Yamamoto, 2007); así como varias de las tendencias propuestas por los trabajos de Gussenhoven (2002) y Fitzpatrick (2000)

Ejemplos de excepciones a estas tendencias se reflejan en las observaciones de idiomas en los que preguntas pueden ser señaladas por tonos bajos (Gandour et al., 2004; Gussenhoven, 2002), y la observación de que los tonos altos no necesariamente señalan cortesía en el japonés (Kamiyama & Yamamoto, 2007). Esto refleja influencia del aprendizaje y probablemente de la cultura.

Entre las características universales de los idiomas que reflejan variaciones debidas al idioma se encuentran el fraseo (Frazier et al., 2006) y marcar el foco (Grice & Baumann, 2007). Estas funciones se presentan en todos los idiomas pero los medios por los que se consiguen varían.

Entre los hallazgos que reflejan un claro efecto del lenguaje y por tanto del aprendizaje y probablemente de la cultura, se encuentran la observación de la existencia de actitudes marcadas prosódicamente al interior de una lengua pero ausentes en otra (Kamiyama & Yamamoto, 2007; de hecho, la cualidad de la voz asociada al *kyoshuku* puede señalar aburrimiento en inglés británico, sospecha en inglés americano y sufrimiento, pena o dolor en ruso; Erickson, 2006); la influencia del lenguaje materno en las producciones (Kamiyama & Yamamoto, 2007; Meng et al., 2009; Zhang et al., 2010) así como en las percepciones (Shochi et al., 2006; Ward & Bayyari, 2010) que se dan en el contexto de segundas lenguas, también ha sido resaltada. Observaciones como la capacidad de percibir pausas donde acústicamente no existen y de que individuos hablantes de idiomas como el francés, con un patrón de *stress* más predecible que otros, tengan dificultades en percibir patrones más complejos como el del inglés (Frazier et al., 2006) también resaltan la influencia del aprendizaje y la cultura. Particularmente interesantes son las observaciones de Grice y Baumann (2007) de que distintas variedades de un lenguaje pueden diferir en cuanto al marcaje prosódico de funciones, como el foco

y el tópic; por ejemplo, la variedad de inglés hablada en la india no deacentúa palabras que otras variedades del inglés deacentúan por ser parte de la información conocida; efectivamente, la pronunciación, incluyendo la entonación es una de las características que cambia más rápido tanto a nivel regional (dando lugar a variación dialectal), como al nivel de lenguaje estándar (Grice & Baumann, 2007). Estos cambios podrían considerarse culturales.

Grice y Baumann (2007) resaltan que es a nivel paralingüístico donde diferentes lenguas parecen presentar más similitudes, pero que es precisamente en este nivel donde se pueden presentar malentendidos culturales. Después de todo, el significado, al no estar gramaticalizado, no está sistematizado y puede ser malinterpretado.

En resumen, se nota que el mismo patrón prosódico puede ser interpretado diferente por diferentes culturas (ver Sochi et al. 2006 y Ward & Bayyari, 2010), que el mismo significado puede ser producido con diferentes patrones prosódicos (ver Gussenhoven, 2002; Fitzpatrick, 2000); y que incluso el mismo patrón prosódico puede indicar diferentes cosas, dependiendo del contexto (ver (Hirschberg, 2002b).

Considero importante el avance en los estudios interculturales en prosodia lingüística. Si se asume una base biológica general para la prosodia lingüística (como propone la teoría de los tres códigos biológicos de Gussenhoven), es posible conceptualizar a la prosodia lingüística como un sistema más o menos universal con algunas características específicas dependiendo de los idiomas.

La idea de Gussenhoven (2002) de que algunos significados de los códigos se gramaticalizan cobra mayor importancia si se observan ejemplos de lenguas donde parámetros como la cualidad de la voz tiene un significado más o menos convencionalizado (*kyoshuku* japonés), y lenguas donde no es un significado convencionalizado (en francés señala diferentes formas de afecto negativo, pero ninguna en particular Shochi et. al., 2006). El grado de convencionalización del significado de parámetros prosódicos varía dependiendo de las lenguas, y,

siguiendo el marco de Gussenhoven (2002), los usos gramaticalizados son usos discretos, mientras que los usos no gramaticalizados presentan características continuas.

Como remarca Fitzpatrick (2000); parece que las teorías exclusivamente biológicas y universalistas van cediendo terreno a las teorías que integran también factores de aprendizaje y cultura. Estas últimas describen mejor el fenómeno de la prosodia lingüística pues permiten integrar contraejemplos en el mismo marco conceptual. Estas teorías han surgido del campo de investigación lingüística; probablemente las teorías de una ciencia que tiene como objeto de estudio al lenguaje tienden (y deben) ser mejores para explicar fenómenos que se dan en el habla y el lenguaje.

Aunque se ha resaltado que la prosodia a veces no “cabe” dentro del programa de investigación lingüístico, ya que en ocasiones es considerada un fenómeno paralingüístico (Kouassi, 2008), cada vez hay más autores que argumentan a favor del estudio de la prosodia y de su gran importancia; algunos incluso proponiéndola como la estructura fundamental del lenguaje: Kouassi (2008) resalta que la prosodia es un verdadero subsistema de la lengua que permite a los enunciadores transmitir un conjunto lingüísticamente estructurado de sentido, con lo que se opone a una visión en la que la prosodia es paralingüística; Meng et al. (2009) y Zhang et al., (2010) resaltan que si bien al aprender segundas lenguas la transferencia de la lengua materna tiene efectos tanto en las características segmentales como suprasegmentales, son los efectos sobre las segundas las que tienen un impacto más fuerte en la comprensión de las producciones de hablantes de lenguas extranjeras; Frazier y colaboradores (2006), discutiendo principalmente la función de fraseo, consideran que la prosodia proporciona el “esqueleto” para poder considerar a las unidades como tales y que permite representar en la memoria las unidades para su procesamiento, considera que la prosodia es central para resolver el problema perceptual de representar las relaciones de unas sílabas con otras; de manera análoga al proceso de unir atributos como la forma, el color y el movimiento de un objeto en la modalidad visual; considera que si se

acepta esta noción es más fácil comprender por qué la prosodia y la entonación interactúan con varios niveles de la gramática; de esta forma la prosodia podría ser pensada como la estructura esquelética de la que depende el resto de la frase (de hecho, en el proceso de planeación motora de la conducta vocal, es programada de manera más temprana como se verá en el capítulo de temas complementarios; Kelm y Beckman, 2009).

Efectivamente, se ha visto que la prosodia interactúa con diversos niveles del lenguaje como se ha resaltado al abordar la determinación prosódica de la modificación de la sintaxis al marcar el foco en español (Gutiérrez Bravo, 2008); función al respecto de la cual se ha señalado que para que pueda darse debe haber información semántica con la que interactúe la prosodia (Bryant & Fox Tree, 2005); por otra parte, en el apartado “distinción léxica” de la sección de funciones de la prosodia se abordó la influencia de características suprasegmentales en el significado, como los contrastes tonales en chino, el *stress* en español e inglés y la duración en japonés y croata (Arciuli & Slowiaczek, 2007; Gandour, Dziedzic, Wong, Lowe, Tong, Hsieh, Satthamnuwong y Lurito, 2003; Kouassi, 2008; Mildner, 2004). La influencia de las características suprasegmentales en la realización fonética de las estructuras segmentales es remarcada en el trabajo de Grice y Baumann (2007). Fitzpatrick (2000) resalta la interacción de la prosodia con varios niveles del lenguaje (como fonética, fonología, sintaxis y semántica), su relación con la estructura de la información, y también su relación con características extralingüísticas como emoción, interés y certidumbre.

Un tema interesante que no ha sido muy ampliamente estudiado es el de la posible superposición de significados prosódicos: significados composicionales que se darían dependiendo del contorno (Hirschberg, 2002a) y de posibles combinatorias de parámetros prosódicos asociados a diferentes significados; por ejemplo Grice y Baumann (2007) consideran que en algunas preguntas de la serie *wh* en alemán podría cambiar el contorno prosódico asociado (normalmente caída de tono al final de la frase) en los casos en los que se desea transmitir insistencia o cortesía de manera simultánea. La observación de Gussenhoven (2002) de que

los códigos (entendidos como los códigos biológicos de su teoría) pueden interactuar entre ellos, dando lugar a tonos diferentes con la misma forma general (como una bajada de tono que puede señalar una oración declarativa pero también fin de turno, pero que a la vez los tonos que señalan declaraciones simples, fin de turno, y declaraciones que se hacen al simultáneamente con finalización de turno pueden variar) se relaciona con la siguiente interrogante: ¿Se puede considerar que hay generatividad (que nuevos significados puedan surgir a partir de algunos primitivos) en significados prosódicos? De hecho, esta es una de las tesis del origen del lenguaje que maneja Hockett (1971), al respecto de la posible generatividad en los gritos de primates primitivos que anteriormente estarían asociados a significados discretos y mediante este proceso habrían permitido que se dieran significados composicionales más complejos.

Al principio del capítulo mencionamos que la división de la prosodia en el campo lingüístico contrastaba con la que se hacía en el campo de la psicología y fisiología; propusimos el enfoque de funciones para mostrar las posibles divisiones de la prosodia y vimos que desde la conceptualización lingüística la función emotiva/emocional se puede considerar al mismo nivel jerárquico de las demás funciones. También resaltamos que algunos de los principales intereses en el campo de la psicología – la prosodia emocional que proviene efectivamente de estados emocionales – es considerada como un fenómeno paralingüístico desde la lingüística; sin embargo, el proponer límites entre qué es lingüístico y qué es paralingüístico ha sido considerado como una empresa poco productiva, siendo en algunas ocasiones un problema irresoluble (Fitzpatrick, 2000).

El enfoque de funciones permite una división más selectiva de la prosodia por lo que apoyamos su uso, consideramos que en lugar de considerar un agregado emotivo/emocional, la mejor conceptualización sería, además de las funciones como fraseo, focus, etc. una función emocional (que proviene de emociones genuinas) y una función emotiva (que se emplea con intenciones comunicativas y en la que no es necesario que la emoción sea genuina), sin atender a la distinción entre lingüístico y paralingüístico. Siguiendo a Campbell (2004), consideramos que

también es importante tomar en cuenta a las funciones sociales de la prosodia, que también han sido objeto de discusión sobre su estatus lingüístico o paralingüístico. También Chen (2007) e indirectamente Kamiyama y Yamamoto (2007) hacen énfasis en el hecho de que las situaciones sociales requieren de estrategias y convenciones prosódicas específicas.

Sobre las variables que se toman en cuenta al estudiar a la prosodia, se nota que el énfasis ha sido puesto en las espectrales, al grado de que varios artículos se refieren más a la entonación que a la prosodia (Fitzpatrick, 2000; Grice & Baumann, 2007; Gussenhoven, 2002; Hirschberg, 2002b; aunque esto no es específico del campo lingüístico, ya que también varios trabajos del área de prosodia emocional hicieron énfasis en la entonación; Bänziger & Scherer, 2005; Grandjean et al., 2006; Ladd et al., 1985; Wildgruber et al., 2005). También las variables temporales de la prosodia son enfatizadas en algunos trabajos (por ejemplo Bryant & Fox Tree, 2005; Frazier et al., 2006). En contraste, las características de cualidad de la voz no han sido tan abordadas, incluso aunque se ha visto que pueden transmitir diversos significados y formar parte del complejo acústico que señala funciones (Campbell, 2004; Kamiyama & Yamamoto 2007; Shochi et al., 2006). Esta tendencia también fue observada en la literatura sobre prosodia emocional. Es probable que la relativa escasez de estudios que incluyen a la cualidad de la voz como variable importante sea porque la definición acústica de estas variables es un poco menos precisa que la de otras características acústicas (por ejemplo recordemos que la definición de cualidad de la voz para *kyoshoku* es referida como “una mezcla de sufrimiento, vergüenza y pena”). Con todo, vale la pena notar que trabajos como el de Campbell (2004) brindan algunas guías para el cálculo de parámetros relacionados con la cualidad de la voz a partir de mediciones acústicas, lo que puede constituir un avance hacia el uso de características de cualidad de la voz en futuros estudios. Sin embargo, también vale la pena recordar que no siempre coinciden las impresiones perceptuales con las características acústicas (se nota por ejemplo en la observación de que es posible percibir pausas aunque no se den en la señal acústica; Frazier et al., 2006).

En el capítulo anterior se formuló una pregunta sobre la posible independencia de la longitud de los estímulos con las características relevantes de la prosodia, ¿Si se normalizaran estímulos de diferentes longitudes que expresaran la misma emoción el contorno resultante sería equivalente? Esta interrogante sigue en pie a falta de estudios empíricos que la aborden, pero cabe resaltar que Grice y Baumann (2007) apoyan la idea de que la forma general de la entonación es más o menos igual para diferentes longitudes; al menos en lo que se refiere a prosodia lingüística.

Una posible área de estudio futuro es la comparación de resultados que provienen de distintos sistemas de transcripción entonacional, por ejemplo el ToBI, el sistema británico y el INTSINT. ¿Qué diferencias puede tener en los resultados y avances teóricos la utilización de marcos de transcripción particulares (incluyendo sus presupuestos teóricos)?

Por otra parte, una de las principales aplicaciones prácticas posibles del estudio de la prosodia lingüística se da en el marco de la enseñanza de segundas lenguas: incrementar la conciencia sobre la importancia de la prosodia en la comprensión de los significados, incluyendo significados pragmáticos y sociales, y sobre la amplia gama de cosas que la prosodia puede codificar, puede ayudar a desarrollar mejores programas de enseñanza de segundas lenguas, tanto para la producción como para la comprensión de las mismas, así como tecnología de apoyo educativo (Grice & Baumann, 2007; Kamiyama & Yamamoto 2007;.Meng et al. 2009; Zhang et al., 2010)

En el siguiente capítulo abordaremos estudios que comparen específicamente a la prosodia emocional y lingüística y algunos de los marcos teóricos alternativos a una distinción categórica de las prosodias.

3.4. Prosodia Lingüística y Prosodia Emocional.

Hasta el momento se han revisado trabajos que se concentran tanto en la prosodia emocional como en la prosodia lingüística, de manera focalizada. La caracterización de cada una de ellas, por separado, ha sido hecha de manera cuidadosa y se espera que ambas hayan sido presentadas de manera adecuada. Se han visto también algunas de las diferencias en el uso de términos, principalmente “afectivo” entre las áreas lingüística y psicológica; así como algunas de las diferencias de interés teórico de los cuerpos de investigación psicológica (principalmente dirigido a la prosodia emocional) y lingüística (principalmente dirigido a prosodia lingüística).

En el presente capítulo se abordarán trabajos dirigidos simultáneamente a la prosodia lingüística y emocional, ya sea con enfoques de comparación o de interacción; así como algunas elaboraciones teóricas que requieren, para su adecuada presentación, ser antecedidas por los capítulos centrados puntualmente en cada una de las prosodias y de los trabajos presentados en el presente capítulo. La literatura que se aborda en el presente capítulo ha explorado la producción, la percepción y la organización funcional de las prosodias.

A partir de este capítulo pretendemos adelantar la respuesta a la interrogante de si las dos prosodias pueden, justificablemente, ser consideradas como distintas entre sí; a partir de los estudios de organización funcional y de características acústicas. Conceptualmente, ha sido asumido que son diferentes, lo que queda ilustrado desde el uso de términos específicos para cada una de ellas así como, simplemente, a partir de la existencia de estudios de comparación entre ellas. La discusión sobre si una conceptualización de la prosodia como perteneciente a las categorías lingüística y emocional es justificable y adecuada iniciará en este capítulo y será concretada al final del presente trabajo de tesis, haciendo uso de los argumentos y evidencias planteados a lo largo de la totalidad de capítulos del mismo.

Efectivamente, son varios los autores que han señalado la capacidad de la prosodia para codificar tanto información lingüística como emocional (por ejemplo,

Grandjean & Scherer, 2006; Klouda et al., 1988; Lakshminarayanan, Ben Shalom, Van Wassenhove, Orbelo, Houde, & Poeppel, 2003; McRoberts, Studdert-Kennedy, & Shankweiler, 1995; Perlman Lorch, Borod, & Koff, 1998; Rymarczyk & Grabowska, 2007; Pihan, Tabert, Assuras, & Borod, 2008).

En el lenguaje natural ambas formas de prosodia se dan juntas. No es difícil imaginar una pregunta que a la vez transmite enojo o tristeza, o una oración con foco marcado prosódicamente que a la vez transmite alegría o asco. Las dos prosodias no están necesariamente separadas ni son mutuamente excluyentes como será ilustrado en la sección de elaboraciones teóricas del presente capítulo.

Varios autores, con argumentos tanto empíricos como teóricos, han apoyado la idea de cierta independencia de la prosodia con respecto a otros aspectos del lenguaje (por ejemplo, Pell, 1999b), por un lado, así como a la relativa independencia de la prosodia lingüística y emocional (por ejemplo, Joannette, Ansaldo, Kahlaoui, Côté, Abusamra, Ferreres & Roch-Lecours, 2008; McRoberts et al., 1995; Pihan et al., 2008) por otro. También hay autores que se oponen francamente a este tipo de distinción (Seddoh, 2002); así como algunos trabajos que proponen divisiones alternativas a la prosodia, por ejemplo, Van Lancker y colaboradores (2006), quienes proponen una división de la prosodia en las categorías emocional, actitudinal, pragmática, lingüística e *indéxica* (que refleja información de identidad y personalidad, por ejemplo). Al respecto de esta última función, otros autores han resaltado la capacidad de la voz (y de la prosodia) para señalar este tipo de información (por ejemplo, Van Santen, Tucker Prud'hommeaux, & Black, 2009), tema que será elaborado ligeramente en el capítulo de temas complementarios. La importancia de la prosodia en la regulación de funciones sociales (Pihan et al., 2008) – que ha sido descrita de manera superficial en la sección de funciones de la prosodia en el capítulo sobre prosodia lingüística – también ha sido resaltada. Estas dos últimas funciones de la prosodia ilustran su importancia en ámbitos que no caben claramente en la categoría emocional ni lingüística, lo que también llama la atención hacia la viabilidad de este contraste conceptual.

Comenzaremos el capítulo con la revisión de los trabajos sobre organización funcional de la prosodia; seguiremos con estudios acústicos y continuaremos con la sección de elaboraciones teóricas, donde revisaremos algunos enfoques importantes y daremos seguimiento a algunos de los temas que hemos presentado anteriormente. Finalizaremos el capítulo discutiendo los hallazgos empíricos y las elaboraciones teóricas planteadas hasta el momento.

3.4.1 Contrastes fisiológicos entre prosodia lingüística y emocional y posturas alternativas.

En los capítulos anteriores revisamos la organización funcional de la prosodia emocional y la prosodia lingüística, por separado. La síntesis de los hallazgos revisados apunta a un papel privilegiado del hemisferio derecho en el procesamiento tanto de la prosodia emocional como lingüística. En la caracterización de la prosodia lingüística, sin embargo, existe un poco más de controversia, ya que algunos autores han llamado la atención hacia la importancia del hemisferio izquierdo. Este contraste indica ya una diferencia en la organización cerebral de la prosodia lingüística, si bien también refleja similitudes. Pell (1999a) considera que un estudio detallado del hemisferio derecho, principalmente mediante la metodología de análisis de lesiones, es importante para poder caracterizar puntualmente los subsistemas que se relacionan al procesamiento de la prosodia; así como el estudio de la interacción de varias distinciones prosódicas (por ejemplo, emocionales y lingüísticas a la vez). También Pihan y colaboradores (2008) se interesan puntualmente en el estudio de efectos interactivos.

En esta sección abordaremos estudios que hacen contrastes puntuales sobre la prosodia lingüística y emocional y presentaremos algunas propuestas alternativas sobre las variables que explican mejor el involucramiento hemisférico.

Efectivamente, existe una postura que apoya la idea de que la prosodia lingüística y emocional, así como su organización cerebral, difieren (McRoberts et al., 1995; Snow, 2000; Van Lancker et al., 2006; Wildgruber, Ackermann, Kreifelts, & Ethofer, 2006). El origen histórico de esta distinción se puede trazar hasta 1915,

con la distinción entre los “modos” emocional y proposicional hecha por Jackson. Posteriormente fue propuesto el término *prosodia* (por Monrad-Krohn en 1947) y desde ese momento se suelen asumir las categorías emocional y lingüística (Snow, 2000).

Por otra parte, propuestas alternativas sobre la organización de la prosodia han sugerido que las variables que determinan el involucramiento hemisférico son atributos físicos (procesamiento diferencial para características temporales y espectrales de la prosodia en el en el hemisferio izquierdo y derecho, respectivamente; (Pell, 1999a; Van Lancker et al., 2006), o atributos temporales (eventos más cortos, incluyendo los fonemas, serían procesados preferencialmente en el hemisferio izquierdo mientras que eventos más largos, en especial los que afectan a las oraciones por completo, serían procesados en el hemisferio derecho; Van Lancker et al., 2006). Inclusive algunos autores consideran a la organización cerebral de la prosodia como mejor caracterizada en conjunto, sin atender ni a la distinción lingüística contra emocional (Snow, 2000). Desde esta perspectiva, la prosodia se procesa principalmente en el hemisferio derecho y luego es asociada a la información lingüística via cuerpo caloso (Lakshminarayanan et al., 2003)

Grandjean y Scherer (2006) presentan dos experimentos en los que exploran la posibilidad de aislar el proceso de decodificación de la prosodia emocional y lingüística a nivel neurológico

Uno de ellos involucró a participantes con daño cerebral del hemisferio izquierdo o derecho y participantes control; los participantes fueron expuestos a estímulos que combinaban distinciones prosódicas lingüísticas (declaraciones, interrogaciones e imperativos) y emocionales (alegría, enojo y tristeza) o que sólo hacían distinción en una de las dos dimensiones en frases de contenido léxico neutro. Los estímulos fueron presentados en su manera original o filtrados para remover el contenido léxico. Los autores encuentran dificultades en el reconocimiento de estímulos que distinguen ambas dimensiones a la vez y dificultades en el reconocimiento de la prosodia lingüística para personas con daño en el hemisferio

izquierdo, en comparación con participantes control. Los participantes con daño en el hemisferio derecho presentaron dificultades en el reconocimiento y discriminación de estímulos simples (sólo prosodia lingüística o emocional), comparados con participantes control.

En el segundo experimento emplearon electroencefalografía en participantes control expuestos a palabras con diferencias fonémicas (es decir, en un solo fonema; específicamente ocuparon las palabras del francés, *ballon*, *talon*, *vallon*, cuyo fonema inicial es /b/, /t/ y /v/ respectivamente) y manipulados para transmitir alegría, tristeza o modo neutral, así como de manera declarativa e interrogativa. Los participantes debían identificar la prosodia emocional, lingüística y las diferencias fonémicas (el sonido inicial) en bloques contrabalanceados. Concluyen que los patrones de actividad reflejan el involucramiento de diferentes redes para la prosodia lingüística y emocional. Dado que los mismos estímulos fueron empleados para las tareas fonémicas y las distinciones tanto emocionales como lingüísticas fueron transmitidas principalmente mediante f_0 , los autores concluyen que el patrón observado no puede ser explicado aduciendo a diferencias acústicas, es decir, el modelo de lateralización hemisférica por parámetros acústicos no se apoya. Consideran que es importante seguir haciendo estudios que aborden no sólo la comparación, sino la interacción de la prosodia lingüística y emocional, y que la metodología de manipulación acústica es adecuada para explorar este tipo de interrogantes.

Gandour, Wong, Dzemic, Lowe, Tong y Li (2003) exploran si el involucramiento de los hemisferios obedece a la significancia de los estímulos (emocional vs lingüístico) o a las características acústicas de los estímulos presentados a los sujetos. Trabajan con el idioma chino, ya que al ser un lenguaje tonal, permite hacer comparaciones entre el procesamiento de la prosodia emocional y lingüística que se dan a nivel de frase contra las distinciones prosódicas léxicas (abordadas en el capítulo de prosodia lingüística), que se dan al nivel de la sílaba; centrándose, por tanto, en las variables tonales transmitidas mediante la variación de f_0 de la prosodia. Mediante la metodología de resonancia magnética funcional y

la exposición de participantes hablantes de chino y de inglés a estímulos que marcaban simultáneamente distinciones lingüísticas (declarativo contra interrogativo) y emocionales (alegría, enojo y tristeza) – además de los contrastes tonales necesarios en el idioma chino – exploran qué variable determina el involucramiento hemisférico. Los autores concluyen que el hemisferio derecho desempeña un papel fundamental para el procesamiento de la prosodia emocional y que al menos algunos aspectos de la prosodia emocional son disociables de la prosodia lingüística. Encuentran diferencias entre grupos al comparar prosodia emocional contra lingüística: actividad en regiones anteriores y posteriores de la región prefrontal del hemisferio derecho, así como actividad en el hemisferio izquierdo sólo para el grupo de hablantes de chino; siguiendo a los autores, este patrón de resultados no puede ser explicado por las teorías de disociación por atributos (tono vs características temporales) ni por la teoría de la extensión temporal de la información prosódica, ya que tanto la prosodia emocional como lingüística se dan en extensiones temporales equivalentes, a nivel de frase. Desde su perspectiva, los resultados reflejan la influencia de efectos *top-down*, modulados por atención, y que dependen de la significancia de los estímulos prosódicos, es decir, de la distinción emocional y lingüística. Cabe subrayar que en este caso los estímulos prosódicos no eran puros, sino compuestos, y que lo que determinó las diferencias observadas fue la tarea: aunque los estímulos eran los mismos, se observó un patrón diferente al poner atención a los contrastes emocionales o lingüísticos. Así, se puede ver cómo lo que determina las diferencias de organización cerebral observada no es necesariamente una característica del estímulo (que la prosodia sea intrínsecamente emocional o lingüística y esto determine la actividad cerebral), sino la contribución del sujeto (la atención focalizada a una de las dimensiones).

Los trabajos de Rymarczyk y Grabowska (2007), Klouda y colaboradores (1988) y Van Lancker y colaboradores (2006) enfatizan la importancia de estructuras subcorticales para el procesamiento de la prosodia en general, las primeras autoras enfatizan el papel de los ganglios basales mientras que los segundos enfatizan el papel del cuerpo calloso (principalmente para la prosodia emocional;

recordemos que este tema se había tocado en el capítulo de prosodia emocional del presente trabajo, al presentar el trabajo de Ross y Monnot, 2008). El tercer equipo aborda principalmente la relación de los hallazgos de dificultades prosódicas subcorticales con los modelos de distribución funcional de la prosodia y será descrito hacia el final de esta sección.

Rymarczyk y Grabowska (2007) consideran que la red neuronal dedicada a la prosodia incluye componentes corticales y subcorticales; donde las áreas prefrontal, temporo-parietal y los ganglios basales serían las estructuras principales, aunque llama la atención a la necesidad de distinguir mejor las similitudes o diferencias entre las estructuras involucradas en el procesamiento de prosodia emocional o lingüística. Después de analizar el desempeño de personas con daño cerebral en diferentes zonas del hemisferio derecho, encuentra que aunque todos los participantes muestran dificultades prosódicas, éstas son más marcadas para la prosodia emocional. La autora considera que la organización de la prosodia emocional y lingüística podría involucrar áreas sobrepuestas, y que la prosodia lingüística, en comparación con la emocional – que es procesada principalmente en el hemisferio derecho – podría estar representada de manera más bilateral. La autora también encuentra un efecto de interacción del sexo de los pacientes y el sitio de la lesión: mujeres con daño en áreas frontales tienen mayores dificultades que hombres con daño en las mismas áreas, mientras que el patrón opuesto se observa para daños en áreas subcorticales; sin embargo, considera que los resultados podrían reflejar influencia de otras variables, por lo que se deben tomar con cautela. Los resultados de su trabajo, que demuestran diferente gravedad en las dificultades prosódicas emocionales y lingüísticas en el marco de lesiones cerebrales, son compatibles con la idea de diferente organización cerebral para las dos prosodias, aunque cabe destacar la propuesta de la superposición de áreas dedicadas a estas funciones; lo que reflejaría, sí, una organización distinta de las diferentes prosodias, pero no completamente independiente.

Klouda y colaboradores (1988), mediante un estudio de caso clínico de lesión en el cuerpo calloso, notan que el desempeño en f0 para la señalización de contrastes prosódicos mejora dependiendo del tiempo posterior a la lesión cerebral, mientras que los parámetros temporales se encontraron preservados en todas las evaluaciones. Los autores consideran que aunque el hemisferio derecho contribuye al control de f0 en la prosodia, esta función puede ser realizada por el hemisferio izquierdo con el tiempo posterior a la lesión. El autor considera que sus resultados son compatibles con la hipótesis de disociación de la prosodia por atributos físicos (tonales contra temporales) y cita otros trabajos que apoyan esta visión. Con todo, también encuentra más déficits para prosodia emocional que para lingüística; esto encaja bien en un modelo en el que el hemisferio derecho participa tanto para la prosodia emocional como lingüística y que su organización, aunque diferente, puede estar sobrepuesta.

Por su parte, Joannette y colaboradores (2008) reportan que las dificultades más características en pacientes con lesiones del hemisferio derecho son las que afectan al procesamiento de la prosodia emocional y prosodia lingüística de modalidad (distinción de actos de habla). También reportan casos de lesiones en el hemisferio derecho en que se producen alteraciones puras del procesamiento de la postura lingüística de modalidad, cita el caso de un paciente que expresaba todas las oraciones de forma declarativa. También observan que las alteraciones de pacientes con lesiones en el hemisferio izquierdo y derecho son cualitativamente diferentes, estando más asociadas a dificultades en el procesamiento de acento léxico. Los autores concluyen que en el procesamiento de la prosodia se observa una cooperación interhemisférica. Pell (1999a) también reporta casos de lesiones que pueden afectar independientemente ya sea a la prosodia lingüística o a la emocional. Hallazgos como estos, que resaltan dificultades asociadas a una, pero no a otra de las prosodias, son compatibles con la hipótesis de lateralización por significancia.

Pell (1999a,1999b) reporta los hallazgos de un estudio de participantes con daño cerebral en el hemisferio derecho y participantes control, con el fin de clarificar el

involucramiento del hemisferio derecho en la prosodia. Reporta de manera separada los hallazgos para f_0 (1999a) y para los parámetros temporales (1999b).

Para su estudio empleó frases que presentaban simultáneamente distinciones prosódicas emocionales y lingüísticas (*stress* enfático (similar a foco) y modalidad (con este término se refieren a lo que nosotros presentamos como actos de habla en el capítulo anterior)).

En cuanto a los resultados para f_0 , encuentra que aunque los participantes con daño cerebral podían producir frases similares a las normales en la forma general (acústicamente, por ejemplo), la variación en f_0 era menor; y sus producciones eran menos efectivas para transmitir énfasis y contrastes emocionales cuando dicha efectividad era evaluada a partir de valoraciones perceptuales por participantes control. Esto resalta la importancia de conducir experimentos basados en la metodología de valoración perceptual, además de análisis acústicos.

Aunque los patrones acústicos de distinción de declaraciones e interrogaciones se encontraba preservada en los individuos con daño cerebral, la capacidad de hacer esta distinción simultáneamente con distinciones emocionales entraña más dificultades: al producir frases que reflejasen tristeza, el aumento terminal de f_0 relacionado a las preguntas fue poco pronunciado; a su vez, este incremento terminal no se alcanzó al producirse en frases que reflejaban alegría. Este efecto de interacción ilustra la importancia de los estudios que exploran ambos tipos de distinciones simultáneamente, ya que son resultados que no podrían ser detectados en estudios dedicados exclusivamente a una de ellas.

Resumiendo sus resultados, Pell considera que la capacidad de producir distinciones categóricas (por ejemplo, énfasis y modalidad) podría estar mejor conservada que la de producir contrastes continuos. La distinción entre contrastes categóricos y continuos está íntimamente relacionada a la distinción entre usos gramáticos y naturales abordada en el capítulo anterior en el marco de la teoría de los tres códigos biológicos de Gussenhoven (2002). Más adelante se presentará

una discusión sobre la distinción entre aspectos categóricos y gradados en la prosodia, y su relación con la gramaticalización.

Pell también reporta restricciones en medidas de variabilidad de la prosodia de manera general, dato que considera podría relacionarse con las observaciones clínicas de *aplanamiento afectivo* del habla en personas con déficits prosódicos. Efectivamente, las distinciones emocionales son de naturaleza continua, por lo que siguiendo la postura de afectamiento diferencial para contrastes continuos y categóricos, no es un resultado incongruente.

Pell propone que los participantes con daño cerebral mantienen el conocimiento acerca de *cómo* señalar distinciones prosódicamente, pero que implementar este conocimiento es lo que entraña dificultades, de este modo, el origen de las dificultades prosódicas podría ser de tipo mecánico o fonético. Recordemos la postura de Ross (2008), presentada en el capítulo de prosodia emocional, en la que la información prosódica del hemisferio derecho se superpondría al programa motor del habla del hemisferio izquierdo, por vías subcorticales, modificándolo.

