

DESCRIPCIÓN PROSÓDICA DE LAS EMOCIONES ACTUADAS EN ESPAÑOL MEXICANO¹

Ofelia Elizabeth González Franco
Universidad Autónoma de Querétaro (México)
ofelia.gonzalez@uaq.edu.mx

Elia Haydée Carrasco Ortiz
Universidad Autónoma de Querétaro (México)
haydee.carrasco@uaq.edu.mx

Eva Patricia Velásquez Upegui
Universidad Autónoma de Querétaro (México)
evapvelasquez@uaq.edu.mx

Recibido: 20/04/2022 - **Aprobado:** 06/09/2022 - **Publicado:** 13/03/2023

DOI: doi.org/10.17533/udea.lyl.n83a02

Resumen: En este artículo se realiza la descripción de los patrones entonativos de cinco emociones —enojado, triste, neutral, tranquilo y emocionado— en enunciados aseverativos de un corpus oral actuado de tres hablantes mujeres de Querétaro, México. Este trabajo se justifica tras una breve revisión de los estudios en torno a la prosodia emocional del español y su relación con la teoría bioinformacional. A partir del análisis de los datos se obtiene una caracterización de la curva tonal por emoción, con valores por arriba de los umbrales de percepción, que pueden utilizarse como referencia para futuros estudios descriptivos o de percepción.

Palabras clave: prosodia; emociones; español mexicano; teoría bioinformacional; frecuencia fundamental-F0.

PROSODIC DESCRIPTION OF ACTED EMOTIONS IN MEXICAN SPANISH

Abstract: This study describes the intonation patterns of five emotions —angry, sad, neutral, calm and excited— in assertive statements. The data was obtained from an oral acted corpus of three female speakers from Querétaro, Mexico. The analysis was based on the Spanish emotional prosody studies and its relationship with the bio-informational theory of emotion. Results shows a characterization of the tonal curve by emotion, with perceptible points of interest as reference for future descriptive or perception studies.

Keywords: prosody; emotions; Mexican Spanish; bio-informational theory; fundamental frequency-F0.

1. Artículo producto de la ponencia presentada en el Tercer Encuentro del Seminario de Prosodia, organizado por el Laboratorio de Fonética de la Universidad de Antioquia y la Universidad Autónoma de Querétaro, realizado en los días 7 y 8 de octubre de 2021.

1. Introducción

La entonación es un componente esencial de la lengua, pues no solamente es un indicador de la información lingüística, sino también de la sociolingüística, especialmente durante la realización particular de una lengua en cada región o grupo social. Asimismo, el tono define aspectos relacionados con la paralingüística, como los sentimientos, los estados de ánimo y las actitudes de los hablantes. Este tipo de significados paralingüísticos se han descrito en diferentes lenguas y en el español de otras regiones, pero no en México. A pesar de ser un componente esencial para cualquier intercambio lingüístico, se trata de un fenómeno difícil de describir, pues combina en sí mismo componentes gramaticales, pragmáticos y no verbales. Además, puede presentar múltiples realizaciones dependiendo de cada contexto y cultura. Por esta razón, es relevante encontrar los aspectos que coinciden entre las características prosódicas y las emociones, sobrepasando estas dificultades.

Este trabajo contribuye al esbozo de una descripción de la prosodia emocional del español en general y en particular del mexicano. Para ello, se lleva a cabo el diseño de un corpus actuado que permite dar cuenta del comportamiento de la curva entonativa de las emociones en una serie de enunciados aseverativos caracterizados por la misma estructura acentual y número de sílabas. Además, se relacionan los datos encontrados con la teoría bioinformacional de las emociones (Lang, Bradley & Cutbert, 1997) que permite conceptualizarlas dentro de un modelo continuo con rasgos de activación y valencia para lograr una mejor caracterización de sus manifestaciones prosódicas, independientemente de las realizaciones particulares de cada hablante.

En el primer apartado de este artículo se presentan los modelos teóricos y analíticos utilizados para la descripción de la prosodia emocional, cuyos elementos se integran en la parte metodológica. A continuación, se expone un breve panorama acerca de los estudios en torno a la prosodia de las emociones del español. Después, se comparten los resultados de este análisis que permiten una caracterización preliminar del perfil melódico de las emociones del español mexicano a partir de los rasgos de activación y valencia. Finalmente, se presentan las conclusiones, en las que se resumen los hallazgos del trabajo, sus alcances y sus limitaciones.

2. Modelos teóricos para el análisis de la prosodia emocional

El análisis de la prosodia ha seguido diferentes modelos a lo largo de la historia, pero pueden resumirse en dos vertientes: la primera, centrada en la descripción de la configuración fonética de las frases entonativas y, la segunda, en un análisis fonológico por niveles que describe ciertos puntos de contraste. Utilizando distintas terminologías, todos los modelos tratan de dar cuenta del contorno melódico que refleja movimientos de ascenso y descenso, pero se centran en distintas unidades de análisis, tomando como referencia a la configuración nuclear o bien prestando atención a toda la melodía y los cambios que ocurren en diferentes puntos del enunciado.

El modelo de análisis por niveles ha tenido diversas interpretaciones, pero en general busca la representación de los contornos a través de un sistema de unidades discretas: acentos, niveles y junturas, que forman a su vez

morfemas tonales. En este contexto, se desarrolló la perspectiva de la fonología autosegmental (Goldsmith, 1976), que representa los elementos fonológicos no de forma concatenada o lineal como se había hecho anteriormente, sino registrando en secuencias paralelas el plano melódico —los segmentos y rasgos— y el prosódico (estructura). Junto con esta, se desarrolló la teoría métrica (Lieberman & Prince, 1977), la cual se enfoca principalmente en la importancia del acento por su relación de prominencia entre núcleos silábicos pertenecientes a unidades mayores y dependientes sintácticamente de la frase entonativa.

A partir de estas perspectivas, surge el modelo de análisis métrico-autosegmental (Ladd, 2008; Pierrehumbert, 1980), el cual busca describir los diferentes contornos melódicos de las lenguas a través de un sistema entonativo que combina diferentes elementos fonológicos contrastivos prosódicamente marcados. En los años noventa este modelo es sistematizado mediante la transcripción prosódica Sp_ToBI —Tones and Break Indices—, utilizado en diferentes lenguas y aplicado al español en el Sp_ToBI por (Beckman, Díaz-Campos, Tevis & Morgan, 2002) con todas las convenciones necesarias para lograr la notación fonémica de la entonación. Este modelo ha tenido varias revisiones, entre las que se destacan las realizadas por Hualde (2003), Estebas y Prieto (2008), Hualde y Prieto (2015).

En el sistema Sp_ToBI se describen cuatro niveles: el ortográfico, el tonal, el fraseo melódico y el misceláneo para fenómenos paralingüísticos. Para la descripción parte de dos unidades fonológicas: los acentos tonales anclados en sílabas tónicas representadas con el diacrítico (*) y los tonos de juntura o límites prosódicos marcados con (–) o (%), dependiendo si son intermedios o finales, respectivamente. Los tonos se representan mediante dos etiquetas L (Low) para los tonos bajos y H (High) para los altos, dependiendo de los valores tonales adyacentes y del tono global del hablante. De acuerdo con el umbral de percepción descrito anteriormente (Pamies, Fernández Planas, Martínez, Ortega & Amorós, 2002), corroborado por Murrieta (2016) y utilizado en otros trabajos como el de Fernández Planas y Martínez Celdrán (2003), las diferencias de al menos 1.5 st se etiquetan como acentos bitonales, mientras que los mayores a 3 st se marcan con el diacrítico (¡) para ascensos y (!) para descensos. Los picos desplazados se marcan con (>) (Estebas-Vilaplana y Prieto, 2008). Combinando estos elementos, se puede definir la curva entonativa e identificar las configuraciones nucleares, partiendo de los valores de la frecuencia fundamental (F0) (Pierrehumbert, 1980).

