

Expresión y contacto: dimensiones de la afectividad en prosodia

Francisco José CANTERO SERENA
Universitat de Barcelona

RESUMEN. En este trabajo, se exploran las distintas dimensiones de la afectividad relacionadas con la prosodia del habla. En primer lugar, se distinguen los distintos códigos verbales que intervienen en el intercambio comunicativo, entre los que hay que distinguir los que están fijados socialmente (códigos estables) y los que deben negociar los interlocutores (códigos semiestables). En segundo lugar, se propone un método objetivo para el análisis acústico de la prosodia en su conjunto: además de la entonación (la melodía del habla), la dinámica (el juego de intensidades dentro del enunciado) y el ritmo (el juego de duraciones). Más adelante, se caracteriza funcionalmente el lenguaje afectivo y los códigos que intervienen en él. Finalmente, se explica cómo la prosodia es el contenedor (o recipiente fónico) de todos los contenidos lingüísticos que se ponen en juego en la comunicación. En conjunto, las dimensiones de la afectividad en prosodia se relacionan con: la expresión de emociones, cortesías y foco, la modalización de la información y el contacto entre los interlocutores.

PALABRAS CLAVE. Prosodia, entonación, intensidad, duración, afectividad, interacción.

ABSTRACT. In this paper, we explore the different dimensions of affectivity that are related to speech prosody. On one hand, we distinguish between the different verbal codes that take part in the communicative exchange: the ones that are socially established (stable codes) and those that must be implicitly agreed or set by the interlocutors (semi-stable codes). On the other hand, we propose an objective method to conduct the acoustic analysis of the speech prosody as a whole: that is to say, taking into account the intonation (speech melody), the dynamics (the relation of intensities within the sentence) and the rhythm (the relation of durations) of the speech. Later, we explain what affective language is and which codes are involved in it. Finally, we explain how prosody is the container (or phonic holder) of all the linguistic contents that are brought into play when we try to communicate. All in all, the dimensions of affectivity in prosody are related to: the expression of emotions, politeness and focus; the modalization of information; and the contact between the interlocutors.

KEYWORDS. Prosody, intonation, intensity, duration, affectivity, interaction.

1. INTRODUCCIÓN: MÁS ALLÁ DEL CÓDIGO LINGÜÍSTICO

Entendemos la entonación como la interpretación lingüística de la melodía del habla, esto es, de la melodía que constituyen los valores tonales sucesivos (especialmente, el tono de las vocales) a lo largo del discurso hablado. En nuestro marco teórico (*cfr.* Cantero Serena 2002, Cantero Serena & Mateo Ruiz 2011), distinguimos tres niveles de análisis de las melodías: un nivel de análisis prelingüístico, un nivel de análisis lingüístico y un nivel de análisis paralingüístico. Estos tres niveles son perspectivas complementarias de análisis que

nos permiten identificar y describir las distintas funciones que cumple la entonación, pues una misma melodía, cada melodía, siempre cumple diversas funciones simultáneamente (en una suerte de *sincretismo melódico*) y nuestro análisis (ya sea fonético o fonológico) debe distinguirlas con toda claridad, para poder comprenderlas.

Desde la perspectiva de análisis prelingüística, la entonación cumple una función de contenedor fónico de las unidades del habla, permite integrar el discurso hablado en unidades reconocibles e inteligibles y es la base de fenómenos como el acento dialectal: cada comunidad geográfica, cada variedad del idioma, tiene un perfil melódico determinado (su «acento») que estructura el discurso y lo cohesiona.

Desde una perspectiva de análisis lingüística, por su parte, la entonación permite distinguir las melodías que aportan significatividad por sí solas al enunciado, al margen de cualquier otra dimensión léxica, gramatical o pragmática. La entonación lingüística no aporta significados concretos, sino que su valor opositivo consiste en modalizar la información que esas otras dimensiones (léxica, gramatical o pragmática) aportan al discurso y en marcar los enunciados como neutros, interrogativos, suspendidos o enfáticos.

Desde una perspectiva de análisis paralingüística, finalmente, la entonación cumple otras funciones comunicativas, más allá de la significatividad estrictamente lingüística. Distinguiamos tres tipos de entonación paralingüística: la entonación emocional, la entonación de cortesía y la entonación de foco (*vid. fig. 1*):

Entonación prelingüística	Entonación lingüística	Entonación paralingüística
↓	↓	↓
Integración del discurso Acento dialectal Acento extranjero	Distinciones lingüísticas: Neutra / Interrogativa / Suspendida / Enfática	Entonación emocional Entonación de cortesía Entonación de foco

Figura 1. Niveles de análisis de la entonación

El nivel prelingüístico constituye un nivel previo y bien diferenciado del nivel lingüístico, por cuanto los perfiles melódicos dialectales (o de acento extranjero), sus rasgos y sus patrones, son exclusivos de comunidades de habla determinadas y no de todos los hablantes de la lengua. Los rasgos de la entonación dialectal, así, constituyen códigos estables, propios de una comunidad, frente al código lingüístico, común a todos los hablantes de la lengua. Por ejemplo, en un enunciado determinado podemos identificar los rasgos que definen la entonación interrogativa en español (que forman parte del código lingüístico del idioma), independientemente de los rasgos dialectales que también identificamos en ese mismo enunciado (y que forman parte del código prelingüístico propio de los hablantes de esa zona).

Por su parte, las entonaciones paralingüísticas constituyen también códigos de la lengua, pero no forman parte del código lingüístico como tal (no son comunes a todos los hablantes del idioma), sino que son característicos de una comunidad de habla específica y constituyen códigos semiestables (*vid.* Cantero Serena 2014a).

Más allá del código lingüístico del idioma, por tanto, las entonaciones prelingüísticas (como los acentos dialectales) constituyen códigos estables, pero parciales (propios solo de una zona, de una variedad del idioma); y las entonaciones paralingüísticas constituyen códigos semiestables (propios, también, de una comunidad de habla).

Nos referimos a códigos de entonación «semiestables» cuando los hablantes de una comunidad de habla pueden negociar los rasgos melódicos que aportan cortesía, por ejemplo, o ironía, y compartirlos: tales rasgos no constituyen un código estable porque la negociación permite (y requiere) una cierta flexibilidad en su concreción, de modo que otros oyentes, ajenos al intercambio en el que esos hablantes están negociando sus formas de cortesía, pueden no comprender su sentido, porque no han participado de la negociación¹.