Su postura indica que la variable que distingue el involucramiento hemisférico es el contraste entre aspectos continuos y categóricos; aunque él mismo considera que se debe realizar más investigación para poder concluirlo. Esto va en contra de la postura en la que la organización funcional de la prosodia emocional y lingüística tengan diferente organización cerebral; siguiendo la perspectiva de Pell (1999a), esos resultados pueden deberse a que las prosodias tienden a tener diferente codificación: categórica o continua, pero que no es necesariamente la función emocional o lingüística la que determina el involucramiento hemisférico.

Por su parte, en cuanto a los hallazgos sobre las características temporales de la prosodia (Pell, 1999b), encuentra una capacidad de codificación generalmente preservada en individuos con daño cerebral. De manera análoga a los resultados para f_0 , encuentran pocas diferencias acústicas, aunque son más notorias al evaluarlas mediante juicios perceptuales de participantes control; encuentra también mayores dificultades para expresar distinciones gradadas en contraste

con las categóricas. En este trabajo refuerza el planteamiento de la preservación del conocimiento de cómo transmitir distinciones de manera prosódica, pero de una dificultad para implementarlo.

Merece particular atención el caso de la prosodia al nivel de palabra, explorada mediante el contraste de énfasis en ambos estudios; dada la buena conservación general de las características tanto temporales como tonales en individuos con daño cerebral. El autor considera que el hemisferio derecho podría tener un rol muy pequeño en este tipo de distinciones (Pell, 1999a, b). Esto apoya la hipótesis de diferente organización cerebral para las funciones emocionales y lingüísticas (al menos cuando nos centramos en la función de énfasis). La caracterización de la prosodia lingüística definitivamente requiere ser analizada más sutilmente, explorando las diferencias de organización cerebral para diferentes funciones lingüísticas: la organización de contrastes prosódicos léxicos (por ejemplo stress y tonos en lenguas tonales), foco (énfasis), actos de habla y funciones pragmáticas podría diferir en cierto grado.

Wildgruber y colaboradores (2006)¹⁹, conducen una serie de experimentos mediante la metodología de resonancia magnética funcional para explorar la organización de la prosodia emocional; entre ellos, un contraste crucial para lo que nos atañe en este capítulo: efectúan una comparación entre prosodia emocional y prosodia lingüística. Concluyen que tanto la prosodia emocional como la prosodia lingüística comparten áreas en el hemisferio derecho para el procesamiento acústico más temprano de los estímulos, pero que el juicio explícito de aspectos lingüísticos de la prosodia está relacionado a áreas del lenguaje en el hemisferio izquierdo, mientras que la evaluación explícita de prosodia emocional se relaciona a áreas orbitofrontales bilaterales. Este patrón de resultados apoyan simultáneamente la idea de una organización diferente para prosodia emocional y lingüística y la idea de redes neuronales sobrepuestas: los estadios iniciales serían compartidos mientras que estadios tardíos (y quizá más específicos) serían

¹⁹ También brindan un modelo de procesamiento de la prosodia emocional por pasos cuya revisión puntual no será hecha por cuestiones de extensión.

diferentes. Los autores llaman la atención hacia las diferencias de procesamiento *implícito* (regiones subcorticales que varían dependiendo de la emoción) contra las de procesamiento *explícito* (regiones orbitofrontales bilaterales) en el caso del procesamiento de la prosodia emocional. Aunque no exploran un contraste análogo para la prosodia lingüística, cabe preguntarse si tareas de evaluación y reconocimiento explícito varían en contraste con tareas pasivas; si bien una tarea pasiva podría reflejar sólo un procesamiento acústico, no necesariamente ocurriría así, en especial si el procesamiento del lenguaje es particularmente automático, recordemos que incluso han sido identificadas áreas compartidas en el procesamiento de lenguaje y de música (Levitin & Menon, 2003; Patel et al., 1998), que podrían reflejar un procesamiento lingüístico automático ante estímulos que no son, de hecho, lingüísticos. Esto será abordado en un capítulo posterior.

Lakshminarayanan y colaboradores (2003) apoyan la hipótesis de lateralización por atributos (tono contra parámetros temporales) al emplear una metodología acústica y no fisiológica (presentaremos este trabajo en la próxima sección).

Perlman Lorch y colaboradores (1998) describen que en afásicos de Broca parece haber disprosodia, dado el ritmo lento y las numerosas interrupciones en el habla (que de hecho son los signos clásicos de la afasia de Broca); sin embargo, encuentra patrones normales de caída y declinación de f_0 . Asimismo, observa que estos pacientes parecen compensar su falta de fluidez produciendo tonos de continuación (abordados en el capítulo de prosodia lingüística) fácilmente perceptibles. Esto contrasta con lo que se observa en personas con lesiones del hemisferio derecho; también observa que personas con daños en el hemisferio izquierdo ocupan la entonación como canal comunicativo de manera más frecuente que aquellos con daño en el hemisferio derecho. Parece que las habilidades emocionales preservadas en estos pacientes les ayudan a la comunicación. Sus observaciones parecen compatibles con el modelo de disociación por atributos: si el procesamiento de tono se realiza principalmente en el hemisferio derecho, intacto en estos pacientes, aún estaría disponible para fines comunicativos.

De manera relacionada, Snow (2000), apoya la idea de la especialización hemisférica por atributos, en oposición a la teoría de disociación por significancia; considera que es mejor tratar a la prosodia emocional y lingüística como una sola con bases neurales compartidas. Además, el autor propone un modelo en el que tanto la prosodia emocional como lingüística tienen una base emocional. Considera que información sobre actitudes (por ejemplo, estar impresionado, indiferencia, protesta, involucramiento, certidumbre, etc.), transmitida prosódicamente, constituye una clase de emociones similar a otras, definidas como afectivas (como alegría, tristeza, enojo). El autor cita algunos ejemplos en los que las actitudes podrían subyacer distinciones lingüísticas; uno de ellos es la marca de preguntas (en inglés, como en español, mediante una subida de tono al final de la oración), que tendría por base una actitud inquisitiva o de sorpresa²⁰, y relaciona estos patrones con emociones que subyacerían a las distinciones lingüísticas; por ejemplo, observaciones de que las vocalizaciones afectivas, en situaciones de incomodidad o arousal, están asociadas a tonos más altos. Considera que la prosodia transmite significados actitudinales, discursivos, y gramaticales; de los cuales, los dos últimos dos son propiamente lingüísticos por su relación con componentes de la gramática. Sin embargo, propone que estos no son independientes de los significados actitudinales (y por extensión, emocionales). El autor concluye que tanto la prosodia emocional como lingüística (ejemplificadas en las tres funciones que propone) son lingüísticas, dado que pueden transmitir significados convencionales y además son inherentemente emocionales. Su base cerebral sería principalmente en el hemisferio derecho. Se aprecia que esta proposición tiene ciertas similitudes con la teoría de los tres códigos biológicos de Gussenhoven (2002) presentada en el capítulo anterior, en el que tanto la prosodia emocional como lingüística derivan de códigos biológicamente determinados. Con todo, al interpretar la propuesta de Snow (2000) es importante recordar el debate sobre la diferencia del uso de términos en los campos lingüístico y psicológico, presentado en el capítulo anterior; asimismo, el autor mismo cita varias evidencias de diferencias de procesamiento cerebral

²⁰ Nuevamente, cabe remarcar que ocupa el término *actitud* y no *emoción* para referirse a la sorpresa.

para la prosodia lingüística y emocional, y argumentos en contra de ellas que permiten que los datos encajen en su propuesta, además de que en ocasiones sus argumentos parecen forzados, su propuesta está aislada. Aún así, es interesante y potencialmente útil para estimular el debate en el área.

Por su parte, Pihan y colaboradores (2008) realizan un estudio electroencefalográfico para explorar el procesamiento de prosodia lingüística en frases neutras o con contenido emocional. Utilizaron estímulos que transmitían a la vez patrones de tono de subida o bajada (asociados con interrogaciones y declaraciones, respectivamente, y hechos a partir de manipulación acústica a partir de la versión declarativa) y reflejando miedo, alegría o un estado neutro. Aunque en su estudio no se requirió una evaluación específica de la prosodia emocional, encuentran efectos de interferencia en la tarea de reconocimiento de prosodia lingüística. En los estímulos que reflejaban alegría y miedo, comparados con la prosodia emocional neutral fue más difícil discriminar un patrón de subida o bajada, lo que refleja un efecto de interacción. Los autores consideran que esto puede deberse a la alta variabilidad tonal que caracteriza a los estímulos que transmiten estas emociones, ya que esta variabilidad podría haber hecho más difícil identificar un patrón claro de subida o bajada de tono. Aunque los autores esperaban que el hemisferio derecho tuviera una actividad marcada, siguiendo la hipótesis de lateralización por atributos, los resultados no reflejaron este patrón; se observó un incremento de la actividad de forma bilateral para los estímulos que reflejaban alegría, pero esto no fue observado para la prosodia que reflejaba miedo ni estado neutro. Los autores interpretan estos resultados en términos de un procesamiento automático emocional, a pesar de que su tarea no lo requería, que interfiere con la percepción de la prosodia lingüística. Resaltan la importancia de investigar la organización cerebral de los posibles efectos interactivos de la prosodia lingüística y emocional (específicamente en la dimensión de variables tonales), y consideran que la metodología de síntesis acústica es adecuada para hacerlo.

Reportan otro efecto de interacción observado en estudios anteriores: en preguntas de la serie wh- (abordadas en el capítulo de prosodia lingüística; hallazgo para el idioma inglés) el patrón de tono esperado es de caída de tono, mientras que en preguntas dicotómicas (sí/no), el patrón esperado es de subida de tono. Ambos tipos de preguntas son percibidas como más corteses y agradables si son acompañadas por este patrón tonal que si son acompañadas por el patrón inverso; lo que sugiere, desde su perspectiva, que los contornos tonales pueden mediar a la vez dimensiones emotivas y pragmáticas. Observaciones similares son referidas por Grandjean y Scherer (2008). Nótese la capacidad de acomodar estas observaciones en la perspectiva extendida (que incluye la noción de implicaturas conversacionales) de la teoría de los tres códigos biológicos de Gussenhoven propuesta por Hirschberg (2002b).

Sus observaciones de organización cerebral no se oponen claramente a la hipótesis de lateralización por atributos, ya que el involucramiento diferencial de los hemisferios ha sido propuesto para diferentes emociones (Achuff, 2001). También Rymarczyk y Grabowska (2007) enfatizan que diferentes áreas del cerebro (haciendo énfasis en el hemisferio derecho) pueden relacionarse con la percepción de diferentes emociones. Esto tampoco se contrapone claramente a la perspectiva de lateralización por significancia, ya que no se hace la comparación con la actividad asociada a la tarea explícita de reconocimiento de prosodia emocional.

Van Lancker y su equipo (2006), reportan estudios de caso de dos individuos con dificultades prosódicas posteriores a daño subcortical. La paciente 1 mostraba una lesión en el globo pálido y putamen de forma bilateral mientras que el paciente 2 mostraba un daño en el putamen, globo pálido y cápsula interna derechos. En la paciente 1, el habla carecía de variación tonal y las expresiones conversacionales espontáneas se habían perdido; en el paciente 2, había una expresividad reducida y su estilo comunicativo era impulsivo; en ambos casos se tenían muestras del habla anteriores a las lesiones, por lo que los autores afirman que estos patrones de expresión pueden adjudicarse a las mismas. En contraste, las habilidades para

la comprensión de prosodia tanto lingüística como afectiva se encontraban preservadas en ambos pacientes. La paciente 1 tenía más dificultades para producir prosodia emocional al ser elicitada que al repetir un modelo, lo que había afectado sus relaciones sociales y ocupacionales; el paciente 2 podía producir frases con prosodia emocional tanto elicitadas como repetidas; sin embargo sus dificultades prosódicas habían afectado su comunicación en ámbitos personales y profesionales. Las dificultades de ambos pacientes no tenían orígenes motores ni estaban restringidas a la prosodia emocional. Los autores resaltan que estos casos (así como otros que reportan al revisar la literatura) no encuentran cabida ni en la hipótesis de lateralización por atributos ni en la hipótesis de lateralización por significancia y llaman la atención hacia el desarrollo de un modelo que integre las contribuciones subcorticales a la prosodia. A su vez, consideran que, dado el estado del arte, deben hacerse más estudios de caso; consideran que a veces las dificultades prosódicas pueden ser parte de un cuadro clínico más amplio y no reflejar un daño específico a la función, en otras palabras, la disprosodia podría ser un síntoma de un síndrome más amplio; por ejemplo, el daño a ganglios basales se ha asociado a abulia, disforia e irritabilidad; la disprosodia podría ser parte de este conjunto de síntomas. Consideran que estudios anteriores pueden ser reinterpretados como referidos a la prosodia, aunque originalmente no hayan hecho referencia explícita a ella, incluyendo literatura sobre motivación, emociones, control motor, enfermedades neurodegenerativas, entre otras, con lo que apoyan su propuesta de ver a la disprosodia como un síntoma dentro de un conjunto. Resaltan asociaciones de daño subcortical con deficiencias en el ritmo del habla y errores de percepción temporal del habla en el marco de la enfermedad de Parkinson, así como asociación de daño en el tálamo y cerebelo con anomalías en el *timing* del habla. También aluden a dificultades prosódicas asociadas al hemisferio izquierdo, conocidas como “melodía del habla alterada” de variedad apráxica. Enfatizan que los modelos de lateralización por atributos y por significancia no son mutuamente excluyentes y pueden coexistir, permitiendo hacer sentido a algunos hallazgos clínicos aparentemente inconsistentes. Asimismo, a partir de estudios de personas congénitamente

carentes de cuerpo calloso, destaca el hecho que ambos hemisferios son importantes para el correcto procesamiento prosódico. Finalmente, cabe destacar su propuesta de la importancia de tomar en cuenta si las lesiones afectan a una o varias funciones de la prosodia (el autor propone además de las funciones lingüística y emocional las funciones actitudinal, pragmática e indéxica; perspectiva que será abordada con más detalle posteriormente), y si afecta a uno o varios parámetros de la prosodia (tono, variables temporales, volumen y cualidad de la voz).

¿Podría ser el involucramiento del hemisferio derecho en el procesamiento general de las emociones lo que determina su importancia para la prosodia emocional? Es una posibilidad. Tanto Perlman Lorch y colaboradores (1988) como Seddoh (2002), llaman la atención a la clarificación de este punto, así como al hecho de descartar dificultades de procesamiento emocional general como el origen de las dificultades prosódicas emocionales. Con todo, cabe recordar que también se ha hecho énfasis en la importancia del hemisferio derecho para la prosodia lingüística. Se requeriría aceptar la postura de Snow (2000), de que la base de la prosodia tanto emocional como lingüística es emocional, para poder mantener un origen general de procesamiento emocional como la base de las dificultades prosódicas observadas tanto en la prosodia emocional como lingüística. Tanto Perlman Lorch (1988) como Joannette y colaboradores (2008) resaltan también la importancia del hemisferio derecho para la conducta social, así como aspectos pragmáticos y discursivos del lenguaje; al haber visto que al menos algunos de los usos de la prosodia son pragmáticos, resta la posibilidad de que el daño al hemisferio derecho, involucrado tanto en las emociones como en otros aspectos de comunicación, sea el origen de las dificultades prosódicas de ambos tipos observadas; es decir, las dificultades podrían no deberse a dificultades de procesamiento prosódico *per se*, sino a dificultades de procesamiento emocional general y de aspectos de sociales y de la comunicación. Esta posibilidad parece remota en vista de lo que hemos revisado a lo largo del presente trabajo, pero es importante mencionarla para presentar un panorama adecuado de las posturas sobre la organización cerebral de la prosodia.

A partir de los estudios reportados en esta sección se puede observar que la organización cerebral de la prosodia lingüística y emocional difiere (basándose en los artículos que apoyan la hipótesis de lateralización por significancia); y que la organización de ambos tipos de prosodia está, al menos en parte, sobrepuesto (Wildgruber et al., 2006). La prosodia emocional y lingüística están fuertemente relacionadas al hemisferio derecho, aunque la organización de la prosodia lingüística se muestra como más distribuida y también asociada al hemisferio izquierdo. Hay más acuerdo para la organización cerebral de la prosodia emocional que para la prosodia lingüística (Lakshminarayanan et al., 2003). Las estructuras subcorticales también son importantes en el procesamiento de la prosodia (Klouda et al., 1988; Rymarczyk y Grabowska 2007; Van Lancker et al., 2006) y, aunque no han sido asociadas a algún tipo de prosodia particular; hay posturas que consideran que debe trabajarse más en la caracterización de la organización funcional de la prosodia, en general (Van Lancker et al., 2006). Finalmente, algunas perspectivas se oponen a una organización cerebral distinta para las dos prosodias, argumentándolo incluso desde una base conceptual (Seddoh, 2002; Snow, 2000).

En la siguiente sección abordaremos trabajos de metodología acústica que abordan a las dos prosodias de manera simultánea.

3.4.2 Estudios Acústicos sobre Prosodia Lingüística y Prosodia Emocional

En esta sección abordaremos trabajos con metodología acústica. En la sección anterior ya hemos reportado algunos resultados acústicos, pero los hemos integrado a la revisión de trabajos de metodologías e intereses eminentemente fisiológicos. En contraste con los análisis fisiológicos, los trabajos acústicos que abordan a las dos prosodias de manera simultánea son más escasos

Uno de los contrastes más importantes entre las características acústicas de la prosodia emocional y lingüística es su representación a manera de tonos asociados a una parte de la oración (por ejemplo, subida de tono al final de la oración para señalar preguntas o acentos tonales para señalar énfasis asociados a una palabra particular) en la prosodia lingüística, mientras que las características

prosódicas asociadas a la prosodia emocional se representan en las oraciones por completo (de ahí la efectividad de medidas globales como promedio y variabilidad de las medidas acústicas para caracterizar a la prosodia emocional, tal como se revisó en el capítulo dedicado a tal prosodia).

También es importante recordar que, dado que la prosodia (lingüística o emocional) se da en el contexto del habla, siempre, por definición, los aspectos segmentales y suprasegmentales se dan de manera conjunta; influyéndose mutuamente (ver noción de tono intrínseco asociado a diferentes vocales en Harwell Rodríguez, 1998; pp. 11). En otras palabras, tan pronto como existe un significado emocional (e incluso un estado emocional neutro, pues se ha visto que las frases neutras también son identificables), los aspectos segmentales y suprasegmentales de la prosodia lingüística²¹ necesarios para transmitir ideas y significado se ven afectados.

Simplemente no hay forma de encontrar instancias de prosodia lingüística o emocional de manera aislada; la única forma de analizarlas por separado es con ayuda de enfoques teóricos y metodologías de análisis que pretenden eliminar la contribución de una de ellas. Ni siquiera los estudios de vocalizaciones emocionales como los de Sauter y colaboradores (2010) reflejan verdadera prosodia emocional, pues, al no darse en el marco del lenguaje no son, en sentido estricto, prosódicas; aunque su enfoque es útil si este tipo de vocalizaciones comparten similitudes acústicas con las emociones expresadas en el lenguaje. Esta dependencia forzosa del lenguaje de la prosodia emocional está en el centro de la crítica que Seddoh (2002) hace a la prosodia emocional, abordada en la sección de elaboraciones teóricas del presente capítulo.

Lakshminarayanan y su equipo (2003) investigaron el efecto de manipulaciones espectrales en la identificación de la prosodia. Emplearon estímulos para prosodia emocional (alegría, enojo y tristeza) y estímulos para prosodia lingüística (declaración, interrogación, marca de continuación) manipulados acústicamente

²¹ Incluso lo más *estándar* o *prototípico* como una declaración simple, pues se ha visto que tiene características acústicas específicas.

para mostrar sólo la información tonal sustituyendo la información segmental por un fonema fijo, filtro de paso de bajas frecuencias (mantiene información de f_0), filtro de paso de altas frecuencias (mantiene sólo un poco de f_0 e información de otras bandas espectrales), un análogo al original sintetizado por un generador de voz que mantiene la información tonal y segmental, además de los estímulos originales. Los participantes debían identificar la emoción o información de prosodia lingüística mediante un paradigma de decisión forzada.

Los autores encuentran un buen desempeño en todas las tareas, así como una diferencia ligera pero significativa, entre el desempeño de la condición de prosodia afectiva contra la de prosodia lingüística (proporción de respuestas correctas 0.87 y 0.83, respectivamente); los tiempos de reacción también fueron significativamente mejores para la tarea de prosodia afectiva. Las manipulaciones acústicas mostraron diferentes efectos dependiendo de la clase de los estímulos: la condición de sustitución de la información segmental por un fonema fijo tuvo efectos perjudiciales en el desempeño de las tres condiciones emocionales y en los estímulos de marca de continuación lingüística, dicho efecto fue menos marcado para los estímulos de declaraciones e interrogaciones. El filtro de paso de bajas frecuencias tuvo un efecto perjudicial sólo para la prosodia emocional de tristeza y sólo para la marca de continuación lingüística.

Los autores enfatizan que se puede obtener un buen desempeño en tareas de reconocimiento de prosodia lingüística en estímulos muy empobrecidos acústicamente, de acuerdo a estudios anteriores, y extienden esta capacidad para la prosodia emocional. Consideran que aunque tradicionalmente se ha visto a f_0 como la pista perceptual más crucial para la prosodia (lo refieren como “hipótesis de f_0 ”), no basta para explicar sus resultados, por lo que se requiere tomar en cuenta y dar importancia a otras variables de la prosodia. Dado que la condición con el peor desempeño fue en la que se modificó la información segmental por un fonema fijo /æ/ (una vocal), los segmentos sordos (sin vibración de cuerdas vocales) fueron presentadas como sonoros (con vibración de las mismas), por lo que la información sobre la silabificación se perdió en esta condición. Desde la

perspectiva de los autores, esto puede ser el origen del pobre desempeño en la condición, aunque no se hace un análisis posterior específico. Enfatizan que en el caso de las condiciones afectivas, la mayoría de los errores se encuentran en la condición de fonema fijo, mientras que en el caso de la prosodia lingüística es más distribuida; siguiendo su argumento anterior, la estructura silábica parecería ser más importante en el caso del procesamiento de la prosodia afectiva. Consideran importante la realización de estudios futuros sobre la estructura silábica y su influencia en los dos tipos de prosodia. Este trabajo refleja, desde una perspectiva acústica, diferencias en el desempeño de prosodia lingüística y emocional en las mismas condiciones, resultados que están de acuerdo con una perspectiva en la que ambas difieren.

Gobl y Ní Chasaide (2003) se interesan por el papel de la cualidad de la voz en la comunicación de las emociones, estados de ánimo y actitudes. Para ello emplearon voz sintetizada con siete diferentes cualidades de voz y analizaron las respuestas de los participantes mediante diferenciales semánticos para los diferentes atributos afectivos (recordar las diferencias del uso del término “*afectivo*” en diferentes áreas discutidas en el capítulo de prosodia lingüística). Para evitar problemas de traducción, empleamos los términos originales de las cualidades de voz empleadas por los autores: *harsh voice*, *tense voice*, *modal voice*, *breathy voice*, *whispery voice*, *creaky voice* y *lax-creaky voice*. Cabe destacar que aún sin tomar en cuenta dificultades de traducción, el uso de términos para referirse a las cualidades de voz no está muy estandarizado y, algunas veces, las descripciones son impresionistas (recordar la definición de *kyoshuku* presentada en el capítulo de prosodia lingüística); el equipo de Gobl ha hecho esfuerzos por la caracterización acústica de las cualidades de voz y por el uso de una terminología basada en el trabajo de Laver (citado en Gobl & Ní Chasaide, 2003), que caracteriza las cualidades de voz en términos de la tensión muscular a diferentes niveles del aparato fonador.

El autor considera que la asociación entre los diferentes atributos afectivos y la cualidad de la voz no es biunívoca; en contraste, las cualidades de voz están asociadas a diferentes posibilidades en el plano afectivo.

De las dimensiones que explora el autor, la que es más claramente diferenciada es *relajado/estresado*. Algunos polos recibieron altas puntuaciones, como *enojado, aburrido, íntimo, contento, formal y confiado*; mientras que otras rara vez fueron percibidas, como *miedoso, amigable, alegre, triste y sin miedo*. Los autores consideran que, desde una perspectiva más amplia, los estímulos para *tense voice* y *harsh voice* se relacionan con términos que reflejan alto *arousal* o alto control sobre las emociones (dominancia). Mientras que *breathy voice, whispery voice, creaky voice* y *lax-creaky voice* se relacionan con el otro polo. La cualidad de *modal voice* (conceptualmente la voz más común y que se emplea en la mayoría de las situaciones) no reflejó neutralidad en su estudio, comportándose de manera similar a *tense voice*, obteniendo resultados en dirección de *confiado, formal y estresado*, aunque no de manera muy marcada. La cualidad de *whispery voice* obtuvo resultados fuertes para *tímido* y relativamente altos, pero sometidos a gran variabilidad entre sujetos, para *miedoso*.

Dado que las cualidades de voz reflejan asociación a diferentes tipos de atributos, relacionados entre sí, aunque no necesariamente, los autores lo discuten en términos de la posible señalización de *arousal*, aunque resaltan que otros autores han considerado su participación en la señalización de valencia. También consideran que la cualidad de la voz puede relacionarse con la comunicación de diferencias sutiles de estados afectivos medios o leves, que no implican necesariamente un componente fisiológico (estados de ánimo y actitudes), mientras que otras variables acústicas podrían relacionarse a la señalización de estados más intensos (como emociones).

Se puede suponer que si bien el papel de la cualidad de la voz tiene gran importancia para la prosodia emocional y, aparentemente en menor medida en la prosodia lingüística (como presentamos en los capítulos anteriores), la

diferenciación hacia una interpretación específica depende de la interacción con otras variables prosódicas.

Aunque los resultados no reflejan un panorama muy claro, cabe destacar el esfuerzo por clarificar el papel de la cualidad de la voz en la prosodia, así como las propuestas para la unificación de uso de términos con base en sus correlatos de producción y el esfuerzo por la descripción acústica. El autor también plantea una serie de dificultades conceptuales y metodológicas – principalmente con respecto a la medición de la cualidad de la voz – cuya presentación puntual rebasa los alcances del presente trabajo.

McRoberts y colaboradores (1995) examinan el rol de f_0 en el señalamiento de énfasis lingüístico y emocional, exaltan efectos de interacción y evidencias sobre una disociación en los dos procesos. El autor identifica un patrón que se da en las preguntas dicotómicas (sí/no) en inglés: las *relaciones de intercambio en entonación* (*Trading relations in intonation* en el original), que consisten en la observación de que al incrementar f_0 para transmitir énfasis en una palabra al interior de una pregunta, el aumento de f_0 al final de la oración para señalar que es interrogativa es menor. El autor se interesó en investigar si esta relación también se da al modificar f_0 para una función diferente que el énfasis lingüístico, para lo que emplea prosodia emocional; sin embargo no encuentra la *relación de intercambio*. A partir de esta observación, sugiere que el mecanismo para el incremento de f_0 al transmitir énfasis difiere del mecanismo usado para transmitir afecto, y que, en general, el mecanismo de control de f_0 difiere en la prosodia lingüística y emocional. El autor resalta que se puede percibir prosodia afectiva y emocional en estímulos cortos (3 sílabas). Considera también que el hecho de que se pueda percibir una pregunta aún cuando el aumento de tono al final de la misma sea muy pequeño, o nulo, refleja un ajuste del sistema perceptual a los patrones de articulación naturales (en tanto que el autor considera que las *relaciones de intercambio* se dan automáticamente por limitaciones fisiológicas del habla).

El autor emplea estímulos codificados por actores; considera que esta metodología es mejor que el uso de modificaciones acústicas ya que las modificaciones acústicas, al ser manipuladas de manera intencional y libre, no reflejan necesariamente las características, tendencias y limitaciones naturales de producción humana (por ejemplo, las *relaciones de intercambio*, diferencias en la duración de las frases, etc. además de la pérdida de algunas características de variabilidad propias del habla humana), lo que lleva a la construcción de estímulos artificiales y de validez ecológica disminuída. Discutiremos brevemente esta propuesta más adelante.

Varios de estos trabajos reflejan el interés por las metodologías y aplicaciones computacionales (Gobl & Ní Chasaide, 2003; Grandjean & Scherer, 2006; Van Santen et al., 2009), por ejemplo a la síntesis de habla, y en su posible aplicabilidad para reflejar actitudes en sistemas de síntesis, por ejemplo cortesía y otros usos, así como diagnóstico de desórdenes del lenguaje y enseñanza de idiomas. El trabajo del equipo de Van Santen y colaboradores (2009) constituye una muestra de este tipo de esfuerzos. En dicho trabajo, ofrecen una metodología para la evaluación automatizada de la prosodia con miras clínicas. El autor considera que el desarrollo de este tipo de metodologías es importante para no confiar únicamente en juicios perceptuales de expertos para evaluar la prosodia y para la identificación de algunos marcadores tempranos de dificultades clínicas que no sean audibles o estén asociadas a variables demasiado complejas que planteen mucha dificultad computacional para un humano.

En esta sección también se ha revisado evidencia que apoya la perspectiva que sostiene que la prosodia lingüística y emocional difieren. Se ha notado el énfasis que pone la investigación acústica en f_0 (Lakshminarayanan et al., 2003; Pell, 1999b) y en menor medida en las variables temporales (Pell 1999b), y se ha señalado la necesidad de tomar en cuenta otras variables como la cualidad de la voz (Gobl & Ní Chasaide, 2003). En futuros estudios se deberían incluir otras variables además del análisis de los efectos de interacción entre ellas.

La revisión principal de estudios empíricos finaliza en este punto; como conclusión provisional, la visión en la que la prosodia lingüística y emocional difieren se ve apoyada tanto acústica como fisiológicamente. En la siguiente sección revisaremos brevemente algunas consideraciones teóricas importantes sobre la prosodia lingüística y emocional y daremos paso al cierre del presente capítulo.

3.4.3. Elaboraciones Teóricas.

En esta sección presentaremos de manera sintética algunas propuestas teóricas importantes sobre la prosodia lingüística y emocional (para una visión más amplia se recomienda consultar Grandjean & Scherer, 2006; Grice & Baumann 2007; Gussenhoven, 1999; Pell, 1999; Scherer & Bänziger, 2004; Seddoh, 2002; Snow, 2000; Van Lancker et al., 2006; Wilson & Wharton, 2006).

3.4.3.1 La perspectiva de Van Lancker

Van Lancker y colaboradores (2006), en contraste con la división entre prosodia lingüística y prosodia emocional plantean una división en 5 funciones: Afectiva o emocional, actitudinal, pragmática, lingüística (que incluye gramática y semántica) e indéxica (que incluye información sobre la identidad, personalidad, edad, entre otras). Apoya su división de funciones con diversos estudios enfocados específicamente en ellas. En la función indéxica se pueden acomodar los estudios sobre acento extranjero y síndrome del acento extranjero que serán presentados en el próximo capítulo.

A su vez, el autor señala que las principales variables que transmiten la información prosódica son la variación del tono, volumen, variables duracionales (ej. sílaba, palabra, frase, longitud de las frases prosódicas, pausas, tempo, ritmo) y cambios en la cualidad de la voz.

También enfatiza que los modelos de lateralización por atributos y por significancia (analizados en detalle en la sección de fisiología del presente capítulo) no son mutuamente excluyentes y son, de hecho, complementarios. Llama la atención hacia la inadecuación de modelos de organización funcional presentes para acomodar los resultados y observaciones sobre dificultades prosódicas producidas por lesiones en zonas subcorticales y enfatiza que la disprosodia puede

considerarse como un síntoma constituyente de diferentes síndromes incluyendo enfermedades neurodegenerativas y lesiones que afectan a procesos involucrados con la prosodia y otras funciones.

Cabe destacar la similitud entre la perspectiva de Van Lancker y la forma general de dividir la prosodia en el campo lingüístico, con más ramas que solamente la distinción lingüística-emocional. Este tipo de clasificaciones, más amplias, tienen la virtud de ser más analíticas y potencialmente más clarificadoras tanto sobre las bases biológicas como las características acústicas asociadas a las diferentes categorías de prosodia. Al respecto considérese, por ejemplo que si bien ha habido perspectivas que intentan clarificar los patrones cerebrales asociados a diferentes emociones, en lo que respecta a la prosodia emocional (Rymarczyk & Grabowska, 2007), no se nota un interés por analizar los patrones cerebrales específicos de las diferentes funciones de la prosodia lingüística (salvo las distinciones entre acento léxico y entonación que ha realizado el equipo de Gandour: Gandour, Dziedzic, Wong, Lowe, Tong, Hsieh, Sathamnuwong & Lurito, 2003; Gandour, Wong, Dziedzic, Lowe, Tong, & Li, 2003; Gandour et al., 2004), empleando simplemente el término “prosodia lingüística” en los títulos de los estudios, sin atender a las múltiples funciones al interior de la misma. En los estudios que abordan a la “prosodia lingüística” (como término general), quizá la función que más se emplea son los actos de habla, específicamente en la distinción entre preguntas y declaraciones. El uso de categorizaciones de la prosodia más amplias, es decir, con más categorías que exclusivamente lingüística y emocional podría ayudar a resolver esta situación.