En este punto, es importante retomar las diferencias conceptuales entre entonación prelingüística, lingüística y paralingüística (Cantero, 2014), que permiten entender la complejidad del fenómeno entonativo, pero a su vez, como menciona Ballesteros (2014) es un «proceso de abstracción útil para el estudio fonético de la melodía», pero que conlleva a un elevadísimo coste teórico en el que se pierde su interdependencia estructural; esta división, aunque permite hacer una clasificación teórica, se debe recordar que «los tres niveles se implican mutuamente» en un fenómeno unitario (pp. 51-52).

Por un lado, la entonación prelingüística comprende los fenómenos de acento, ritmo y melodía, cuya función es delimitar e integrar las unidades discursivas por el contenido lingüístico que contienen en un perfil melódico.

En este nivel se incluye la melodía de la pronunciación asociada a una lengua específica o a un acento dialectal.

Por otro lado, la entonación lingüística incluye los rasgos melódicos, que permiten distinguir los tonemas o unidades funcionales de la entonación que son fonológicamente significativos y pueden asociarse a una forma fónica definida. Estos tonemas brindan una base, pero pueden variar dependiendo de la entonación paralingüística o estilística que se le añada, dando lugar a diferentes tipos de exclamaciones, interrogaciones o aseveraciones. Las variantes melódicas paralingüísticas expresan factores pragmáticos como la personalidad, el grado de seguridad y las emociones, que son el centro de esta investigación.

La expresión de las emociones en la voz puede ser muy variable, dependiendo de la función a la que acompañe y las motivaciones intrínsecas del hablante. Es por ello que para el estudio de la prosodia emocional puede partirse de enunciados aseverativos, caracterizados por una entonación no marcada o neutra, para dejar la expresividad al componente emocional, el cual se espera que modifique la construcción prosódica (Martín Butragueño, 2015). Al respecto, cabe recordar que este puede presentarse en cualquier tipo de acto de habla: directivos, declarativos, expresivos, interrogativos, etc.

De forma ideal, el análisis prosódico debe estar basado en el habla en uso, que promueve una mayor espontaneidad (Martín Butragueño, 2015), pero en ocasiones se utilizan emisiones más controladas para mejorar la calidad del audio, la facilitación del análisis y la sistematización de los datos a partir de patrones comparables entre sí (Face, 2010).

2.1. El modelo bioinformacional y la teoría del código biológico

Las emociones son una característica inherente a los seres humanos que acompaña la mayoría de las acciones diarias y determina la interacción con los otros. Durante la comunicación verbal se transmite de forma implícita gran parte del contenido emocional que acompaña al mensaje explícito, lo complementa, modifica o, en ocasiones, cambia totalmente su significado. La naturaleza de las emociones ha generado interés desde muchas perspectivas y campos de estudio. Ya Darwin (1872) habla de la voz como una reacción física portadora de señales afectivas relacionadas con la lucha por la supervivencia y cómo estos patrones, al igual que otros comportamientos humanos, han evolucionado para resolver ciertos problemas de la especie.

En la literatura pueden encontrarse múltiples definiciones de las emociones como estados de la mente, al detectar reacciones físicas que ocasionan a su vez variaciones acústicas en la voz (Banse & Scherer, 1996). En ese sentido, James (1884) define las emociones como la forma en que se perciben los cambios físicos ante un hecho excitante. Estos cambios involucran cambios coordinados en varios componentes como: activación neuropsicológica, expresión motriz, sentimientos subjetivos y, posiblemente, tendencias a la acción y procesos cognitivos (Scherer, 2005) con una duración determinada y una intensidad decreciente. La emoción es, pues, una disposición para la acción, no un acto en sí mismo, que tiene como base un componente corporal y una sensación consciente, el cual cuenta con una determinada dirección (Chronaki, Wigelsworth, Pell, & Kotz, 2018).

En esta misma línea, se desprenden un sinnúmero de trabajos en los que se afirma que la expresión de las emociones es muy similar en todos los seres humanos, a partir de la cual se han descrito «emociones básicas» presentes en diferentes contextos socioculturales. Las emociones han sido estudiadas desde distintos enfoques que pueden agruparse en modelos discretos o continuos. Los enfoques discretos asumen que existen emociones básicas bien caracterizadas, de las que se generan las demás (Ekman, 1972). Sin embargo, se ha encontrado que no existe una correspondencia unívoca entre emociones básicas y conductas o áreas cerebrales específicas (Scherer, 2005), por lo que no hay un acuerdo general sobre cuáles serían las emociones básicas.

Por otro lado, están los enfoques continuos que describen las emociones dentro de un plano multidimensional, relacionándose entre sí con propiedades subjetivas (Russell, 1980; Russell & Mehrabian, 1977). Este da cuenta de aquellas características que varían dependiendo de las reacciones fisiológicas en tres dimensiones: valencia (agradable o desagradable), activación (excitado o calmado) y dominancia (alta y baja). Los modelos discretos y continuos están relacionados, ya que las categorías emocionales discretas pueden ser descritas a través del modelo continuo.

El modelo bioinformacional (Bradley & Lang, 1994) resulta útil para el estudio de las emociones más allá de las expresiones visibles —gestos, emociones, movimientos, etc.—, entre las que se encuentra el habla «evento acústico que refleja el funcionamiento del sistema nervioso central y por lo tanto acarrea información sobre el estado emocional de un individuo» (Espinosa & Reyes, 2010). De este modelo se rescatan las propiedades de activación y valencia, pues en ellas se presenta la mayor variabilidad y se ha observado mayor relación con las propiedades acústicas de la voz. La *valencia*, también llamada placer, describe qué tan negativa o positiva es una emoción específica, mientras que la *activación*, también llamada intensidad, describe la excitación o alerta interna de un individuo (ver Figura 1). Otra de las dimensiones es la de *dominación*, la cual describe el grado de control del individuo sobre la situación. No obstante, no se ha encontrado relación evidente de esta con patrones suprasegmentales de la emoción.

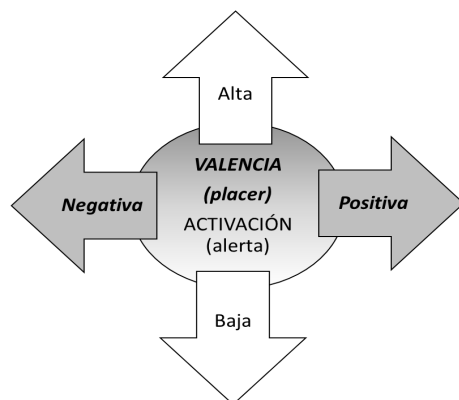


Figura 1. *Propiedades de activación y valencia en el modelo bioinformacional.* Fuente: Bradley & Lang (1994).

Las propiedades de activación y valencia han sido utilizadas como marco de múltiples estudios en torno a las propiedades suprasegmentales de las emociones (Charpentier, Kovarski, Houy-Durand, Malvy, Saby, Bonnet-Brilhault, Latinus, & Gomot, M. , 2018; Gilbers, Fuller, Gilbers, Broersma, Goudbeek, Free, & Başkent, 2015; Luo, Fu, & Galvin, 2007), pues parten del supuesto de que estas van acompañadas de cambios fisiológicos que afectan la respiración, la fonación y la articulación y de patrones o configuraciones acústicas específicas según el tipo de emoción (Scherer, 2003).