Tales códigos semiestables (que pueden describirse, por tanto, solo hasta cierto punto) afectan a la entonación de cortesía y descortesía, a la entonación de foco (que informa a fenómenos como la ironía) y a un tipo de entonación emocional convencionalizada socialmente (como la que emplean los actores, por ejemplo, para expresar emociones: actores que son eficaces, precisamente, porque hay una cierta convención en este tipo de entonaciones); otras entonaciones emocionales, idiosincrásicas, entendemos que pueden estar no codificadas, sino que serían expresión genuina, no controlada, de la emoción del hablante (*vid.* fig. 2):

Entonación prelingüística		<i>Códigos estables</i> (de cada variedad del idioma)	
Entonación lingüística		<i>Código estable</i> (del idioma)	
Entonación paralingüística	De cortesía		<i>Códigos semiestables</i> (no del idioma, sino de cada comunidad de habla)
	De foco		
	Emocional	Convencional	
		Genuina	<i>Entonación idiosincrásica no codificada</i>

Figura 2. Códigos de la entonación

Los códigos prelingüísticos, por su carácter meramente integrador del discurso, no aportan ningún nivel de significatividad propia al mensaje: es decir, los acentos dialectal o extranjero no son significativos en sí mismos, no «quieren decir» nada ni aportan una intencionalidad al mensaje² sino que son meros portadores del discurso; en cambio, los códigos

¹ Nótese que cuando hablamos de *negociación* nos referimos a un concepto que (aunque lo incluye) va más allá del concepto clásico de *cooperación*: en este, nos referimos a las estrategias con que los interlocutores facilitan y regulan el intercambio; en aquel, a las estrategias que permiten compatibilizar los códigos que cada interlocutor emplea y ajustar cada uno de sus aspectos (las pronunciaciones, las melodías, los significados de las palabras, las formas flexivas, las estructuras sintácticas, las intenciones que cada interlocutor proyecta: cada rasgo y cada aspecto de cada código es objeto de esa negociación tácita).

² Excepto, tal vez, en sus usos humorísticos o irónicos, en cuyo caso serían ya entonaciones paralingüísticas: por ejemplo, en los casos de «cambio de código» (*code switch*: *vid.* Heller 1988) entre dialectos,

paralingüísticos (por ejemplo, la cortesía o la ironía verbal) sí que son totalmente significativos e intencionales, pero específicos de un grupo cultural o identitario: códigos que se comprenden no en todo el dominio lingüístico del idioma, sino en determinadas comunidades de habla (geográficas, de estrato social o de registro, incluyendo los grupos culturales, profesionales o incluso familiares).

Un fenómeno crucial en este juego de códigos es la confusión entre los códigos prelingüísticos y paralingüísticos: para los hablantes con un acento dialectal determinado, o con un acento extranjero, su acento (su entonación prelingüística) no es intencional y no aporta ningún significado ni ninguna modalidad: sencillamente, como hemos dicho, es su forma de hablar. Un hablante andaluz, o un hablante vasco, tienen un acento dialectal característico del que tal vez ni siquiera son conscientes; del mismo modo, un hablante de español brasileño o un italiano tienen un acento extranjero no intencional, cuya causa es la transferencia de los patrones entonativos de su lengua propia.

Sin embargo, para un oyente de otra región (o para un oyente nativo, ante los hablantes con acento extranjero), es muy común confundir ese acento dialectal (o extranjero) con una intención comunicativa determinada. Así, para un español del norte una afirmación neutra de un andaluz puede parecerle una entonación expresiva, de sorna o de alegría; para un levantino, una afirmación neutra de un vasco puede parecerle un énfasis de cortesía, una apelación directa a su interlocutor: entonces, los andaluces pueden parecer gente divertida (o chusca) y los vascos pueden parecer gente confiable y campechana. Del mismo modo, a un español peninsular la entonación de un brasileño puede parecerle de una extrema cortesía, seguramente no confiable, mientras que la entonación de un italiano puede parecerle muy personal, muy emocional, muy seductora.

Tomar el acento dialectal o extranjero (que, para el hablante, es neutra) como una entonación marcada afectivamente suele ser una confusión que tal vez está en la base de no pocos malentendidos entre grupos sociales y culturales, y no pocos problemas de rechazo al «otro».

Esta confusión entre diversos tipos de código (prelingüísticos y paralingüísticos) se incrementa cuando los códigos que manejan los hablantes no son del todo bien conocidos. Como hemos dicho, los códigos paralingüísticos son códigos semiestables y parciales: no los comparten todos los hablantes de la lengua, y no son códigos bien fijados, sino que deben negociarse. Esta negociación es sencilla entre hablantes de una misma comunidad de habla, pero ardua entre hablantes de comunidades diferentes, más aún cuando se trata de comunidades identitarias.

Como veremos, este tipo de negociación se relaciona con las paredes del «recipiente fónico» (*vid. infra*, apartado 5) que contiene el discurso, que son paredes prosódicas: la melodía, la dinámica y el ritmo del habla.

un hablante puede remedar la entonación de otro dialecto para enfatizar su mensaje (por ejemplo, para ironizarlo): entonces, esa entonación dialectal, en ese contexto, constituye un fenómeno de foco (foco ancho, que afecta a todo el enunciado) y, por tanto, no sería una entonación prelingüística, sino paralingüística. Sobre el cambio de código entre variedades de la lengua: *vid. Cantero & De Arriba (1996)*.

2. MÁS ALLÁ DE LA ENTONACIÓN: DIMENSIONES DE LA PROSODIA

Cuanto hemos dicho sobre los niveles de análisis de la entonación y los códigos que constituyen, puede decirse no solo de la entonación sino de los demás fenómenos suprasegmentales del habla: de la prosodia en su conjunto. A menudo, nos referimos a la prosodia del habla centrándonos exclusivamente en su entonación, pero la prosodia abarca más fenómenos, algunos de los cuales son factores clave en la transmisión de información y, sobre todo, de afectividad.

Como hemos visto, la parte material de la entonación (es decir, su forma fonética) es la *melodía* del habla: la relación de intervalos que constituye la sucesión de tonos perceptibles a lo largo del enunciado. Pero la prosodia es más que su entonación.

Además de la melodía, también es relevante el juego de intensidades: su *dinámica*. Cada vocal tiene una frecuencia fundamental (F_0 , que percibimos como un tono, constituyente de la melodía del enunciado) y también tiene un pico de intensidad. De hecho, las vocales son voz que sale libremente a su paso por los resonadores, y la voz es energía; frente a las vocales, las consonantes son constricciones en la salida del aire, interrupciones (totales o parciales) en la señal acústica. Las relaciones entre los picos de intensidad de las vocales constituyen también una suerte de «melodía de intensidades» que caracteriza y afecta al enunciado. Esa dinámica es una parte importante de la prosodia del discurso, por más que apenas está codificada lingüísticamente: afecta, más que a la modalización de la información, a la expresión de la afectividad.