3.4.3.2 La perspectiva de Relevancia y Prosodia de Wilson y Wharton

Interesados principalmente en la parte pragmática del lenguaje, Wilson y Wharton (2006) notan que la prosodia puede transmitir tanto información “natural” (como la prosodia emocional genuina) e información propiamente lingüística (basada en códigos y convenciones de significado). En el segundo caso enfatizan que aunque no cambian propiamente el significado de las oraciones, ponen énfasis y contribuyen a una interpretación específica entre un conjunto de interpretaciones posibles para una frase (ver funciones de foco y desambiguación en el capítulo de

prosodia lingüística). En todos los casos, desde su perspectiva, la prosodia interactúa con el contexto.

Hacen énfasis en el hecho de que hay funciones propias de la prosodia (codificadas e inherentemente comunicativas) y funciones inferenciales (significados a los que se llega mediante un proceso de inferencia, no de decodificación), así como instancias en las que ambos procesos toman parte.

Dado que hay instancias comunicativas en las que es deseable o esperado que se oculten ciertas marcas reveladas en la voz (por ejemplo, aburrimiento o enojo), evitar ocultarlas o incluso realzarlas en un contexto donde se espera lo contrario permite que el significado sea distinto, se transmite más de lo que se dice (Recordemos la postura de implicaturas conversacionales de Hirschberg, 2002b, basada en el marco de Grice).

Sin embargo los autores consideran que una perspectiva basada en el marco de Grice falla en capturar algunas sutilezas de la comunicación, por lo que recurren a la *Teoría de la Relevancia*. Dicha teoría está basada en una definición de relevancia y dos principios generales, uno cognitivo y uno comunicativo. Desde esta perspectiva, la *relevancia* se caracteriza en términos de costos y beneficios, en relación a los requerimientos de procesamiento de *inputs*; los beneficios se consideran *efectos cognitivos positivos* mientras que los costos *esfuerzo de procesamiento*.

De acuerdo a esta teoría, el *principio de relevancia* se refiere a la tendencia humana a maximizar la relevancia; se asume de base evolutiva y basada en habilidades como el uso de heurísticos y empleo de teoría de la mente. En esta teoría un *input* (de cualquier clase) se considera relevante cuando brinda *efectos cognitivos positivos* en un cierto contexto y de acuerdo a ciertos presupuestos. Este principio subyacería a las percepciones, de manera difusa, de atributos como simpatía, condescendencia, nerviosismo, etc, en la voz. La capacidad compartida por los humanos de inferir la relevancia de los *inputs* les permitiría saber qué

estímulos particulares serían relevantes en ciertos contextos, lo que les permitiría explotar este conocimiento para transmitir significado de manera implicada.

De acuerdo al segundo principio, el *principio comunicativo de la relevancia*, una frase u otro tipo de acto manifiestamente comunicativo crean el presupuesto de relevancia, en contraste con los demás tipos de inputs. En contraste con las frases y los actos manifiestos, todo otro posible estímulo no es inequívocamente relevante, mientras que la persona a la que va dirigida una frase u acto tendería a asumir que dichos *inputs* deben ser lo suficientemente relevantes como para merecer su atención, buscando una interpretación que permita cumplir este presupuesto.

Finalmente, la teoría propone un *mecanismo de comprensión* en el que se busca una interpretación que plantee el menor *esfuerzo de procesamiento* que se detiene sólo hasta que la expectativa de relevancia se cumple. La idea es que los *efectos cognitivos de procesamiento* deben justificar el *esfuerzo de procesamiento* adicional que se requiere para comprender una cierta frase o acto.

También hacen distinciones entre comunicación *fuerte* y *débil* así como entre implicaturas *fuertes* y *débiles*. Una implicatura se considera *fuerte* cuando debe ser derivada en el curso de una proposición para poder satisfacer el presupuesto de relevancia, y *débil* cuando al derivarse ayuda a la construcción de una interpretación satisfactoria pero no es inequívoca dado que la frase de la que deriva la implicatura provee información para la inferencia de diferentes implicaturas relacionadas, de las cuales cualquiera podría satisfacer la expectativa de relevancia; de este modo, las implicaturas *débiles* se asocian a impresiones difusas. Los autores proponen que las frases, de manera oral, transmiten implicaturas de ambos tipos y que la prosodia afectiva suele contribuir a las implicaturas *débiles*.

Los autores proponen que estos son los principios en los que se basaría la prosodia, de manera general, para transmitir la diferente gama de funciones que ha sido descrita en el presente trabajo.

3.4.3.3 Continuos y gradación

Hay varios autores que han discutido la naturaleza continua y gradada de la prosodia o aspectos de ella; se ha propuesto la contribución de la prosodia a continuos de significado; continuos que corren desde la prosodia lingüística a la prosodia emocional; la capacidad de la prosodia para transmitir grados dentro de la misma variable (ej. enojo leve, moderado e intenso); entre otros. Aquí presentaremos algunas de estas ideas.

La idea de gradación no es innovadora en el campo de las emociones; de hecho, está en la base de algunos instrumentos de medición de las mismas, como el SAM (Bradley & Lang, 1994). En contraste, la idea de contrastes categóricos es fundamental en el campo lingüístico (Gussenhoven, 1999). Hockett (1971), dada la propiedad de arbitrariedad en la asociación entre los significados y significantes en el lenguaje humano, ilustra la imposibilidad de la existencia del lenguaje si los contrastes fueran continuos: si cada significado estuviera rodeado de otros significados arbitrarios en la vecindad de la codificación del habla, sería imposible lograr una codificación (en el sentido de asociación de significado y significante), al menos en la forma en la que lo hace el lenguaje; el mismo autor describe también un sistema de comunicación análogo que sí presenta la característica de ser un sistema gradado (la danza de las abejas), pero no permite la generatividad (surgimiento de nuevos significados). De hecho, se nota una cierta reticencia a postular continuos en el campo lingüístico, en especial relacionados a variaciones acústicas (ej. Stockwell, 1963).

La idea de la capacidad de la prosodia para transmitir emociones en diversos grados coincide con la conceptualización de las emociones y no será discutida. En contraste, la idea de continuos en la prosodia lingüística o que la involucran son más escasos y dan lugar a discusión. Incluso, algunas veces, al encontrarse con un fenómeno gradado se le da el estatus de paralingüístico o extralingüístico como ilustran los trabajos de Gussenhoven (2002) y Fitzpatrick (2000).

Con todo, ha habido algunas propuestas que proponen la capacidad de la prosodia para transmitir significados gradados; por ejemplo, Erickson (2006),

resalta que una palabra puede transmitir diferentes significados, dependiendo de sus características acústicas, lo que revela una interacción entre prosodia y semántica; Seddoh (2002), enfatizando el estatus lingüístico de la entonación, resalta que de manera diferente a la información segmental, tiene propiedades gradadas mediante las cuales son expresadas sutilezas de significado.

Gussenhoven (1999) elabora extensamente sobre la gradación en los contrastes de entonación. El autor propone que, si bien muchos de los contrastes prosódicos son discretos, parte del “espacio fonético” no es usado y permite que surjan significados gradados. Ya que en su teoría los hablantes son activos, explotan estos espacios para poder transmitir significados, que se asocian naturalmente a la variación *no lingüística* del tono. Sin embargo, ya que los significados gradados y no lingüísticos pueden ser similares a los significados lingüísticos (asociados a categorías de contraste fonológico) la distinción entre unos y otros puede ser difícil de discernir. También considera que los contrastes entonacionales (lingüísticos y discretos) surgen a partir de la *discretización* de contrastes que eran originalmente gradados. Más tarde elabora la teoría de los tres códigos biológicos (Gussenhoven, 2002) que recupera muchas de las nociones de este trabajo.

Cabe resaltar la similitud de la proposición del uso del “espacio fonético no utilizado” con nuestra *hipótesis de grados de libertad* adelantada en el capítulo de prosodia emocional del presente trabajo.

Por su parte, Grice y Baumann (2007) notan la coexistencia de dos continuos: el continuo de lingüístico-paralingüístico y el continuo categórico-gradado. Los polos extremos de estos continuos estarían relacionados: el polo “lingüístico” con el polo “categórico” y el polo “paralingüístico” con el polo “gradado”. Lo importante aquí es la relación de diferentes funciones de la prosodia con diversos puntos en esos continuos; por ejemplo las funciones emocionales, actitudes, cortesía, etc. estarían relacionadas al polo paralingüístico/gradiente, las funciones léxicas y morfológicas en lenguajes tonales estarían en el polo lingüístico/categórico y los actos de habla y estructura de la información estarían en puntos medios.

Wilson y Wharton (2006), en el marco de su teoría de prosodia y relevancia proponen el continuo entre mostrar y significar. Éste tiene relación con la idea de un continuo que va de más a menos lingüístico o de más universal a más específico del lenguaje. Como se recordará, en su teoría los hablantes hacen uso de su conocimiento de la relevancia para la comunicación; pues bien, dependiendo de qué tanto una frase o acto sea empleado para transmitir significado intencional o no intencionalmente, se le ubica en un continuo que va del polo “mostrar” (menos intencionalmente) a “significar” (más intencionalmente). A diferencia de los continuos universal-específico o más a menos lingüístico, el continuo que proponen Wharton y Wilson también admite estímulos no hablados, que interactúan con el lenguaje en actos comunicativos (por ejemplo, un bostezo).

Resaltan que continuos que van de “más universal” a “más similar al lenguaje” también han sido propuestos para gestos. También Gussenhoven (1999) trabaja con gestos de lenguajes de señas al proponer los argumentos sobre diferencias gradadas; esto, aunado a evidencias de características similares a la prosodia en los lenguajes de señas, nos llevará a abordar el tema en un capítulo posterior.

Pell (1999a), por su parte, resalta que aún en contrastes que se asumen categóricos (como el contraste entre declaraciones e interrogaciones) se puede transmitir el grado en el que se transmite interrogación, así como el grado del énfasis en la función de foco. Recordemos también la propuesta de que las lesiones en el hemisferio derecho podrían afectar preferencialmente a los contrastes prosódicos continuos, en oposición a los categóricos, lo que reflejaría un involucramiento preferencial del hemisferio derecho en los contrastes categóricos continuos (Pell, 1999a, b).

Finalmente, cabe resaltar la perspectiva de Gobl y Ní Chasaide (2003) sobre la cualidad de la voz. Los autores consideran que la misma varía de forma continua, pudiendo haber diferentes grados de cada una de las cualidades de voz que proponen. Cabe recordar que como su definición de cualidades de voz incorpora la actividad (también continua) en diferentes partes del aparato fonador, se puede

considerar que en las cualidades de voz – que son continuas –, interactúan varias variables que también son continuas.

3.4.3.4 Indisociabilidad de la prosodia lingüística y emocional.

En la sección de fisiología presentamos la perspectiva de Snow (2000), donde la prosodia lingüística tiene también una base emocional y discutimos esta propuesta en relación con el diferente uso de términos en el campo lingüístico y psicológico; debate presentado en el capítulo de prosodia lingüística. En su perspectiva, los sentimientos y actitudes del hablante constituyen la base para el significado que tendrían los diferentes componentes prosódicos, como los acentos tonales. Esto ilustraría una indisociabilidad de la prosodia lingüística y emocional.

3.4.3.5 Efectos de jale y empuje de Scherer

En el capítulo de prosodia emocional abordamos la perspectiva de efectos de *jale* y *empuje* de Scherer y Bänziger (2004). En el trabajo de Grandjean y Scherer (2006), se retoma y amplía esta perspectiva. Recordemos que los efectos de *empuje* se referían a “las externalizaciones biológicamente determinadas de procesos que ocurren naturalmente en el organismo, particularmente procesamiento de la información y preparación de la conducta”, mientras que los efectos de *jale* a “las normas o *modelos* determinados socioculturalmente sobre las características de la señal requeridas por los códigos compartidos socialmente para la comunicación de estados internos e intenciones conductuales”. En esta perspectiva ambas prosodias son sujetas a ambos tipos de efectos, aunque la prosodia emocional está sometida principalmente a los efectos de *empuje* mientras que la prosodia lingüística a los de *jale*. Grandjean y Scherer (2006) proponen dos principios generales que subyacen la codificación de la información prosódica en habla: la *covariación* y la *configuración*. El principio de *covariación* asume una relación continua, pero no necesariamente lineal entre algunos aspectos de las respuestas emocionales y algunas variables acústicas; por ejemplo, a mayor *arousal*, mayor F0. En contraste, la *configuración* implica que un significado específico es transmitido por una frase e inferido por el oyente a partir de la configuración prosódica total; es decir, es un contraste categórico que se puede asociar a la prosodia lingüística, principalmente. Sin embargo los contrastes

categoricos permitirían aún así transmitir significados emocionales al desviarse de lo esperado en un cierto contexto donde se espera un contorno específico; este es el principio que subyacería a las observaciones de interacción de patrones tonales de subida y bajada con preguntas dicotómicas y de la serie *wh*- reportado en la sección de fisiología del presente capítulo (Pihan et al., 2008). Los autores consideran que las variables continuas podrían estar principalmente ligadas a efectos de *empuje* mientras que las variables discretas a los efectos de *jale*; también consideran que el origen evolutivo de la *covariación* es más antiguo y relacionado a los mecanismos biopsicológicos que afectan a las vocalizaciones en diferentes especies, mientras que el principio de *configuración* estaría asociado al surgimiento del lenguaje y transmite mensajes específicos, determinados por códigos. Nótese la relación de esta perspectiva con la postura de los tres códigos biológicos de Gussenhoven (2002), que propone que hay usos gramaticalizados y no gramaticalizados de sus códigos.

3.4.3.6 Hipótesis de carga prosódica

Esta hipótesis supone que entre más contrastes simultáneos se requiera marcar en una oración, mayor es la “carga prosódica” que representa; por ejemplo, si se pide que se haga énfasis en una palabra (mediante f_0) y a la vez se transmita una pregunta (que requiere elevación de f_0), se dejan de producir distinciones emocionales en las palabras enfatizadas; aun así, las marcas en f_0 asociadas a las emociones al nivel de frase, así como otros atributos como el ritmo de habla aún se producen y permiten desambiguar la información afectiva. En los casos en los que una cierta producción requiera reflejar más atributos de los que permiten los mecanismos de producción del habla, se producen efectos de compromiso (*trade-off effects*), en el que se asignan los recursos de producción de la manera que maximice el efecto y minimice el esfuerzo sin requerir un gran esfuerzo perceptual y cognitivo en el oyente (Bryant y Fox Tree, 2005; Pell, 1999).

3.4.3.7 La falta de sentido de la división entre prosodia lingüística y emocional: La perspectiva de Seddoh

Seddoh (2002) se opone de manera férrea a la distinción entre prosodia lingüística y prosodia emocional. En su trabajo se centra en la entonación (dado el énfasis

que se ha puesto en esta variable al estudiar la prosodia), no en la prosodia de manera general y considera que la distinción entre las funciones emocional y lingüística para la entonación puede ser un marco teórico que dificulte el progreso del entendimiento de los desórdenes de la entonación. Considera que hay poca evidencia empírica que apoye esta distinción (aunque deben recordarse las evidencias que aquí se describen, algunas de ellas posteriores al trabajo de este autor).

Conceptualmente, considera que el uso de los términos es incorrecto; llamar “entonación lingüística” a un fenómeno que se ya se asume lingüístico sería redundante y no aportaría nada al término y sería análogo a términos como “veterinario de animales”. A su vez, en el caso de la entonación emocional, considera que si bien las pistas perceptuales características podrían tener un origen no lingüístico, considerarla como un fenómeno no lingüístico (dada su oposición con el término “entonación lingüística”) implicaría – desde su punto de vista – que los contornos asociados a esta categoría están determinados por completo por factores no lingüísticos. Sin embargo esto no podría ser viable, ya que la entonación opera en combinación con la información segmental a niveles de frases y oraciones; estas últimas entidades serían unidades sintácticas, y, dadas observaciones de que la sintaxis influye a la entonación tanto en contextos emocionales como no emocionales, la prosodia emocional estaría influida por variables lingüísticas. A su vez, cita estudios en los que se observa que los contornos asociados a la entonación lingüística son influidos por variables lingüísticas y por variables fisiológicas no lingüísticas. Por otra parte, considera que si la “entonación emocional” fuese un fenómeno independiente del lenguaje, sería posible elicitarla en contextos completamente no lingüísticos; sin embargo, al ser la entonación una entidad que se da al nivel de frases y oraciones, esto no sería, por definición, posible (aunque recuérdese la hipótesis de características acústicas similares en expresiones vocales no lingüísticas de la emoción y la prosodia emocional que presentamos en el capítulo de prosodia emocional). También resalta efectos de interacción entre la entonación, el contenido semántico y la categoría sintáctica de las oraciones para determinar el significado emocional,

así como la posible dificultad de hablantes de lenguajes tonales (en comparación con lenguajes no tonales) de señalar estados emocionales mediante la entonación (la prosodia en el caso de lenguajes tonales será abordada en un capítulo posterior); este argumento puede ser válido al solo tomar en cuenta la entonación, pero es importante recordar que la prosodia se sirve de diferentes variables (recuérdese la perspectiva de carga prosódica).

Como alternativa a la distinción entre entonación lingüística y emocional propone el concepto de “*entonación emocionalizada*”; desde esta perspectiva, la entonación siempre estaría determinada por factores lingüísticos pero sería influida por las variables fisiológicas asociadas a las emociones, modificando las características acústicas de las mismas y permitiendo así brindar esta información. Estas variaciones acústicas tendrían un origen no lingüístico, basado en áreas emocionales del cerebro.

Según el autor, esta conceptualización tendría la ventaja de mantener el estatus lingüístico del término entonación y permitiría integrar la conceptualización de algunos lingüistas de que a diferencia de la información segmental, la entonación permite transmitir significados gradados sin dejar de ser una entidad lingüística (recordemos que la idea de gradación en entidades lingüísticas no goza siempre de apoyo en el área). Así como la ventaja – de evidente orientación clínica – de mantener el interés en descartar que el origen de dificultades emocionales prosódicas posteriores a lesiones cerebrales tenga una base en dificultades de procesamiento emocional en general; prueba que, comenta, no se realiza siempre en el trabajo clínico.

Quizá una de las mayores contribuciones de este autor es que llama la atención al uso correcto de los términos prosodia y entonación, siendo prosodia un término supraordenado a entonación y no un sinónimo, como se le maneja en algunas ocasiones.

3.4.3.8 Hipótesis de Grados de Libertad

En el capítulo de prosodia emocional planteamos la *Hipótesis de Grados de Libertad* (la principal contribución del presente trabajo): “La prosodia emocional

debe hacer uso de las características no fundamentales para la comprensión de un mensaje, es decir, la señal acústica se puede modificar mientras no afecte a las diferencias contrastivas en un lenguaje (tanto segmentales como suprasegmentales)". Las características contrastivas serían necesarias e inmodificables, mientras que todo cambio en la señal acústica que no afectase a las características acústicas esenciales podría ser considerado como "*grados de libertad*". Un razonamiento análogo podría aplicar a otros niveles de la gramática, expandiendo el alcance de la hipótesis.

La hipótesis podría aplicar también a los lenguajes de señas; aunque en estos casos la señal no es acústica, sino visual, ya que también existen características contrastivas necesarias y variaciones permisibles. Cabe resaltarlo porque los lenguajes de señas también presentan prosodia; cuestión que será abordada en el capítulo de temas complementarios.

En el trabajo de Gussenhoven (1999) encontramos una idea similar en su proposición del uso del "espacio fonético" no utilizado para poder transmitir significados gradados, como las emociones.

Nuestra hipótesis encuentra apoyo documental indirecto, empírico y teórico como ilustraremos a continuación.

- La observación de Seddoh (2002) de que hablantes de lenguas tonales tienen más dificultad para codificar información emocional en la entonación reflejaría que hay menos grados de libertad tonales en estas lenguas, ya que las diferencias contrastivas tonales especifican en gran medida la forma de usar este canal.
- El planteamiento general de la carga prosódica (Pell, 1999; Bryant & Fox Tree, 2005), es compatible con nuestra *hipótesis de grados de libertad*: a mayor carga prosódica, menos grados de libertad.
- La observación de Gutiérrez Bravo (2008) de que las diferencias en la libertad de la posición en las oraciones de la información focalizada en inglés y en español tiene base prosódica, es compatible con nuestra

hipótesis: en inglés hay mayor libertad prosódica y la información focalizada puede aparecer en cualquier lugar de la oración; en español, el acento nuclear debe aparecer al final de la oración (menos grados de libertad prosódicos), lo que lleva a modificaciones de la sintaxis para que la información focalizada se dé al final de la oración (reflejando, en este caso, *grados de libertad sintácticos*).

- Kouassi (2008) observa diferencias de género en la producción de tonos que requieren una bajada y una subida; aunque las producciones de ambos géneros muestran el patrón de bajada y subida, en las mujeres el patrón temporal muestra una bajada más temprana. El hecho de que sean percibidos como el mismo patrón de tono (bajada y subida), a pesar de las diferencias en el patrón temporal, indica que no son afectadas las diferencias contrastivas e indica la presencia de variación libre (o grados de libertad); dicho de otra manera, la diferencia en producción entre hombres y mujeres es fonética y no fonológica. De hecho, el autor resalta que “la prosodia es un sistema lingüístico que permite al sujeto enunciante no solamente expresar su subjetividad al coenunciante, sino mantener al mismo tiempo el enunciado perfectamente objetivo”. Esta capacidad de transmitir ambas cosas simultáneamente puede interpretarse en el marco de la *hipótesis de los grados de libertad*: el enunciado transmite la información objetiva mediante el contenido proposicional y obligado y la subjetividad mediante las variaciones que caben en los grados de libertad.
- Frazier et al. (2006) presentan una perspectiva en la que la prosodia puede ser conceptualizada desde un continuo entre gramaticalizada y no gramaticalizada. En el caso de estar gramaticalizada, habría algunas reglas sobre la pronunciación permisible de una oración en ciertos contextos, aunque podría existir una variación considerable en las realizaciones de dichas pronunciaciones; las variaciones constituirían los grados de libertad; las variaciones no permisibles lo serían por oponerse a las diferencias contrastivas requeridas. Asimismo, el autor enfatiza que algunos aspectos de la prosodia son obligatorios (diferencias contrastivas) y otros estilísticos

(grados de libertad). El autor enfatiza, sin embargo, la importancia de la relatividad de la prosodia consigo misma; el mismo tamaño en una pausa podría ser una pronunciación aceptable para una persona que habla a un ritmo X, pero violar los requisitos de una variación aceptable para una persona que habla con un ritmo Y.

Siguiendo la idea del uso del “espacio fonético” libre de Gussenhoven (1999), cabe preguntarse si otros significados gradados, por ejemplo las actitudes, pueden considerarse similares a la prosodia emocional (genuina, en oposición a las *reglas de exposición*, presentadas en el capítulo de prosodia emocional), en el sentido de que hacen uso de los grados de libertad: que no afecten las diferencias contrastivas en un lenguaje. Si bien parecen no ser diferencias contrastivas (en el sentido de cambio de significado) podrían reflejar un contraste más ligero, pero también codificado (la discusión entre codificación o no de información en la prosodia será abordada en el siguiente apartado), lo que reflejaría un posible continuo entre lo contrastivo y lo no contrastivo, en lugar de una división categórica.

Aparentemente esta hipótesis sólo podría aplicarse a la prosodia emocional (o a todas las prosodias distintas de la prosodia lingüística, siguiendo el marco de Van Lancker, 2006; incluyendo la función indéxica), ya que si postulamos que las diferencias contrastivas son los atributos requeridos forzosamente por el sistema, por lo tanto, gramáticos y por lo tanto dependientes del lenguaje, se dificulta pensar que se pudiera aplicar la *hipótesis de grados de libertad* a la prosodia lingüística. Con todo, podría aplicarse diferencialmente a diferentes atributos prosódicos, siguiendo la conceptualización del continuo que va de lingüístico a paralingüístico de Grice y Baumann (2007).

La *hipótesis de grados de libertad* podría guiar y justificar la realización de estudios que, sin una base teórica, probablemente no serían realizados, entre los cuales consideramos al menos dos:

Estudios diacrónicos sobre la variación en la prosodia lingüística: Si bien por el momento se asume a la prosodia emocional como un fenómeno más o menos universal, por lo que no habría de esperar cambios históricos en sus características (al menos no cambios en periodos cortos de tiempo, como décadas, que no deberían permitir efectos de adaptación por presiones evolutivas), considerando la *hipótesis de los grados de libertad* y las observaciones de Grice y Baumann (2007) de que la pronunciación cambia de manera relativamente rápida tanto a nivel regional como de lenguaje estándar, estos cambios afectarían las señales acústicas obligatorias (diferencias contrastivas) a lo largo del tiempo, y con ellas los grados de libertad; así, si la hipótesis es correcta, se podría esperar que la prosodia emocional varíe a lo largo del tiempo para ajustarse a los grados de libertad disponibles.

El uso agregado de diferentes realizaciones prosódicas (emocionales, estilísticas, de género, etc.) para una misma frase podría revelar cuáles son las diferencias fundamentales; la convergencia entre distintas realizaciones prosódicas serían las diferencias contrastivas requeridas y las variaciones reflejarían los grados de libertad.

Siguiendo la teoría de los tres códigos biológicos de Gussenhoven (2002), la totalidad posible de grados de libertad podría estar constreñida biológicamente; recordemos que el autor plantea que algunos usos de los códigos biológicos se gramaticalizan, y al hacerlo se sujetan a los procesos de variación y cambio lingüístico, pero que jamás podrían revertir completamente la relación biológica de los códigos en los que se basan, si bien se pueden mover en cierto rango. La parte inamovible de los códigos biológicos jamás podría formar parte de los grados de libertad de la señal.

3.4.3.9 ¿Qué codifica la prosodia?

Al pensar sobre la prosodia puede ser muy útil remitirse a la distinción que hacen Wilson y Wharton (2006) sobre la información que es inherentemente comunicativa y que se halla codificada y la información que no tiene como función principal ser comunicada pero puede serlo mediante la inferencia del interlocutor.

Como hemos visto, la prosodia puede transmitir información de muchas clases: emocional, lingüística, indéxica, entre otras, sin contar las ramificaciones de cada una. La distinción entre información codificada y no codificada es muy útil para no perderse en la inmensidad de información que podría transmitir. De la información que transmite, toda la que se comprende de manera inferencial debería tener por base procesos psicológicos distintos a los de una decodificación prosódica propiamente dicha. Una decodificación presupone un código compartido, ya sea biológico o cultural; mientras que la información comprendida de manera inferencial podría estar servida por procesos como el razonamiento. Con todo es importante recordar que en la perspectiva de Wilson y Wharton (2006), los tipos no son puros: se puede, y de hecho se da en gran cantidad de ocasiones, una combinación de decodificación e inferencia (en especial para la prosodia lingüística).

¿Si las variaciones prosódicas son signos lingüísticos, son signos de *qué*?

Quizá lo más sencillo es pensar en los usos léxicos de la prosodia (ej. la distinción entre papá y papa y la distinción entre escuela preparatoria /ko:ko:/ y aquí /koko/ en japonés). Después vendría la distinción de actos de habla, por ejemplo, entre preguntas o declaraciones, aunque como se ha visto, a este nivel ya se empieza a ver la influencia de otras variables y efectos de interacción. Hasta aquí es relativamente sencillo asociar ciertas características prosódicas (ej. duración, incremento de tono al final de las oraciones) con su función, o, dicho de otra manera, determinar de *qué* son signo las variaciones prosódicas. Sin embargo, más allá de estas dos funciones de la prosodia, empieza a volverse más difícil saber de *qué* son signos las variaciones prosódicas.

Wilson y Wharton (2006) plantean que hay dos tipos de codificación: codificación conceptual (ej. [perro] codifica al concepto *perro*) y codificación de información procesual: por ejemplo en “*Tracy DANCED that night*” (donde las mayúsculas indican foco), el énfasis pretende guiar el proceso de inferencia por parte del oyente, acortando el rango de posibles interpretaciones de una frase; ayudando a que una hipótesis entre diferentes hipótesis posibles sea más saliente y

umentando la probabilidad de que sea relevante (recordar la postura de relevancia y prosodia de estos autores). En este segundo tipo de casos, la prosodia estaría codificando no información conceptual sino información procesual que debería seguir el interlocutor para lograr una comunicación efectiva.

El estudio de Ward y Bayyari (2010) ilustra excelentemente esta perspectiva; recordemos que realizan un experimento intercultural sobre las pistas prosódicas que indica que se espera una *respuesta de involucramiento* en árabe. La información es procesual: no necesariamente aporta algo al significado del mensaje, pero indica al interlocutor que se espera una respuesta de involucramiento. La prosodia se puede considerar como subyacente por un código, dado el mejor desempeño de árabes en comparación con hablantes de inglés (que también tiene pistas sobre la espera de *respuestas de involucramiento*, aunque el patrón prosódico es distinto). Hablantes de inglés sin entrenamiento de árabe confundían esta pista prosódica con un índice de afecto negativo (presumiblemente mediante un proceso inferencial); en contraste con hablantes de árabe. Hablantes de inglés que habían sido expuestos al idioma árabe y entrenados sobre las pistas prosódicas de expectativa de *respuestas de involucramiento* no mostraban la tendencia a percibir afecto negativo; presumiblemente, al conocer el código después del entrenamiento, la información que transmite fue captada por un mecanismo de codificación en contraste con inferencia en ausencia de código.

En la siguiente sección ampliaremos el tratamiento de algunos temas y cerraremos el capítulo.

3.4.4 Resumen y discusión.

En esta sección resaltaremos brevemente algunos temas importantes que han surgido hasta este momento.

3.4.4.1 Variables de la prosodia. La cualidad de la voz

Si bien se ha hecho énfasis principalmente en las variables tonales como f_0 y temporales (Lakshminarayanan et al., 2003; Pell, 1999b; Snow, 2000; Van Santen et al. 2009), son muchos los autores que han resaltado la importancia de tomar en

cuenta otras variables prosódicas. La cualidad de la voz ha recibido considerable atención y a pesar de las dificultades conceptuales y metodológicas, se sigue estudiando y argumentando a favor de su importancia al estudiar la prosodia (Gobl & Ní Chasaide, 2003). En algunas lenguas, la cualidad de la voz es usada en conjunto con otras variables para señalar diferencias contrastivas incluso a nivel segmental (Gussenhoven, 1999). Los esfuerzos del equipo de Gobl son importantes al tratar de generar mejores metodologías de medición y definición de la cualidad de la voz basada en la actividad de diferentes niveles del aparato fonador (Gobl & Ní Chasaide, 2003).

3.4.4.2 Metodología de Síntesis de habla y modificación acústica.

Varios trabajos se han apoyado en tecnología de síntesis de habla y modificación acústica de estímulos orales (por ejemplo, Lakshminarayanan et al., 2003; Pihan et al., 2008). Si bien esto permite manipular independientemente las variables de interés, proveyendo de un buen control de variables, es importante recordar que las variables que componen la prosodia no son necesariamente ortogonales. El trabajo de McRoberts y colaboradores (1995) hace énfasis en esto y argumenta que el uso de la síntesis de habla y modificación acústica podría generar estímulos que no se comportan como lo haría el habla genuina. Esto no invalida los resultados de experimentos que emplean la metodología de síntesis y manipulación acústica (los cuales serían, de hecho, difíciles de implementar de otra manera), pero llama la atención a considerar el argumento de McRoberts y su equipo al interpretar estudios. Las comparaciones de resultados entre estímulos genuinos y estímulos por síntesis y manipulación acústica pueden aclarar la fortaleza o debilidad del argumento de McRoberts y colaboradores. El avance de las aplicaciones de tecnología computacional del habla también tendrá un efecto positivo en el uso de la misma como metodología de desarrollo de estímulos.

3.4.4.3 Especificidad de la prosodia neutra.

La prosodia emocional neutra tiene características acústicas específicas (Pell, Paulmann, Dara, Alasseri, & Kotz, 2009), tal como lo indica la capacidad de las personas de reconocerla en pruebas perceptuales. En los casos de prosodia neutra se ha discutido que, dado que la prosodia se emplea para señalar actitudes

también se emplea para indicar lo más claramente posible que no se tiene ninguna actitud particular, siempre transmite cierta información actitudinal (aunque sea, precisamente, señalar que no se tiene ninguna actitud particular; Seddoh, 2002). También se ha sugerido que la prosodia emocional neutral puede estar asociada a una intención comunicativa en la que se trata de sonar objetivo o con dominio sobre lo que se está diciendo (Pell, Paulmann, Dara, Alasseri, & Kotz, 2009). De cualquier modo, la prosodia emocional neutral, debe tener las características acústicas de algún tipo de acto de habla (declarativo, interrogativo, imperativo, etc). ¿Será posible que existe un estado emocional o afectivo asociado a la intención de sonar objetivo? Esta pregunta quizá sería relevante en el caso de un marco teórico como el de Snow (2000); de otra manera, simplemente se puede suponer que la prosodia emocional neutral es más cercana a la prosodia “ideal”, aquella que sólo contiene las diferencias contrastivas necesarias y no tiene mucha carga prosódica.