Este modelo puede relacionarse directamente con la propuesta de Gussenhoven (2010) sobre la explicación del sistema paralingüístico a través de tres códigos biológicos: el de frecuencia, el de producción y el de esfuerzo. Cada uno de estos códigos tiene una relación fisiológica-funcional que está biológicamente determinada, pero se manifiesta dependiendo de las necesidades pragmáticas del hablante. Para este trabajo se rescata especialmente el *código de esfuerzo*, definido como el nivel o rango de energía alcanzado por la fuerza que se le imprime a la emisión. Esto produce movimientos más marcados con funciones de foco, pero también comparte las propiedades de activación y valencia descritas anteriormente, en el plano de la producción del habla y la interpretación afectiva.

3. Panorama de los estudios de prosodia emocional del español

Como se ha visto anteriormente, el estudio de las emociones como fenómeno paralingüístico presenta dificultades conceptuales, pues, dependiendo del modelo, es posicionado en diferentes niveles del proceso enunciativo; además, el concepto y la clasificación de las emociones también son caracterizados de múltiples formas. A pesar de ello, se han realizado estudios en diversas lenguas, incluida el español, especialmente el peninsular, como en

los trabajos de Hidalgo Navarro (2020) y Rodero (2011) por mencionar algunos. A continuación, se realiza un breve resumen de algunas descripciones útiles para contextualizar el presente trabajo.

Las primeras descripciones del español de Navarro Tomás (1944) y otros estudios más recientes describen los enunciados declarativos típicamente con un final descendente, partiendo del núcleo hasta la juntura terminal L* L %, con variaciones dialectales en la sílaba nuclear (Prieto & Roseano, 2018). Este tipo de enunciado se toma como punto de referencia para la mayoría de las descripciones y comparaciones realizadas en torno a las emociones.

En otro estudio de español peninsular (Hidalgo Navarro, 2020) se describió la alegría, el enfado, el miedo, la sorpresa y la tristeza a partir de las variaciones de F0 en la curva tonal, además de considerarse la intensidad y la velocidad de la entonación; cabe apuntar que, en el caso de esta investigación, estos factores fueron comparados con el tono neutro. El autor realiza una descripción muy detallada en porcentaje de aparición de los diferentes tipos de curva y reporta similitudes en la velocidad entre alegría y tristeza, mientras que alegría y enfado tienen en común una F0 alta. Algunas de las conclusiones de este estudio es que las emociones de alegría y miedo solo se diferencian en la intensidad, y que pueden identificarse parámetros prosódicos coincidentes entre enfado y sorpresa.

En Venezuela, Martínez y Rojas (2011) realizaron el análisis de los correlatos acústicos de F0 e intensidad asociados a emociones actuadas —alegría, tristeza y rabia— en una misma oración, emitida por 5 hablantes. Encontraron que la tristeza se comporta de forma similar a los enunciados neutros, mientras que la rabia y la alegría tienen ascensos y descensos más marcados en la F0, pero se diferencian entre sí por la duración y velocidad, ya que en la rabia son menores.

Para el español de Colombia se encuentra un estudio (Velásquez, Garzón & Soto, 2018) en el que se describe la producción de 4 enunciados actuados en tono neutro, con tristeza, alegría y enojo. La curva entonativa de los enunciados analizados es representada mediante el sistema Sp_ToBI. Como resultado del análisis, se representa el tono neutro como plano y sostenido —H* H* H* M %—. La tristeza, por su parte, presenta una curva similar, pero con un descenso al final —H* H* H L* L%—. La alegría, con un patrón H*L+>H*H*L%, encuentra un ascenso progresivo desde la sílaba tónica hasta el punto máximo en el núcleo y termina en un tono descendente. El enojo se representa por L+;H*L+>H*H+L*L% por sus movimientos ascendentes en el segmento prenuclear y nuclear, con un descenso en la juntura terminal. Al comparar la diferencia de F0, la única emoción que se diferencia de las demás es la alegría.

En cuanto a investigaciones relacionadas con el español de México, solo se encuentra la aproximación que hace Martín Butragueño (2016) a los actos de habla expresivos, a partir de una taxonomía continua de ánimo-desánimo. En este trabajo, describe que el pico inicial de los enunciados expresivos analizados generalmente es ascendente y que la valencia negativa de un enunciado se manifiesta en el pico prenuclear de forma descendente. Entretanto, el énfasis o fuerza ilocutiva alta se muestra en un acento nuclear ascendente, pero no integra las emociones como tal.

También existen estudios que modelan las emociones para integrarlas en sistemas de reconocimiento de texto a voz y viceversa. En un ejemplo de este tipo de estudios (Iriando Guaus, Rodríguez, Lázaro, Montoya, Blanco, Bernadas, Oliver, Tena, & Longhi, 2000), se realiza un análisis suprasegmental de los patrones de voz de un corpus simulado de 7 emociones, el cual es validado a través de pruebas de percepción. Como resultado, se logra modificar la prosodia de un texto convertido a voz mediante técnicas de síntesis digital de audio. Aunque el objetivo de este y otros trabajos de este tipo tienen la finalidad de modelar una emoción artificialmente es importante rescatar los elementos que utilizan para lograrlo: el rango y variabilidad tonal, la velocidad de inflexiones tonales, la estabilidad en la intensidad y la duración de los silencios.

En otro trabajo (Montero, Gutiérrez-Arriola, Colás, Enríquez & Pardo, 1999) se validó y creó una base de datos de emociones en español simulado, incluyendo cuatro emociones: alegría, tristeza, enojo y sorpresa como parte del proyecto VAESS (Voices Attitudes and Emotions in Synthetic Speech). Este estudio refleja la importancia de la calidad de voz en el reconocimiento de emociones por medio de la voz, y se encontró que la tristeza y la sorpresa se reconocen mejor por sus cualidades suprasegmentales, en comparación con el enojo y la alegría.

Otro estudio en español peninsular es el de modelización acústica (Rodríguez, Lázaro, Montoya, Blanco, Bernadas, Oliver, & Longhi, 1999), en el que se creó un corpus de habla simulada con 8 actores y 7 emociones básicas: alegría, deseo, rabia, miedo, sorpresa, tristeza y asco, producidas con 3 niveles diferentes de intensidad cada una. A partir de este, se eligieron los 34 discursos mejor evaluados en una prueba de percepción. En esta investigación, se destaca que el análisis se realizó promediando parámetros acústicos concretos de entonación, intensidad, ritmo y timbre, no solo entre sujetos, sino también intrasujeto, con el fin de poder modelizar las emociones.

Al revisar todas estas aportaciones se detecta que en la mayoría se utilizan modelos discretos para la clasificación de las emociones y se utiliza el enunciado «neutro» como punto de referencia para hacer la comparación. También se observa que para su análisis se han utilizado los parámetros prosódicos de F0, intensidad y velocidad, así como la descripción de la curva tonal y los movimientos en el tonema. En ocasiones, parten de corpus espontáneos, pero en muchas otras de corpus actuados, ya sea con múltiples enunciados o con uno solo.