Junto al juego de tonos (la melodía) y de intensidades (la dinámica), también es relevante el juego de duraciones de las unidades del enunciado, que constituye su *ritmo*. No es fácil determinar la duración de las unidades del discurso, porque los «segmentos» (las vocales y las consonantes) a menudo están solapados: en realidad, los segmentos no son «segmentos», sino haces de rasgos acústicos que percibimos conjuntamente como el «timbre» de un sonido. Percibimos el timbre de la vocal, la identificamos con naturalidad, pero a menudo no podemos determinar su punto de inicio o su punto final. Esta dificultad para medir las duraciones ha impedido, durante mucho tiempo, realizar un análisis objetivo del juego de duraciones que caracteriza la prosodia de un enunciado. Sin embargo, el ritmo es un fenómeno crucial en la prosodia, responsable de buena parte de la carga afectiva de los enunciados. Nosotros hemos optado por medir las distancias entre los picos de intensidad de cada vocal: los pies rítmicos, que nos ofrecen un valor objetivo y seguro.

La *melodía*, la *dinámica* y el *ritmo* son las tres dimensiones fonéticas de la prosodia del habla, fenómenos que hasta ahora no se han estudiado sistemáticamente, en parte por considerarse, tradicionalmente, como ajenos a la doble articulación del lenguaje (es decir, fuera del código lingüístico, relacionados más con el intercambio de afectividad que con la transmisión de información) y, en parte, por la propia dificultad que entraña su análisis acústico.

Tomando como punto de partida nuestro método de Análisis Melódico del Habla (AMH)³, que consiste en identificar los centros tonales del enunciado, relativizar sus valores, determinar sus intervalos y establecer la melodía que constituyen, en Cantero Serena (2019a) hemos propuesto un método de Análisis Prosódico del Habla (APH) que permite realizar también el análisis dinámico y el análisis rítmico de los enunciados: un método de análisis acústico que permite describir y comparar la melodía del enunciado, su dinámica y su ritmo.

Veamos un ejemplo. En la fig. 3, mostramos el sonograma y las curvas de F_0 y de intensidad de un enunciado de nuestro corpus de habla espontánea en español peninsular (*vid.* Cantero Serena 2016): «¿Qué recuerda con mayor agrado de... de esa época?». El locutor, en este enunciado, está transmitiendo un cierto desagrado sobre lo que pregunta (tal vez, para orientar la respuesta a su interlocutor): se trata, por tanto, de un enunciado con una clara marca de emocionalidad.

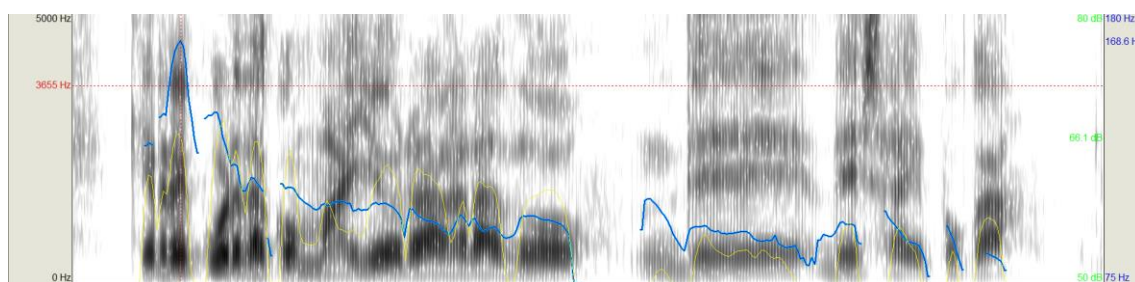


Figura 3. Sonograma del enunciado: «¿Qué recuerda con mayor agrado de... de esa época?»

Para realizar el análisis prosódico, tomamos los valores en Hz de la F_0 las vocales del enunciado, los valores en dB de sus picos de intensidad y los valores de tiempo de esos picos de intensidad. El método APH consiste en relativizar los valores absolutos en Hz y dB y generar una curva estándar que exprese la melodía y la dinámica del enunciado; paralelamente, se procesan los valores de tiempo, midiendo la distancia entre los picos de intensidad (esa distancia será la duración de cada pie rítmico) y se relativizan, de modo que el gráfico resultante exprese las relaciones de duración entre los pies rítmicos.

En la fig. 4 se muestra el gráfico resultante del análisis prosódico completo.

En el gráfico, puede observarse la curva melódica del enunciado: un primer pico en «re», un cuerpo descendente hasta el titubeo en «de...», y la inflexión final, resituada al alza después del titubeo.

Superpuesta a ella, aparece la curva dinámica del enunciado, que expresa las relaciones entre los picos de intensidad de cada vocal. Puede observarse cómo hay una mayor intensidad en las vocales tónicas de cada palabra «recuerda», «mayor», «agrado», «época».

³ Véase una visión de conjunto del método, actualizada, en Cantero Serena & Font-Rotchés (2020).

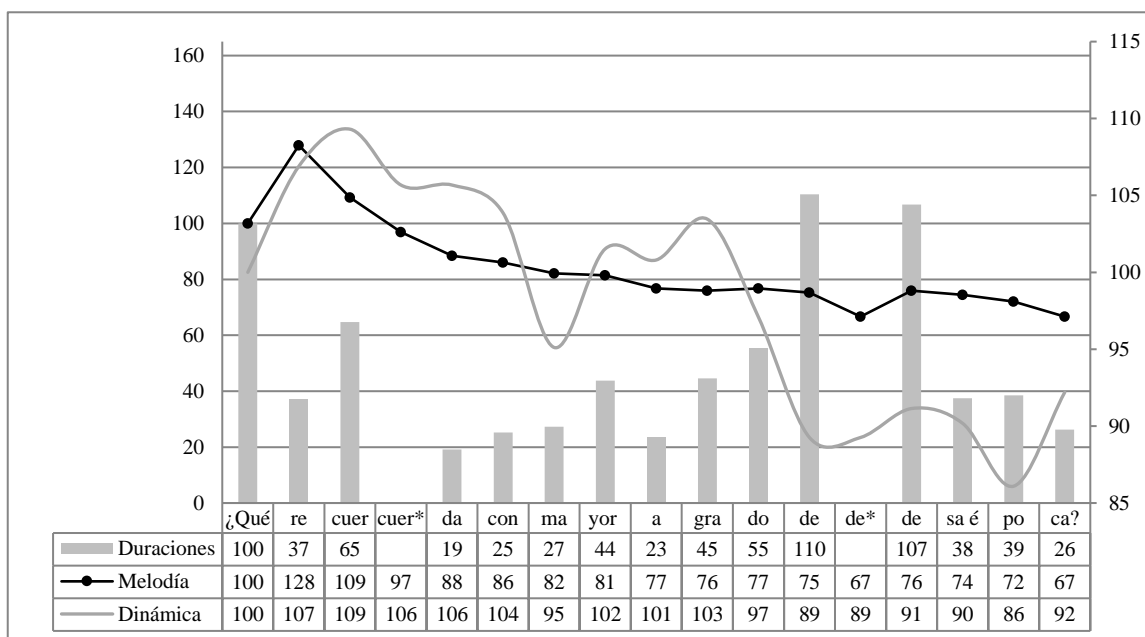


Figura 4. Análisis prosódico del enunciado: «¿Qué recuerda con mayor agrado de... de esa época?»