3.4.4.4 Importancia de estudios tanto acústicos como perceptuales

Al estudiar la prosodia, no basta con tomar en cuenta los resultados acústicos o perceptuales de los experimentos; es importante usar ambas medidas. En algunos estudios vimos como a nivel acústico las producciones parecían normales, pero a nivel perceptual eran menos discriminables (Pell, 1999a, b). Bryant y Fox Tree (2005) resaltan que en algunas condiciones de filtrado acústico, las personas siguen oyendo una frecuencia que, de hecho, no está presente; Frazier et al (2006) resaltan que a hablantes de francés les es difícil percibir patrones de *stress* de lenguas como el inglés, de distribución más libre en comparación con su idioma, donde el principal acento se encuentra al final de la frase normalmente. En resumen, es importante el uso de ambos enfoques al investigar sobre la prosodia (ver Klouda et al., 1998; McRoberts et al. 1995; Pell 1999a, b)

3.4.4.5 Importancia de distinguir el dominio temporal en el que se extiende la prosodia

Hasta el momento hemos visto que la prosodia se refiere a los aspectos suprasegmentales del habla; es decir, después de los fonemas, todo contraste es prosódico. Sin embargo esto es un dominio amplio, pues va desde la sílaba (y la mora, en lenguajes como el japonés) hasta las oraciones completas. Es

importante tener en cuenta esto y distinguir el dominio temporal en el que se da la prosodia (Lakshminarayanan et al., 2003), en especial dada la existencia de diferente lateralización de f_0 en los contrastes léxicos (nivel de sílaba) y entonacionales (frases y oraciones) que ha estudiado el equipo de Gandour (Gandour, Dzemidzic, Wong, Lowe, Tong, Hsieh, Sathamnuwong y Lurito, 2003). Siguiendo la perspectiva de Wilson y Wharton (2006), incluso estarían codificando diferente clase de entidades: la prosodia al nivel de sílaba (al menos en chino) estaría haciendo distinciones de información conceptual mientras que la prosodia a otros niveles, codificaría información procesual. Para la mejor comprensión de la prosodia, es fundamental hacer la distinción en el rango que afecta.

3.4.4.6 Bases neurales para diferentes funciones de la prosodia lingüística.

Si bien se ha reconocido una gran variedad de funciones de la prosodia, no se han hecho comparaciones de la organización funcional entre ellas, salvo la investigación de Gandour, Dzemidzic, Wong, Lowe, Tong, Hsieh, Sathamnuwong y Lurito (2003; y otras del equipo de Gandour), que compara la organización cerebral al nivel de sílabas y frases. Con todo, el contraste no está hecho por funciones, en sentido estricto: si bien la información al nivel de sílabas contribuye a la función de distinciones léxicas, no se hace comparación con otras funciones como tal, sino con prosodia en otro nivel temporal. En contraste, sí se ha propuesto diferente organización funcional para la prosodia relacionada a diferentes emociones. (Rymarczyk & Grabowska, 2007).

3.4.4.7 Análisis conceptual

A través del uso de los términos, se ha visto que la principal característica que contrasta entre prosodia lingüística y prosodia emocional es la función que transmite; es decir, se considera prosodia emocional cuando transmite información emocional y prosodia lingüística cuando transmite información lingüística. De esta manera se aprecia como el principal contraste entre estos términos es uno funcional. Si bien se han visto contrastes acústicos (prosodia emocional relacionada a variables globales y contrastes continuos y prosodia lingüística a partes específicas, contrastes categóricos y contornos), y fisiológicos

(organización cerebral de las dos categorías de prosodia), estas no son parte del nivel conceptual.

Algunos autores han puesto de relieve también el nivel de lo que podría llamarse “control voluntario” en la distinción entre las prosodias. Específicamente, en la tradición lingüística, donde se distinguen los términos emocional y emotivo en el sentido de que la prosodia puede transmitir emociones genuinas, sin el control voluntario de esta transmisión (y que se considera un fenómeno extralingüístico), o transmitir impresiones emocionales (no necesariamente genuinas) con fines comunicativos. En estos casos, se podría considerar parte de la prosodia emocional las instancias de emociones genuinas transmitidas involuntariamente y parte de la prosodia lingüística las emociones (no necesariamente genuinas) transmitidas voluntariamente con fines comunicativos.

Así, el grado de voluntad de transmisión y el estatus de genuinas en las emociones podrían ser otro nivel que contrastara entre la prosodia lingüística y emocional en casos de especial ambigüedad. Entre estos casos de especial ambigüedad considérese, por ejemplo, el caso de las intenciones comunicativas (pragmáticas, sociales, etc) de emociones fingidas (Ver Makarova & Petrushin, 2002; Shochi, Aubergé & Rilliard, 2006; Wilson & Wharton, 2006). En el Anexo 1 se incluyen definiciones de prosodia emocional y lingüística; se observa la presencia de más definiciones específicas de la prosodia emocional que de la prosodia lingüística, quizá porque en el campo de la psicología y neurociencias se tiene más interés en la primera, mientras que, como fue discutido con anterioridad, el término “prosodia lingüística” puede ser poco informativo en el campo lingüístico.

3.4.4.8 Funciones que no son ni lingüísticas ni emocionales: Las funciones indécicas de la prosodia.

La propuesta de Van Lancker y colaboradores (2006) incluye las funciones indécicas de la prosodia, que transmiten información sobre la identidad, edad, personalidad, etc. Esta propuesta es particularmente importante considerando que el acento extranjero transmite información sobre el origen de la persona, y lo hace

por medios prosódicos principalmente, tal como revisaremos en el siguiente capítulo. Lo interesante de esta propuesta es que permite incluir como función de la prosodia información que no es ni emocional ni lingüística. También Wilson y Wharton (2006) han destacado la posible existencia de marcas prosódicas que no sean ni lingüísticas ni emocionales, sino culturales; similares a gestos como el mantener el dedo medio extendido mientras todos los demás se mantienen cerrados para insultar.

3.4.4.9 Importancia del marco teórico

Hasta el momento hemos presentado diferentes marcos teóricos para estudiar la prosodia; además del marco general en el que se inscribe el presente, en el que se distinguen la prosodia lingüística y emocional, hemos abordado el marco de Van Lancker y colaboradores (2006) donde hay una división más amplia; el marco general de división de funciones en el campo lingüístico, donde la prosodia emocional está en el mismo nivel jerárquico que otras funciones de la prosodia; el marco de Snow (2000), donde tanto la prosodia emocional como lingüística tienen una base emocional; el marco de Seddoh (2002), en el que la distinción entre prosodia emocional y lingüística carece de sentido, pero se puede hablar de *entonación emocionalizada*; entre otras. También hemos avanzado *la hipótesis de grados de libertad*, con algunas posibles aplicaciones e implicaciones teóricas. Todos los marcos conceptuales tienen sus ventajas y desventajas, sus aportaciones y limitaciones.

Definitivamente, el marco conceptual desde el que se trabaja influye el tipo de estudios que se realizan y los intereses en un programa de investigación. Además de la investigación, pueden influir en la práctica clínica o de otras áreas. Joannette y colaboradores (2008) describen cómo los cambios en las elaboraciones teóricas del lenguaje, en general, han proveído la posibilidad de adentrarse en otras funciones y niveles del lenguaje; a saber, el advenimiento de elaboraciones teóricas del lenguaje que incluyen los niveles pragmático y discursivo ha influido el trabajo neurológico en el habla. En contraste, en trabajos anteriores a dichas elaboraciones teóricas el estudio del lenguaje se centraba en los niveles de articulación, fonología, semántica léxica, y morfosintaxis; en ese tiempo, eran

precisamente estos niveles los que estaban en el centro de las consideraciones teóricas sobre el lenguaje. Los autores consideran que, dada la definición más amplia del lenguaje en la actualidad (que incluye niveles discursivos y pragmáticos), las dificultades prosódicas podrían considerarse como un tipo de afasia (ya que se define afasia como una alteración en el lenguaje debida a una lesión cerebral, sin hacer referencia a la etiología o localización). Es decir, al haberse modificado la concepción del lenguaje, las dificultades prosódicas satisfacen los criterios de la definición para ser considerados un tipo de afasia. Los autores resaltan que aunque esta consideración teórica no contribuye a comprender las causas del fenómeno, podría tener repercusiones en la práctica clínica: los pacientes con lesiones que afecten la prosodia (o las habilidades pragmáticas y discursivas), se beneficiarían de la misma atención y cuidado que reciben los pacientes que sufren afasias (en su conceptualización tradicional).

Desde mi perspectiva, un análisis teórico adecuado es fundamental para el progreso de la investigación y de las aplicaciones derivadas de ella, por lo que nunca debe pasarse por alto. Qué marcos conceptuales son los que mejor capturan el conocimiento empírico actual, así como qué capacidad predictiva de los resultados y qué capacidad de orientar el curso de investigaciones futuras son atributos fundamentales de los marcos teóricos. Estos no necesariamente se contraponen y es importante conocer varios, así como sus ventajas y desventajas y los campos de sinergia.

La teoría y los hallazgos empíricos se influyen mutuamente en un proceso constante de retroalimentación (y deben hacerlo). Tanto en la interpretación de resultados (teoría a hallazgos empíricos) como en la modificación y adecuación de teorías (hallazgos empíricos a teoría), entre otras influencias mutuas, como es, definitivamente, el curso de las investigaciones a lo largo del tiempo.

Dentro de las teorías, el uso preciso de términos es fundamental, tal como resalta Seddoh (2002) al discutir sobre el uso indebido de los términos prosodia y entonación como sinónimos. En un capítulo posterior discutiremos las definiciones

de prosodia y trataremos de brindar una que capture de manera parsimoniosa las características fundamentales de dicha entidad.

3.4.4.10 Aplicabilidad

Si bien el presente trabajo es eminentemente teórico, también podemos identificar algunas áreas donde el estudio de la prosodia puede tener aplicaciones importantes. Los campos más destacados en este sentido son el computacional, tanto en sistemas de producción y síntesis como de identificación del habla, que pueden constituir interfaces de comunicación más cercana a lo natural entre hombres y máquinas (Luengo et al., 2005; Makarova & Petrushin, 2002; Martínez Sánchez et al., 2002; Saíz Ruiz & Talavera Martín, 1997; Seppänen, Väyrynen, & Toivanen, 2003; Van Santen et al., 2009), tal como resaltan varios de los trabajos computacionales que se abordaron hasta el momento; el campo de la enseñanza de lenguas extranjeras, dada la gran importancia de la prosodia en la inteligibilidad, que se abordó en el capítulo de prosodia lingüística (Meng, Tseng, Kondo, Harrison, & Viselgia, 2009; Zhang, Li, Lo, & Meng, 2010); en el campo clínico, en la identificación temprana de desórdenes del lenguaje (Snow, 2000; Van Santen et al., 2009) y en el desarrollo de programas de intervención y rehabilitación tanto para personas con dificultades prosódicas como para que personas que mantienen estas habilidades intactas las exploren con fines comunicativos (Joanette et al., 2008; Perlman Lorch et al., 1998).

3.4.4.11 Campos relacionados

Consideramos que el estudio de campos que implican a la prosodia desde diversas perspectivas es importante para terminar de clarificar las contribuciones biológicas, psicofisiológicas y culturales a la prosodia del lenguaje, así como su nivel conceptual. Por tanto, en el siguiente capítulo abordaremos brevísimamente algunos campos relacionados: el acento extranjero y síndrome del acento extranjero; la relación de la prosodia y la música; el desarrollo de la prosodia, la prosodia en lenguajes tonales y la prosodia en lenguajes de señas.

El tratamiento de cada tema será muy breve y sólo abordará de manera sintética algunas consideraciones que contribuyan al objetivo principal de la tesis.

3.5 Temas complementarios

En el presente capítulo se abordan temas con el potencial de ampliar la visión presentada en los capítulos anteriores y que permiten aislar variables, tal como fue justificado en la introducción. Al final del presente trabajo, se mostrará como un capítulo fundamental, sin el cual habría sido imposible plantear y justificar algunas propuestas.

3.5.1 Acentos y Síndrome del Acento Extranjero

En esta sección se abordará brevemente el tema de los acentos extranjeros y el síndrome del acento extranjero, de fuerte base prosódica. En capítulos anteriores hemos visto como la prosodia es importante en la inteligibilidad de hablantes de lenguas extranjeras (Meng et al., 2009; Zhang et al., 2010); que hablantes de lenguas extranjeras pueden mantener patrones prosódicos de la lengua materna en funciones como la toma de turnos (Ward & Bayyari, 2010) y la segmentación silábica (Dromey et al., 2005), y que las diferencias prosódicas entre dos variantes del mismo idioma (por ejemplo inglés británico y americano) pueden ser tan grandes como las diferencias entre lenguas distintas (Van Santen et al., 2009).

González-Álvarez y colaboradores (2003) resaltan que la conducta motora verbal es quizá la más compleja de las conductas motoras humanas, y que requiere un control temporal fino de diversos músculos, controlados por diversos nervios y pares craneales, que permita emitir hasta 30 segmentos sonoros por segundo, estos patrones son fijados en automatismos, difíciles de cambiar, por lo que el habla con un acento extranjero natural sería producto de estos automatismos motores que reflejan características prosódicas y fonéticas de la primera lengua.

El acento extranjero cabe en las funciones indécicas propuestas por Van Lancker y colaboradores (2006). Aunque desde la perspectiva de Wilson y Wharton (2006) no sería una función propia de la prosodia (de inherente base comunicativa). La contribución cultural a los acentos (y biológica en el caso del síndrome del acento extranjero), al afectar principalmente a la prosodia, forma parte del interés del presente trabajo. Desde nuestra hipótesis de los *grados de libertad* la posibilidad

de que se den acentos extranjeros forma parte de las diferencias no contrastivas y no fundamentales del lenguaje.

Las contribuciones más importantes que ofrece la literatura sobre el acento extranjero al estudio de la prosodia residen en conceptos como las “configuraciones fonéticas” y en el estudio de la interacción entre la información segmental y suprasegmental, lo que lleva a pensarlas no como categorías mutuamente excluyentes y con una línea divisoria clara, sino como parte, quizá, de un continuo.

Los acentos extranjeros pueden ser detectados con relativa facilidad, no se requiere alguna instrucción especial para hacerlo y se puede hacer inclusive cuando los estímulos se modifican considerablemente, desde una palabra e inclusive con sólo milisegundos de habla (Miller et al., 2006; Munro et al., 2010).

Se ha considerado a los acentos extranjeros desde diferentes perspectivas: por ejemplo, que la sensibilidad a los acentos se ha desarrollado en los humanos como un mecanismo que fomenta la selección de parejas al interior de un mismo grupo; que sirve como una ventaja evolutiva al poder determinar perceptualmente el grupo al que pertenece un interlocutor (hipótesis de la contraseña; Munro et al., 2010); aunque también se ha propuesto que es, simplemente, un epifenómeno (Munro et al., 2010).

Más allá de esta discusión, se ha visto que puede servir para señalar la identidad a los demás y a uno mismo (Scott et al., 2006).

Hablar con acentos extranjeros puede tener consecuencias positivas y negativas a varios niveles. Entre las consecuencias positivas: las personas nativas que oyen a alguien hablar con acento extranjero pueden ajustar sus patrones de habla para mejorar la comprensión si asumen que su interlocutor tiene una habilidad limitada en esa segunda lengua (Munro, Derwing, & Burgess, 2010); la existencia de estereotipos positivos asociados con ciertos acentos podría también tener consecuencias positivas para quienes los manifiesten en determinadas circunstancias. Entre las negativas, aún siendo inteligible el habla de alguien con

acento extranjero, puede haber intentos de evitar conversaciones con esas personas, no por prejuicios, sino por falta de confianza de hablantes nativos en su habilidad para entender habla con acento; también pueden influir estereotipos negativos, relacionándose con discriminación en empleo, hospedaje y servicios. (Munro, Derwing, & Burgess, 2010; piénsese, por ejemplo, en la situación de una persona con acento alemán en un territorio de ocupación militar nazi; acentos asociados a clases socioeconómicas bajas o a grupos discriminados, como hispanohablantes en algunas zonas de Estados Unidos). Este tipo de efectos, entre muchos otros, están basados tanto en la capacidad perceptual elevada de distinguir acentos extranjeros como en efectos de categorización social (si se desea ver algunos trabajos que ilustran consecuencias sociales y psicológicas de los acentos, así como el tipo de variables, como atributos de personalidad, credibilidad, atractivo y otras que se pueden inferir de los mismos pueden consultarse Cavanaugh, 2005; Colín Rodea, 1992; González-Álvarez, et al., 2003; Lev-Ari & Keysar, 2010; Méar-Crine & Leclerc, 1976; Miller et al., 2006; Moen, 2000; Pipitone & Gallup Jr, 2008; Ramos, 1997; Young, 2003).

En los acentos extranjeros hay variaciones tanto a niveles prosódicos como segmentales (Scott et al., 2006; y aunque algunos errores en sintaxis u otros niveles podrían influir, estos no son del dominio de la fonética y fonología). De éstas, las prosódicas han sido enfatizadas a partir de trabajos tanto con acentos naturales como en la literatura sobre síndrome del acento extranjero (González-Álvarez et al. 2003; Moen, 2006; Munro et al., 2010; Piske, MacKay & Flege, 2001)

El síndrome del acento extranjero es una alteración en la cual se dan alteraciones principalmente en la prosodia (González-Álvarez et al. 2003; Moen, 2006) que, aunque no propias del lenguaje original, sean posibles en algún lenguaje natural; de esta manera, se basa en gran medida en la percepción de terceras personas, que han de juzgarlo no como un habla patológica, sino extranjera (Blumstein & Kurowsky, 2006). Cabe destacar que en este síndrome el patrón de habla no se puede modificar a voluntad por lo que es muy difícil, o imposible, de evitar (González-Álvarez et al, 2003; Miller et al, 2006). También se ha puesto en

entredicho la existencia del síndrome como tal, ya que más del 60% de las veces se da en conjunto con otros síntomas como disartria, aprosodias o apraxias (Blumstein & Kurowski, 2006). Con todo, también pueden existir dificultades a otros niveles del lenguaje, como segmental y gramatical (González-Álvarez et al., 2003; Miller et al., 2006; Moen, 2006; Varley et al., 2006).

Tal como señalaron Grice y Baumann (2007), la literatura sobre acento extranjero ha hecho énfasis en la codependencia de los niveles segmental y suprasegmental de la lengua, señalando también que la distinción entre estos niveles no siempre es clara (Piske et al., 2001), ya que efectos en un nivel pueden tener influencia en el otro; por ejemplo, Piske y colaboradores (2001) enfatizan que programas de entrenamiento en idiomas centrados en el nivel suprasegmental tienen efectos positivos en el nivel segmental.

En el marco de la falta de distinción clara entre los niveles segmental y suprasegmental se puede observar nuevamente la influencia del marco teórico desde el cual se trabaja, como se verá a continuación.

En una aproximación específica a la fonología – la *fonología gestual*²² – existe un contraste fundamental con aproximaciones fonológicas basadas en el nivel segmental: estas últimas no hacen referencia al dominio temporal. Las aproximaciones basadas en el nivel segmental proponen un conjunto de entidades discretas que son ordenadas de manera serial para producir los mensajes (estas entidades son los segmentos), el tiempo sólo entra como noción en estas teorías, mientras que el cálculo del tiempo real, del tiempo físico, entra en el ámbito de la fonética o de modelos de control motor del habla; se ve, así, cómo el tiempo específico para los segmentos no está especificado. En contraste, la fonología gestual se basa en los movimientos articulatorios que varían en el tiempo, lo que implica necesariamente el dominio temporal; al estar basados en los movimientos articulatorios, las unidades en estos modelos no pueden ser entidades discretas, ya que existe (inherentemente en el proceso del habla) sobreposición articulatoria, por lo que las unidades difieren dependiendo del contexto donde se produzcan

²² *Gestural phonology* originalmente.

(Moen, 2006). En otras palabras, la distinción entre segmental y suprasegmental (así como entre fonética y fonología) que surge a partir de la concepción teórica es más marcada en teorías basadas en segmentos y más difusa en teorías como la fonología gestual. La aproximación como la de Moen (2006) realza la importancia de distinguir el dominio temporal en el que se da la prosodia (que discutimos en el capítulo anterior; por ejemplo entre el nivel de sílaba y de frase), ya que simplemente distinguir segmental vs suprasegmental es más difícil en esta teoría; a cambio, permite analizar con más detenimiento y especificidad diferentes atributos ya que el modelo – como otros modelos fonológicos no lineales – se basa en el uso de capas que representan diferente información (Moen, 2000). Si bien el modelo de Moen (2006) emplea diferentes capas para representar el movimiento y actividad de diferentes partes del aparato fonador, hay modelos que incluyen capas para contrastes tonales, rítmicos, entonacionales, etc. Modelos con estas características podrían ser útiles en el estudio de la prosodia; a la vez, se infiere que – de entre la multiplicidad de aproximaciones a la fonología que existen en la lingüística – la mayoría de la investigación que hemos revisado hasta el momento se basa en los modelos segmentales.

Otro concepto teórico importante que surge de la literatura de investigación en acento extranjero es el de las *configuraciones fonéticas*²³ (Moen, 2000). Este concepto se refiere a “la tendencia de diferentes partes del tracto vocal a ser sujetas a ajustes musculares específicos de largo plazo durante el habla (Moen, 2000). El concepto de configuraciones fonéticas (y como términos más específicos, configuraciones laríngeas y supralaríngeas) cabe dentro del concepto de cualidad de la voz. La cualidad de la voz, resultado de configuraciones de largo plazo, está distribuida a través de muchos segmentos (Munro et al., 2010).

Lo importante de este concepto para el estudio del acento extranjero es que configuraciones fonéticas específicas están relacionadas a idiomas específicos (Moen, 2000), y a la percepción del acento extranjero (Munro et al., 2010), y, para el objetivo de este trabajo, en general, si estos patrones musculares influyen en la

²³ Originalmente *phonetic settings*.

prosodia, y son influidos por la cultura, identificamos una contribución con un componente tanto biológico y cultural: ajustes musculares que dependen de la cultura.

3.5.1.1 Puntos importantes

En la presente sección se abordó brevemente el acento extranjero y su relación con la prosodia.

Específicamente en el marco de la *hipótesis de grados de libertad* que se ha planteado, el acento extranjero constituye un buen ejemplo de diferencias acústicas que son permisibles y que no afectan a las diferencias contrastivas requeridas por el sistema. Aunque no constituye una diferencia contrastiva, termina por capturar información sobre el origen de la persona. También vimos que algunos atributos del acento extranjero también se dan a nivel segmental y de atributos como la sintaxis. El análisis específico de qué tipo de errores en el habla extranjera son los que hablantes nativos “perdonan” más fácilmente nos podría dar una idea de que tan nucleares o no (que tanto forman parte de grados de libertad) diferentes atributos. Corros Mazón (2007) indica que hablantes de inglés que habían estudiado polaco como segunda lengua eran evaluados de manera más positiva si sus errores solamente eran errores de pronunciación que cuando había además errores de morfología y sintaxis. Cabe destacar también la propuesta que hacen Piske y colaboradores (2001) sobre la generación de vocales intermedias en hablantes de lenguas extranjeras; entre más diferentes fuesen, más probable sería que se formara una nueva categoría para los nuevos sonidos, mientras más similares, sería más probable que se formara una categoría compuesta entre los sonidos nuevos y los sonidos más similares de la primera lengua, con una forma final de articulación en algún punto de los dos extremos²⁴. Este proceso podría ser válido tanto a nivel segmental como suprasegmental, hipótesis que podría explorarse empíricamente.

²⁴ Por ejemplo, en el caso del fonema /ʁ/ del francés (la “r” francesa), podría ser asimilado como /x/ (la “j” mexicana, como en México o Jóven), y producirse de alguna manera intermedia entre estos dos sonidos.

La noción de configuraciones fonéticas podría ser particularmente útil al estudiar la cualidad de la voz (tal como ha sido abordada en capítulos previos), ya que las primera están basadas en patrones musculares mientras que la segunda no siempre está definida en dichos términos (pero recuérdese la aproximación de Gobl y Ní Chasaide, 2003).

Si hay una organización diferente para la prosodia asociada a funciones indécicas y aquella asociada a otro tipo de funciones es un tema que no ha sido estudiado. Los hallazgos sobre la organización cerebral del síndrome del acento extranjero podrían estar relacionados, pero dada la lejanía del vínculo con el tema principal de la tesis y cuestiones de extensión, no serán abordados.

La literatura sobre acentos y el síndrome del acento extranjero ilustra influencias biológicas y culturales a la prosodia que no caben propiamente ni en la prosodia lingüística ni en la prosodia emocional, aunque se acomodan muy bien a la categoría indécica de Van Lancker y colaboradores (2006). Esto llama la atención hacia la revisión del marco teórico de la división de la prosodia lingüística y emocional. La breve revisión sobre la aproximación de la fonología gestual también llama la atención hacia la importancia del marco teórico desde el que se trabaje. La prosodia, o algunos atributos de ella, pueden transmitir información más allá de distinciones lingüísticas y emocionales; tal como el origen geográfico, el estado físico, la salud, edad y otras que no serán abordadas en detalle (Erickson, 2006; varias referencias al respecto pueden ser encontradas en el trabajo de Van Lancker et al., 2006).

La prosodia también puede considerarse en algunas ocasiones (además de las evidentes en la prosodia lingüística) como una forma de comunicación intencional. Si bien las palabras en un idioma y sus reglas morfológicas y sintácticas están dadas, la prosodia, al no constituirse necesariamente como una diferencia contrastiva (al menos no en su totalidad, pues parte de ella entra dentro de los *grados de libertad*), permite adaptaciones personales; por ejemplo un habla más dulce, más “ruda”, más femenina o masculina, etc. De hecho, en la sección de desarrollo veremos como se puede diferenciar por el habla a niños de niñas, aún

cuando no hay diferencias significativas en el tracto vocal. Este tipo de atributos también puede ser marcado a nivel de palabras; desde las versiones más o menos simples, como la flexión de femenino para hablar de uno mismo en español, como otras más sofisticadas, por ejemplo la frecuencia de uso de diminutivos en español y portugués, a veces asociada a lo femenino, o incluso el uso de ciertas palabras, o partículas, como el empleo de の /no/ y ぞ /zo/ o ぜ /ze/ al final de una oración en japonés, también empleadas por mujeres u hombres, respectivamente. Este tipo de variables permiten comunicar matices y dan brindan una flexibilidad comunicativa, manteniendo el uso de un código compartido. Esto ilustra la importancia de los *grados de libertad* en la expresión oral.

3.5.2 Música y Prosodia

El interés en el estudio de la música y la prosodia (y el lenguaje en general) ha sido más aparente entre las personas dedicadas al estudio de la música que del lenguaje. Una de las relaciones que más fácilmente saltan a la vista es la de los atributos compartidos en ambas; el tono, volumen, ritmo, y cualidad de la voz son atributos compartidos entre la prosodia del lenguaje y la música (en el caso de cualidad de la voz, un equivalente en música es el *timbre*). Efectivamente son varios los autores que enfatizan estas similitudes (Ilie y Thompson, 2006; Patel & Daniele, 2003; Patel et al., 1998; Thompson, Schellenberg & Husain, 2004). Sin embargo, el sistema de la música no se limita a estas variables, e incluye otras como las escalas, el énfasis estructural, la armonía, el modo; variables que no tienen su equivalente en el habla (Juslin y Laukka, 2003; Patel et al., 1998), tal como el sistema del lenguaje no se limita a la prosodia.

Estas similitudes entre la prosodia y la música incluso han sido notadas desde las definiciones de prosodia que algunos autores utilizan; por ejemplo, la definición de Van Lancker y colaboradores (2006; "*prosody, the melody of speech is made up of fluctuations in pitch...*") y la de Thompson y colaboradores (2004) "*Speech prosody refers to the musical aspects of speech, including its melody (intonation) and its rhythm (stress and timing)*".

El interés marcado de los estudiosos de la música por las relaciones en el lenguaje está relacionado, quizá, con la hipótesis de que el lenguaje hubiera inspirado los orígenes de la música, y hubiera proveído algunos códigos – principalmente emocionales – para la expresión musical (Curtis & Bharucha 2010; Juslin y Laukka, 2003; Livingstone & Thompson, 2006; Thompson et al., 2004). Aunque también hay posturas que postulan que el lenguaje se originó de la música o que tanto música y lenguaje se desarrollaron de un sistema que presentaba características de ambos sistemas (Curtis & Bharucha 2010; Juslin y Laukka 2003; Livingstone & Thompson, 2006) o a partir de las vocalizaciones no lingüísticas (Scherer, 1995).

Se han realizado varios estudios que comparan similitudes entre la prosodia y la música tanto con intereses puntuales en la prosodia emocional (Ilie y Thompson, 2006) como en la prosodia lingüística (Patel et al., 1998), así como aquellos interesados tanto en las similitudes y diferencias en características acústicas y perceptuales (Ilie y Thompson, 2006) como en la organización funcional que subyace a música y prosodia (Patel et al., 1998).

Las aportaciones teóricas más importantes que se pueden derivar de este cuerpo de literatura para el tema de interés, son visiones alternativas en las que la organización funcional de la prosodia (e inclusive del lenguaje, en general) no son exclusivas para esta función (Levitin & Menon, 2003). A la vez, encontramos similitudes y diferencias con la música en las características acústicas de la prosodia emocional (por ejemplo las observaciones de Curtis y Bharucha, 2010, sobre el ajuste de la prosodia emocional para señalar tristeza a las formas de transmitirla estudiadas en música), como – de alguna manera – de la prosodia lingüística (por ejemplo en observaciones en las que tanto en música y prosodia se agrupan elementos en frases y los finales de estas frases suelen ir acompañados de una disminución en la velocidad, y/o en el tono Patel et al., 1998; Thompson, 2004).

A continuación se presentan sintéticamente algunos estudios sobre música y prosodia.

Levitin y Menon (2003) realizaron un estudio mediante la metodología de resonancia magnética funcional comparando la respuesta cerebral a música clásica y versiones modificadas para eliminar la coherencia temporal (pero respetando características como tono, volumen y timbra) de las mismas piezas musicales. Al comparar la activación cerebral asociada al procesamiento perceptual de ambas tareas encuentran activación en el área 47 de Brodmann, estructura asociada desde hace tiempo al procesamiento de estructura lingüística en lenguajes hablados y de señas, así como en la zona análoga en el hemisferio derecho. Los autores proponen que estas áreas podrían estar relacionadas al procesamiento de estímulos con estructuras finas que se desarrollan en el tiempo y tienen significados (en el caso de la música, significados generales dependientes de la estructura de la misma), y no exclusivamente con el procesamiento de estímulos lingüísticos, y tampoco requieren ser auditivos, como demuestra el involucramiento del área en lenguajes de señas. De manera más amplia, los autores consideran que esta área podría formar parte de un sistema perceptual más amplio que se dedica a analizar unidades perceptuales que forman parte de una estructura que se desenvuelve en el tiempo, permitiendo determinar qué aspectos, en ese flujo perceptual, tienen relación entre ellos. Los autores resaltan también la importancia del ordenamiento temporal de eventos común tanto a la música (notas, acordes, etc.) como al lenguaje (tanto en lenguajes orales como en lenguajes de señas); hallazgos previos sobre procesamiento similar de secuencias de acordes y lenguaje, e hipótesis propuestas de procesamiento común a la sintaxis de la prosodia y de la música también sugieren mecanismos comunes de procesamiento entre lenguaje y música; aunque hace falta más atención en las relaciones específicas entre la prosodia, como parte del lenguaje, y la música .

Patel y colaboradores (1998) realizan un estudio en el que comparan la capacidad de dos pacientes con amusia para discriminar prosodia lingüística (actos de habla, foco y fraseo) y patrones musicales derivados de la prosodia lingüística con el objetivo de explorar la relación cerebral en el procesamiento de estas dos funciones. Dadas las similitudes en las características acústicas entre la prosodia y la música, los autores consideran que es evidente que exista alguna conexión

entre ellas, en contraste con la posición general en la neuropsicología en la que el procesamiento de lenguaje y música es independiente en general (aunque recordemos el problema de las conceptualizaciones de lenguaje que discutimos anteriormente). Si bien es evidente la relación entre música y prosodia a niveles tempranos (ejemplo, la cóclea), y su disyunción más tardía (ejemplo, procesamiento semántico para el lenguaje y procesamiento de relaciones tonales para la música), los autores están interesados en explorar un nivel intermedio y no trivial en el que ambos procesos tengan organización funcional sobrepuesta. En ambos participantes, el desempeño fue similar para música y prosodia, lo que sugiere recursos cerebrales compartidos para los dominios prosódico y musical desde la perspectiva de los autores. Específicamente, los autores consideran que estos dos procesos podrían compartir los recursos neurales para mantener patrones auditivos en la memoria de trabajo, específicamente en la corteza primaria auditiva izquierda y la corteza prefrontal derecha.