4. Metodología

4.1. Elaboración del corpus y elicitación de datos

Los datos que se utilizaron para este análisis fueron creados *ad hoc* para un instrumento de percepción de emociones. En este contexto fue importante evitar enunciados con una carga expresiva, controlar la duración, la estructura silábica y acentual. Teniendo en cuenta esto, se crearon 41 enunciados, todos ellos con vocabulario cotidiano, incluido en pruebas dirigidas a población infantil (Mansilla, 2005). Las oraciones fueron creadas

manteniendo el mismo número de sílabas y ubicando los acentos léxicos en la misma posición para observar las prominencias tonales en una estructura similar del tipo: $\sigma\sigma^*\sigma\sigma^*\sigma\sigma^*$ (σ = sílaba).

En cuanto a las emociones, se eligieron una para cada extremo del modelo bioinformacional, para tener el contraste de activación + ACT - ACT y valencia + VAL - VAL, además de la emoción neutra NEU como punto de referencia. Se consideran extremas, pues en los estudios previos de Russell (1980), Bradley y Lang (1994), las emociones seleccionadas para este estudio se han ubicado en este plano dimensional más lejanas del centro y en un punto medio entre los valores de activación y valencia de su cuadrante.

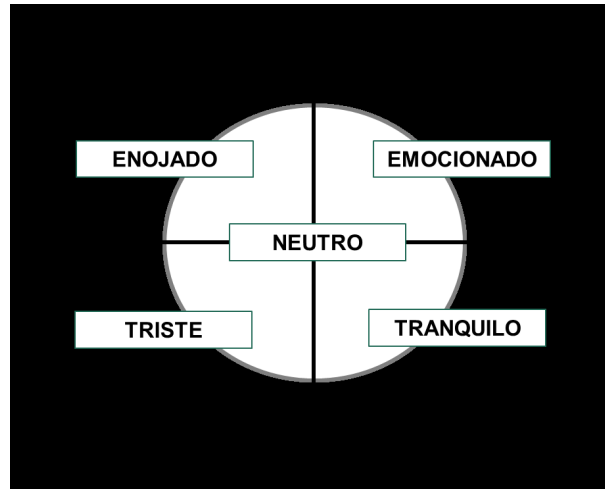


Figura 2. Emociones seleccionadas dentro del modelo bioinformacional

La muestra fue tomada de 3 hablantes nativas de español —en adelante, identificadas como M1, M2 y M3—, con residencia en Querétaro, México, ciudad cercana geográficamente a la Ciudad de México, pero que también colinda con otros estados considerados más cercanos al norte del país. Las tres informantes son adultas, con una edad promedio de 39 años ($DE = 11.1$), con estudios de licenciatura, profesionistas, familiarizadas con el manejo de la voz y la actuación, y en interacción constante con niños en el área laboral. En cuanto al tono de voz se refiere, la tesitura de las tres hablantes es similar y mantiene patrones suprasegmentales marcados de habla dirigida al niño.²

Para preparar a las hablantes se simulaban situaciones en las que se dirigieran a un niño, poniendo el contexto con situaciones y apoyos visuales —fotografías con el gesto de cada emoción—. Se les pidió que evocaran la emoción mentalmente con una situación y posteriormente repitieran las 41 oraciones 3 veces con la entonación prototípica,

2. Esto se hizo así, pues las oraciones utilizadas para este estudio forman parte de un trabajo de percepción de prosodia emocional. En ese sentido, se ha encontrado en trabajos previos que las emociones con habla dirigida al niño se reconocen mejor que las de habla adulta (Fu, Chinchilla, Nogaki, & Galvin, 2005; Paredes-Gallardo, Innes-Brown, Madsen, Dau, & Marozeau, 2018), inclusive son perceptibles desde el vientre materno inteligiblemente (Kuhl, 2004).

basada en un estereotipo de nivel o expresión modelo (Scherer, 2003) para cada emoción: emocionado, enojado, neutro, tranquilo y triste. Todas las grabaciones se hicieron en una cabina sonoamortiguada y con una grabadora Tascam DR-100MKII, a una distancia de 30 cm. Las oraciones fueron revisadas acústica y visualmente, y se realizó una segunda toma de aquellas que tuvieran algún tipo de pausa, fraseo intermedio o función expresiva, con el fin de que todas ellas estuvieran en el mismo estatus informativo. En todos los casos no se forzó una realización prosódica determinada para la emoción, pero sí se les solicitó que dijeran la frase de corrido, sin énfasis en ningún elemento del enunciado.

4.2. Procesamiento de los datos

Una vez obtenidas las grabaciones, se utilizó el programa Praat 6.01.05 (Boersma & Weenik, 2019) para su análisis y segmentación. Utilizando el programa Audacity 2.3.2 (Mazzoni, 2019) todos los audios fueron normalizados a 65 dB SPL y se eliminó el ruido, manteniendo las variaciones de intensidad dentro de cada estímulo. Esta decisión metodológica se tomó sabiendo que la intensidad relativa es un elemento importante para el reconocimiento de emociones, pero no el más saliente, por lo que se discrimina menos entre emociones que el tono (Chatterjee, Kulkarni, Christensen, Deroche, & Limb, 2015) en el que se enfoca este trabajo.

El proceso de etiquetado se hizo de forma intuitiva. Primero, se realizó la transcripción ortográfica del enunciado y la segmentación a nivel silábico y vocálico, identificando las vocales con acento léxico con (*). También se hizo una revisión independiente del *pitch*, restaurando los valores de la F0 donde se perdían o distorsionaban por presencia de ruido, microprosodia o voz laringizada. Por último, la curva tonal general es descrita mediante el sistema Sp_ToBI (Estebas-Vilaplana & Prieto, 2008).

5. Análisis de los datos

Se analizaron instrumentalmente las 1845 frases que conforman el corpus de estudio, extrayendo los promedios para cada vocal de: frecuencia fundamental (F0) en Hz, semitonos (st) e intensidad en decibeles (dB). A partir de estos valores, se llevaron a cabo varios análisis comparativos, tanto por emoción como por los niveles de activación y valencia, los cuales se describen a continuación. Para el análisis estadístico y la representación gráfica de resultados se utilizó el programa R-Studio v1.4.1717 (RStudio Team, 2021).

5.1. Parámetros generales

Al analizar la configuración global tonal de las tres hablantes (Tabla 1), se observa que tienen un promedio

similar en todas las oraciones; solo se observa un mayor rango en st para M2. En cuanto a la realización específica de las emociones, se observa una reducción en promedio y rango de F0 —en el orden *emocionado* > *enojado* > *tranquilo* > *triste* > *neutro* para M2—, mientras que para M1 y M3 *triste* tiene promedios más altos, similares a los de *emocionado* y *enojado*.