Las barras del gráfico expresan las relaciones de duración entre los pies rítmicos del enunciado. Puede observarse cómo las vocales tónicas tienden a ser más largas (al margen, claro está, del titubeo en «de... de»).

De un vistazo, podemos comprobar cómo la melodía y la dinámica no acaban de coincidir: los picos de intensidad y los picos tonales no casan, ni siquiera en el primer pico del enunciado (en la palabra «recuerda», hay una prominencia tonal en «re», que no tiene ninguna relevancia dinámica; y hay prominencia dinámica en «cuer», que no tiene ninguna relevancia melódica). Esa no coincidencia podríamos considerarla, tal vez, un rasgo de énfasis, de foco ancho, que afecta a todo el enunciado.

También observamos dos detalles importantes: un alargamiento inesperado en la última sílaba de «agrado», y una subida de intensidad en la última sílaba de «época», la última palabra del enunciado. Ambos rasgos (alargamiento en un caso, dinámica en el otro), reclaman claramente la atención del interlocutor, focalizando sobre esas palabras y mostrando su propio desagrado en la palabra «agrado» y en la palabra «época».

Evidentemente, con un simple análisis de la melodía no podríamos determinar cómo se transmite ese desagrado sobre lo que está preguntando el locutor; en cambio, añadiendo el análisis de la dinámica y de las duraciones de los pies rítmicos, podemos identificar con total objetividad los rasgos prosódicos que transmiten esa afectividad, esa emoción negativa sobre el tema de conversación.

3. EL LENGUAJE AFECTIVO

La emocionalidad, la ironía o la cortesía (por poner los ejemplos más obvios) son fuente constante de malentendidos entre hablantes de comunidades distintas, y cuando un hablante tiene una intención A, su interlocutor (de otra comunidad) puede comprender una intención B, por lo que actuará haciendo Y (en vez de hacer X, como correspondería a la intención A); lo que, a su vez, provocará en el primero una reacción Z, etc. La confrontación, entonces, puede ser inevitable. Los hablantes partían de la idea de que manejaban el mismo código lingüístico (y así es, en cierto modo, porque ambos son hablantes del mismo idioma); pero la realidad es que no son hablantes de los mismos códigos afectivos (relacionados con el nivel de análisis paralingüístico), que son códigos semiestables, no fijados sino negociables.

Para comprender de qué modo la afectividad tiñe toda la comunicación oral y poder identificar su relación con la prosodia, podemos relacionar los tres niveles de análisis (prelingüístico, lingüístico y paralingüístico, que vimos en el primer apartado), con las funciones del lenguaje de Jakobson (1963). Según este autor,

- el lenguaje cumple la llamada *función fática* cuando el mensaje se centra en el canal, en el propio establecimiento de contacto entre los interlocutores,
- la *función expresiva* cuando el mensaje se centra en el propio emisor,
- la *función apelativa* cuando el mensaje se centra en el receptor,
- la *función poética* cuando el mensaje se centra en la forma del propio mensaje,
- la *función metalingüística* cuando el mensaje se centra en el código en que este está cifrado
- y la *función referencial* cuando el mensaje se centra en el contexto (es decir, cuando el mensaje se centra en alguna otra cosa que no es el propio acto comunicativo).

Para empezar, la función fática se relaciona con el nivel de análisis prelingüístico: en este nivel, como vimos, la entonación es el elemento estructurador del habla, permitiendo la integración del discurso en unidades inteligibles. Pero también, y como veremos más adelante, la prosodia es el factor que permite establecer el contacto entre los interlocutores, identificando la compatibilidad entre los códigos que ponen en juego los participantes del acto comunicativo.

La negociación entre los códigos de los hablantes no es trivial, sino que requiere un amplio despliegue de competencias comunicativas de tipo estratégico⁴. Solo es posible negociar nuestros códigos si son lo suficientemente compatibles entre ellos: por ejemplo, si los códigos dialectales que se ponen en juego forman parte de un mismo idioma, o bien si los códigos semiestables pertenecen a comunidades de habla muy cercanas. Así, sabemos que los rasgos de la entonación de cortesía son muy similares en el español peninsular y en catalán⁵ (pues sus hablantes pertenecen a una misma comunidad cultural, aunque sean hablantes

⁴ Sobre la teoría de la competencia comunicativa, *vid.* Cantero Serena (2008); en Cantero Serena (2014b) se expone la dimensión fónica de las distintas competencias comunicativas.

⁵ *Cfr.* Devís Herraiz (2011), Devís Herraiz & Cantero Serena (2014).

de dos idiomas distintos), mientras que la entonación de cortesía es muy distinta en las distintas variedades del español de América (distintas comunidades culturales, un único idioma).

Como veremos, cada hablante ha desarrollado a lo largo de su vida una competencia fónica determinada, cuya forma externa es, justamente, su entonación y su prosodia (lo que, metafóricamente, llamaremos las *paredes del recipiente fónico*, que son paredes prosódicas).

Por su parte, las funciones expresiva, apelativa y poética del lenguaje se relacionan con el nivel de análisis paralingüístico:

- la función expresiva se relaciona, prioritariamente, con la entonación emocional,
- la función apelativa, con la entonación de cortesía,
- la función poética, con la entonación de foco.

Estas relaciones, obviamente, no son unívocas, sino que únicamente indican su centro de gravedad: cuando el mensaje se centra, prioritariamente, en expresar el estado, o los deseos, etc., del propio hablante, es común que la entonación también sea prioritariamente emocional; cuando el mensaje se centra en su receptor, es común que la entonación empleada sea exhortativa y/o de cortesía o descortesía; cuando el mensaje se centra en el propio mensaje, es común que, entonces, aparezcan los rasgos prosódicos de énfasis que focalizan una parte (foco estrecho) o todo el enunciado (foco ancho). Ya hemos visto, en el ejemplo del apartado anterior, cómo dos rasgos de foco estrecho (como el alargamiento de un pie rítmico, o la intensidad de una sílaba átona al final de la frase) permitían caracterizar el perfil prosódico emocional del enunciado.