Por su parte, Ilie y Thompson (2006) investigan la capacidad de la música y la prosodia de inducir emociones, en términos de *arousal* y *valencia*. Expusieron a sus participantes a extractos de 7 minutos de música o habla, manipulados en relación al ritmo, tono y volumen. Encuentran capacidad, tanto de la música como de la prosodia de inducir emociones (evaluadas mediante autoreporte): estímulos con volumen más alto se asociaron a mayor *arousal* y tensión; estímulos con ritmo más rápido fueron asociados a mayor *arousal* y energía; estímulos con un tono más alto fueron asociados a *valencia* más positiva. Los autores concluyen que, dado que sus resultados muestran que ambos tipos de estímulos fueron capaces de inducir emociones (y reportan otros en cuanto a la *percepción* de emociones), éstos indican que la música y la prosodia están fuertemente relacionadas psicológica y neuralmente, y que las experiencias afectivas están asociadas con los circuitos neurales compartidos entre ambas clases de estímulos; más aún, resaltan que características acústicas pueden tener un significado afectivo, procesado por áreas cerebrales que no diferenciarían el tipo de estímulo procesado; esto podría poner en entredicho que el procesamiento de la prosodia emocional sea procesamiento de la prosodia emocional y no procesamiento de

atributos acústicos en general para inferir significado afectivo. Aunque la visión contraria también podría resultar plausible: que el procesamiento de la música no sea procesamiento de la música sino procesamiento de características acústicas de manera general, que derivaría de la capacidad (de base evolutiva) de procesar la prosodia emocional. Esta segunda interpretación parece ser más adecuada.

También resaltan otros estudios (Juslin & Laukka, 2003; Gabrielsson & Lindstrom; citados en Ilie & Thompson, 2006; Livingstone & Thompson, citados en Ilie & Thompson, 2006) que estudiando de manera independiente a la música y a la prosodia han encontrado relaciones entre atributos como el tempo (ritmo, a grandes rasgos) , el tono y el volumen con las emociones tanto para prosodia (emocional) como para la música (ver también Caballero-Meneses & Ménez, 2010); con todo, señalan que aunque los hallazgos para música y prosodia han mostrado cierta superposición, no son idénticos; de hecho, en su estudio reportan un efecto del ritmo en la valencia para el habla, pero no para la música y una diferencia similar para la “energía”, variable que no discutiremos en detalle.

Thompson y colaboradores (2004) analizan en una serie de experimentos el efecto del entrenamiento en música en el desempeño en la identificación perceptual de prosodia emocional. Encuentran ventajas para las personas con entrenamiento musical para discriminar la intención emocional de frases neutras a a partir de la prosodia (tanto en el propio lenguaje como en un lenguaje desconocido) y secuencias de tonos basadas en la prosodia. Enfatizan que los resultados positivos del entrenamiento musical (teclado) en uno de sus experimentos son equivalentes a los resultados del entrenamiento en una disciplina que presumiblemente ayuda a entrenar específicamente la expresión de emociones en la voz (entrenamiento en drama). Los autores consideran que el entrenamiento en música podría ayudar a desarrollar y afinar las habilidades de asociación de tonos y patrones temporales con las emociones. También resaltan similitudes en la organización cerebral de la prosodia y la música, así como estudios que encuentran áreas de procesamiento específico para la música, pero no para la prosodia.

Por su parte, Curtis & Bharucha (2010) realizan una investigación basada explícitamente en las posturas sobre el origen de la música basado en el lenguaje. En música ha habido un estudio intensivo sobre las características que se pueden emplear para transmitir ciertas emociones; los autores buscan evidencia acústica y perceptual en producciones bisilábicas para determinar si las características empleadas para transmitir emociones corresponden con aquellas encontradas en música. Su hallazgo principal es el uso de la tercera menor (se refiere a las diferencias entre notas de una escala musical, denominadas intervalos, y pueden ser definidos como proporciones entre frecuencias fundamentales) en la voz para comunicar tristeza; también encuentran relación entre otros intervalos propuestos en música con el enojo, con emociones negativas en general (valencia negativa) y con emociones positivas de manera general (valencia positiva). De manera interesante, los intervalos en la voz no son exactos, en contraste con los intervalos en música, basados en notas precisas, pero la percepción parece ajustarse categóricamente al intervalo musical más cercano. Los autores consideran sus hallazgos como evidencia a favor de un código acústico compartido para comunicar tristeza (y emociones en general) entre música y prosodia. Los autores interpretan la asociación de la tercera menor y la tristeza como un código que podría no ser arbitrario, sino que las respuestas fisiológicas asociadas a la tristeza que influyen la producción vocal podrían incrementar la probabilidad de ocurrencia espontánea de la tercera menor.

Patel y Daniele (2003) exploran si el lenguaje de los compositores puede influir en sus composiciones. Comparan medidas rítmicas asociadas a la música compuesta por autores franceses y británicos (cuyas lenguas difieren ampliamente en sus patrones rítmicos: el francés es una lengua de ritmo basado en sílabas y el inglés lo es de ritmo basado en estrés), así como medidas rítmicas de los lenguajes. Encuentran diferencias significativas en las comparaciones de ritmo en música y en lenguaje que muestran la misma dirección.

Finalmente, presentamos el trabajo de Juslin y Laukka (2003). Los autores realizan una revisión extensiva sobre trabajos que exploran la influencia de

variables acústicas en la expresión de emociones de manera vocal, por un lado, y en la música, por otro, buscando similitudes y diferencias. Resaltan la llamada *Ley de Spencer*: la perspectiva de que las emociones influyen los procesos fisiológicos, lo que a su vez influye las características acústicas de la voz y el canto. Encuentran patrones comunes para la expresión de emociones en ritmo de habla/tempo; intensidad de la voz/nivel de sonido (y su variabilidad); y frecuencia alta (en los casos en los que se dan los términos divididos por una diagonal, el primer término se refiere a la voz y el segundo a su análogo en la música). Sin embargo, también encuentran que algunas características difieren; por ejemplo, aunque el miedo se asocia a alta intensidad en el habla, se asocia a volumen bajo en música. También encuentran relación de la regularidad y variabilidad de los parámetros con las emociones: las emociones positivas están asociadas a mayor regularidad (menor variabilidad) mientras que el patrón opuesto se observa para las emociones negativas. También resalta estudios neurológicos en los que se ha visto que los contornos tonales en la voz y la música se procesan en las mismas áreas. Concluyen enfatizando que la pista más importante es el ritmo/tempo, en especial en medidas de promedio y no de variabilidad. Finalmente, consideran que parte del gusto por las ejecuciones en los instrumentos musicales se debe a su capacidad de sobrepasar los límites de la articulación de la voz, siendo percibidas en general como *voces superexpresivas*.

3.5.2.1 Puntos importantes

Esta sección nos ha permitido resaltar perspectivas en las que el procesamiento de la prosodia (y probablemente del lenguaje en general), es llevado a cabo por áreas dedicadas al procesamiento general de sus componentes. Es interesante la propuesta de que el lenguaje y la música compartan orígenes y que los códigos, principalmente afectivos, de la música pudieran estar derivados de la prosodia.

Aún más interesantes son las observaciones de similitudes de la música con la prosodia lingüística, así como áreas compartidas de procesamiento de música y lenguaje, en general. En la introducción adelantamos que el estudio de la música y la prosodia brindaba una oportunidad única en la que se podría aislar la prosodia emocional de la prosodia lingüística, ya que ésta última debería ser inexistente en

la música. Si bien funciones lingüísticas como la distinción léxica son inexistentes en la música, hemos visto como algunas funciones, como el fraseo, tienen algunas analogías en la música. Queda claro que la música y la prosodia tanto lingüística como emocional, comparten algunas características, por lo que el estudio conjunto de estas dos variables podría ser fructífero en el futuro. Al respecto de esto último, Curtis y Bharucha (2010) enfatizan la adecuación de la escala musical para medir la prosodia (ya que está basada en una escala logarítmica, tal como el sistema perceptual para tonos) y su potencial para hacer que estudios en los ámbitos prosódico y musical sean comparables.

La existencia de códigos comunes con la música parece ser más evidente para la prosodia emocional; y cabe considerar que la música, como los lenguajes, tiene evidentes influencias culturales, por lo que es entendible la dificultad de encontrar códigos comunes entre prosodia lingüística y música, pero hace aún más intrigantes los hallazgos sobre similitudes y áreas cerebrales compartidas para su procesamiento.

Efectivamente, algunos autores han resaltado algunas aplicaciones clínicas que derivan de la literatura revisada en esta sección; Curtis y Bharucha (2010) consideran que sus observaciones sobre la relación de la tercera menor con la expresión prosódica de la tristeza podría implicar que la identificación espontánea de la tercera menor en el habla de personas podría ser un indicador de depresión, permitiendo hacer una evaluación tanto del estado emocional como de una posible psicopatología, sin hacer uso de medidas de autoreporte. También Patel y colaboradores (1998) enfatizan la necesidad de probar dificultades perceptuales relacionadas con los cuadros clínicos propuestos; por ejemplo, antes de diagnosticar *amusia pura*, se deberían de hacer pruebas para descartar dificultades en campos como la prosodia.

Otra posible aplicación, fuera de la psicología del estudio conjunto de música y prosodia y considerando la hipótesis del desarrollo de la música a partir de la prosodia, es la posibilidad de hacer programas para composición musical expresiva basados en los conocimientos en prosodia. De la manera inversa,

también muchos de los estudios en música sobre la forma de transmitir emociones, a veces con distinciones sutiles, podrían ser útiles para desarrollar expresividad emotiva en prosodia tanto con miras a los campos de actuación como de retórica.

Si bien no se asume un componente semántico en la música, algunos autores han discutido sobre la semanticidad en la misma (Livingstone & Thompson, 2006). A diferencia de la semanticidad en el lenguaje, proponen que podría tener un significado no consensual. También ha sido propuesto que la música transmite un significado general sobre las emociones en forma dinámica, no verbalizable (Levitin & Menon, 2003). Esto se relaciona con las propuestas de Juslin y Laukka (2003) y Curtis y Bharucha (2010) sobre la capacidad de la prosodia para distinguir emociones dada su urgencia, lo que impide distinguir diferencias sutiles pero permite una excelente discriminabilidad entre categorías distintas.

Tanto Juslin y Laukka (2003) como Scherer (1995) hacen varias menciones sobre el canto y la forma de transmitir emociones en él, relevante para la perspectiva de *grados de libertad* en al menos dos sentidos. Primero, si bien en las partituras se especifican el tono, el ritmo, entre otros aspectos acústicos que deben ser dados en una interpretación, los intérpretes aún disponen de algunos grados de libertad (por ejemplo, cualidad de la voz) para transmitir la intención deseada, permitiendo imprimir su estilo propio sin violar los requerimientos de la pieza. En este sentido es particularmente importante la observación de Juslin y Laukka (2003) de la redundancia del “significado” de las características acústicas, a pesar de que se pierda la capacidad de variar una característica, se pueden hacer distinciones mediante otra. Segundo, al ser el canto un uso no cotidiano, la comparación de las posibilidades de variación en el canto en los diferentes lenguajes puede hablar sobre los grados de libertad, lo que a su vez permitiría identificar las características de la señal acústica requeridas por el sistema; en otras palabras, en el canto se pueden variar gran cantidad de atributos pero se deben respetar los esenciales para transmitir el significado. Una manera de ilustrarlo es la posibilidad de alargar un sonido consonántico (aunque sonoro y relativamente cercano a los

sonidos vocálicos), el sonido /n/ en las canciones japonesas, pero la imposibilidad de hacerlo en el idioma español: en japonés bien se puede cantar /se:ka.n:::/, mientras que sería imposible cantar /kan'sjo.n:::/ en español, sólo siendo posible la forma /kan'sjo::n/. Asimismo, puede suponerse que en los cantos de lenguas tonales no se puede afectar el tono léxico, ya que esto cambiaría el significado. Asimismo, en el campo del canto se ha hecho énfasis en el control de variables de calidad de la voz, por ejemplo, al desarrollar diferentes impostaciones, como el falseto y también el uso de recursos expresivos como el *vibrato* y el *staccato*; esto podría ayudar también al campo prosódico tanto para la codificación de estímulos por parte de personas con experiencia en la variación de aspectos vocales como en el intercambio teórico entre las dos áreas. Scherer (1995) apoya la posibilidad de acercar el estudio del canto y el habla en el futuro.

Precisamente por la redundancia en la información que transmiten diferentes atributos acústicos que señalan Juslin y Laukka (2003) recomiendan el uso de la metodología de síntesis para poder variar ortogonalmente los atributos y desentrañar su contribución a la percepción conjunta.

Otras consideraciones importantes, breves, incluyen la observación de Livingstone y Thompson (2006), que consideran a la música, junto con el lenguaje, como un sistema simbólico. Así como la propuesta de Thompson et al. (2004), que consideran que los procesos compartidos de interpretación de la información emocional en prosodia y música podrían ser compartidos aún por otras modalidades, como la visual. Efectivamente, en una sección posterior analizaremos la prosodia en el lenguaje de señas.

Finalmente cabe señalar la discusión que presentan Juslin y Laukka (2003) sobre la relación de los parámetros acústicos con las emociones. Según los autores se puede interpretar de 5 maneras: 1. – que la relación es simplemente una coincidencia; 2. – que las similitudes se deben a una tercera variable que influye tanto a la voz como a la música; 3. – Que los hablantes basan su expresión vocal de emociones en la forma en la que los ejecutantes hacen la música. 4. – Que tanto la música como el lenguaje evolucionaron de forma paralela sin que uno

precediera al otro, y 5. – Que los músicos basan sus ejecuciones sobre la base de la expresión vocal de las emociones. Los autores brindan contraargumentos para las cuatro primeras interpretaciones y apoyan la última de ellas.

3.5.3 Lenguajes Tonales y Prosodia

El término de *lenguajes tonales* se ha empleado tradicionalmente para referirse a aquellos lenguajes que utilizan el tono para hacer distinciones léxicas. Como muchos lenguajes lo utilizan para hacer distinciones gramáticas ha sido propuesto que se consideren lenguajes tonales a aquellos que utilizan el tono para hacer tanto distinciones léxicas como gramáticas (Goldsmith, 1994). Éste término se contrapone contra los términos “lenguajes de entonación” (como el español y el inglés), donde el tono es sólo una característica postléxica relevante sólo a nivel de frase y a las llamadas “lenguas de acento tonal”, donde algunas sílabas requieren ser acentuadas mediante un tono específico, en contraste con las lenguas tonales donde casi cada sílaba requiere un contraste tonal (Grice & Baumann, 2007). Con todo, se ha discutido sobre la dificultad de separar las lenguas tonales de las lenguas de acento tonal (Grice & Baumann, 2007), y en general las lenguas tonales de otro tipo de lenguas (ver Duanmu, 2004 para críticas y una teoría alternativa a la división tipológica). El origen de los tonos en los lenguajes tampoco es claro. Como otras características de los lenguajes, varían de manera histórica, pudiendo surgir nuevos tonos o perderse tonos en el curso de la evolución del lenguaje. Abramson (2004) discute sobre la plausibilidad de diferentes teorías del origen de los tonos.

Las lenguas tonales abundan en la región sudeste de Asia, así como en el oeste y sur de África, en América existen lenguas tonales en México (mazateco, otomí, tlapaneco, trique, zapoteco), en la India también existen diversas lenguas tonales (Narasimhan, Litha, B. Shahid, & Rohini, 2010).

Como se revisó anteriormente, el procesamiento del tono a nivel de sílaba (contraste léxico) y de frase parece estar lateralizado diferencialmente, principalmente al hemisferio izquierdo y derecho, respectivamente (Gandour et al., 2004; Gandour, Dziedzic, Wong, Lowe, Tong, Hsieh, Sathamnuwong y Lurito

2003; Narasimhan et al., 2010; ver también Chen, Xue, Mei, Chen, & Dong, 2009). Este tipo de contrastes no ocurren en lenguajes no tonales, y la evidencia de diferente lateralización de la prosodia lingüística dependiendo de su función (o el dominio temporal) es importante para el estudio de la prosodia.

La principal utilidad de estudiar la prosodia en lenguas tonales para el objetivo del presente trabajo está centrada en el apoyo y ampliaciones que puede brindar a la *hipótesis de grados de libertad* y al posible contraste con algunos hallazgos sobre la prosodia, ya que la mayoría de la investigación en el campo ha sido realizada utilizando lenguas de entonación.

Narasimhan y colaboradores (2010) exploran las características acústicas del *stress* en mizo, un lenguaje tonal de la India. Encuentran que es señalado y percibido mediante el aumento de la duración de la palabra o sílaba. La duración del tono es la principal pista perceptual para señalar *stress* en esta lengua, de acuerdo a lo encontrado en otras lenguas tonales. En otras lenguas puede ser señalado, además, mediante la intensidad o el tono (por ejemplo, en inglés y francés la principal pista perceptual para señalarlo es F0; aunque en lenguas no tonales, como el italiano, la principal, aunque no única, pista perceptual que señala el *stress* también es la duración), pero estos parámetros no son tan libres en lenguas tonales. Esto ilustra que las características acústicas para señalar *stress* son dependientes del lenguaje (influencia cultural), y está de acuerdo a la *hipótesis de los grados de libertad*, ya que en estas lenguas, dado que los tonos son requeridos para distinciones léxicas, el *stress* se manifiesta principalmente mediante la duración.

Yang (2010) explora la interacción entre prosodia y diferentes funciones de marcadores de discurso en mandarín, implicando procesos cognitivos y emocionales en el hablante, el contexto y la interacción con el interlocutor. Analiza las características prosódicas de dos marcadores del discurso comunes en chino: *ranhou* (equivalente a luego, en sus usos temporal y como marcador de consecuencia lógica) y *jiushi* (equivalente a “esto es” y/o “es decir”). *Ranhou* ocurre en puntos donde los hablantes están siguiendo el discurso sobre el mismo

tópico. Los autores encuentran que el grado de incertidumbre sobre el grado de entendimiento mutuo entre los dos interlocutores y su función (temporal o como marcador de consecuencia lógica) influye en la prosodia con la que es dicho: cuando el hablante asume un buen nivel de entendimiento mutuo entre él y su interlocutor, es dicho de manera muy rápida y con una articulación de los tonos poco marcada, e implicaría poca dificultad cognitiva en presentar ideas en conexión; cuando es utilizado en su forma de conector temporal, la duración incrementa y hay una articulación más marcada de los contrastes tonales requeridos; además, con respecto a este último caso, y dado que el patrón tonal requerido en la segunda sílaba de *ranhou* es una caída de tono, esta caída léxica parece interactuar con una caída entonacional para transmitir certeza, originando un tono que cae de manera aún más marcada. En lo que toca al marcador *jiushi*, liga explícitamente una idea anterior a la idea que está a punto de ser presentada, es decir, se asume que la oración que sigue a *jiushi* hará más específica una idea anterior; en el discurso surge cuando la idea presentada se considera incompleta o inadecuada y requiere una resolución. Al igual que en el caso de *ranhou*, al estar asociada a certidumbre por parte del hablante es de duración corta y poca articulación tonal (de hecho la segunda sílaba puede presentarse plana, en lugar de una caída de tono requerida habitualmente), funcionando principalmente como un marcador de continuación; las ocurrencias de *jiushi* más largas están asociadas a un contenido más pragmático o semántico, y se asocian a una articulación más precisa de los contrastes tonales.

El autor resalta la utilidad de estudiar lenguajes tonales para probar hipótesis sobre los atributos universales que han sido propuestos para las lenguas. También resalta que preguntas como "¿De qué manera un lenguaje tonal, como el mandarín, comunica estados emocionales y actitudes?", "¿Cómo señala prominencia y estructura del discurso?" son temas importantes para investigar, ya que son funciones comunicadas prosódicamente en otros lenguajes y el mandarín parece tener restricciones en el movimiento del tono, a causa de los tonos léxicos.

Por otra parte, nota que en mandarín, una articulación más pronunciada de los tonos léxicos sirve para llamar la atención y enfatizar el significado léxico (en este caso de los marcadores de discurso), y que esto representa una oportunidad expresiva – más que una limitación – para el uso de la prosodia en lenguajes tonales de la que carecen otras lenguas. Enfatiza que en mandarín, si bien los tonos léxicos restringen algunos patrones prosódicos, proveen de oportunidades para la prosodia mediante el uso de contrastes más grandes de lo normal; cuando además los tonos coinciden con significados universales de la prosodia (por ejemplo la convergencia entre una caída de tono léxica y una caída de tono para señalar certidumbre), el efecto es más marcad. Por otra parte, un significado universal podría interferir con la realización completa de un tono léxico; por ejemplo, para señalar continuación se ha propuesto la utilización de tonos altos, por lo que en casos en los que co-ocurre un tono léxico de caída y un tono alto de continuación, podría ser forzado un tono alto en la realización prosódica. El autor concluye que este tipo de observaciones ilustran que el uso de tonos en lenguajes como el mandarín no se mantiene de manera rígida, sino que fluctúa de acuerdo a influencias prosódicas universales y varía dependiendo de usos pragmáticos, intencionales, estados cognitivos y características de la interacción.

Siguiendo en la línea de la importancia del estudio de lenguas tonales para probar hipótesis sobre características prosódicas universales propuestas, Yuen (2007), analiza la interacción de los tonos léxicos en cantonés con el fenómeno universal de la declinación (tendencia a la caída de tono a lo largo de las frases), fenómeno documentado tanto para idiomas de entonación como de acento tonal y tonales. Los autores proponen que los oyentes “construyen” una línea base sobre la declinación, de modo que el mismo tono absoluto puede ser percibido como distinto dependiendo de su posición en la frase: para que dos tonos en diferente posición sean percibidos como el mismo, el tono más tardío debe ser más bajo que el primer tono. También los tonos temporalmente cercanos influyen en la interpretación de los tonos adyacentes. Los autores concluyen que los escuchas compensan los efectos de la declinación al asignar un cierto tono real a un tono léxico categórico.

Liang y Van Heuven (2007) exploran la percepción de tonos léxicos por aprendices de Chino (dialecto de Beijing), y evalúan si los antecedentes lingüísticos de los aprendices (hablantes nativos de lenguas tonales o no tonales) influyen en su capacidad para distinguir tonos léxicos y prosodia a nivel de frase. Encuentran que los hablantes nativos de lenguas no tonales tienen dificultades en la tarea de distinción de tonos léxicos en contraste con hablantes de lenguas tonales (tanto chino, como otras) y que, de manera interesante, los hablantes nativos de lenguas no tonales superan en desempeño tanto a aprendices con lengua materna tonal y a los propios hablantes nativos de Chino. En otras palabras, f_0 es percibido principalmente a nivel léxico por hablantes de lenguas tonales y a nivel de frase por hablantes de lenguas de entonación. Dado que la misma dimensión es empleada para distinciones léxicas y de entonación, esto tiende a disimular la presencia de la dimensión de la entonación en chino, al grado de que a veces se discute la propia existencia de la entonación en este idioma. Por ejemplo, para distinguir declaraciones y preguntas en chino, los tonos léxicos interactúan con la entonación. De este modo, dependiendo del tono léxico al final de las frases interrogativas, el contorno final cambia: si el tono léxico final es de subida la pregunta terminará en un contorno que sube, mientras que se obtendrá un contorno de bajada cuando el tono léxico final baja; esto contrasta con idiomas como el español, donde las preguntas, por regla general, suben al final. Por otra parte, las interrogaciones en chino se señalan mediante el uso de un registro tonal más alto en general (Gandour et al., 2003).

Resalta que, en las lenguas del mundo, existe una tendencia al uso ya sea de tonos en las palabras o entonación a lo largo de la frase; de este modo, en una lengua como el chino, que tiene un inventario importante de tonos léxicos, se utilizará en menor medida la variable tonal de la prosodia para señalar contrastes a nivel de frase (además, cabe destacar la existencia de una partícula de pregunta, *ma*, para señalar interrogaciones además de las diferencias prosódicas). Así, resalta el autor, ha sido propuesto que si una cierta dimensión fonética es explotada por la fonología para codificar un aspecto de la gramática no será utilizada en la misma medida para codificar otro aspecto. Esto sería válido no sólo

para el caso de los tonos sino otras dimensiones; por ejemplo, en lenguas donde vocales largas tienen una función contrastiva (como el japonés), la duración no puede ser manipulada con tanta libertad. Nótese la similitud de esta perspectiva con la *hipótesis de los grados de libertad*. Si bien la propuesta original (tal como se presentó en el capítulo de prosodia emocional) se centraba en el uso de los aspectos no contrastivos para transmitir la prosodia emocional, este tipo de observaciones extienden la aplicabilidad de la propuesta, en el sentido de que características contrastivas en dos dominios distintos (por ejemplo, léxico y de distinción de actos de habla), también requieren hacer uso de las características fonéticas libres, no utilizadas por el otro dominio.

3.5.3.1 Puntos importantes

La presente sección ha permitido explorar la prosodia en lenguas tonales, que tienen características propias y diferencias importantes con respecto a otras lenguas donde se ha hecho la mayoría de la investigación en prosodia (principalmente lenguas de entonación; Grice & Baumann, 2007). También permite apoyar y extender la *hipótesis de grados de libertad*. Efectivamente, varios de los trabajos revisados dejan ver nociones similares a dicha hipótesis, y también similares a la *hipótesis de la carga prosódica* presentada en el capítulo de prosodia emocional y lingüística.

La lateralización diferente del procesamiento del tono para su función léxica y entonacional remarca la importancia de distinguir estos dos niveles al abordar, de manera general, a la prosodia.

Asimismo, se ha resaltado también la interacción de los tonos con la prosodia emocional (Pell, 1999a)

El estudio de Yang (2010) resalta contribuciones presumiblemente psicológicas a la prosodia, al relacionar los grados de certidumbre y facilidad de procesamiento cognitivo con diferentes marcadores de discurso y sus características prosódicas, con todo, la aproximación hacia los términos de certidumbre y procesamiento cognitivo es muy laxa. Se requieren otros estudios con una definición más precisa

y mediciones psicológicas para poder evaluar la influencia de procesos psicológicos a la prosodia.

Finalmente, estudios como el de Glanz (1999) y Delogu, Lampis y Olivetti Belardinelli (2006) ilustran la conexión entre la prosodia y los temas más “alejados” que presentamos. Glanz reporta conexión entre el desarrollo de oído absoluto y los lenguajes tonales y Delogu y colaboradores reportan una mayor capacidad de aprendizaje de lenguas tonales por hablantes nativos de lenguas no tonales relacionada a su habilidad melódica. Ambos estudios muestran en conjunto, efectos en ambas direcciones entre las habilidades musicales y prosodicas en el contexto específico de lenguas tonales.

3.5.4 Prosodia en Lenguajes de Señas

En esta sección se revisará la prosodia en lenguajes de señas. Primero se abordará de manera resumida la organización funcional de los lenguajes de señas y posteriormente se comentarán algunos hallazgos específicos sobre la prosodia en estos lenguajes. El estudio de lenguajes de señas es importante considerando posturas en las que se ha teorizado que una comparación de las áreas que se emplean en el lenguaje oral y en el lenguaje de señas ayudará a especificar cuáles son las áreas “fundamentales” del lenguaje, independientes de la modalidad en que se expresen; un razonamiento análogo aplica al caso específico de la prosodia.

Los lenguajes de señas comparten las características primordiales de los lenguajes humanos, difieren únicamente en la modalidad a través de la cual se expresan (Newman, et al., 2010). También son susceptibles de ser afectados por lesiones cerebrales, resultando en afasias (Carlson, 2006a; Rosenzweig & Leiman, 1992). No son un sistema de símbolos universal, como algunas veces se había sugerido, ya que existen muchas variedades y presentan todos los niveles de otros lenguajes: fonología, morfología, sintaxis, semántica, pragmática, inclusive prosodia (Bavelier et al., 1998; Dachkovsky & Sandler, 2009). Por supuesto, en estos lenguajes el nivel fonológico no se refiere al sonido, sino que

se manifiesta por la posición y forma de las manos relativas al cuerpo²⁵. Asimismo, son sujetos a procesos de cambio lingüístico y a contrastes sociolingüísticos, como los lenguajes orales. Cabe resaltar que personas que adquieren una segunda lengua de signos, presentan un acento extranjero (Hickok et al., 2001). A partir de esto se puede observar que son lenguajes, en toda la extensión del término, y que difieren principalmente en la modalidad que se emplea para la comunicación.

Si bien es evidente el hecho de que la modalidad en la que se produce y recibe el lenguaje de señas difiere del lenguaje hablado, cabe preguntarse si la organización cerebral también difiere. En general, el hemisferio derecho procesa con ventaja la información espacial, también aventaja al izquierdo en el procesamiento de movimientos del cuerpo, posiciones y percepción de rostros (componentes importantes en lenguajes de señas). Recordemos que el hemisferio derecho también muestra una ventaja en el procesamiento de información prosódica, pragmática y de coherencia entre frases en lenguajes orales. Siendo esto así, se podría suponer que parte del área equivalente a Wernicke en personas que emplean lenguajes de señas estaría en el hemisferio derecho. Sin embargo, se ha encontrado que lesiones en las regiones donde se localizan las áreas de Broca y Wernicke en personas que hablan un lenguaje oral producen dificultades análogas en personas que emplean lenguajes de signos: A saber, una lesión en el lóbulo frontal, en el área motora suplementaria, produce dificultades de articulación de signos (y, como en la afasia de Broca, no se debe a dificultades periféricas de control motor ni a problemas generales de capacidad de movimiento), mientras que una lesión en el lóbulo temporal provoca dificultades en la comprensión del lenguaje de signos (Campbell et al., 2008).

En cuanto a la producción de los lenguajes de señas, se muestra – como en la producción de lenguajes orales – una marcada lateralización hacia el hemisferio izquierdo y en general patrones generales similares a los que se encuentran en

²⁵ Si bien “fonológico”, incluso etimológicamente, refiere al sonido, se ha empleado como término en los lenguajes de señas en atención a mantener los términos para referirse a niveles del lenguaje: fonología, morfología, sintaxis, semántica, pragmática.

lenguajes orales (Campbell et al., 2008; Kitson, Frackowiak & Frith, 1997; McGuire, Robertson, Thaker, David, Kitson, Franckowiak & Frith, 1997).

En lo que toca a la percepción de lenguajes de señas, se encuentran estudios más diversos. Se ha propuesto que los lenguajes de señas requirieran más de la participación del hemisferio derecho, en comparación con los lenguajes orales por el involucramiento de éste último en el procesamiento de información espacial. Efectivamente, se ha observado una mayor participación del hemisferio derecho al procesar lenguajes de signos (Bavelier et al., 1998). Sin embargo, también se ha observado que algunos pacientes se pueden seguir comunicando en lenguaje de signos aún con lesiones en hemisferio derecho que tengan consecuencias importantes en las habilidades espaciales. Por ejemplo, pacientes que presentaban heminegligencia pero podían seguirse comunicando en lenguaje de señas usando las áreas negligidas, o personas con grandes dificultades para copiar un dibujo, pero pudiendo comunicarse efectivamente en lenguaje de señas (Bavelier et al., 1998; Hickok et al., 2001). Esto supone una participación particular del hemisferio derecho en los lenguajes de señas, ya que las áreas que participan en ellos parecen ser hasta cierto punto independientes de las que procesan otro tipo de información espacial y otro tipo de información, como los movimientos de significación biológica y de reconocimiento de formas (Bavelier, Corina, & Neville, 1998).

Análisis más finos de los problemas en lenguajes de señas (realizados principalmente en el lenguaje de señas americano) secundarios a lesiones en el hemisferio derecho sugieren problemas análogos a los que se presentan en lenguajes orales, tales como dificultades de mantener un discurso organizado; además, dadas las características gramaticales del lenguaje de señas americano, donde al hablar de varios personajes se requiere indicar en el espacio diversos puntos (que simbolizan a los personajes), se ha encontrado que a veces se presenta dificultad de mantener coherencia en estas locaciones (Hickok et al., 2001).

En cuanto a la estructura cerebral, se ha encontrado que aunque los cerebros de personas que oyen y que no oyen no parecen diferir en cuanto al volumen de sustancia gris, podrían existir ligeras diferencias a nivel de sustancia blanca. Las regiones que conectan áreas de audición y perisilvanas en personas que oyen parecen ser algo más pronunciadas. (Emmorey et al., 2003). En dicho estudio compararon personas que oían contra personas con sordera congénita; todos diestros. Encuentran además que el volumen de materia gris en los lóbulos temporales no difiere entre los dos grupos, lo que indica que la privación de información sensorial auditiva no lleva a una poda neuronal en estas áreas, y que en ambos grupos se encuentra un plano temporal más pronunciado en el hemisferio izquierdo, lo que indica que esta característica morfológica no depende de la experiencia auditiva.