Hablante	Emoción	Máx F0	Mín F0	Media F0	Rango F0	Media st	Rango_st
M1	Emocionado	444.4	247.9	346.1	196.5	21.4	11.3
M1	Enojado	350.4	183.3	266.8	167.2	16.9	8.2
M1	Tranquilo	377.7	165.7	271.7	212.0	16.6	7.1
M1	Triste	371.1	190.8	280.9	180.3	17.5	5.8
M1	Neutro	325.1	167.0	246.0	158.1	15.0	3.0
	M1 promedio	373.7	190.9	282.3	182.8	17.5	7.1
M2	Emocionado	582.0	266.6	424.3	315.5	25.0	19.6
M2	Enojado	475.6	214.7	345.1	260.9	21.3	15.4
M2	Tranquilo	408.8	156.3	282.6	252.5	17.3	11.7
M2	Triste	391.3	161.5	276.4	229.8	16.9	9.9
M2	Neutro	347.4	143.2	245.3	204.1	14.8	8.5
	M2 promedio	441.0	188.5	314.7	252.6	19.1	13.0
M3	Emocionado	500.3	293.1	396.7	207.2	23.8	12.1
M3	Enojado	329.0	189.5	259.3	139.5	20.5	5.3
M3	Tranquilo	361.2	156.5	258.8	204.7	20.7	6.9
M3	Triste	383.2	153.7	268.5	229.5	22.0	10.1
M3	Neutro	328.9	151.9	240.4	177.0	19.9	7.5
	M3 promedio	413.5	188.7	301.1	224.8	20.1	10.9

Tabla 1. Medias de la F0 por hablante y por emoción

El rango tonal que hay entre el piso y el techo tonal por emoción permite ver la configuración global de las emociones (Tabla 2). Se observa que las emociones de activación alta como *emocionado* y *enojado* tienen un promedio de F0 mayor, pero se diferencian en el rango que es más elevado en *emocionado*. En cuanto a las emociones de activación baja, *tranquilo* y *triste*, se observa un rango tonal más bajo, pero no alcanzan al tono neutro. En la diferencia de rango, además, se puede observar que las emociones de valencia positiva tienen una diferencia mayor a las de valencia negativa. Entretanto, el tono neutro se encuentra por debajo de todos los valores asignados a las otras emociones. Al realizar el análisis estadístico, se encuentra que las emociones tienen una diferencia significativa en el promedio de F0 en Hz — $F(4,160) = 3045.8; p < 0.01$ — marcadas con * en la Tabla 2.

Emoción	Máxima F0	Mínima F0	Media F0	Rango F0
Emocionado	508.9	269.2	389.0*	239.7
Enojado	385.0	195.8	290.4*	189.2
Tranquilo	382.6	159.5	271.0*	223.1
Triste	381.9	168.7	275.3*	213.2
Neutro	333.8	154.0	243.9*	179.7

Tabla 2. Rango tonal y medias de F0 por emoción

5.2. Inicio del enunciado

Las emociones pueden ser reconocidas con muy poca información acústica. El punto de inicio de una oración puede diferenciar su intención y en este caso la emoción (Castiajo & Pinheiro, 2019). Es por eso que la medida en st de la sílaba inicial permite reconocer si la diferencia es perceptible desde el primer momento. Para analizar la diferencia en el inicio de las oraciones, se realizaron varios ANOVA y se aplicó el ajuste *post hoc* de Bonferroni. El promedio de inicio de cada emoción fue el siguiente: 22.9 st para *emocionado*, 18.5 st para *enojado*, 13.04 st para *neutro*, 13.5 st para *tranquilo* y 15.6 st para *triste*, en comparación con la media total del tono neutro, que fue de 12.7 st. El promedio de inicio de cada emoción fue así: 22.9 st para *emocionado*, 18.5 st para *enojado*, 13.04 st para *neutro*, 13.5 st para *tranquilo* y 15.6 st para *triste*, en comparación con la media total del tono neutro, que fue de 12.7st. El análisis estadístico de estos datos muestra una asociación significativa por el punto de inicio — $F(4,160) = 1858.8; p < 0.01$ —, mientras que en la prueba *post hoc* se encontró que todas son significativamente diferentes, excepto *neutro* y *tranquilo* que, como puede verse en la Figura 3, comparten la misma tendencia de inicio, por lo que solo podrían diferenciarse aquellas que están a los extremos.

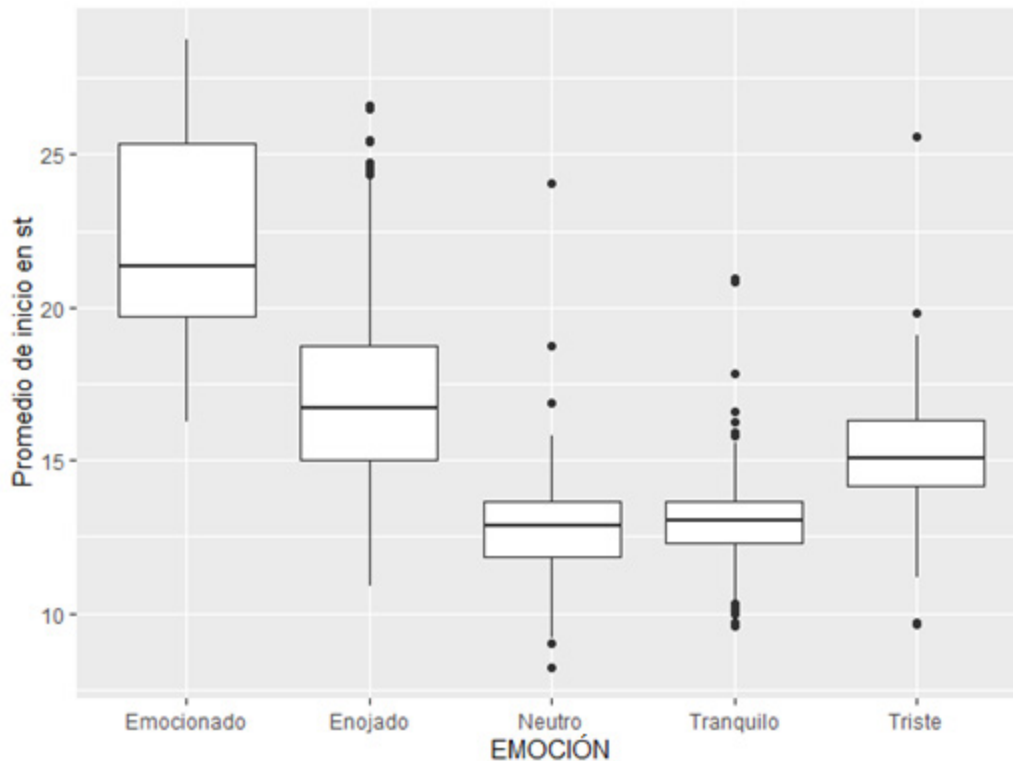


Figura. 3. Diagrama de cajas y bigotes de la distribución de F0 en la vocal de inicio por emoción

En este caso se aplicó la clasificación propuesta dentro del modelo bioinformacional de las emociones seleccionadas a partir de su nivel de activación y valencia. La diferencia entre los dos niveles de activación y valencia puede observarse en las figura 4 y 5. El análisis estadístico para el factor de activación en sus tres niveles —alta, baja y neutra— muestra un efecto significativo en el promedio de inicio — $F(2,80) = 3287.58; p < 0.01$ —. También se realizó el análisis estadístico para el factor de valencia, en el que se encuentra una diferencia significativa entre los tres niveles —positiva, neutra y negativa— — $F(2,80) = 1923.45; p < 0.01$ —. Esto implica que los tres niveles son significativamente diferentes en la altura tonal de su sílaba inicial y, por ende, la prueba *post hoc* confirmó la diferencia significativa entre los niveles $p < 0.01$, tanto para los niveles de activación como de valencia.