Finalmente, la función referencial del lenguaje se relaciona, directamente, con el nivel de análisis lingüístico.⁶

Así, podríamos decir que tanto la función fática (relacionada con los códigos prelingüísticos) como los usos del lenguaje expresivos, apelativos o poéticos (relacionados con los códigos paralingüísticos) son «afectivos», en el sentido de que no se refieren al «mundo» exterior, sino al propio acto comunicativo y a los elementos que lo componen (el canal, el emisor, el receptor, el propio mensaje): todo aquello que no es referencial, diríamos, es afectivo (*vid.* fig. 5).

LENGUAJE AFECTIVO	Códigos prelingüísticos		<i>Función fática</i>
	Códigos paralingüísticos	De cortesía	<i>Función apelativa</i>
		De foco	<i>Función poética</i>
		Emocionales	<i>Función expresiva</i>
Códigos lingüísticos			<i>Función referencial</i>

Figura 5: Códigos del lenguaje afectivo

⁶ Tal vez la función metalingüística no esté al mismo nivel que las demás, teniendo en cuenta que no es una actividad muy frecuente entre los hablantes, que rara vez se cuestionan los códigos que emplean; tal vez los enunciados metalingüísticos cumplan, sencillamente, una función referencial.

Por tanto, llamamos *lenguaje afectivo* al uso no referencial del lenguaje⁷.

4. LA PROSODIA MODALIZA AFECTIVAMENTE LA INFORMACIÓN

En la comunicación oral puede haber mucha o poca información, pero siempre será más importante «todo lo demás», todo lo que no es información: lo que no es información es afectividad, y la afectividad se concreta en el timbre de la voz y en la prosodia. Sobre todo, en la prosodia: el lenguaje afectivo se vehicula a través de los códigos prelingüísticos y, especialmente, paralingüísticos (emocionales, de cortesía, de foco), y se concreta en las tres dimensiones de la prosodia: la melodía de los enunciados, su dinámica y su ritmo.

De hecho, en nuestras relaciones personales pocas veces transmitimos información exclusivamente, porque siempre añadimos una pátina de afectividad a nuestros intercambios. No parece deseable (pero es que, seguramente, tampoco es posible), un discurso hablado desprovisto de toda afectividad: es lo que ocurría con los primeros ingenios de conversión texto-voz, con esa voz sintética desprovista de emoción, tan desagradable, tan poco humana.

La afectividad no es un rasgo más, o un rasgo extra, en la comunicación humana, sino su rasgo más importante. Nos relacionamos con los demás para relacionarnos, no solo para obtener algún beneficio: el beneficio es la propia relación. La comunicación (la comunicación afectiva) no es un lujo, sino una necesidad biológica de nuestra especie, y una necesidad imperiosa de cada uno de nosotros.

Incluso en los intercambios comunicativos en los que (además de una dimensión afectiva) hay una transmisión efectiva de información, de información “lógica”, esta información también está modalizada afectivamente. De hecho, el valor de significatividad de la entonación lingüística consiste, precisamente, en modalizar la información léxico-gramatical, como vimos, marcando los enunciados como neutros, enfáticos, suspendidos o interrogativos.

Por ejemplo, una entonación «interrogativa» no quiere decir nada por sí misma: no quiere decir «pregunta», sino que es una forma melódica opositiva (esto es: lingüística, distintiva) que modaliza el enunciado y lo opone a otros enunciados no interrogativos. El enunciado modalizado por una melodía interrogativa puede ser que sea una pregunta, en efecto, en el contexto adecuado y con el contenido léxico-gramatical adecuado («¿estás seguro?»: en este caso, la melodía interrogativa permite distinguir una pregunta absoluta de una afirmación); pero, en otro contexto, o con otros contenidos léxico-gramaticales, podría ser una advertencia («¿venga, valiente? ¿atrévete? ¿acércate? ¿pégame?»): en este caso, la melodía interrogativa permite distinguir el desafío de una mera invitación o una orden); y aun en otros contextos esa misma melodía podría modalizar el enunciado para que funcione como una

⁷ Los términos «afectivo» y «emocional» a veces se confunden entre sí. En nuestra propuesta, distinguimos claramente *afectivo*, como el «uso no referencial del lenguaje», de *emocional*, que sería una de las diversas manifestaciones de la afectividad (junto con la cortesía o la ironía, por ejemplo, que son lenguaje *afectivo*, pero no *emocional*).

cortesía atenuadora («¿síéntese?»: en este caso, la orden está atenuada por la melodía interrogativa).

La misma melodía, la misma entonación, puede servir para distinguir una pregunta de una afirmación, un desafío de una invitación, o un ruego de una orden, dependiendo del contexto. El valor lingüístico de esa melodía reside en su capacidad para distinguir un valor discursivo de otro: pregunta o afirmación, desafío o invitación, orden o ruego, según el contexto. Aunque se denomine, tradicionalmente, entonación «interrogativa», no sirve solo para hacer preguntas; el término «interrogativa» puede llevarnos a engaño (como si esa melodía solo pudiera cumplir esa función), en efecto, desde una perspectiva muy limitada en la que «lo lingüístico» actúa solo en el nivel oracional, sin contar con el contexto. Este ejemplo muestra bien a las claras que el código lingüístico (el objeto de estudio de la lingüística, tradicionalmente muy centrada en la sintaxis oracional) actúa más allá de las unidades gramaticales: la entonación lingüística modaliza la información que contiene la frase, pero su ámbito de actuación es el *acto de habla* en su conjunto⁸.

En definitiva, la prosodia tiñe de afectividad todos los discursos, añadiendo una dimensión expresiva (emocional), apelativa (de cortesía) o poética (de foco), simultáneamente a la propia modalización lingüística de los enunciados.

No es posible hablar sin prosodia: si la pronunciación es la forma material de la comunicación, la prosodia es su principio estructurador y su vehículo de expresión afectiva. Todos los contenidos léxicos, gramaticales o pragmáticos están contenidos en un contenedor prosódico, que actúa como un verdadero «recipiente».

5. LA METÁFORA DEL RECIPIENTE FÓNICO

Según distintos autores (*vid.*, p. ej., Macken & Ferguson, 1987), en un primer estadio en la adquisición fónica de la lengua, y después de recibir grandes cantidades de *input* fónico, hacia los seis años de vida ya se elabora una suerte de «teoría» del sistema de la lengua, que se ajusta gradualmente y culmina hacia los doce años de edad (cuando se considera que ya se tiene la competencia de un adulto). Durante todo ese tiempo (entre los seis y los doce años), el individuo sigue incorporando reglas y estructuras de su lengua, pero apenas modifica, en lo esencial, su competencia fónica; es decir, continúa su adquisición léxico-gramatical de la lengua, así como su adquisición de las condiciones de uso, pero no su adquisición fónica, que es previa.