Por su parte, MacSweeney y colaboradores (2002) estudian la activación cerebral que se presenta en personas expuestas a lenguaje de señas y lenguaje oral (este último presentado en video, por lo que también se hace uso de la información visual). Encuentran que, en general, no hay diferencias en los patrones de activación cerebral. En ambos se muestra actividad en las áreas de Broca y Wernicke. Sólo se encuentran diferencias que atribuyen a la modalidad de la presentación del lenguaje: más actividad en el lóbulo temporal para quienes emplean lenguaje oral y más actividad occipito-temporal (área V5) para quienes emplean lenguaje de señas. Los mismos autores, al hacer comparaciones en lenguaje de señas entre personas bilingües/bimodales (que emplean lenguaje oral y de señas) y personas sordas que solo empleaban lenguaje de señas, encuentran que las personas sordas muestran una mayor activación en el giro temporal superior izquierdo. Lo discuten en términos de que esta área podría estar predispuesta al procesamiento del lenguaje en modalidad sonora, pero que al no existir esta modalidad en personas sordas, podría tener la capacidad de procesar información visual.

Todo esto revela una organización del lenguaje similar entre los lenguajes de señas y orales. Aunque hay matices distintivos de activación y organización en

cuanto a la percepción de los lenguajes en estas dos modalidades, éstos podrían reflejar patrones que no son propiamente lingüísticos – como la actividad diferencial en corteza occipito-temporal mencionada –, sino sólo diferencias de percepción dependientes de la modalidad empleada para el intercambio de la información y no necesariamente del procesamiento lingüístico como tal.

También hay observaciones sobre la división prosódica de las oraciones, sobre el énfasis que se da a algunos signos, y sobre el empleo de “tonos” contrastivos (Dachkovsky & Sandler, 2009). Por supuesto, todo presentado en la modalidad visual.

Algunos de los parámetros prosódicos en la modalidad visual que se emplean incluyen la mirada, las acciones de articuladores no manuales, movimiento de la cabeza y expresión facial. Por ejemplo, en el lenguaje de señas americano, se ha observado un levantamiento de cejas para marcar el cambio entre tópicos y para indicar preguntas dicotómicas (sí/no), bajarlas para preguntas de la *serie wh-*, o un gesto similar a un asentimiento con la cabeza al final de frases o para enfatizar (Tyrone, Nam, Saltzman, Mathur, & Goldstein, 2010). Nótese que estas funciones corresponden a aquéllas discutidas anteriormente para la prosodia en lenguajes orales.

Newman y colaboradores (2010) analizan si la organización neuroanatómica de este tipo de recursos expresivos del lenguaje es similar entre personas que emplean lenguaje de signos y personas que emplean lenguajes orales. Para ello, presentaron videos de frases en lenguaje americano de señas con marcadores prosódicos, enfáticos y emocionales (registro narrativo), y los compararon contra las mismas frases producidas con muy pocos de estos elementos (registro no narrativo). Mantuvo los marcadores faciales de preguntas, tópicos y adverbiales que son requeridos por este lenguaje de señas. Aunque hubo ligeras variaciones en el orden de los signos y algunas impuestas por las variaciones estilísticas en los dos registros (narrativo y no narrativo), los autores mantuvieron constante el número de referentes y la complejidad sintáctica. Reportan que ambas condiciones muestran una fuerte activación de las áreas clásicas relacionadas con

el lenguaje del hemisferio izquierdo como las áreas de Broca y Wernicke, el surco temporal superior y las zonas análogas en el hemisferio derecho, pero, además, encuentran una mayor actividad en el surco temporal superior y giros frontales inferiores derechos en la condición que incluía características narrativas del discurso. Este patrón de activación es similar al que se encuentra en manipulaciones análogas en lenguajes orales (Kotz et al., 2003; Newman et al. 2010). También encuentran activación de áreas relacionadas con la percepción de rostros y movimiento biológico, lo que en opinión de los autores puede reflejar un papel específico de estas áreas en el procesamiento lingüístico de personas que usan lenguajes de señas como primera lengua.

Corina, Bellugi y Reilly (1999) exploran las diferencias entre los usos afectivos y lingüísticos de las expresiones faciales en los lenguajes de señas. Tienen un interés específico en el control neural de funciones lingüísticas y no lingüísticas. Encuentran activación asociada al hemisferio derecho para las expresiones faciales emocionales y activación asociada al hemisferio izquierdo para las expresiones faciales lingüísticas. Los autores enfatizan que los resultados son particularmente importantes considerando que el sistema muscular que sirve a ambas funciones es el mismo. También hacen referencia al procesamiento facial en personas que no utilizan lenguajes de signos, lateralizado principalmente al hemisferio derecho; en contraste, en personas que sí emplean lenguajes de signos, el procesamiento se asocia diferencialmente a los hemisferios dependiendo de la función lingüística o afectiva en clara analogía con lo que hemos presentado para la prosodia emocional y lingüística en lenguajes orales, aunque cabe destacar que el estudio de estos autores no fue específicamente sobre expresiones faciales prosódicas.

Nespor y Sandler (1999) exploran similitudes y diferencias entre la prosodia en lenguajes de señas y lenguajes orales. Encuentran la presencia de constituyentes como frase fonológica y frase entonacional en lenguajes de señas, y proponen que estos niveles podrían ser universales de la estructura prosódica de los lenguajes. También encuentran algunas diferencias en la prosodia en las dos modalidades: al

haber dos potenciales articuladores análogos en los lenguajes de signos (las dos manos), esto tiene un impacto en su estructura prosódica. Aún más, hay muchos articuladores independientes, por ejemplo en el rostro, lo que permite la transmisión de componentes prosódicos simultáneos y secuenciales.

Otros estudios han abordado temas como la prosodia entre signantes nativos y personas en proceso de adquisición del lenguaje de señas (Boyes Braem, 1999, al respecto del ritmo), consideraciones teóricas y metodológicas (Kingston, 1999), entre otros.

3.5.4.1 Puntos importantes

El estudio de la prosodia en lenguajes de señas permite resaltar algunos puntos, principalmente la similitud tanto en funciones como en procesamiento cerebral que se da en la prosodia para lenguajes orales y signados (además de una organización general similar).

Por otra parte, también se vieron propuestas de características universales de la prosodia, como la presencia de frases fonológicas y entonacionales. Con todo, también se encontraron particularidades en las capacidades prosódicas de lenguajes de señas dada la presencia de múltiples articuladores que brindan la posibilidad de emplear características prosódicas simultáneas. Cabe mencionar que esto también puede hacerse en lenguajes orales (dada la variación de diferentes atributos y la posibilidad de señalar simultáneamente distinciones lingüísticas y emocionales), aunque, considerando las observaciones de redundancia de las variables presentadas anteriormente y la carencia de dos articuladores análogos (como las manos en el caso de lenguajes de señas), la capacidad de hacer distinciones prosódicas simultáneas en lenguajes orales podría ser más limitada. Este tipo de hallazgos, resaltan Nespor y Sandler (1999), podrían implicar que puede haber más de lo que se ha atribuido a la estructura lingüística abstracta, y que es determinado por la modalidad del lenguaje. Definitivamente, el estudio de lenguajes que emplean otra modalidad para establecer la comunicación aumenta los horizontes y posibilidades, y permite probar algunas hipótesis y generar otras teorías, más amplias, que permitan

englobar las observaciones y particularidades de lenguajes en distintas modalidades.

Recuérdese también, la perspectiva de MacSweeney y colaboradores (2002), quienes, en atención a sus resultados que reflejan de actividad en el giro temporal superior izquierdo en personas que emplean lenguajes de señas, consideran que a pesar de que esta región puede estar predispuesta al procesamiento del lenguaje en la modalidad sonora, al no existir esta modalidad en personas sordas, podría adaptarse para procesar información visual.

Por otra parte, los hallazgos de Corina y colaboradores (1999) sobre diferencias en la lateralización que asociadas a expresiones faciales afectivas o lingüísticas, específicamente lateralización al hemisferio derecho de las expresiones afectivas y hacia el izquierdo de las lingüísticas, son muy similares a los patrones de lateralización entre prosodia emocional y lingüística que hemos presentado en capítulos anteriores. Con todo, el análisis no fue hecho comparando expresiones prosódicas, sino lingüísticas en general. Estudios que abordaran específicamente expresiones faciales prosódicas ayudarían a apoyar la hipótesis de lateralización dependiendo de significancia que se presentó anteriormente, incluso en una modalidad distinta, lo que haría que teorías como la de lateralización por atributos pudieran ser descartadas, en su caso, ya que, en sentido estricto, no existe un nivel espectral en la señal lingüística signada, aunque sí existirían atributos temporales.

Reflexionando sobre los hallazgos en los que se ha encontrado una organización cerebral similar en lenguajes que emplean dos modalidades (visual y auditiva), cabe preguntarse: si se creara un lenguaje en otra modalidad (por ejemplo táctil), ¿se podría esperar una organización cerebral relativamente similar? El estudio de los lenguajes de señas nos permite acercarnos más a la comprensión del lenguaje, y muestra soluciones de comunicación distintas de las que muestran las lenguas orales del mundo. En éstas, hay una gran variabilidad en las estructuras fonológicas, gramaticales y sintácticas. En los lenguajes de señas, hay, además, variación en la modalidad en la que se presenta en lenguaje, pero, a la vez, y de

manera resaltable, se presentan similitudes tanto en la prosodia como en otros niveles del lenguaje.

3.5.5 Desarrollo de la Prosodia

En esta sección revisaremos brevemente algunos hallazgos sobre el desarrollo de la prosodia; desde influencias prenatales hasta la vejez.

El oído es un sentido activo ya desde el útero, y se han reportado efectos tanto en la producción vocal como en la percepción de la misma en recién nacidos, cuya explicación aduce a la experiencia intrauterina.

Hopkin (2009) reporta diferencias en el llanto de bebés cuyos padres hablan diferentes lenguas, y lo más importante, que los patrones acústicos son semejantes a las lenguas de sus padres. A su vez, Ross y Monnot (2008) reportan efectos detrimentales en la comprensión de la prosodia provocados por exposición a alcohol en el útero, aún en individuos que no presentaban signos para el diagnóstico de síndrome alcohólico fetal o efectos de alcohol fetal.

Por su parte, Mastropieri y Turkewitz (1999) reportan capacidad de diferenciar emociones a partir de expresiones vocales (y principalmente de la prosodia) en neonatos de entre 12 y 72 horas de nacidos (evaluada a partir de conductas espontáneas asociadas a diferentes emociones), y de manera más importante, que este efecto se da al oír expresiones vocales de las emociones en la lengua materna y no en una lengua extranjera. Los autores proponen un mecanismo de aprendizaje asociativo en el que cambios fisiológicos asociados al ambiente intrauterino serían asociados a las expresiones vocales de la emoción, audibles desde el útero, permitiendo establecer vínculos de manera temprana entre las emociones y sus correlatos vocales. Resaltan que aunque las respuestas diferenciales no indican necesariamente que los neonatos están percibiendo emociones, la discriminación de patrones acústicos sería un precursor necesario para el desarrollo de la percepción emocional. Reportan también que en neonatos (4 días de nacidos), se observa una discriminación del habla en el idioma propio o en uno extranjero, y que se sigue observando al filtrar el contenido segmental, es decir, conservando solamente la información prosódica. También comentan

hallazgos previos sobre preferencia de escucha al idioma materno en bebés de 2 días de nacidos. Resaltan que los hallazgos sobre la sensibilidad a características prosódicas de la voz han sido reportados previamente y forman la base de hallazgos como la preferencia a escuchar el idioma materno, sin embargo, su estudio extiende estos hallazgos a la prosodia emocional. Consideran que comparaciones entre neonatos cuyas lenguas maternas contrasten en ser o no tonales, podrían arrojar diferencias entre grupos.

Finalmente, Snow (2000) reporta que la lateralización de la prosodia no podría estar tan diferenciada en niños pequeños, dado que lesiones cerebrales prenatales o perinatales, independientemente de que afecten al hemisferio izquierdo o derecho, pueden acarrear dificultades en la percepción y expresión de prosodia lingüística, pero que sólo los individuos con daño al hemisferio derecho presentaban también déficits en la prosodia afectiva.

Avanzando un poco en el curso del desarrollo, se ha visto que los bebés pueden comunicar emociones antes de adquirir lenguaje, precisamente a través de la entonación, dentro de la prosodia emocional (entre otros mecanismos; Bänziger & Scherer, 2005). También se ha visto que muestran una ventaja en el desempeño al presentar estímulos al oído derecho cuando estos son sonidos del habla, pero no cuando son sonidos diferentes al habla (Mehler, Dupoux, Pallier, & Dehaene-Lambertz, 1994). Mehler y colaboradores comentan que si bien el hemisferio izquierdo podría procesar el habla, sin importar el lenguaje en el que se dé, progresivamente podría especializarse para procesar el lenguaje materno (o los lenguajes maternos en contextos multilingües). En el mismo sentido, reportan que neonatos de 4 días de nacidos pueden discriminar ya sea el lenguaje materno de uno extranjero (incluso a partir de estímulos con información segmental filtrada) o dos lenguajes extranjeros entre sí, pero que hacia los dos meses de edad, pueden discriminar sólo entre el materno y uno extranjero, lo que, en su opinión, indica que se ha establecido un patrón a partir de su propio lenguaje, contra el cual se contrasta cualquier *input* lingüístico. Más aún, la discriminación se pierde al alterar la estructura prosódica de las frases, lo que resalta la centralidad de la prosodia

para estos hallazgos. DePaolis y colaboradores (2000) resaltan al respecto de este tipo de hallazgos que son las variables temporales, como el ritmo, las principales responsables de los efectos de discriminabilidad.). Esto contrasta con la capacidad para distinguir secuencias fonémicas válidas en un idioma, que surge hacia los 9 meses de edad (y como veremos más adelante, se ha propuesto que la prosodia ayuda al desarrollo de la producción de consonantes). En otras palabras, la adquisición de la prosodia – al menos de parte de ella – es previa a la de la información segmental (Cortés Moreno, 2000; Huang & Jun, 2011), aunque el proceso de adquisición de la misma podría ampliarse hasta la pubertad (Cortés Moreno, 2000).

Un cuerpo de investigación relacionado, en el que no se abundará, es el del habla dirigida a bebés (*motherese*), ésta ha sido observada a través de diferentes culturas; los bebés prefieren escuchar este tipo de habla, específicamente por la contribución de la entonación y por la sensibilidad temprana a la relación entre el tono y las emociones. Esta preferencia se desarrolla por aprendizaje entre los 1 y 4 meses de edad. Se caracteriza por un tono general más alto, cambios tonales más pronunciados, un ritmo más lento, frases cortas y pausas largas, entre otras características. Ayuda a la adquisición del lenguaje al enfatizar palabras y estructuras lingüísticas importantes (Bryant & Barret, 2008; Snow, 2000; Thompson et al., 2004).

DePaolis y colaboradores (2008) hacen un estudio en balbuceo reduplicativo en niños pequeños (10 a 18 meses) en proceso de adquirir su lengua materna. Exploran si sus balbuceos difieren de acuerdo a las características de su idioma materno. Efectivamente, encuentran diferencias entre las producciones de los diferentes niños, especialmente en el ámbito de la duración de sílabas. También resalta el curso del desarrollo temprano de la prosodia dividido por atributos acústicos²⁶. Para f_0 , encuentran la capacidad de imitar el tono a los 3 meses, así como un patrón general de caída de tono de 3 a 9 meses, de acuerdo a las observaciones de declinación expuestas anteriormente (tendencia a la caída de

²⁶ Consultar el artículo para datos más completos.

tono por la reducción de la presión subglotal), sin embargo otros estudios no han confirmado los efectos de declinación y estudios de comparación entre bebés aprendiendo diferentes lenguajes encuentran efectos del lenguaje materno en los patrones de caída de f_0 (o la ausencia de los mismos), lo que resalta influencias culturales ya en estas épocas. En cuanto a las variables de duración, se ha encontrado poco control, y un aumento progresivo del mismo, por ejemplo en la elongación de las sílabas al final de las frases y contrastes como las consonantes geminadas (consonantes que hacen distinción por duración, como en el japonés *kita* (来た venir, tiempo pasado) vs *kitta* (切った cortar, tiempo pasado)) en el primer año de vida; hacia los 18 a 20 meses se ve un mayor control que coincide con la producción de palabras (los autores consideran que antes de esta edad podría haber una restricción cognitiva que impidiera integrar patrones segmentales y suprasegmentales). En general, consideran que el ambiente tiene poca influencia en los patrones temporales del habla en la etapa prelingüística. En cuanto al volumen, encuentran que no es similar al de la producción adulta entre los 7 y 11 meses de edad, sin embargo, parece ser precisamente ésta la variable mediante la cual se señala el *stress* a edades tempranas (aunque es importante resaltar que las características más importantes para señalar este último difieren entre lenguajes y que la mayoría de los estudios se realizan para el idioma inglés)

Mehler y colaboradores (1994), reportan activación en el hemisferio derecho en bebés de 20 meses (la época en la que empieza la producción de palabras) al oír palabras desconocidas; en adultos se encuentra también activación en este hemisferio (específicamente en la corteza prefrontal) al oír palabras posibles en su lengua, pero que no existen. DePaolis y colaboradores (2008) resaltan que para esta edad, los bebés ya cuentan con cierta habilidad para controlar f_0 , la intensidad y algunas variables duracionales; mencionan que la adquisición de la prosodia de una forma similar a la adulta se consigue hasta que se obtiene una masa considerable de palabras (y de hecho, las funciones léxicas de la prosodia requieren, por definición, de la presencia de las mismas). Kelm y Beckman (2009) estudian la influencia de la prosodia, específicamente la estructura métrica (moras, sílabas, frases, etc.), y consideran que la prosodia ayuda a coordinar diferentes

gestos (recordemos la postura de la fonología gestual) a diferentes escalas temporales, por ejemplo, los segmentos y las sílabas dentro de la frase en su conjunto. Más aún, consideran que en el proceso de producción, las praxias para la consecución de porciones grandes en el lenguaje (por ejemplo, frases) son programadas de manera temprana, antes de los gestos necesarios para la producción específica de palabras; en el proceso de desarrollo se observa algo análogo: los gestos para la producción de consonantes parecen darse sólo en el contexto de “frases” completas (desde la etapa de balbuceo) y poco a poco se puede controlar, primero a nivel de sílaba y después de manera más autónoma. Los autores consideran que la frase entonacional y la sílaba (de estructura de complejidad máxima CCVCC y cuyas consonantes iniciales y finales son menos sonoras que las intermedias y que la vocal central) son unidades universales que se dan en el curso normal de desarrollo, mientras que una diferenciación más allá de este punto es influida por el ambiente, es decir, por el lenguaje al que está expuesto el niño. En este sentido, resaltan que no se encuentran diferencias en la duración de las “sílabas” entre bebés inmersos en ambientes con distintas lenguas entre los 4 y 6 meses, pero que para la época en la que comienza el balbuceo reduplicativo ya se ve influencia de la lengua materna y que para la época en la que se producen las primeras palabras reconocibles, ya se ven reflejadas las particularidades prosódicas de su idioma, a pesar de las restricciones universales de complejidad motora propias de las primeras frases.

Pasando a la infancia, se ha encontrado que participantes que aprenden un lenguaje extranjero (por cambio de residencia) al ser niños (5 a 9 años), en contraste con adolescentes (12-17 años) y adultos (20 a 26 años) logran producciones prosódicas similares a las nativas, mientras que los adolescentes, aunque muestran similitudes con los hablantes nativos, muestran algunas desviaciones en las producciones prosódicas y este efecto es más marcado en los adultos, incluyendo la percepción de acento extranjero (Huang & Jun, 2011). Cortés Moreno (2000) enfatiza que si bien a una edad temprana los niños ya tienen un cierto dominio de algunos patrones de entonación, el aprendizaje de la prosodia puede seguir ocurriendo hasta la pubertad (o incluso después), ya que

está relacionada a una amplia gama de funciones. Especialmente en el campo de las funciones pragmáticas, los niños no requieren hacer uso de muchas de ellas, por ejemplo contextos de reuniones sociales, profesionales, usos de cortesía, entre otros; si bien los niños saben preguntar, declarar y exclamar prosódicamente, no tienen práctica en muchos otros aspectos y convenciones prosódicas. A partir de este tipo de observaciones, este autor enfatiza la necesidad de poner atención en la prosodia para la enseñanza de lenguas extranjeras y para ayudar a explotar las capacidades expresivas y comunicativas de diferentes patrones prosódicos. Otro hallazgo importante en niños es que se pueden diferenciar las voces dependiendo del género (tanto en habla como en canto), a pesar de que no existen diferencias generales en la anatomía de los aparatos vocales de niños y niñas (Haunch, sin año; Urberg Carlson & Munson, 2008).

En cuanto al desarrollo de la prosodia en la etapa adulta, es un tema que no se abordará, ya que a lo largo del presente trabajo la mayoría de las investigaciones reportadas han analizado, precisamente, a participantes en esta etapa. Cabe resaltar, sin embargo, las diferencias de género encontradas principalmente en el campo de prosodia emocional, y que éstas pueden incluir influencias tanto biológicas como culturales en su desarrollo.

Finalmente, en el caso de la vejez, Zellner (2006) reporta cambios en la estructura y función de las cuerdas vocales, así como pérdida de elasticidad y fuerza en músculos, en especial faciales y respiratorios, por lo que el control neuromuscular de la voz, que depende de la coordinación de elementos pulmonares, laríngeos y articulatorios se ve afectado, siendo uno de los cambios más pronunciados un ritmo más lento. En cuanto al oído, también se ve afectado de manera periférica o central, inclusive ambas, lo que puede llevar a la pérdida de la capacidad para escuchar frecuencias altas así como a déficits en la resolución temporal y de frecuencia, lo que hace más difícil separar la señal lingüística del ruido que la acompaña, con todo, los individuos lo pueden hacer; sin embargo, el esfuerzo cognitivo adicional para la percepción de la señal puede interferir con la

comprensión y memoria sobre dicha señal. La autora resalta en cuanto a la articulación y tempo (término similar a *ritmo*, ver Caballero-Meneses & Menez, 2010, para un ejemplo del uso del término en el campo de la música) del habla, un ralentizamiento de la misma, así como mayor variabilidad del ritmo del habla, aunque no correlaciona con la edad cronológica; encuentra dificultades para percibir ritmos rápidos de habla, presumiblemente por las alteraciones en la percepción auditiva y no por un procesamiento más lento. En cuanto a f_0 y la entonación encuentra un decremento de f_0 en mujeres, con una caída marcada alrededor de la menopausia; en hombres se observa un decremento ligero seguido por un aumento marcado en f_0 comenzando alrededor de los 50 años; la estabilidad de f_0 es afectada por la salud (recordemos las funciones indécicas de la prosodia); cabe destacar que la salud presenta alta variabilidad entre hablantes de la misma edad. En cuanto a la intensidad, se ve un decremento asociado a la edad, probablemente por cambios en los sistemas respiratorio y fonador, aunque la magnitud de este decremento no alcanza para que sea distinguida en el habla normal para conversar; asimismo, se ve un “temblor” en la intensidad de vocales sostenidas, lo que constituye un gran predictor de edad cronológica y percibida. También destaca patrones de habla dirigidos a las personas de la tercera edad que se dan espontáneamente, análogos al habla dirigida a niños (*motherese*) mencionada anteriormente; estos patrones son llamados *Elderspeak* y *Oldiespeak*; la autora describe sus características e implicaciones psicosociales.

En cuanto a la prosodia emocional, se ha visto una habilidad conservada para la producción (medida a través de f_0 ; Zellner, 2006), pero una disminución en la capacidad para la comprensión (Fonseca, et al., 2007; Zellner, 2006). Este tipo de efectos no serían explicables por pérdida de audición asociada a la edad ni por efectos del envejecimiento sobre la cognición (Ross & Monnot, 2011). Fonseca y colaboradores (2007) reportan también algunos déficits en tareas de prosodia lingüística para comprensión, pero no para repetición; con todo, resalta que si bien se observan diferencias significativas con respecto a adultos jóvenes, el desempeño sigue siendo alto (más de 90% de aciertos en la tarea). Consideran que la presencia de déficits más marcados para la prosodia emocional que para la

lingüística podría tener que ver con las variaciones melódicas, aparentemente más sutiles (recordemos que hemos visto su asociación a variables continuas y distribuidas en toda la frase, en contraste con las lingüísticas, muchas veces categóricas y concentradas en ciertas partes del contorno melódico) en la prosodia emocional.

Finalmente, Ross y Monnot (2011), explorando la prosodia emocional, resaltan efectos de edad en la distribución de errores al categorizar emociones. Encuentran diferencias en la distribución de errores al juzgar prosodia asociada a emociones primarias (especialmente alegría) pero no al juzgar emociones sociales. Consideran que este patrón estaría mediado por la atribución cognitiva, que provocaría una alteración en la *saliencia* de diferentes estímulos afectivos prosódicos al aumentar la edad (por cambios en lóbulos frontales); así como al posible deterioro de las funciones del hemisferio derecho asociado al envejecimiento (aunado a la hipótesis de lateralización diferencial de emociones).

3.5.5.1 Puntos importantes

Esta sección nos permite resaltar, principalmente, los hallazgos sobre una organización cerebral de la prosodia a edades tempranas posiblemente menos diferenciada; así como una diferenciación progresiva.

Se observa también que aunque el desarrollo de la prosodia se aprecia de manera más temprana que el desarrollo de los contrastes segmentales (incluso apoya al desarrollo de estos últimos), también se prolonga durante un tiempo considerable: en la pubertad se siguen adquiriendo algunos usos pragmáticos de la misma.

Entre las influencias culturales y biológicas a la prosodia, tema eje del presente trabajo, se pueden resaltar las influencias culturales a temprana edad (e incluso en el útero), como demuestran los hallazgos de preferencia por la lengua materna y la capacidad de diferenciar la lengua materna de lenguas extranjeras, así como la influencia de variables biológicas a edades avanzadas, como los cambios en f_0 alrededor del tiempo de menopausia en mujeres y alrededor de los 50 años en hombres, así como las observaciones de cambios en variables rítmicas.

Se aprecia también convergencia y complementariedad entre ésta y las demás secciones del capítulo – y en general de ellas entre sí –, por ejemplo, al analizar lenguas extranjeras en conjunción con el desarrollo. También surgen algunas preguntas; por ejemplo, si bien se ha visto que existen etapas de “balbuceo” en lenguajes de señas, cabe preguntarse si hallazgos como la presencia intercultural de habla dirigida a niños (motherese), con su prosodia característica, también se da en lenguajes de señas y qué características presenta.

3.5.6 Finalización del núcleo empírico

Definitivamente, haber realizado una revisión extensa sobre la prosodia lingüística y emocional, así como de los estudios que las analizan en conjunto, seguida por una revisión breve de otros temas “menores” nos ha permitido tener una visión amplia sobre la prosodia, y sobre las contribuciones biológicas, psicofisiológicas y culturales, tal como el objetivo de la tesis planteó. Efectivamente, artículos sobre didáctica de lenguas extranjeras, por poner solo un ejemplo, han hecho énfasis en usos pragmáticos de la prosodia, interesándose asimismo en el curso de desarrollo de los mismos, tema que difícilmente habría surgido en una tradición como la de la psicología y neurobiología.

En este punto se da por finalizada la revisión empírica; en el siguiente capítulo se presenta el análisis conceptual sobre el término “prosodia” para, enseguida, dar paso a las conclusiones. Se ha reservado el análisis conceptual para el final del trabajo para tener como contexto y antecedente toda la información empírica revisada antes de discutir la base misma de los conceptos que han guiado la investigación en el campo.

3.6 Análisis Conceptual

El trabajar a partir de conceptos claros que guíen la investigación es fundamental para el avance en cualquier campo. En el caso específico del término “prosodia”, existe una gran cantidad y diversidad de definiciones. En el anexo 1 se incluye una lista con definiciones tomadas de diversos autores a partir de las cuales se trabajó el presente análisis conceptual, algunos de esos autores tomaron sus definiciones, a su vez de otros autores, pero esta información no se cita para mantener la simplicidad.

Efectivamente, llegar a una definición específica del término prosodia es una labor con dificultades, como han resaltado en el pasado Shattuck-Hufnagel y Turk, (1996) y Batista Antunes (2007). Diversas definiciones de la prosodia se han referido a diversos atributos, muchas otras la toman como un sinónimo del término entonación (aunque se ha enfatizado lo erróneo de esta aproximación) y que la entonación se debe tomar como un componente de la prosodia (un término subordinado) y no como un sinónimo de la misma (Seddoh, 2002); en diversas ocasiones también se le iguala con el término “suprasegmental” (Ej. Batista Antunes, 2007; Pell & Skorup, 2008; Péron et al., 2010); aunque también se ha cuestionado este uso, ya que existen algunos fenómenos suprasegmentales que no son considerados prosódicos, como la coarticulación, asimilación, entre otros (Batista Antunes, 2007). Llama particularmente la atención la definición de Joannette y colaboradores (2008), en la que el término prosodia refiere al “procesamiento cognitivo necesario para comprender o expresar intenciones comunicativas usando aspectos suprasegmentales del habla, tales como las variaciones de la entonación, las pausas y las modulaciones de la intensidad vocal”, ya que en esta definición, el término refiere al procesamiento cognitivo y no directamente a las características suprasegmentales, como la mayoría de las otras definiciones. Una definición explícita del término “suprasegmental” ha sido dada por Narasimhan y colaboradores (2010): “El término ‘suprasegmental’ se refiere a aquellas propiedades de las frases que no pertenecen a ningún segmento aislado. Las propiedades suprasegmentales son el estrés, tono, entonación, duración y

organización de los segmentos en sílabas”. También la cualidad de la voz se ha tomado en cuenta entre las características suprasegmentales

A partir del análisis de las diferentes definiciones del Anexo 1, con sus similitudes y diferencias, y de la revisión que se ha realizado hasta el momento, se podría definir prosodia de la siguiente manera:

“Las características suprasegmentales del habla que sirven para transmitir información sobre el estado emocional, distinciones lingüísticas²⁷, pragmáticas, actitudinales e información general sobre la persona”

Esta definición aclara en una primera instancia que el dominio en el que se da la prosodia es suprasegmental, implicando que los fenómenos segmentales caen fuera de su alcance; inmediatamente alude a las funciones de la prosodia: emocional, lingüística, pragmática, actitudinal e indéxica, de acuerdo con la postura de Van Lancker y colaboradores (2006). Finalmente, al emplear el término “características suprasegmentales” implica las variables que hemos estudiado a lo largo del presente trabajo: tono, características temporales, volumen y cualidad de la voz; pero no se limita a ellas, en atención a estudios posteriores que podrían tomar en cuenta otras variables. Se considera importante tomar en cuenta a la cualidad de la voz al leer el término “características suprasegmentales”(aunque no es mencionada en varias definiciones que enumeran específicamente las variables que consideran) dada la gran cantidad de estudios en los que se enfatiza su importancia.

Esta definición es aplicable a las características prosódicas de distintas lenguas; al no especificar exactamente las funciones lingüísticas a las que alude, permite incluir las funciones que desempeña en distintos lenguajes: por ejemplo la distinción tonal en lenguas tonales y la distinción por stress en lenguas como el español y el inglés (de entonación).

²⁷ Las que hemos abordado en el presente trabajo, pero no limitadas a ellas, ya que en el futuro podrían ser abordadas otras distinciones lingüísticas transmitidas por medios prosódicos.

Se ha preferido el uso del término “emocionales” a “afectivas” dado el uso que se ha dado al término “afectivo” en la literatura revisada, ya que algunas veces engloba a las actitudes; con todo, es importante interpretar “emocionales” no en el sentido psicológico *fuerte* del término, que incluye entre sus criterios definitorios una duración corta, sino englobando fenómenos afectivos (*afectivos* en el uso psicológico) más extendidos en el tiempo, como los estados de ánimo. En el capítulo de prosodia lingüística se abordó el problema del diferente uso de términos en distintas áreas.

La inclusión del término “pragmáticas” podría obviarse, ya que algunas definiciones del lenguaje incluyen el nivel pragmático, lo que implicaría que el término “distinciones lingüísticas” conlleva la posibilidad de distinciones pragmáticas; pero dado que otras definiciones del lenguaje no contemplan el nivel pragmático explícitamente, se ha decidido incluirlo en la definición.

La inclusión de la transmisión actitudinal es necesaria al no emplear el término “afectivo” en el sentido que lo incluiría y dados los estudios de las funciones de la prosodia revisados.