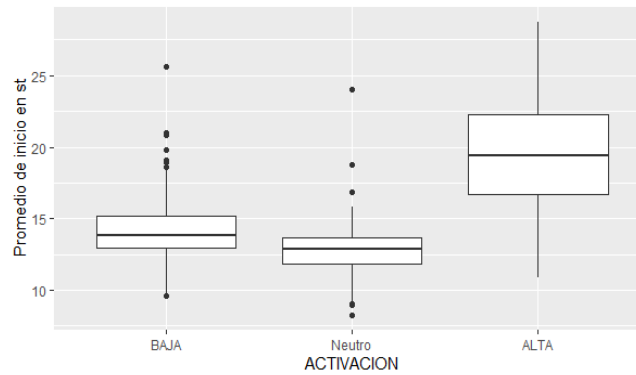


Figura 4. *Altura de la F0 de sílaba inicial en promedio de st por niveles de activación*

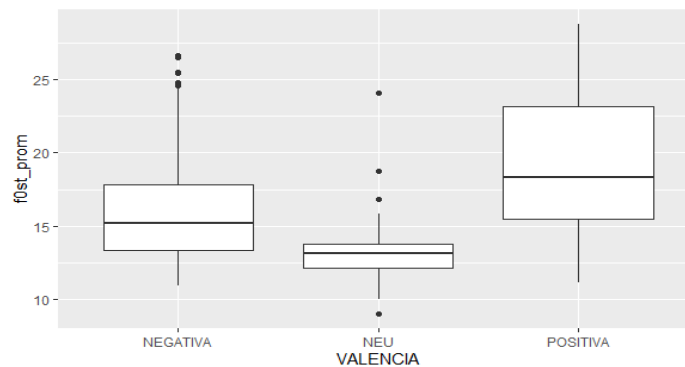


Figura 5. *Altura de la F0 de sílaba inicial en promedio de st por niveles de valencia*

5.3. Curva tonal

Para la descripción de la curva tonal se extrajeron los valores absolutos de todas las vocales para observar el movimiento general y establecer así los movimientos más amplios. Después, se hizo un análisis relativo de las variaciones de st entre segmentos y se identificaron aquellos en los que las emociones presentaban diferencias significativas entre sí. A partir de la medición en semitonos, se calculó la diferencia entre vocales, para determinar aquellos movimientos mayores a 1.5 st, tomando en cuenta el umbral de percepción descrito anteriormente. Con estos mismos valores se obtuvo una curva que representa el tipo de movimiento, ya sea ascenso o descenso que hay entre las sílabas (Figura 6). Esta representación gráfica es de desarrollo propio y va de la mano de la propuesta metodológica del análisis en contexto de la diferencia entre sílabas acentuadas. En un primer momento, puede observarse que la diferencia se hace más significativa cuando el código de esfuerzo es mayor, lo que corresponde a una activación alta como en *emocionado* y *enojado*, aunque la configuración general no cambia y se asemeja a la configuración esperada para enunciados declarativos.

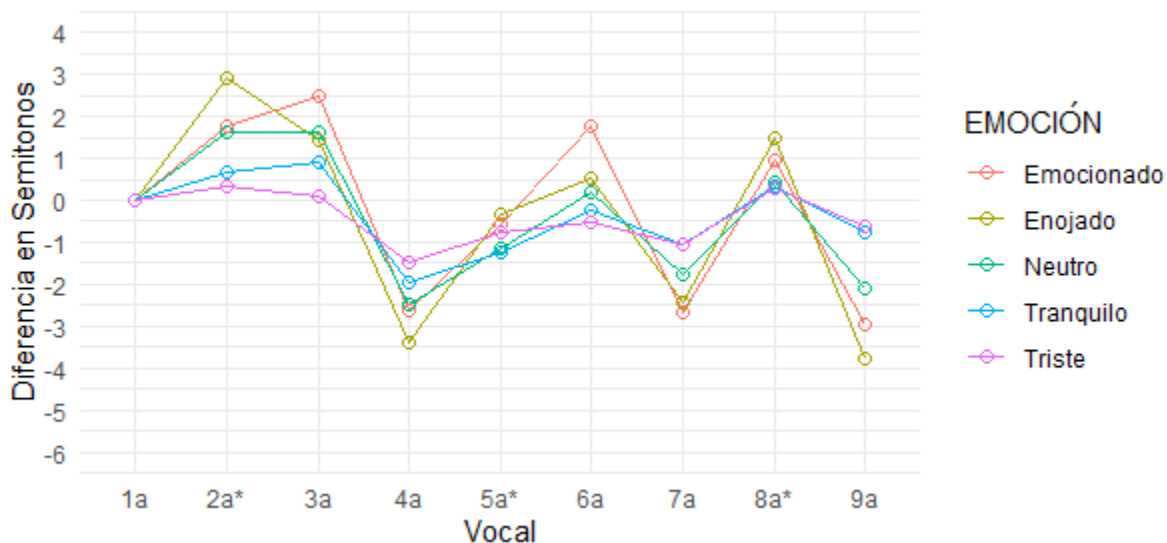


Figura 6. Diferencias en st para cada emoción

En los enunciados neutros se observa un inicio acorde a la descripción de un enunciado aseverativo del español (Navarro Tomás, 1944; Quilis, 1997), caracterizado por un comienzo en tono medio que se eleva progresivamente hasta la sílaba tónica y encuentra su pico tonal alineado a la derecha en la sílaba postónica L+>H* (Estebas Vilaplana & Prieto, 2008), así como con un patrón circunflejo en el tonema, como el descrito para los enunciados aseverativos de foco amplio del español de la Ciudad de México (Martín Butragueño, 2017).

El segmento prenuclear se caracteriza por ser bitonal en emociones de activación alta —*emocionado* y *enojado*— con un ascenso mayor a 3 st, mientras que las emociones de activación baja —*triste* y *tranquilo*— son monotonales. En este segmento es interesante observar que las emociones de valencia positiva —*emocionado* y *tranquilo*— presentan un pico desplazado, mientras que las de valencia negativa sí se definen en la sílaba tónica.

En los enunciados con activación alta, la diferencia en estos movimientos tonales es mayor, observando ascensos mayores a 1.5 st en todas las sílabas y, por lo tanto, más evidente su contraste con los de activación baja. En los enunciados de tristeza, se observa también el inicio en un tono alto que, a diferencia de lo reportado en otros trabajos (Velásquez, Garzón, & Soto, 2018) no desciende en el tonema, sino que se mantiene M %. En la Tabla 3 puede observarse un resumen de la configuración general del tono por sílaba en cada una de las emociones.

Emoción	2a*	5 ^a *	8 ^a *	9a	Activación	Valencia
Emocionado	L+>̣H*	L+>̣H*	L+̣H*	!L%	+	+
Enojado	L+̣H*	L+>̣H*	L+̣H*	!L%	+	-
Neutro	L+H*	L+>H*	L+H*	L%		
Tranquilo	H*	L+ >H*	H*	M%	-	+
Triste	H*	L+>H*	H*	M%	-	-

Tabla 3. Resumen de configuraciones tonales por sílaba para cada emoción

Nota: En negritas se destacan los picos desplazados mayores a .3 st

5.4. Configuración nuclear por emoción

La configuración nuclear es fundamental, pues es en este segmento en donde se consolida el sentido de unidad y significado del enunciado y en donde se realizan variaciones que pueden proyectar la intención del hablante. Para este análisis se observa solamente la diferencia en st del segmento prenuclear, nuclear y el posnuclear (Figura 7). En esta gráfica se puede observar el tonema con el movimiento circunflejo descrito anteriormente del habla de la zona central de México (Martín Butragueño, 2017) para el tono neutro y con un descenso más pronunciado en el caso de las emociones de activación alta !L %, mientras que las de activación baja se mantienen M %.