Que la adquisición de la competencia fónica sea previa a la adquisición del resto de competencias y conocimientos lingüísticos no es trivial: no es imaginable, por ejemplo, la adquisición de ningún vocabulario si antes el niño no es capaz de percibir la pronunciación de dicha palabra clara y distintamente; para que el niño pueda distinguir una pronunciación concreta en

⁸ En este sentido, en Hidalgo Navarro (2019) se ofrece por primera vez una propuesta exhaustiva y rigurosa que plantea un estudio de la entonación plenamente integrado en un enfoque pragmático: el modelo *interactivo-funcional*. Considero este trabajo (que tuve la fortuna de prologar) el punto de referencia teórico desde el que partirá en el futuro el estudio de la entonación, a medida que vayan abandonándose, por irrelevantes, los estudios que la afrontan como un producto derivado de la sintaxis oracional.

el maremágnum de sonidos que lo rodean, ha de haber sistematizado, al menos, los rasgos fónicos mínimos que le permitan una percepción categorial; solo entonces puede pronunciar las palabras de una manera suficientemente inteligible para sus interlocutores adultos.

Cualquier enunciado lingüístico implica su forma sonora, y cualquier discurso, por mínimo que pueda parecer (una sola palabra, un monosílabo), implica su integración fónica: esa es la función *prelingüística* que cumple la prosodia, su carácter de «contenedor» del discurso. Del mismo modo que no hay habla sin entonación, en la comunicación oral no hay palabras, ni estructuras, sin su forma sonora; por tanto, no hay, ni puede haber, adquisición de los componentes léxico-gramatical y pragmático sin una competencia fónica previa. Así, durante el periodo de adquisición fónica del niño, su desarrollo léxico-gramatical está siempre condicionado por el nivel de competencia fónica conseguido: solo cuando ya tiene una competencia prosódica suficiente aparecen las palabras-frase (hacia los dos años de vida); solo cuando ha desarrollado una percepción categorial suficiente se dispara la adquisición léxica (entre los tres y los cuatro años); etc.

En Cantero Serena (1994) (*cf.* también: Bartolí Rigol 2005), hemos comparado la competencia fónica con un «recipiente», un *recipiente fónico* que se construye durante los primeros años de vida, dentro del cual se introducen paulatinamente los elementos, las reglas y las estructuras léxico-gramaticales que el individuo adquiere a lo largo de su vida.

La competencia fónica, desde luego, no es una «cosa» (como sugiere la comparación con un recipiente) sino una «capacidad de hacer»: de pronunciar los elementos léxicos adquiridos, de integrar el discurso codificado en la lengua y de hacerlo comprensible, de entender otros discursos codificados en la misma lengua, de reconocer a otros hablantes del mismo código, etc. Así, hablar de «recipiente fónico» es una metáfora, cuya materialidad permite formular con toda claridad el carácter previo y condicionador de la competencia fónica con respecto a las demás competencias lingüísticas y comunicativas.

MacNeilage & Davis (1990) proponen una metáfora similar para explicar las características del balbuceo y las primeras emisiones lingüísticas del bebé: la hipótesis *Frames, then Contents* («primero los esquemas, después los contenidos»). En cierto modo, la *Frame Dominance* o «predominancia de los esquemas» sobre los contenidos que estos autores proponen (Davis & MacNeilage 1995) sería extrapolable a nuestra *metáfora del recipiente fónico* (*vid.* fig. 6): el individuo construye, en su primera infancia (y como primer estadio en su adquisición de la lengua), un recipiente fónico («*Frame*») dentro del cual va introduciendo, a lo largo de su vida, todos sus conocimientos léxico-gramaticales («*Contents*»). Tales conocimientos, abstractos, toman forma (sonora) en el recipiente fónico que los acoge, y a partir de entonces solo existen y solo pueden actualizarse materializados en la forma sonora que impone el recipiente fónico que les da sentido.

En este caso, el *Frame* estaría formado por el filtro perceptivo-productivo (fonológico y fonético) de la competencia fónica, que incluye también un cierto filtro emocional: la capacidad de codificar y reconocer emociones por medio de la entonación (tal vez, el primer hito dentro de la adquisición fónica), pero también la capacidad de *reconocer* a otro hablante de la misma comunidad lingüística, de *extrañar* otros códigos fónicos ajenos (e irritarse con ellos), etc.

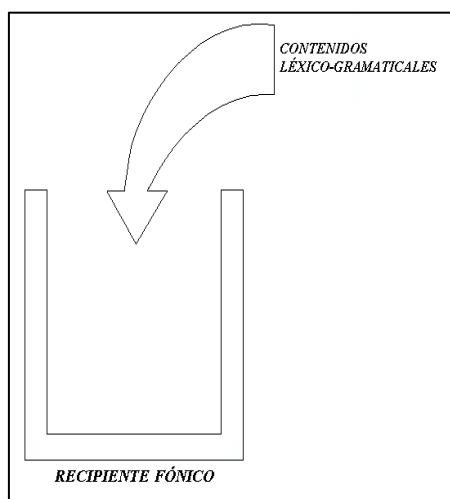


Figura 6. La adquisición fonética es previa

Nuestra metáfora explica numerosos fenómenos relacionados con la adquisición fonética de una LE: el *factor edad* y el condicionamiento fonético de la L1 (*interferencias y transferencias*), la ineficacia de la corrección fonética tradicional, la *fosilización* masiva en la pronunciación de los alumnos, la persistencia del *acento extranjero*, etc.

Los conocimientos léxico-gramaticales se introducen y acomodan en un recipiente fonético (relativamente) acabado. Normalmente, tales contenidos coinciden con el recipiente fonético, *porque constituyen un mismo código*: es decir, las palabras y las estructuras de la L1 del individuo que este escucha en sus interlocutores coinciden y son compatibles con los patrones fonéticos que puede percibir y que es capaz de producir espontáneamente, pues pertenecen al mismo idioma. Así, el aprendizaje de contenidos léxico-gramaticales de la propia L1 es espontáneo: apenas activado el léxico nuevo, o la estructura gramatical nueva, se incorporan a la competencia del individuo y perduran en la memoria. En este caso, cuando los contenidos léxico-gramaticales y el recipiente fonético que los contiene pertenecen a la misma lengua, hablamos de un *hablante nativo*.

A menudo, la enseñanza de una LE está centrada, fundamentalmente, en sus aspectos léxico-gramaticales, y el aspecto fonético queda relegado a un segundo o a un tercer plano, al margen de si la metodología empleada es tradicional, nocional-funcional o comunicativa: pues en cualquier caso importan las palabras empleadas, su orden sintáctico, la propiedad de su uso, etc., antes que su integración fonética o su forma sonora. Por añadidura, la mediación entre el alumno y la LE meta a menudo es exclusivamente (o casi exclusivamente) lectoescritora (*vid. Cantero Serena 2020*).