La transmisión de la información indéxica (“[...]información general sobre la persona[...]”) se ha decidido incluir dados los hallazgos de la importancia de la prosodia para la señalización del acento extranjero y de su importancia para la señalización del sexo (aún en niños, cuyos aparatos fonadores aún son similares biológicamente), entre otros atributos generales de la persona.

Con todo, es importante aclarar que este tipo de definición asume una distinción entre los niveles segmental y suprasegmental, que en teorías como la de la fonética gestual podría no ser tan clara (Moen, 2006).

Sin embargo, aún queda un cabo suelto, el de la prosodia en el lenguaje de señas.

En el caso de los lenguajes de señas, se ha visto el uso de articuladores no manuales como rostro y sus componentes para hacer distinciones prosódicas ¿de

qué manera se puede llegar a una definición que incluya las particularidades prosódicas de los lenguajes de señas, que no se basan en el sonido?

Una definición que incluya a la prosodia de los lenguajes orales y de señas podría alcanzar un nivel de abstracción más alto que aquella que define sólo la prosodia en lenguajes orales, pero pierde, al mismo tiempo, particularidad. La siguiente definición representa un esfuerzo de englobar la prosodia de lenguajes orales y de señas (y, potencialmente, de lenguajes que se transmitiesen en otras modalidades).

“Características de los lenguajes que hacen uso de los grados de libertad disponibles en la modalidad en la que se expresan, sin afectar a los símbolos lingüísticos menores que expresan conceptos o funciones y que se pueden superponer o presentar concomitantemente a ellos para transmitir información sobre el estado emocional, distinciones lingüísticas, pragmáticas, actitudinales e información general sobre la persona”.

Resalta primero el uso del término “*grados de libertad*”, término teórico desarrollado en el presente trabajo. El siguiente componente, “[...] en la modalidad en la que se expresan [...]” permite englobar a la modalidad oral-auditiva de los lenguajes orales y kinética-visual de los lenguajes de señas; “[...] sin afectar los símbolos lingüísticos menores que expresan conceptos o funciones[...]

” permite que la definición aplique ya sea a las palabras (irrespectivamente de que sean palabras contenido o función) en el caso de lenguajes orales o a los signos (y a la forma de expresar sus relaciones) en el caso de los lenguajes de señas; “[...]y que se puedan superponer o dar concomitantemente a ellos[...]

” permite que sea válida tanto para la superposición acústica de las características prosódicas y segmentales en lenguajes orales o simultaneidad (como el uso de articuladores no manuales en lenguajes de señas); la última parte de la definición “[...]para transmitir información sobre el estado emocional, algunas distinciones lingüísticas, pragmáticas, actitudinales e información general sobre la persona [...]” coincide con la definición brindada para lenguajes orales dado que se asume que las funciones de la prosodia en ambos tipos de lenguajes son similares.

Una aparente debilidad de esta definición es la parte en la que se propone “[...]sin afectar los símbolos lingüísticos menores[...]”, ya que, al menos en lenguajes orales, los símbolos lingüísticos menores son las palabras, que a veces contrastan entre sí por medios prosódicos (por ejemplo, por tonos en lenguajes tonales y por estrés en lenguajes como el español), por lo que la definición negaría la posibilidad de la prosodia para hacer contrastes léxicos; sin embargo, si se asume que en un cierto contexto – donde puede haber variaciones prosódicas – el símbolo lingüístico menor está dado, esto no tiene ninguna importancia; por ejemplo en la frase “La papa es deliciosa”, “papa” es un símbolo dado, por lo que “*La papá es deliciosa” es imposible desde esta definición, ya que afectaría al signo lingüístico menor “papa” para poder ser realizada. Este problema también podría resolverse agregando a la definición, justo después de “[...] sin afectar a los símbolos lingüísticos menores que expresan conceptos o funciones [...]” el término “(excepto para distinciones léxicas)”.

En fin, la segunda definición propuesta tiene la capacidad de englobar a la prosodia de lenguajes orales y de señas e inclusive a lenguajes hipotéticos en otras modalidades, pero pierde especificidad con respecto a la primera definición (considerando que la mayor parte de los estudios se hacen en lenguajes orales). Posiblemente tendría mayor aplicación en el caso de los estudios de comparación de lenguajes orales y de señas. En todo caso, sirve como un esfuerzo para llegar a un nivel de abstracción más alto sobre la prosodia en lenguajes conocidos (orales y de señas) e hipotéticos (kinético-táctiles, químico-olfatorios, entre otros).

Los estudios sobre lenguajes de señas específicamente (al menos los aquí revisados) no suelen brindar una definición específica de prosodia para la modalidad kinético-visual (excepto por el trabajo reciente de Brentari, González, Seidl, & Wilbur, 2011). En el presente trabajo no se intentará formular una definición específica de prosodia para lenguajes de señas ya que rebasa los alcances del mismo.

4. Conclusión

Al introducir el trabajo de tesis se plantearon los objetivos fundamentales del mismo:

- Identificación de variables biológicas, psicofisiológicas y culturales tanto en la prosodia emocional como prosodia lingüística.
- Análisis conceptual y empírico sobre la distinción lingüística-emocional para la prosodia.
- Análisis conceptual del término “prosodia”.

Para lograrlos se propuso un enfoque para analizarlas independientemente, que consistió en la revisión de investigaciones con interés puntual en una de ellas y de estudios que las comparan, así como temas un poco más “lejanos” pero aún con interés en la prosodia por su capacidad para aclarar algunas interrogantes y por ofrecer posibilidades de aislamiento de variables de manera particular. Estos temas fueron: acentos y síndrome del acento extranjero, música y prosodia, lenguajes tonales y prosodia, prosodia en lenguajes de señas y desarrollo de la prosodia. Al momento de proponer el enfoque, no se sabía si permitiría una consecución adecuada de los objetivos, a partir de este momento se puede evaluar su utilidad en el desarrollo del trabajo.

En cuanto al primer objetivo, la influencia de variables biológicas, psicofisiológicas y culturales a la prosodia se encuentran:

Para la prosodia emocional, de manera sintética, las siguientes: los factores fisiológicos que subyacen a las emociones, con sus consecuencias en la conducta, especialmente en la fonación, probables efectos evolutivos dadas las similitudes interculturales en la prosodia emocional, probables efectos de género a partir de diferencias cerebrales anatómicas, efectos de las *reglas de exposición culturales*, de las características de los lenguajes, de hablar un idioma particular (en especial el efecto de ventajas al interior del grupo) y efectos de hablar varios idiomas.

Mientras que para la prosodia lingüística, se encuentra entre las contribuciones biológicas el fenómeno de declinación (caída de f_0 determinada por la caída de la presión subglotal) y de ralentización de articuladores hacia el final de las frases, presente en muchos idiomas y de base fisiológica. Se encuentran posibles influencias biológicas matizadas por la cultura, por ejemplo, las tendencias universales, como la asociación de tonos altos a cortesía y a pregunta; funciones lingüísticas de la prosodia, como el fraseo y el foco, que aunque se dan en todas las lenguas, los medios por los que se consiguen varían (por ejemplo, en algunos idiomas existen marcas morfosintácticas obligatorias). Las influencias claramente culturales están ilustradas por la existencia de actitudes marcadas prosódicamente al interior de una lengua pero ausentes en otra, diferente asociación de características prosódicas con actitudes en diferentes lenguajes, influencia del lenguaje materno en la producción y percepción prosódica, capacidad de percibir pausas donde acústicamente no existen, variaciones dialectales en la prosodia (entre diferentes variedades del mismo lenguaje).

Ambos conjuntos de hallazgos fueron ampliamente abordados en los capítulos de prosodia emocional y prosodia lingüística, respectivamente, y resumidos en la sección de “Resumen y discusión” al interior de cada capítulo. Cabe destacar, de la sección de desarrollo de la prosodia en el capítulo de temas complementarios, los hallazgos de influencias culturales tempranas, incluso a nivel prenatal, y los de influencias biológicas tardías, como los cambios en la prosodia alrededor del tiempo de la menopausia, en mujeres, y de los 50 años en hombres.

En cuanto al análisis conceptual y empírico de las diferencias entre la prosodia emocional y lingüística, fue abordado principalmente en el capítulo “*Prosodia Lingüística y Prosodia Emocional*”. Si bien la caracterización empírica tanto a nivel acústico como perceptual y a nivel de organización cerebral se inició en los capítulos destinados a cada una de las prosodias por separado, en dicho capítulo se abordaron específicamente los temas sobre la diferenciación de la prosodia lingüística y emocional a nivel conceptual y a nivel acústico y de organización cerebral.

Se encuentra, efectivamente, una diferente distribución cerebral para la prosodia lingüística y emocional, aunque superpuesta en algunas áreas. Ambas prosodias están lateralizadas principalmente al hemisferio derecho, pero la prosodia lingüística está más distribuida e incluye al hemisferio izquierdo en mayor medida que la emocional. Acústicamente, si bien ambas prosodias hacen uso de los mismos atributos (tono, variables temporales, volumen, cualidad de la voz), la prosodia emocional está asociada principalmente a características generales en las frases (promedios y variación de las variables, por ejemplo) y al comportamiento continuo de las variables, mientras que la prosodia lingüística está asociada a cambios locales (por ejemplo final de las frases para distinciones de actos de habla, al menos en algunos idiomas, así como en palabras específicas para el foco o a sílabas específicas para funciones léxicas) y a contrastes categóricos en las variables.

En cuanto al análisis conceptual de la distinción de la prosodia en las categorías emocional y lingüística se encontró que el principal contraste está dado por la función: la prosodia emocional transmite información relacionada con las emociones mientras que la prosodia lingüística con contrastes lingüísticos. En el mismo capítulo se abordaron posturas alternativas a la división emocional y lingüística, de las cuales es importante destacar la de Van Lancker y colaboradores (2006), ya que se retomó al hacer el análisis conceptual del término “prosodia” y al proponer una definición propia.

Si bien el contraste empírico muestra que la distinción emocional-lingüística en la prosodia está reflejada tanto por la organización cerebral como a nivel acústico, también muestra la influencia del marco teórico a partir del cual fue hecha; las categorías emocional y lingüística se consideran *a priori*, lo que lleva a una buena cantidad de estudios y a la identificación de diferencias. Sin embargo también hay otras perspectivas teóricas que emplean categorías diferentes a la distinción emocional-lingüística que pueden ser útiles en el futuro, y que también gozan de apoyo empírico, aunque están menos extendidas. Al hacer el análisis conceptual y al incluir hallazgos de los temas “lejanos”, hemos decidido basarnos en la

perspectiva de Van Lancker y colaboradores (2006), que distingue la prosodia en 5 categorías: emocional, actitudinal, pragmática, lingüística e indéxica; ya que hace una distinción más fina (por ejemplo al no incluir actitudinal y pragmática dentro de lingüística, como es preciso hacerlo en la perspectiva de la división dicotómica) así como la función indéxica. Además, esta división es un poco más cercana a la división por funciones que se maneja en el campo de la lingüística.

Al respecto de la integración de estudios provenientes de diversos campos, se ha resaltado la influencia de la terminología y de diferencias conceptuales de términos en campos ajenos, como los términos “afectivo”, “emocional” y “emotivo” en los campos psicológico y lingüístico; lo que ilustra la importancia de la claridad conceptual y de la identificación de diferencias al integrar diversas perspectivas, y, de manera más general, la importancia del análisis conceptual, ya que este brinda claridad a los términos a partir de los cuales se integra la base teórica de los campos, que entre más clara sea permite mayor solidez y alcance de la teoría.

El uso de un análisis empírico sobre la distribución cerebral de la prosodia emocional y lingüística, aunado al análisis conceptual de los términos, con el objetivo de explorar la adecuación de esta distinción en el campo, fue una idea planteada desde el inicio, con resultados que podrían considerarse fructíferos: aunque la distinción podría ser clara a nivel conceptual, podría no gozar de apoyo empírico; por otra parte, evidencia empírica podría provenir de términos muy distintos a nivel conceptual, lo que no permitiría integrarlos adecuadamente, a pesar de que aparentemente se refiriesen a lo mismo.

Como resultado del análisis conceptual, se propusieron dos definiciones de la prosodia, una particular para los lenguajes orales y una que integra la prosodia en lenguajes de señas y potencialmente en lenguajes que se diesen para otras modalidades. Evidentemente la segunda definición no habría sido posible, ni siquiera concebible, sin la revisión de los temas “lejanos”, específicamente el tema

de la prosodia en lenguajes de señas. Estas definiciones están disponibles en el capítulo de análisis conceptual.

A partir de nuestra definición de prosodia, que integra la división en 5 categorías de Van Lancker y colaboradores (2006), se podría avanzar en la investigación, guiando estudios y buscando apoyo empírico. Esta división en 5 categorías podría probarse más fructífera que la división dicotómica (lingüística vs emocional) en el futuro. Uno de los resultados más directos del uso de este marco sería una mejor caracterización de la organización cerebral de la prosodia lingüística (tal como es conceptualizada en el marco de la división emocional-lingüística). Explicando: si bien se han explorado las áreas cerebrales relacionadas al procesamiento prosódico dependiendo de la emoción particular, las diferencias en la organización cerebral de las diferentes funciones de la prosodia lingüística es menos entendida. Esto sucede porque en la división dicotómica se incluyen las funciones pragmática y actitudinal dentro de la lingüística. Estudios que integrasen la división en 5 funciones, al requerir contrastes y caracterizaciones específicas para cada una de ellas, contribuirían a la mejor comprensión de la prosodia lingüística tal como es conceptualizada tradicionalmente (es decir, desde la perspectiva dicotómica).

De esta manera, los objetivos planteados al inicio del presente trabajo han sido cumplidos. En cuanto al enfoque de revisión amplia y de temas “lejanos”, se ha mostrado adecuado, satisfactorio y ha permitido ir más allá, contribuyendo a planteamientos imposibles si se hubiese omitido (como la definición extendida de prosodia).

Consideramos que el uso del enfoque de revisión amplia y de temas relacionados que propusimos en el presente trabajo podría ayudar al desarrollo de diferentes campos que caen en el dominio de distintas disciplinas, como la toma de decisiones (estudiada por la economía, la psicología, las neurociencias), conducta alimentaria (psicología, fisiología, nutriología), entre muchos otros.

A lo largo del presente trabajo se han hecho algunas proposiciones y se han planteado diversas ideas, expuestas principalmente en las secciones de “Resumen y discusión” al interior de los tres primeros capítulos y en la parte final de cada sección en el capítulo de temas complementarios.

Entre ellas, la principal es la “*hipótesis de grados de libertad*” que se planteó originalmente en el capítulo de prosodia emocional y se elaboró y amplió a lo largo del trabajo, también se identificó evidencia que la apoya y propuestas teóricas similares (ver sección 3.4.3.8), planteadas previamente, como la propuesta del uso del “espacio fonético” no utilizado para contrastes discretos (Gussenhoven, 1999), la hipótesis de la carga prosódica de Pell (Bryant y Fox Tree, 2005; Pell, 1999). A grandes rasgos, la hipótesis propone que las características de la señal acústica (o, por extensión, de la señal signada, en el caso de lenguajes de señas) que no son requeridas por el sistema para marcar diferencias contrastivas constituyen los grados de libertad, y es a partir del uso de éstos que se puede transmitir información emocional, indéxica, entre otras. Se puede considerar que es principalmente la prosodia lingüística la que determina las características requeridas por el sistema, sin embargo, al ser de carácter discreto, aún es posible cierta variación. Por ejemplo, si bien para marcar una pregunta se hace uso de una subida de tono en español, tanto el momento en el que se comienza a subir el tono como la altura final del mismo permiten variación, es decir, son *grados de libertad*. La hipótesis de *grados de libertad* permite acomodar los hallazgos de la prosodia emocional, los hallazgos sobre acentos, las diferencias en la prosodia en lenguas tonales y lenguas de entonación y concuerda con algunas propuestas teóricas anteriores.

También se han ofrecido algunas propuestas de investigación a partir de esta hipótesis, como la de la posible existencia de variaciones diacrónicas en la prosodia emocional. Esta idea sería difícilmente concebible sin integrar información amplia y de campos diversos, como se ha venido haciendo a lo largo del presente trabajo.

Otros temas e ideas a resaltar son (1) la importancia de la cualidad de la voz en la prosodia en general (de hecho, formó parte de la definición de prosodia propuesta para lenguas orales), así como en la prosodia emocional y lingüística; (2) la importancia de la realización de estudios acústicos y perceptuales, dada la carencia de correspondencia biunívoca entre estos dos niveles de análisis, y, más generalmente, (3) la importancia del análisis y uso de diferentes metodologías y de estudios guiados por diferentes cuerpos teóricos, postura que se encuentra en el núcleo del presente trabajo.

Dos temas planteados, que no se logró confirmar o rechazar, pero se ha sentado la base para la posterior exploración son el de la *independencia de longitud* y la de los posibles *significados composicionales de la prosodia*. La independencia de longitud se refiere a si las medidas globales (promedios, variaciones) y dependientes de tiempo (contornos) son similares independientemente de la longitud de las frases (en el caso de los contornos, al ser normalizadas con respecto al tiempo). Los significados composicionales de la prosodia se refieren a la interacción de diferentes funciones prosódicas y a la posible generatividad en cuanto a los “significados” de ciertas variaciones prosódicas. Este tipo de proceso se encuentra en la base de la teoría del origen del lenguaje de Hocket (1971). La realización de estudios que se dirijan a resolver estas interrogantes es un objetivo que se podría seguir en el futuro.

Entre las posibles aplicaciones prácticas del estudio de la prosodia se encuentran principalmente la enseñanza de idiomas, al identificar que el nivel suprasegmental tiene una gran influencia en la inteligibilidad y percepción de acento extranjero; clínica, en el desarrollo de nuevas metodologías diagnósticas, como la identificación temprana de problemas de lenguaje (ya que las características prosódicas van surgiendo antes que las segmentales) y terapéuticas, especialmente en el caso de síndrome de acento extranjero.

En el marco de la enseñanza de idiomas, resaltan los esfuerzos de Kamiyama y Yamamoto (2007) de crear un sistema de transcripción prosódica simple y didáctico, y que capture incluso características de cualidad de la voz. También

Graber Rosenberg (1998) ofrece un trabajo muy interesante de diseño de una fuente tipográfica apta para capturar variaciones prosódicas.

Algunos temas que requieren mayor clarificación en el futuro son el de el dominio temporal de la prosodia: ya que se considera prosodia a todo lo suprasegmental, esto es un dominio muy amplio, moras, sílabas, frases intermedias, frases completas. Al ver algunas rupturas en la organización cerebral (principalmente en el nivel de sílabas y frases, ilustrado al estudiar lenguas tonales), podría ser importante distinguir de manera más sutil el dominio temporal al que se refiere e incluso acuñar términos específicos para los distintos dominios temporales, así como el de los usos sociales de la prosodia, incluyendo las variables que determinan el uso de ciertos patrones y la exploración de analogías en estas variables a través de las culturas. La prosodia en lenguajes de señas y la capacidad de entender a la prosodia independientemente de la modalidad del lenguaje (oral, de señas) a la que esté asociada, es un camino que debería ser más explorado en el futuro. La prosodia vista como una función compuesta por procesos (por ejemplo, percepción de tiempo y percepción de tono) es una que también merece ser ampliada, dada la buena comprensión que hay de los procesos propuestos como constitutivos y de hallazgos como dificultades en tareas de percepción de tiempo y prosódicas (específicamente las que involucran variables temporales) en personas con Parkinson.

A partir de lo revisado, y tal como lo enfatizan Chen y su equipo (2009), se puede observar la importancia de la cultura en la organización cerebral, ilustrada en el caso de la prosodia del lenguaje y particularmente en los hallazgos específicos sobre lenguajes tonales. En el futuro, vale la pena seguir explorando la influencia de la cultura en la organización cerebral, en general.

Entre las limitaciones del presente trabajo se pueden mencionar la falta de una revisión metodológica exhaustiva y de análisis sobre las ventajas y desventajas de diferentes aproximaciones metodológicas. Otra limitación es la falta de inclusión

de ejemplos en audio sobre las diferentes funciones de la prosodia, limitación que comparte con gran cantidad de estudios en el área. La amplitud de los temas abordados puede considerarse una limitación, ya que no permite profundizar mucho en los temas; esto se observa principalmente en los temas “lejanos” que fueron revisados; a cambio, este mismo hecho provee la posibilidad de un enfoque integrador y la inclusión de muy diversas perspectivas y estudios. Una explicación detallada de algunos de los términos empleados en algunos momentos también fue omitida, esto con el fin de no desviarse demasiado del tema principal y de no hacer innecesariamente amplias algunas secciones del trabajo.

En fin, la prosodia es un fenómeno central en el lenguaje, con contribuciones biológicas, psicofisiológicas y culturales, de naturaleza relativa al todo, ya que sus características no se pueden identificar a niveles aislados; ni el ritmo, ni los cambios de tono, ni la duración corta o larga o el stress se pueden identificar en un solo segmento, sino en relación con otros, y que ha sido identificado como central por muchos autores en el lenguaje, a pesar de que no siempre se le da su importancia (Joanette et al., 2008); algunos autores incluso enfatizan que la prosodia constituye parte de la gramática de los lenguajes (Moen, 2006; Shattuck-Hufnagel & Turk, 1996). Su división en funciones es principalmente conceptual, por lo que su división dicotómica (emocional-lingüística) o en más niveles, es dependiente de la teoría, aunque encuentra apoyo empírico dado el uso de un cierto marco teórico para guiar la investigación. La división por características acústicas (tono, ritmo, volumen, cualidad de la voz), a pesar de tener una base física, tampoco es óptima, ya que no coincide necesariamente con la percepción y porque aunque para una cierta función una cierta característica acústica tenga mayor peso, es la influencia de la prosodia (con sus diferentes variables e interacciones entre ellas) en su totalidad la que permite la transmisión de la información, por lo que es importante caracterizarla en su conjunto.

En retrospectiva, las dos aportaciones más importantes del presente trabajo son la proposición del uso de un enfoque de revisión extenso y la proposición de la *hipótesis de los grados de libertad*. Ésta última, particularmente importante por su

capacidad para acomodar diversos hallazgos (en prosodia emocional, en acentos, en el contraste entre lenguas tonales y de entonación, etc.), porque concuerda con proposiciones teóricas anteriores, y porque ha permitido plantear propuestas de investigación empírica que no tendrían cabida en otras perspectivas teóricas, como la búsqueda de variaciones diacrónicas en prosodia emocional.

5. Referencias

- Abramson, A. S. (2004). The plausibility of phonetic explanations of tonogenesis. En: G. Fant, H. Fujisaki, J. Cao, Y. Xu, & (Eds.), *From traditional phonology to modern speech processing: Festschrift for Professor Wu Zongji's 95th birthday* (pp. 17-29). Beijing: Foreign Language Teaching and Research Press.
- Achuff, P. (4 de Noviembre de 2001). The Lateralization of Emotion. Extraído el 7 de Julio de 2011, desde *Brain & Mind*:
<http://www.cerebromente.org.br/n14/mente/lateralization.htm>
- Arciuli, J., & Slowiaczek, L. M. (2007). The where and when of linguistic word-level prosody. *Neuropsychologia*, 45, 2638-2642.
- Bänziger, T., & Scherer, K. R. (2005). The role of intonation in emotional expressions. *Speech Communication*, 46, 252-267.
- Batista Antunes, L. (2007). *O papel da prosódia na expressão de atitudes do locutor em questões*. Belo Horizonte, Brasil: Tesis de Doctorado.
- Bavelier, D., Corina, D. P., & Neville, H. J. (1998). Brain and Language: a Perspective from Sign Language. *Neuron*, 21, 275-278.
- Belin, P., Fecteau, S., & Bédard, C. (2004). Thinking the voice: neural correlates of voice perception. *TRENDS in Cognitive Sciences*, 8 (3), 129-135.
- Benus, S., Gravano, A., & Hirschberg, J. (2007). Prosody, emotions, and... 'whatever'. En: *Proceedings of International Conference on Speech Communication and Technology* (pp. 2629-2632). Belgium.
- Besson, M., Magne, C., & Schön, D. (2002). Emotional Prosody: Sex differences in sensitivity to speech melody. *TRENDS in Cognitive Sciences*, 6 (10), 405-407.

- Blumstein, S. E., & Kurowski, K. (2006). The foreign accent syndrome: a perspective. *Journal of Neurolinguistics*, 19, 346-355.
- Boyes Braem, P. (1999). Rhythmic temporal patterns in the signing of deaf early and late learners of Swiss German sign language. *Language and Speech*, 42 (2-3), 177.
- Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25, 49-59.
- Breitenstein, C., Van Lancker, D., Daum, I., & Waters, C. H. (2001). Impaired Perception of Vocal Emotions in Parkinson's Disease: Influence of Speech Time Processing and Executive Functioning. *Brain and Cognition*, 45, 277-314.
- Brentari, D., González, C., Seidl, A., & Wilbur, R. (2011). Sensitivity to visual prosodic cues in signers and nonsigners. *Language and speech*, 54 (Pt 1), 49-72.
- Bryant, G. A., & Barret, H. C. (2008). Vocal Emotion Recognition Across Disparate Cultures. *Journal of Cognition and Culture*, 8, 135–148.
- Bryant, G. A., & Fox Tree, J. E. (2005). Is there an Ironic Tone of Voice? *Language and Speech*, 48 (3), 257-277.
- Buchanan, T. W., Lutz, K., Mirzazade, S., Specht, K., Shah, N. J., Zilles, K., et al. (2000). Recognition of emotional prosody and verbal components of spoken language: an fMRI study. *Cognitive Brain Research*, 9, 227-238.
- Caballero-Meneses, J. A., & Menez, M. (2010). Influencia del tempo de la música en las emociones. *Revista Colombiana de Psicología*, 19 (1), 37-44.
- Caffi, C., & Janney, R. W. (1994). Toward a pragmatics of emotive communication. *Journal of Pragmatics*, 22, 325-373.

- Campbell, N. (2004). Accounting for Voice-Quality Variation. *Speech Prosody 2004*, 217-220.
- Campbell, R., MacSweeney, M., & Waters, D. (2008). Sign Language and the Brain: A Review. *Journal of Deaf Studies and Deaf Education*, 13 (1), 4-20.
- Carlson, N. R. (2006a). Comunicación humana. En: N. R. Carlson, *Fisiología de la conducta* (pp. 531-569). Madrid: Pearson Educación.
- Carlson, N. R. (2006b). Audición, sentidos somáticos y sentidos químicos. En: N. R. Carlson, *Fisiología de la Conducta* (pp. 223-268). Madrid: Pearson Educación.
- Cavanaugh, J. R. (2005). Accent matters: Material consequences of sounding local in northern Italy. *Language & Communication*, 127-148.
- Chen, C., Xue, G., Mei, L., Chen, C., & Dong, Q. (2009). Cultural neurolinguistics. *Progress in Brain Research*, 178, 159-171.
- Chen, C.-M. (2007). The Functions of Prosody in Discourse Analysis - A Case Study of Tags in Mandarin Talk Shows. *Nouveaux cahiers de linguistique française*, 28, 327-333.
- Cid Uribe, M. E., & Maluenda Parraguez, L. (2005). El alargamiento segmental en el habla pública de Chile: Comportamiento Prosódico-Discursivo. *Onomázein*, 1, 43-55.
- Colín Rodea, M. (1992). El acento extranjero como marcador de origen y símbolo de poder a partir de las diferencias étnicas y psicolingüísticas. *Estudios de Lingüística Aplicada*, (15/16), 213-220.
- Corina, D. P., Bellugi, U., & Reilly, J. (1999). Neuropsychological studies of linguistic and affective facial expressions in deaf. *Language and Speech*, 42 (2-3), 307.
- Corros Mazón, F. J. (2007). La sociolingüística y el lenguaje gestual en la enseñanza del ELE: Algunas consideraciones. Las destrezas orales en la

enseñanza del español L2-LE: *XVII Congreso Internacional de la Asociación del Español como lengua extranjera (ASELE)* (pp. 431-444). Logroño: Universidad de La Rioja.

Cortés Moreno, M. (2006). Análisis acústico de la transferencia de rasgos del sistema tonal chino al habla en español como lengua extranjera. *Estudios de fonética experimental*, (15), 45-65.

Cortés Moreno, M. (2000). Sobre la adquisición de la prosodia en lengua extranjera: Estado de la cuestión. *Didáctica (Lengua y Literatura)*, 12, 91-119.

Cruz Aldrete, M. (2009). Gramática de la lengua de señas mexicana. Extraído el 2 de Diciembre de 2010, desde *Estudios de Lingüística del Español (ELiEs)*: Volumen 28: <http://elies.rediris.es/>

Curtis, M. E., & Bharucha, J. J. (2010). The Minor Third Communicates Sadness in Speech, Mirroring Its Use in Music. *Emotion*, 10 (3), 335-348.

Dachkovsky, S., & Sandler, W. (2009). Visual Intonation in the Prosody of a Sign Language. *Language and Speech*, 52 (2/3), 287-314.

Delogu, F., Lampis, G., & Olivetti Belardinelli, M. (2006). Music-to-language transfer effect: may melodic ability improve learning of tonal languages by native nontonal speakers? *Cognitive Processing*, 7 (3), 203-207.

DePaolis, R. A., Vihman, M. M., & Kunnari, S. (2008). Prosody in production at the onset of word use: A cross-linguistic study. *Journal of Phonetics*, 36, 406-422.

Devillers, L., & Vasilescu, I. (2004). Reliability of lexical and prosodic cues in two real-life spoken dialog corpora. En: *Proceedings of LREC*. Lisboa, Portugal.

Devillers, L., & Vidrascu, L. (2006). Real-life emotions detection with lexical and paralinguistic cues on Human-Human call center dialogs. En: *Proceedings of the Interspeech 2006* (pp. 801-804).

- Dromey, C., Silveira, J., & Sandor, P. (2005). Recognition of affective prosody by speakers of English as a first or foreign language. *Speech Communication*, 47, 351-359.
- Duanmu, S. (2004). Tone and Non-Tone Languages: An Alternative to Language Typology and Parameters. *Language and Linguistics*, 5 (4), 891-924.
- Emmorey, K., & McCullough, S. (2009). The bimodal bilingual brain: Effects of sign language experience. *Brain and Language*, 109 (2-3), 124-132.
- Emmorey, K., Allen, J. S., Bruss, J., Schenker, N., & Damasio, H. (2003). A morphometric analysis of auditory brain regions in congenitally deaf adults. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 10049-10054.
- Engelhardt, P. E., Patsenko, E. G., & Ferreira, F. (2007). Pupillometric Indices of Visual and Prosodic Information on Spoken Language Comprehension. En: D. S. McNamara, J. C. Trafton, & (Eds.), *Proceedings of the 29th Annual conference of the Cognitive Sciences Society* (pp. 971-976). Nashville: Cognitive Science Society.
- Erickson, D. (2006). Some gender and cultural differences in perception of affective expressions. En: *Proceedings of the Speech Prosody 2006 Conference* (p. paper 029). Dresden, Alemania: Disponible en: http://www.isca-speech.org/archive/sp2006/sp06_029.html.
- Faurot, K., Dellinger, D., Eatough, A., & Parkhurst, S. (1999). Lenguaje de signos mexicano. La identidad como lenguaje del sistema de signos mexicano. Extraído el 2 de Diciembre de 2010, desde *Sil.org*: <http://www.sil.org/Mexico/lenguajes-de-signos/G009e-Identidad-mfs.pdf>
- Fitzpatrick, J. (2000). On intonational typology. En: P. Siemund, & (ed.), *Methodological Issues in Language Typology. Sprachtypologie und Universalienforschung* (pp. 88-96).

- Fonseca, R. P., Wagner, G. P., Rinaldi, J., & Parente, M. A. (2007). O envelhecimento influencia as habilidades pragmáticas, léxico-semânticas e prosódicas do hemisfério direito? *Estudos interdisciplinares envelhecimento*, 12, 53-97.
- Frazier, L., Carlson, K., & Clifton Jr, C. (2006). Prosodic phrasing is central to language comprehension. *TRENDS in Cognitive Sciences*, 10 (6), 244-249.
- Gandour, J., Dzemidzic, M., Wong, D., Lowe, M., Tong, Y., Hsieh, L., Sathamnuwong, N. & Lurito, J. (2003). Temporal integration of speech prosody is shaped by language experience: An fMRI study. *Brain and Language*, 84, 378-336.
- Gandour, J., Tong, Y., Wong, D., Talavage, T., Dzemidzic, M., Xu, Y., Li, X. & Lowe M. (2004). Hemispheric roles in the perception of speech prosody. *Neuroimage*, 23, 344-357.
- Gandour, J., Wong, D., Dzemidzic, M., Lowe, M., Tong, Y., & Li, X. (2003). A Cross-Linguistic fMRI Study of Perception of Intonation and Emotion in Chinese. *Human Brain Mapping*, 18, 149-157.
- García Toro, M. (1997). Fisiología de la prosodia: nivel fonoarticulatorio. En: M. García Toro, *Aprosodia Expresiva en la Voz de Pacientes con Depresión* (pp. 23-24). Alcalá de Henares: Facultad de Medicina Universidad Alcalá de Henares.
- Glanz, J. (5 de Noviembre de 1999). Study Links Perfect Pitch to Tonal Language. Extraído el 28 de Julio de 2011, desde *New York Times*: <http://www.nytimes.com/1999/11/05/us/study-links-perfect-pitch-to-tonal-language.html>
- Gobl, C., & Ní Chasaide, A. (2003). The role of voice quality in communicating emotion, mood and attitude. *Speech Communication*, 40, 189-212.