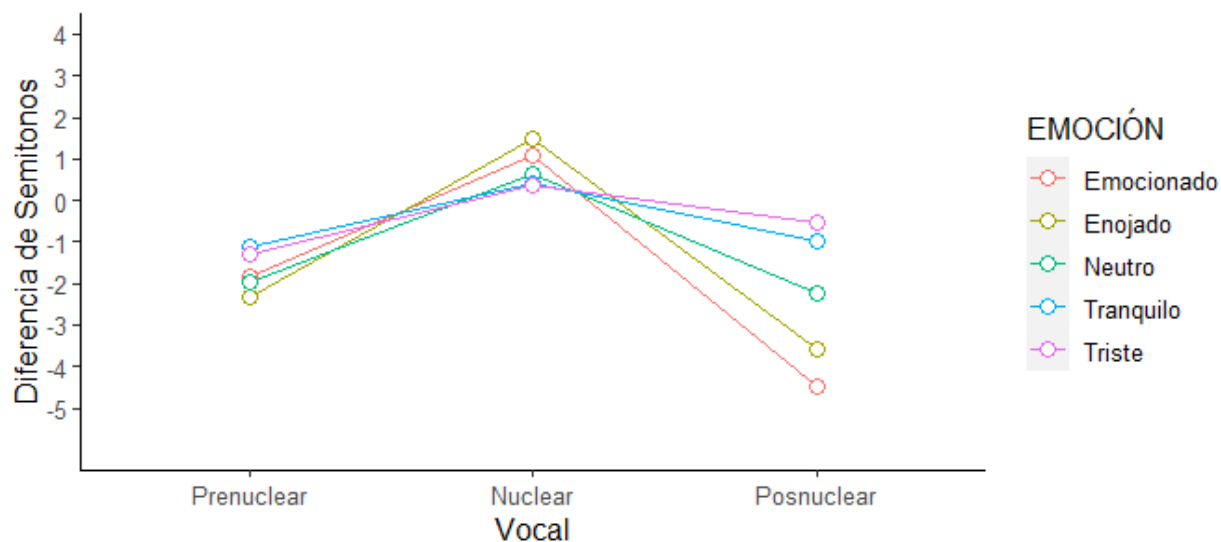


Figura 7. Media de las diferencias de F0 en el tonema

6. Conclusiones

Con base en el marco metodológico del que parte este estudio, se evidencia la razón de la carencia de estudios suficientes en torno a este elemento paralingüístico. La propuesta de análisis en este trabajo es valiosa en sí misma por tratarse de un primer acercamiento a la descripción de la prosodia de las emociones en el español mexicano. Además, la propuesta específica de este trabajo, que consiste en la agrupación de las emociones en torno a las dimensiones de activación y valencia, provee una posibilidad de caracterización que no se ha utilizado en el español. También resulta valioso el uso de la medida en st, pues al comparar con estudios previos (Rodríguez, Lázaro Montoya, Blanco, Bernadas, Oliver, & Longhi, 1999), en los que las variaciones tonales se reportan con porcentajes, es difícil comparar las variaciones tonales entre emociones, pues se necesita convertir a una medida relativa que las haga equiparables, como la propuesta en st. Al traducir las medidas absolutas, tenemos información que, inclusive, permite agrupar las emociones en rasgos como la activación y la valencia para analizar su comportamiento tonal.

A partir de esta propuesta de análisis se encuentran más regularidades asociadas a este código de esfuerzo, y pueden ser una guía para la caracterización de las emociones en contornos diferenciados y la agrupación por las propiedades mismas de las emociones. Por el momento, los resultados observados en la configuración del inicio y nuclear brindan la posibilidad de caracterizar las emociones por su nivel de activación. Aunque la F0 no permite identificar la valencia, sí se observó el pico desplazado en la sílaba inicial de los enunciados con valencia positiva, mientras que en las de valencia negativa el pico está en el núcleo de la juntura inicial. Los datos que proveen este corpus aún presentan la posibilidad de análisis de factores como la duración y la intensidad relativa, que aunque juegan un papel importante en la producción de las emociones, sobrepasan el alcance de este estudio.

Los valores obtenidos en las curvas por emoción presentan valores similares a los que se comparten en el trabajo de Montero, Gutiérrez-Arriola, Colás, Enríquez y Pardo (1999) en cuanto al tono de la sílaba inicial de las emociones. También se encuentra similitud con el trabajo de Hidalgo Navarro (2020), pues caracteriza la alegría y el enfado con una F0 alta, relacionada con las excursiones tonales de la alegría y el enojo en el corpus de este trabajo. Mientras que los enunciados tristes y neutros que Martínez y Rojas (2011) encuentran similares, aquí se encontró que son diferentes en cuanto a la sílaba inicial y tampoco comparten la configuración del tonema. Al contrastar con el español de Colombia (Velásquez, Garzón & Soto, 2018) se observan configuraciones tonales distintas, pues la variante del español en México presenta el patrón circunflejo, excursiones tonales más pronunciadas y acentos bitonales recurrentes.

Aunque este trabajo tiene limitaciones metodológicas en cuanto al número de hablantes y espontaneidad de los estímulos, la sistematicidad con la que se crearon los enunciados y la manera como se realizó el análisis dan cuenta real de variaciones que pueden ser utilizadas como referencia para estudios posteriores. Sería muy importante integrar a este análisis la variabilidad entre hablantes hombre-mujer y las mismas oraciones, con el

propósito de que puedan servir para comparar con otras variaciones dialectales dentro del mismo español mexicano. También sería conveniente controlar la coda de la sílaba final, pues en el análisis se observaron fenómenos de ensordecimiento o elisión de fonemas, aunado al debilitamiento natural del habla queretana.

Además, será importante probar estas distinciones frente a la capacidad de reconocerlas en una tarea perceptual, para integrar el fenómeno de la emoción desde el punto de vista de la interacción, en donde la realización de las emociones en la voz realmente tiene su fundamento. La revisión de la descripción propuesta en este trabajo, tomando en cuenta los enunciados con las formas acústicas emocionales más perceptibles, también será una forma de afinar con más detalle los rasgos acústicos que permiten reconocer las emociones y agruparlas por sus niveles de activación y valencia de una forma aún más clara.