Es decir, los contenidos léxico-gramaticales de la LE aprendida se introducen y se acomodan, inadvertidamente, en el único recipiente fonético de que dispone el alumno: el de su L1. Obviamente, tales contenidos no constituyen un mismo código con dicho recipiente fonético, sino que forman parte de códigos distintos. Este podría ser, sin duda, el obstáculo principal en el

aprendizaje de una LE: el individuo no es sensible, ni reconoce, ni incorpora espontáneamente a su competencia comunicativa elementos léxico-gramaticales extraños a su recipiente fónico. Cuando los contenidos léxico-gramaticales y el recipiente fónico que los contiene pertenecen a lenguas distintas, hablamos de un *hablante extranjero*: emplea palabras y estructuras de la LE aprendida, pero las actualiza y las integra en su L1, esto es, *habla en su L1 con palabras y estructuras de la LE*. Algo más que meras «interferencias» o «transferencias» fónicas.

El factor de la edad, por tanto, estaría relacionado con la creación del recipiente fónico: antes de los seis años, la tarea principal del niño es ir creando el recipiente fónico de su L1, por lo que aún puede crear un recipiente nuevo, o varios recipientes (por otra parte, antes de los seis años el niño no sabe leer ni escribir, por lo que todo el *input* en una segunda lengua ha de ser, necesariamente, oral). A partir de la edad escolar, sin embargo, el individuo está alfabetizado y la instrucción formal suele establecerse inequívocamente en los parámetros (tradicionales) de la enseñanza mediatizada por la lengua escrita. El factor clave, por tanto, no es la edad en sí misma, sino el estilo de aprendizaje: si está mediatizado o no por el lenguaje escrito o, dicho de otro modo, si ha permitido o no crear un recipiente fónico adecuado a la lengua meta.

Los esfuerzos por «corregir» la pronunciación del alumno *a posteriori* difícilmente pueden cumplir su objetivo, pues el alumno *no pronuncia la LE* (aunque dé esa sensación, porque emplea palabras de la LE), sino que pronuncia, inevitablemente, su propia L1. Después de un trabajo sistemático de corrección fonética puede aspirarse, como mucho, a crear una suerte de «apartado especial» dentro del recipiente fónico del alumno (*vid. fig. 7*):

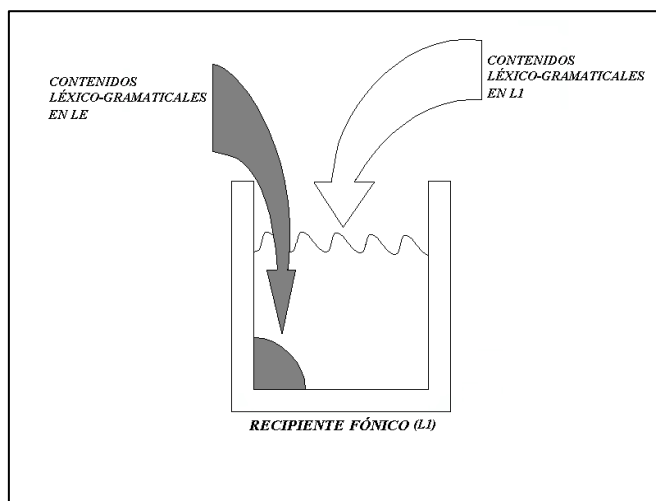


Figura 7. Un apartado especial en LE dentro del recipiente fónico en L1

Este «apartado especial», incómodo en un recipiente fónico de otra lengua, es la mejor manera de conseguir una «pronunciación» inmóvil, impostada, *fossilizada*, que a veces no es más que una «traducción» de las reglas de pronunciación de la ortografía según las normas de la LE aprendida. O, lo que es lo mismo, a menudo es la propia corrección fonética la que impone una fossilización masiva en la pronunciación de los alumnos extranjeros.

El recipiente fónico, para completar la metáfora, tiene una textura fonológica (que, como una esponja o un filtro, absorbe solo el *input* que coincide con sus restricciones funcionales y deja literalmente «sordo» al individuo para discriminar cualquier otra diferencia) y las paredes son prosódicas: no en vano, en nuestra adquisición fónica lo primero que se construye es la competencia entonativa, las paredes del recipiente, el contenedor del discurso. Así, la textura del «apartado especial» de la LE en el recipiente fónico de la L1 puede llegar a ser distinta de la textura del resto del recipiente, pero nunca tendrá unas paredes distintas: el *acento extranjero* del alumno será persistente, en cualquier caso, aunque su «pronunciación» de los sonidos aislados sea aceptable.

6. LA PROSODIA REGULA LA COMPATIBILIDAD ENTRE CÓDIGOS

Como hemos dicho, el recipiente fónico (o recipientes, si ha creado más de uno, en distintas lenguas o en distintos dialectos) de cada individuo le da la capacidad de *reconocer* a otro hablante de la misma comunidad lingüística y de *extrañar* otros códigos fónicos ajenos (e incluso irritarse con ellos).

Al identificar los códigos fónicos de nuestro interlocutor (especialmente, los códigos prelingüísticos: los que estructuran el discurso en unidades inteligibles y con un perfil melódico determinado), identificamos la compatibilidad de ese código con los nuestros y abrimos (o no) la posibilidad de negociar con él: de establecer contacto, de abrir el canal de comunicación.

Entre códigos compatibles, la posibilidad de intercomprensión oral es muy elevada, y la negociación entre los interlocutores es fácil: entre dos dialectos muy cercanos, por ejemplo, puede que los interlocutores ni siquiera se den cuenta de que están negociando su entonación y es posible que, de pronto, sus entonaciones se mezclen (es decir, que se contagien de la entonación del otro). Justamente es la prosodia la que permite identificar la compatibilidad entre los códigos prelingüísticos, abrir el canal de comunicación, establecer contacto con el otro y contagiarse con él, a medida que se establece un espacio común de intersubjetividad y se negocian todos los aspectos del intercambio (desde la misma pronunciación hasta los significados y las intenciones)⁹.

La comunicación oral consiste, precisamente, en establecer contacto con los demás (es decir, en relacionarse, sin más, con el único beneficio de la propia relación), y la prosodia (en su nivel de análisis prelingüístico) es la puerta de entrada al otro o el muro infranqueable que impide la comunicación; es el fenómeno que permite la negociación entre los interlocutores y la regula.

⁹ En Cantero Serena (2019b) se ofrece un esquema definido de las condiciones de la negociación entre los interlocutores, que de entrada no comparten ni los mismos códigos (que cada uno de ellos ha ido elaborando a lo largo de su vida), ni los mismos contextos (pues cada uno tiene un modelo del mundo propio, compatible pero diferente del de los demás).