- Goldsmith, J. (1994). Tone Languages. En: Asher R. & Simpson, J. (Eds.) *Encyclopedia of Language and Linguistics*, 4626-4628, Oxford, Nueva York: Pergamon Press
- Gomez, P., & Danuser, B. (2007). Relationships Between Musical Structure and Psychophysiological Measures of Emotion. *Emotion*, 7, 377-387.
- González-Álvarez, J., Parcet-Ibars, M. A., Ávila, C., & Geffner-Sclarsky, D. (2003). Una rara alteración del habla de origen neurológico: el síndrome del acento extranjero. *Revista de Neurología*, 36 (3), 227-234.
- Graber Rosenberg, T. M. (1998). *Prosodic Font*. Massachusetts: MIT, Program in Media Arts and Sciences, Tesis de maestría.
- Grandjean, D., & Scherer, K. R. (2006). Examining the neural mechanisms involved in the affective and pragmatic coding of prosody. *SP-2006* (p. paper 268). Dresden, Alemania: Disponible en: <http://www.isca-speech.org/archive>.
- Grandjean, D., Bänzinger, T., & Scherer, K. R. (2006). Intonation as an Interface between Language and Affect. *Progress in Brain Research*, 156, 235-247.
- Grice, M., & Baumann, S. (2007). An Introduction to Intonation - Functions and Models. En: T. Jürgen, G. Ulrike, & (eds.), *Non-Native Prosody Phonetic Transcription and Teaching Practice* (pp. 25-52). Berlín y Nueva York: Mouton de Gruyter.
- Grimshaw, G. M., Séguin, J. A., & Godfrey, H. K. (2009). Once more with feeling: The effects on emotional prosody on hemispheric specialisation for linguistic processing. *Journal of Neurolinguistics*, 22, 313-326.
- Gussenhoven, C. (1999). Discreteness and Gradience in Intonational Contrasts. *Language and Speech*, 42 (2-3), 283-305.
- Gussenhoven, C. (2002). Intonation and Interpretation: Phonetics and Phonology. En: B. Bel, & I. Marlien, *Proceedings of Speech Prosody 2002* (pp. 47-57). Aix-en-provence, Francia.

- Gussenhoven, C. (2001). Suprasegmentals. En: N. J. Smelser, & P. B. Baltes, *International Encyclopedia of the Social and the Behavioural Sciences* (pp. 15294-15298). Oxford: Pergamon.
- Gutiérrez Bravo, R. (2008). La identificación de los tópicos y los focos. *Nueva revista de filología hispánica*, 56 (2), 363-401.
- Harwell Rodriguez, R. (1998). Chapter 1 Introduction and Literature Overview. En: R. Harwell Rodriguez, *Acoustic and Perceptual Comparisons of Imitative Prosody in Kindergartners with and without speech disorders* (pp. 2-48). Tampa, Florida: Tesis de Maestría.
- Haunch, B. (sin año). Research into Boys and Girl's Voices. Extraído el 5 de Agosto de 2011, desde *Campaign for the Traditional Cathedral Choir*.
<http://www.ctcc.org.uk/musical.htm>
- Hickok, G., Bellugi, U., & Klima, E. S. (2001). Sign language in the Brain. *Scientific American*, 58-65, Junio.
- Hirschberg, J. (2002a). Communication and prosody: Functional aspects of prosody. *Speech Communication*, 31-43.
- Hirschberg, J. (2002b). The Pragmatics of Intonational Meaning. En: B. Bel, & I. Marlien, *Proceedings of the Speech Prosody 2002 conference* (pp. 65-68). Aix-en-provence, Francia.
- Hockett, C. (1971). El puesto del hombre en la naturaleza. En: C. Hockett, *Curso de lingüística moderna* (pp. 547-576). Buenos Aires: Editorial Universitaria.
- Hopkin, K. (6 de Noviembre de 2009). Babies Already Have an Accent. Extraído el 5 de Agosto de 2011, desde *Scientific American*:
<http://www.scientificamerican.com/podcast/episode.cfm?id=babies-already-have-an-accent-09-11-06>

- Hualde, J. I. (2001). Los sonidos de la lengua: Fonética y Fonología. En: J. I. Hualde, *Introducción a la lingüística hispánica* (pp. 45-71). Cambridge: Cambridge University Press.
- Huang, B. H., & Jun, S.-A. (2011). Age Effect on the Acquisition of Second Language Prosody. Extraído el 5 de Agosto de 2011, desde *Boston University*: www.bu.edu/buclid/files/2011/05/33-Huang.pdf
- Ilie, G., & Thompson, W. F. (2006). A comparison of the effects of music and speech prosody on three dimensions of affective experience. *Proceedings of the 9th International Conference on Music Perception and Cognition*, (pp. 20-27). Boloña.
- Joanette, Y., Ansaldo, A. I., Kahlaoui, K., Côté, H., Abusamra, V., Ferreres, A. & Roch-Lecours, A. (2008). Impacto de las lesiones del hemisferio derecho sobre las habilidades lingüísticas: perspectivas teórica y clínica. *Revista de Neurología*, 46 (8), 481-488.
- Juslin, P. N., & Laukka, P. (2003). Communication of Emotions in Vocal Expression and Music Performance: Different Channels, Same Code? *Psychological Bulletin*, 129 (5), 770-814.
- Juslin, P. N., & Västfjäll, D. (2008). Emotional Responses to Music: The Need to Consider Underlying mechanisms. *Behavioral and Brain Sciences*, 559-621.
- Kamiyama, T., & Yamamoto, Y. (2007). Visual representation of prosody for tactful communication skills - the case of request in Japanese as a Foreign Language taught to French university students. *Phonetics Teaching and Learning Conference 2007*. Londres, Reino Unido.
- Kelm, A., & Beckman, M. E. (2009). Prosodic structure and consonant development across languages. En: M. Vigário, S. Frota, M. J. Freitas, & (Eds.), *Interactions in phonetics and phonology* (pp. 109-136). Amsterdam/Filadelfia: John Benjamins.

- Kingston, J. (1999). Extrapolating from spoken to signed prosody via laboratory phonology. *Language and Speech*, 42 (2-3), 251.
- Klouda, G. V., Robin, D. A., Graff-Radford, N. R., & Cooper, W. E. (1988). The Role of Callosal Connections in Speech Prosody. *Brain and Language*, 35, 154-171.
- Kotz, S. A., Meyer, M., Alter, K., Besson, M., Cramon, D. Y., & Friederici, D. Y. (2003). On the lateralization of emotional prosody: an event-related functional MR investigation. *Brain and Language*, 86, 366-376.
- Kouassi, R. R. (2008). La construction suprasegmentale du sens. *Revue du Itml*, (1), 1-25.
- Ladd, D. R., Silverman, K. E., Tolkmitt, F., Bergmann, G., & Scherer, K. R. (1985). Evidence for the independent function of intonation contour type, voice quality and F0 range in signaling speaker affect. *Journal of the Acoustical Society of America*, 78, 435-444.
- Lakshminarayanan, K., Ben Shalom, D., Van Wassenhove, V., Orbelo, D., Houde, J., & Poeppel, D. (2003). The effect of spectral manipulations on the identification of affective and linguistic prosody. *Brain and Language*, 84, 250-263.
- Lastra, Y. (1992). El campo de la sociolingüística. En: Lastra, Y. *Sociolingüística para Hispanoamericanos. Una Introducción*. México: El Colegio de México.
- Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology*, 46, 1093-1096.
- Levitin, D. J., & Menon, V. (2003). Musical structure is processed in "language" areas of the brain: a possible role for Brodmann Area 47 in temporal coherence. *NeuroImage*, 20, 2142-2152.

- Li, W., & Yang, Y. (2009). Perception of Prosodic Hierarchical Boundaries in Mandarin Chinese Sentences. *Neuroscience*, 158, 1416-1425.
- Liang, J., & van Heuven, V. J. (2007). Chinese tone and intonation perceived by L1 and L2 listeners. En: C. Gussenhoven, & T. Riad, *Experimental Studies in Word and Sentence Prosody* (pp. 27-61). Berlín, Nueva York: Mouton de Gruyter.
- Livingstone, S. R., & Thompson, W. F. (2006). Multimodal affective Interaction: A comment on musical origins. *Music Perception*, 24 (1), 89-94.
- Luengo, I., Navas, E., Hernáez, I., & Sánchez, J. (2005). Reconocimiento automático de emociones utilizando parámetros prosódicos. *Procesamiento del Lenguaje Natural*, 35, 13-20.
- Machado, A., & Silva, F. J. (2007). Toward a Richer View of the Scientific Method. *American Psychologist*, 62 (7), 671-681.
- MacSweeney, M., Woll, B., Campbell, R., McGuire, P. K., David, A. S., Williams, S. C., Suckling, J., Calvert, G. & Brammer, M. (2002). Neural systems underlying British Sign Language and audio-visual English processing in native users. *Brain*, 125, 1583-1593.
- Makarova, V., & Petrushin, V. A. (2002). RUSLANA: A Database of Russian Emotional Utterances. En: *ICSLP-2002* (pp. 2041-2044).
- Martínez Sánchez, F., Montero Martínez, J. M., & de la Cerra, J. (2002). Sesgos cognitivos en el reconocimiento de expresiones emocionales de voz sintética en la alexitimia. *Psicothema*, 14 (2), 344-349.
- Mastropieri, D., & Turkewitz, G. (1999). Prenatal Experience and Neonatal Responsiveness to Vocal Expressions of Emotion. *Developmental Psychobiology*, 35 (3), 204-214.

- McGuire, P. K., Robertson, D., Thaker, A., David, A. S., Kitson, N., Frackowiak, R. S. & Frith, C. D. (1997). Neural correlates of thinking in sign language. *Neuroreport*, 8 (3), 695-698.
- McRoberts, G. W., Studdert-Kennedy, & Shankweiler, P. (1995). The role of fundamental frequency in signaling linguistic stress and affect: Evidence for a dissociation. *Perception & Psychophysics*, 57 (2), 159-174.
- Méar-Crine, A., & Leclerc, T. (1976). Attitudes des adolescents canadiens-français vis-à-vis du franco-québécois et du français académique. *Cahier de linguistique*, 6, 155-170.
- Mehler, J., Dupoux, E., Pallier, C., & Dehaene-Lambertz, G. (1994). Cross-linguistic approaches to speech processing. *Current Opinion in Neurobiology*, 4, 171-176.
- Meng, H., Tseng, C.-y., Kondo, M., Harrison, A., & Viselgia, T. (2009). Studying L2 Suprasegmental Features in Asian Englishes: A Position Paper. *INTERSPEECH-2009*, 1715-1718.
- Mildner, V. (2004). Hemispheric asymmetry for linguistic prosody: A study of stress perception in Croatian. *Brain and Cognition*, 55, 358-361.
- Miller, N., Lowit, A., & O'Sullivan, H. (2006). What makes acquired foreign accent syndrome foreign? *Journal of Neurolinguistics*, 19, 385-409.
- Moen, I. (2006). Analysis of a case of the foreign accent syndrome in terms of the framework of gestural phonology. *Journal of Neurolinguistics*, 19, 410-423.
- Moen, I. (2000). Foreign accent syndrome: A review of contemporary explanation. *Aphasiology*, 14 (1), 5-15.
- Munro, M. J., Derwing, T. M., & Burgess, C. S. (2010). Detection of nonnative speaker status from content-masked speech. *Speech Communication*, 52, 626-637.

- Narasimhan, S. V., Litha, R. J., B. Shahid, H., & Rohini, C. (2010). Acoustic Correlates of Stress in Mizo, a Tonal Language. *Language in India*, 10, 224-235.
- Nespor, M., & Sandler, W. (1999). Prosody in Israeli sign language. *Language and Speech*, 42 (2-3), 143.
- Neville, H. J., Mills, D. L., & Lawson, D. S. (1992). Fractionating Language: Different Neural Subsystems with Different Sensitive Periods. *Cerebral Cortex*, 2 (3), 244-258.
- Newman, A. J., Supalla, T., Hauser, P. C., Newport, E. L., & Bavelier, D. (2010). Prosodic and narrative processing in American Sign Language: An fMRI study. *NeuroImage*, 52, 669-676.
- Notimex en Guadalajara. (1 de Diciembre de 2010). Destacan importancia del Diccionario de mexicanismos . Extraído el 2 de Diciembre de 2010, desde *La crónica*: http://www.cronica.com.mx/nota.php?id_nota=547456
- Paradis, M. (1998). The other side of language: Pragmatic competence. *Journal of Neurolinguistics*, 11 (1-2), 1-10.
- Patel, A. D., Peretz, I., Tramo, M., & Labreque, R. (1998). Processing Prosodic and Musical Patterns: A Neuropsychological Investigation. *Brain and Language*, 61, 123-144.
- Patel, A., & Daniele, J. R. (2003). An empirical comparison of rhythm in language and music. *Cognition*, 87, B35-B45.
- Pell, M. D. (1999a). Fundamental Frequency Encoding of Linguistic and Emotional Prosody by Right Hemisphere-Damaged Speakers. *Brain and Language*, 69, 161-192.
- Pell, M. D. (2006). Implicit recognition of vocal emotions in native and non-native speech. *Speech Prosody* (Paper 259). Dresden, Germany: Disponible en: http://www.isca-speech.org/archive/sp2006/sp06_259.html.

- Pell, M. D. (2002). Surveying emotional prosody in the brain. En: B. Bel, I. Marlien, & (Eds), *Proceedings of the Speech Prosody 2002 Conference* (pp. 77-82).
- Pell, M. D. (1999b). The temporal organization of affective and non-affective speech in patients with right-hemisphere infarcts. *Cortex*, 35, 455-477.
- Pell, M. D., & Skorup, V. (2008). Implicit processing of emotional prosody in a foreign versus native language. *Speech Communication*, 50, 519-530.
- Pell, M. D., Monetta, L., Paulmann, S., & Kotz, S. (2009). Recognizing Emotions in a Foreign Language. *Journal of Nonverbal Behavior*, 33, 107-120.
- Pell, M. D., Paulmann, S., Dara, C., Allasseri, A., & Kotz, S. A. (2009). Factors in the recognition of vocally expressed emotions: A comparison of four languages. *Journal of Phonetics*, 37, 417-435.
- Pell, M. (2002). Surveying emotional prosody in the brain. En: B. Bel, I. Marlien, & (Eds), *Proceedings of the Speech Prosody 2002 Conference* (pp. 77-82).
- Pérez Espinosa, H., & Reyes García, C. A. (4 de Mayo de 2010). Reconocimiento de Emociones a Partir de Voz Basado en un Modelo Emocional Continuo. Extraído el 22 de Febrero de 2011, desde *INAOE Ciencias Computacionales*: <http://ccc.inaoep.mx/portalfiles/file/CCC-10-005.pdf>
- Perlman Lorch, M., Borod, J. C., & Koff, E. (1998). The role of emotion in the linguistic and pragmatic aspects of aphasic performance. *Journal of Neurolinguistics*, 11 (1-2), 103-118.
- Péron, J., Grandjean, D., Le Jeune, F., Sauleau, P., Haegelen, C., Drapier, D., Rouaud, T., Drapier, S. & Vérin, M. (2010). Recognition of emotional prosody is altered after subthalamic nucleus deep brain stimulation in Parkinson's disease. *Neuropsychologia*, 48, 1053-1062.
- Pihan, H., Tabert, M., Assuras, S., & Borod, J. (2008). Unattended emotional intonations modulate linguistic prosody processing. *Brain and Language*, 105, 141-147.

- Pipitone, R. N., & Gallup Jr, G. G. (2008). Women's voice attractiveness varies across the menstrual cycle. *Evolution and Human Behavior*, 29 (4), 268-274.
- Piske, T., MacKay, I. R., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics*, 29, 191-215.
- Ramos, J. M. (1997). Avaliação de dialetos brasileiros: o sotaque. *Revista de estudos da linguagem*, 6 (5), 103-125.
- Roberson, D., & Hanley, J. R. (2010). Relatively speaking: An account of the relationship between language and thought in the color domain. En: B. C. Malt, P. Wolff, & (eds.), *Words and the mind: How words capture human experience* (pp. 183-198). Nueva York: Oxford University Press.
- Rosenzweig, M. R., & Leiman, A. I. (1992). *Psicología fisiológica*. Mexico: McGrawHill.
- Ross, E. D., & Monnot, M. (2011). Affective prosody: What do comprehension errors tell us about hemispheric lateralization of emotions, sex and aging effects, and the role of cognitive appraisal. *Neuropsychologia*, 49, 866-877.
- Ross, E., & Monnot, M. (2008). Neurology of affective prosody and its functional-anatomic organization in right hemisphere. *Brain and Language*, 104, 51-74.
- Russ, J. B., Gur, R. C., & Bilker, W. B. (2008). Validation of affective and neutral sentence content for prosodic testing. *Behavior Research Methods*, 40 (4), 935-939.
- Rymarczyk, K., & Grabowska, A. (2007). Sex differences in brain control of prosody. *Neuropsychologia*, 45, 921-930.
- Saíz Ruiz, J., & Talavera Martín, J. A. (1997). *Aprosodia expresiva en la voz de pacientes con depresión*. Facultad de Medicina. Universidad de Alcalá de Henares: Tesis doctoral.

- Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proceedings of the National Academy of Sciences*, 107 (6), 2408-2412.
- Scherer, K. R. (1995). Expression of Emotion in Voice and Music. *Journal of Voice*, 9 (3), 235-248.
- Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication*, 40, 227–256.
- Scherer, K. R., & Bänziger, T. (2004). Emotional Expression in Prosody: A Review and an Agenda for Future Research. *SP-2004* (pp. 359-366). Nara, Japon: Disponible en: <http://www.isca-speech.org/archive>.
- Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion Inferences from Vocal Expression Correlate Across Languages and Cultures. *Journal of Cross-Cultural Psychology*, 32 (1), 76-92.
- Schirmer, A., Kotz, S. A., & Friederici, A. D. (2002). Sex differentiates the role of emotional prosody during word processing. *Cognitive Brain Research*, 14, 228-233.
- Scott, S. K., Clegg, F., Rudge, P., & Burgess, P. (2006). Foreign accent syndrome, speech rhythm and the functional neuroanatomy of speech production. *Journal of Neurolinguistics*, 19, 370-384.
- Seddoh, A. S. (2002). How discrete or independent are "affective prosody" and "linguistic prosody"? *Aphasiology*, 16 (7), 683-692.
- Seppänen, T., Väyrynen, E., & Toivanen, J. (2003). Prosody-based classification of emotions in spoken Finnish. En: *Proceedings of Eurospeech '03* (pp. 717-720). Ginebra.
- Shah, A. P., Baum, S. R., & Dwivedi, V. D. (2006). Neural substrates of linguistic prosody: Evidence from syntactic disambiguation in the productions of brain damaged patients. *Brain and Language*, 96, 78-89.

- Shapiro, L. P., & Nagel, N. H. (1995). Lexical Properties, Prosody, and Syntax: Implications for Normal and Disordered Language. *Brain and Language*, 50, 240-257.
- Shattuck-Hufnagel, S., & Turk, A. E. (1996). A Prosody Tutorial for Investigators of Auditory Sentence Processing. *Journal of Psycholinguistic Research*, 25 (2), 193-247.
- Shochi, T., Aubergé, V., & Rilliard, A. (2006). Comment les attitudes prosodiques sont parfois de « faux-amis » : les affects sociaux du japonais vs. français. Extraído el 24 de noviembre de 2011 desde *Les Journées d'Étude sur la Parole*: <http://jep2006.irisa.fr/openconf/author/final/final-87.pdf>
- Snow, D. (2000). The Emotional Basis of Linguistic and Nonlinguistic Intonation: Implications for Hemisphere Specialization. *Developmental Neuropsychology*, 17 (1), 1-28.
- Stockwell, R. P. (1963). Review: Generality, gradience and the all-or-none. *Language*, 39 (1), 87-91.
- Szymanowski, F., Szymanowski, F., Kotz, S. A., Schröder, C., Rotte, M., & Dengler, R. (2007). Gender Differences in Processing Emotional Prosody. *Clinical Neuropsychology*, 118 (4), e102-e103.
- The Ohio State University Department of Linguistics. (1999). ToBI. Extraído el 13 de febrero de 2011, desde *The Department of Linguistics at The Ohio State University*: <http://www.ling.ohio-state.edu/~tobi/>
- Thompson, W. F., & Balkwill, L. L. (2006). Decoding speech prosody in five languages. *Semiotica*, 158, 407-424.
- Thompson, W. F., Schellenberg, E. G., & Husain, G. (2004). Decoding Speech Prosody: Do Music Lessons Help? *Emotion*, 4 (1), 46-64.

- Thönnessen, H., Boers, F., Dammers, J., Chen, Y.-H., Norra, C., & Mathiak, K. (2010). Early sensory encoding of affective prosody: Neuromagnetic tomography of emotional category changes. *NeuroImage*, 50, 250-259.
- Toivanen, J., Väyrynen, E., & Seppänen, T. (2004). Automatic Discrimination of Emotion from Spoken Finnish. *Language and Speech*, 47 (4), 383-412.
- Toledo, G. (2010). Métricas rítmicas en microdiscursos. *Onomázein*, 21 (1), 71-95.
- Tyrone, M. E., Nam, H., Saltzman, E., Mathur, G., & Goldstein, L. (2010). Prosody and Movement in American Sign Language: A Task-Dynamics Approach. *Proceedings of the Speech Prosody 2010*. Chicago.
- Urberg Carlson, K., & Munson, B. (2008). Breathiness and Perceived Sex Typicality. *Poster presentado en la convención ASHA 2008*, (pp. 20-22). Chicago.
- Van Lancker, D., Pachana, N., Cummings, J. L., & Sidtis, J. J. (2006). Dysprosodic speech following basal ganglia insult: Toward a conceptual framework for the study of the cerebral representation of prosody. *Brain and Language*, 97, 135-153.
- Van Santen, J. P., Tucker Prud'hommeaux, E., & Black, L. M. (2009). Automated assessment of prosody production. *Speech Communication*, 51, 1082-1097.
- Varley, R., Whiteside, S., Hammill, C., & Cooper, K. (2006). Phases in speech encoding and foreign accent syndrome. *Journal of Neurolinguistics*, 19, 356-369.
- Ward, N. G., & Bayyari, Y. A. (2010). American and Arab Perceptions of an Arabic Turn-Taking Cue. *Journal of Cross-Cultural Psychology*, 41 (2), 270-275.
- Wildgruber, D., Ackermann, H., Kreifelts, B., & Ethofer, T. (2006). Cerebral Processing of Linguistic and Emotional Prosody: fMRI Studies. *Progress in brain research*, 156, 249-268.

- Wildgruber, D., Riecker, A., Hertrich, I., Erb, M., Grodd, W., Ethofer, T. & Ackermann, H. (2005). Identification of emotional intonation evaluated by fMRI. *NeuroImage*, 24 (4), 1233-1241.
- Wilson, D., & Wharton, T. (2006). Relevance and prosody. *Journal of Pragmatics*, 38, 1559-1579.
- Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences*, 104 (19), 7780–7785.
- Worchel, S., Cooper, J., Goethals, G. R., & Olson, J. (2002). Prejuicios y estereotipos. En: S. Worchel, J. Cooper, G. R. Goethals, & J. Olson, *Psicología Social*. México: Thomson.
- Wymer, J. H., Lindman, L. S., & Booksh, R. L. (2002). A Neuropsychological Perspective of Aprosody: Features, Function, Assessment, and Treatment. *Applied Neuropsychology*, 9 (1), 37-47.
- Yang, L.-c. (2010). Harmony and Tension in Mandarin Chinese Prosody: Constraints and Opportunities of Lexical Tones in Discourse Markers. *Speech Prosody 2010*, 11-14.
- Young, C. E. (2003). College Student's Reactions to Accents of L2 Learners of Spanish and English. *Selected Proceedings of the First Workshop on Spanish Sociolinguistics* (pp. 107-111). Massachusetts: Cascadia Proceedings Project. Disponible en: www.lingref.com, document #1013.
- Yuen, I. (2007). Declination and tone perception in Cantonese. En: C. Gussenhoven, & T. Riad, *Experimental Studies in Word and Sentence Prosody* (pp. 63-78). Berlin, Nueva York: Mouton de Gruyter.
- Zakzanis, K. K. (1999). Ideomotor prosodic apraxia. *Journal of Neurology, Neurosurgery & Psychiatry*, 694-695.

Zellner Keller, B. (2006). Ageing and Speech Prosody. *Speech Prosody 2006*, Paper 1.

Zhang, S., Li, K., Lo, W.-K., & Meng, H. (2010). Perception of English suprasegmental Features by Non-Native Chinese Learners. *Speech Prosody 2010* (pp. 1-4). Chicago: Disponible en:
<http://aune.lpl.univ-aix.fr/~sprosig/sp2010/papers/100942.pdf>.

6. Anexos

6.1 Anexo 1: Tabla de definiciones.

Se presentan definiciones de diversos autores a partir de las cuales se realizó el análisis conceptual. A la izquierda se incluye la referencia, en la columna del centro el tipo de definición (prosodia (G)eneral,(L)ingüística o (E)mocional y en la columna de la extrema derecha se incluye la definición en su idioma original.

Referencia	Tipo	Definición
Arciuli & Slowiaczek (2007)	G	All the world's languages exhibit rhythm or 'prosody' by incorporating supra-segmental features such as stress and/or tonal variation. These features can be marked by acoustic parameters including fundamental frequency (F0), duration and amplitude.
Bänziger & Scherer (2005)	G	The prosody of connected speech may be analysed and described in terms of the variation of a large number of prosodic features. There are, however, three features which are most consistently used for linguistic purposes either singly or jointly. These three features are pitch, length, and loudness.
Batista Antunes (2007)	G/G/G	Como sentido amplo, muitos estudos utilizam, então, o termo prosódia como sinônimo do supra-segmental, ou seja, tudo que encontra-se além do nível do segmento; mas, como nos lembra Couper-Kuhlen (1986), é preciso cautela nessa associação, pois há alguns fenômenos supra-segmentais que não são considerados prosódicos, como a coarticulação, a assimilação, a nasalização progressiva ou regressiva, etc. / O conceito restrito de prosódia define a prosódia como sinônimo de entonação / Neste trabalho optamos pelo sentido amplo de prosódia, ou seja, não a tomaremos apenas como sinônimo de melodia da fala, ligada somente ao parâmetro físico da frequência fundamental, mas englobando outros aspectos, tais como, duração de sílabas e pausas, além da interação entre a frequência fundamental e a duração. O parâmetro intensidade não pôde ser levado em consideração.
Besson et al., (2002)	E	Emotional prosody is defined as the ability to express emotions through variations of different parameters of the human speech, such as pitch contour, intensity and duration.
Blumstein & Kurowski (2006)	G	Broadly construed, prosody encompasses those properties of speech relating to pitch, stress, and quantity.
Bryant & Barret	G	Prosodic features in speech (pitch, loudness, duration, and

(2008)		spectral properties) [...] / [...] global prosodic information (i.e., pitch, loudness, duration and spectral information) [...]
Buchanan et al. (2000)	G/E	The term prosody describes the non-propositional cues that may be passed along through language, including intonation, stresses, and accents. / Emotional prosody involves the expression of emotion through the intonation of spoken language.
Campbell (2004)	E (G?)	“The term affective prosody refers to the combination of attitudinal and emotional prosody.
Carlson (2006a)	G	Uso de los cambios de entonación y énfasis para añadir significado al habla aparte del especificado por las palabras concretas; es un medio importante de comunicación de emociones.
Dromey et al. (2005)	E	The non-verbal acoustic cues that convey the emotional valence of a speaker’s communication are often referred to as affective prosody.
Engelhardt, Patsenko, & Ferreira (2007)	G	Prosody is the stress, timing, and intonation of an utterance, and is related to syntactic structure
Erickson (2006)	E	Acoustic cues of affective prosody include F0, intonation pattern, duration/speech rate, loudnes, voice quality and a combination of all of these.
Harwell Rodriguez (1998)	G	the “suprasegmental features of fundamental frequency, duration and intensity that contribute to...the melody of speech production”
Joanette et al. (2008)	G	La prosodia es el componente del lenguaje que refiere al procesamiento cognitivo necesario para comprender o expresar intenciones comunicativas usando aspectos suprasegmentales del habla, tales como las variaciones de la entonación, las pausas y las modulaciones de la intensidad vocal.
Klouda et al. (1988)	G	Prosody refers to variations in three acoustic properties of the speech wave (duration, amplitude, and fundamental frequency) which can alter the meaning of spoken language.
Kouassi (2008)	G	Les unités “suprasegmentales” ou prosodiques constituent un ensemble de faits qui se posent sur les structures segmentales pour, ensemble, produire la signification. Ces unités ne sont pas encore unanimement reconnues comme faisant parties des phénomènes de langue [...] L’intonation est « <i>un processus linguistique associant une forme acoustique (F0, durée, énergie)</i>

		à un sens »
Li & Yang (2009)	G	Prosody comprises intonation, accentuation and rhythmic patterns, conveying both linguistic and affective information to listeners by variations in acoustic–phonetic parameters such as fundamental frequency (F0), intensity, duration, timbre and spectral characteristics.
Pell & Skorup (2008)	G	[...] the suprasegmental content of speech or <i>speech prosody</i> (negritas añadidas)
Pell, Monetta, Paulmann, & Kotz (2009)	E	During speech communication, listeners attend to changes in pitch, loudness, rhythm, and voice quality (emotional prosody) to form an impression about the speaker’s emotion state in conjunction with linguistic decoding
Péron et al. (2010)	E	Emotional prosody is defined as modifications in segmental and supra-segmental speech parameters during an emotional episode
Russ et al. (2008)	G	“that faculty of speech which conveys different shades of meaning by means of variations in stress and pitch—irrespective of the words and grammatical construction”
Rymarczyk & Grabowska (2007)	E/L	Affective (emotional) prosody is a neuropsychological function that encompasses non-verbal aspects of language that are necessary for recognizing and conveying emotions in communication, whereas non-affective (linguistic) prosody [...] mainly serves to remove syntactic ambiguities and to clarify the interactive character of a sentence; that is, whether a sentence is a question, an order, or a neutral statement. It can also distinguish between phrases and noun compounds (e.g. green house versus greenhouse)
Seddoh (2002)	G	Together with tone, pitch accent, stress, length, and rhythm, intonation is an element of prosody, the nonsegmental component of spoken language
Shattuck-Hufnagel & Turk (1996)	G	As a working definition, we specify prosody as both (1) acoustic patterns of F 0, duration, amplitude, spectral tilt, and segmental reduction, and their articulatory correlates, that can be best accounted for by reference to higher-level structures, and (2) the higher-level structures that best account for these patterns. We subscribe to Beckman's (1996) observation that prosody is "a complex grammatical

		structure that must be parsed in its own right"; it is "the organizational structure of speech."
Snow (2000)	G	Speech prosody involves the "musical aspects of language - variations in the pitch, length, and loudness of spoken utterances".
Thompson & Balkwill (2006)	G	Prosody refers to vocal qualities of speech, as distinct from verbal content, and includes intonation, stress, and timing. It signals points of emphasis, indicates a statement or question, and conveys emotional connotations
Thompson et al. (2004)	G	<i>Speech prosody</i> refers to the musical aspects of speech, including its melody (intonation) and its rhythm (stress and timing).
Van Lancker et al. (2006)	G	Prosody, the melody of speech, is made up of fluctuations in pitch or fundamental frequency, variations in loudness (or intensity), a number of durational features (e.g., syllable, word, phrase, and a breath group length, pausing, phrase final lengthening, tempo, and rate) and changing voice quality.
Wildgruber et al. (2005)	E	At the perceptual level, emotional tone is characterized by variations of pitch, syllable durations, loudness, and voice quality across utterances (suprasegmental features) imposed upon segmental verbal information encoded in phonetic/phonological units
Zakzanis (1999)	G	Prosody is a non-verbal or suprasegmental feature of language that conveys various levels of information to the listener, including linguistic, affective (attitudinal and emotional), dialectical, and idiosyncratic data. The acoustical features underlying prosody include pitch, intonation, melody, cadence, loudness, timbre, tempo, stress, accent, and pauses.
Wymer, Lindman, & Booksh (2002)	G	The nonlinguistic melodic line of speech produced by variations of pitch, rhythm, and stress of pronunciation, resulting in semantic (conceptual or associative) or emotional meaning.