7. CONCLUSIONES: DIMENSIONES DE LA AFECTIVIDAD EN PROSODIA

La prosodia, por tanto, permite identificar la compatibilidad entre los códigos de los interlocutores. Esta es la primera dimensión de la afectividad en prosodia: establecer contacto con el otro, permitir la comunicación, regular la relación con la otra persona desde «todo lo demás» (todo aquello que no es pura información).

Por otra parte, los códigos paralingüísticos (de cortesía, de foco y de emoción), que son códigos semiestables, hay que ir negociándolos a lo largo del intercambio. También la prosodia es la forma material que regula esa negociación, no solo identificando su compatibilidad (que permitió abrir el canal y establecer contacto con el otro) sino modificándose paulatinamente para elaborar un espacio común de intersubjetividad, construyendo conjuntamente el propio código, como un juego cuyas reglas se establecen a medida que se juega: así son los códigos paralingüísticos.

Con los códigos de emoción, los interlocutores se expresan como individuos, de una manera más o menos convencional o genuina. Con los códigos de cortesía, los interlocutores se interpelan, cada uno apela al otro, y establecen las fórmulas que regularán su relación. Con los códigos de foco, los interlocutores llaman la atención sobre una parte de su mensaje o sobre su totalidad, enfatizándolo para dirigir la atención del otro al aspecto focalizado.

Incluso la mera transmisión de información (en su función referencial) está modalizada por la prosodia (por la entonación lingüística): es decir, en la comunicación oral incluso la pura información lógica está teñida de afectividad. La comunicación oral es afectiva, en su conjunto.

Estas son las dimensiones de la afectividad en prosodia: la *expresión* (emocional, apelativa o poética, incluyendo la modalización lingüística de la información) y el *contacto* (que permite identificar la compatibilidad entre los códigos y abrir la puerta al intercambio y la negociación).

REFERENCIAS BIBLIOGRÁFICAS

- BARTOLÍ RIGOL, M. (2005): "La pronunciación en la clase de lenguas extranjeras". *Phonica* 1, 1-27. En línea: <<http://revistes.ub.edu/index.php/phonica/article/view/5565/7353>>.
- CANTERO SERENA, F. J. (1994): "La cuestión del *acento* en la enseñanza de lenguas". En J. Sánchez Lobato & I. Santos Gargallo (eds.): *Problemas y métodos en la enseñanza del español como lengua extranjera*. Madrid: S.G.E.L., 247-55.
- CANTERO SERENA, F. J. (2002): *Teoría y análisis de la entonación*. Barcelona: Edicions de la Universitat de Barcelona.
- CANTERO SERENA, F. J. (2008): "Complejidad y competencia comunicativa". *Revista Horizontes de Lingüística Aplicada* 7/1, 71-87.
- CANTERO SERENA, F. J. (2014a): "Códigos de la entonación y entonación emocional". En A. Díaz *et al.* (eds.): *Actas del 31 Congreso Internacional AESLA*. La Laguna: Universidad de La Laguna, 618-29.

- CANTERO SERENA, F. J. (2014b): “Adquisición de competencias fónicas”. En Y. Congosto *et al.* (eds.): *Fonética Experimental, Educación Superior e Investigación. Vol. II. Adquisición y aprendizaje de lenguas / Español como lengua extranjera*. Madrid: Arco/Libros, 29-55.
- CANTERO SERENA, F. J. (2016): “Corpus de habla espontánea para el estudio de la entonación”. En A. M. Fernández Planas (ed.): *53 reflexiones sobre aspectos de la fonética y otros temas de lingüística*. Barcelona: Laboratori de Fonètica de la Universitat de Barcelona, 151-60.
- CANTERO SERENA, F. J. (2019a): “Análisis prosódico del habla: más allá de la melodía”, en M. R. Álvarez Silva, A. Muñoz Alvarado & L. Ruiz Miyares (eds.): *Comunicación Social: Lingüística, Medios Masivos, Arte, Etimología, Folclor y otras ciencias afines*. Vol. II. Santiago de Cuba: Ediciones Centro de Lingüística Aplicada, 485-98.
- CANTERO SERENA, F. J. (2019b): *El arte de no enseñar lengua*. Barcelona: Octaedro.
- CANTERO SERENA, F. J. (2020): “Didáctica de la pronunciación: de la corrección fonética al enfoque oral”. En F. J. Cantero Serena & M. Giralte Lorenz (coords.): *Pronunciación y enfoque oral en lenguas extranjeras*. Barcelona: Octaedro, 11-48.
- CANTERO SERENA, F. J. & C. DE ARRIBA (1996): “El cambio de código: contextos, tipos y funciones”. En J. L. Otal, I. Fortanet & V. Codina (eds.): *Estudios de Lingüística Aplicada*. Castellón: Publicacions de la Universitat Jaume I. 587-96.
- CANTERO SERENA, F. J. & D. FONT-ROTCHÉS (2020): “Melodic Analysis of Speech (MAS): Phonetics of Intonation”. En J. Abasolo, I. de Pablo & A. Ensunza: *Contributions on education*. Bilbao: Universidad del País Vasco, 20-47.
- CANTERO SERENA, F. J. & MATEO RUIZ, M. (2011): “Análisis melódico del habla: complejidad y entonación en el discurso”, *Oralia*14, 105-27.
- DAVIS, B. L. & P. F. MACNEILAGE (1995): “The articulatory basis of babbling”. *Journal of Speech Language and Hearing Research* 38/6, 1199-211.
- DEVÍS HERRAIZ, E. (2011): “La entonación de (des)cortesía en el español coloquial”. *Phonica* 7, 36-79.
- DEVÍS HERRAIZ, E. & F. J. CANTERO SERENA (2014): “The intonation of mitigating politeness in Catalan”. *Journal of Politeness Research* 10/1, 127-49.
- HIDALGO NAVARRO, A. (2019): *Sistema y uso de la entonación en español hablado. Aproximación interactivo-funcional*. Santiago de Chile: Ediciones Universidad Alberto Hurtado.
- HELLER, M. (ed.) (1988): *Code-switching*. Berlin: Mouton de Gruyter.
- JAKOBSON, R. (1963): “Linguistique et poétique”. En *Essais de linguistique générale*. Paris: Minuit. Tr. esp: “Lingüística y poética”, en *Ensayos de lingüística general*. Barcelona: Seix-Barral, 1975, 347-95.
- MACKEN, M.A. & C. A. FERGUSON (1987): “Phonological Universals in Language Acquisition”. En G. Ioup & S. H. Weinberger (eds.): *Interlanguage Phonology*. Cambridge: Newbury House, 110-29.
- MACNEILAGE, P. F. & B. L. DAVIS (1990): “Acquisition of speech production: Frames then contents”. En M. Jannerod (ed.): *Attention and performance 13: Motor representation and control*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc., 453-76.