Referencias bibliográficas

1. Ballesteros, M. P. (2014). Hacia una teoría compleja de la entonación y análisis de la melodía del habla de Navarra. *RILCE*, 30(1), 44-72. <https://hdl.handle.net/10171/36992>
2. Banse, R., & Scherer, K. R. (1996). Acoustic Profiles in Vocal Emotion Expression. *Journal of Personality and Social Psychology*, 70(3), 614-636. <https://psycnet.apa.org/doi/10.1037/0022-3514.70.3.614>
3. Beckman, M. E., Díaz-Campos, M., Tevis, J., & Morgan, M. T. A. (2002). Intonation across Spanish, in the Tones and Break Indices Framework. *Probus*, 14(1), 9–36. <https://doi.org/10.1515/prbs.2002.008>
4. Boersma, P., & Weenik, D. (2019). PRAAT. <http://www.praat.org/>
5. Bradley, M., & Lang, P. J. (1994). Measuring Emotion: The Self-Assessment Manikin and the Semantic Differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1), 49–59. [https://doi.org/10.1016/0005-7916\(94\)90063-9](https://doi.org/10.1016/0005-7916(94)90063-9)
6. Cantero, F. (2014). *Códigos de la entonación y entonación emocional*. Congreso Internacional de La Asociación Española de Lingüística Aplicada ACTAS, 618-629, AESLA.
7. Castiajo, P., & Pinheiro, A. P. (2019). Decoding Emotions from Nonverbal Vocalizations: How Much Voice Signal Is Enough? *Motivation and Emotion*, 43(5), 803–813. <https://doi.org/10.1007/s11031-019-09783-9>
8. Charpentier, J., Kovarski, K., Houy-Durand, E., Malvy, J., Saby, A., Bonnet-Brilhault, F., Latinus, M., & Gomot, M. (2018). Emotional Prosodic Change Detection in Autism Spectrum Disorder: An Electrophysiological Investigation in Children and Adults. *Journal of Neurodevelopmental Disorders*, 10(28), 1–16. <https://doi.org/10.1186/s11689-018-9246-9>
9. Chatterjee, M., Kulkarni, A. M., Christensen, J. A., Deroche, M. L., & Limb, C. J. (2015). Voice Emotion Recognition and Production by Individuals with Normal Hearing and with Cochlear Implants. *The Journal of the Acoustical Society of America*, 137(4), 2205–2205. <https://doi.org/10.1121/1.4920020>
10. Chronaki, G., Wigelsworth, M., Pell, M. D., & Kotz, S. A. (2018). The Development of Cross-Cultural Recognition of Vocal Emotion During Childhood and Adolescence. *Scientific Reports*, 8(8659), 1–17. <https://doi.org/10.1038/s41598-018-26889-1>
11. Darwin, C. (1872). *The Expression of the Emotions in Man and Animals*: John Murray Editor.
12. Ekman, P. (1972). *Universals and Cultural Differences in Facial Expressions of Emotion*. In Nebraska Symposium on Motivation, 19, 207–282. Lincoln: University of Nebraska Press.
13. Espinosa, H. P., & Reyes, C. A. (2010). *Reconocimiento de emociones a partir de voz basado en un modelo emocional continuo*. [Reporte de tesis doctoral]. <https://ccc.inaoep.mx/portalfiles/file/CCC-10-005.pdf>
14. Estebas-Vilaplana, E., & Prieto, P. (2008). La notación prosódica del español: una revisión del Sp-ToBI. *Estudios de Fonética Experimental*, 17, 263-283. http://prosodia.upf.edu/home/arxiu/publicacions/estebas/estebas_notacion_prosodica_espanol.pdf

15. Face, T. L. (2010). The Necessity of Both Naturally-Occurring and Elicited Data in Spanish Intonational Phonology. *Studies in Hispanic and Lusophone Linguistics*, 3(2), 485–500. <https://doi.org/10.1515/shll-2010-1083>
16. Fernández Planas, A. M., & Martínez Celdrán, E. (2003). El tono fundamental y la duración: dos aspectos de la taxonomía prosódica en dos modalidades del habla (enunciativa e interrogativa) del español. *Estudios de Fonética Experimental*, 12, 166-200. <https://raco.cat/index.php/EFE/article/view/140007>
17. Fu, Q.-J., Chinchilla, S., Nogaki, G., & Galvin, J. J. (2005). Voice Gender Identification by Cochlear Implant Users: The Role of Spectral and Temporal Resolution. *The Journal of the Acoustical Society of America*, 118(3), 1711–1718. <https://doi.org/10.1121/1.1985024>
18. Gilbers, S., Fuller, C., Gilbers, D., Broersma, M., Goudbeek, M., Free, R., & Başkent, D. (2015). Normal-Hearing Listeners' and Cochlear Implant Users' Perception of Pitch Cues in Emotional Speech. *I-Perception*, 6(5), 1–19. <https://doi.org/10.1177/0301006615599139>
19. Goldsmith, J. (1976). Autosegmental Phonology [Doctoral Dissertation]. <http://oastats.mit.edu/bitstream/handle/1721.1/16388/03188555-MIT.pdf?sequence=1&isAllowed=y>
20. Gussenhoven, C. (2010). *The Phonology of Tone and Intonation*: Cambridge University Press.
21. Hidalgo Navarro, A. (2020). Rasgos melódicos de la emoción: estudio de un corpus conversacional. *Phonica*, 16, 36–53. <https://doi.org/10.1344/phonica.2020.16.36-53>
22. Hualde, J. I. (2003). El modelo métrico y autosegmental. In P. Prieto (Ed.), *Teorías de la Entonación* (pp. 155–184). Ariel Lingüística.
23. Hualde, J. I., & Prieto, P. (2015). Intonational Variation in Spanish: European and American Varieties Get Access Arrow. In S. Frota & P. Prieto (Eds.), *Intonation in Romance* (pp. 350-391). Oxford Academic. <https://doi.org/10.1093/acprof:oso/9780199685332.003.0010>
24. Iriondo, I., Guaus, R., Rodríguez, A., Lázaro, P., Montoya, N., Blanco, J. M., Bernadas, D., Oliver, J. M., Tena, D., & Longhi, L. (2000). Validation of an Acoustical Modelling of Emotional Expression in Spanish Using Speech Synthesis Techniques. *Speech and Emotion. ISCA Tutorial and Research Workshop*, 161–166. http://www.isca-speech.org/archive_open/speech_emotion/spem_161.html
25. James, W. (1884). What is an Emotion? *Mind Association*, 9(34), 188–205. <https://www.jstor.org/stable/2246769>
26. Kuhl, P. K. (2004). Early Language Acquisition: Cracking the Speech Code. *Nature Reviews Neuroscience*, 5(11), 831–843. <https://doi.org/10.1038/nrn1533>
27. Ladd, R. (2008). *Intonational Phonology*. Cambridge University Press.
28. Lang, P., Bradley, M., & Cutbert, B. N. (NIMH C. for the S. of E. and A. (1997). International Affective Picture System (IAPS): Technical Manual and Affective Ratings. *Psychology*, 77, 1–5.
29. Liberman, M., & Prince, A. (1977). On Stress and Linguistic Rhythm. *Linguistic Inquiry*, 8(2), 249–336. <http://www.jstor.org/stable/10.2307/4177987>

30. Luo, X., Fu, Q. J., & Galvin, J. J. (2007). Cochlear Implants Special Issue Article: Vocal Emotion Recognition by Normal-Hearing Listeners and Cochlear Implant Users. *Trends in Amplification*, 11(4), 301–315. <https://doi.org/10.1177/1084713807305301>
31. Mansilla, T. (2005). *Oraciones en Formato Abierto*. OFA-N.
32. Martín Butragueño, P. (2015). Hacia una prosodia basada en el uso: actos de habla en el español mexicano. *Normas*, 5, 97–115. <https://doi.org/10.7203/normas.5.6824>
33. Martín Butragueño, P. (2016). A veces lloro mis lágrimas: acercamiento multivariable a la prosodia de los actos de habla expresivos en el español de México. *Estudios de Lingüística Aplicada*, 34(63), 59-102. <https://doi.org/10.22201/enallt.01852647p.2016.63.626>
34. Martín Butragueño, P. (2017). Contacto dialectal entonativo. Estudio exploratorio. In L. Orozco & A. Guerrero (Coords.), *Estudios de Variación Lingüística* (pp. 139–168). Instituto Nacional de Antropología e Historia.
35. Martínez, H., & Rojas Avendaño, D. (2011). Prosodia y emociones: datos acústicos, velocidad de habla y percepción de un corpus actuado. *Lengua y Habla*, 15(1), 59-72. <https://dialnet.unirioja.es/servlet/articulo?codigo=3979257>
36. Mazzoni, D. (2019). Audacity 3.1.0. www.audacityteam.org
37. Montero, J. M., Gutiérrez-Arriola, J., Colás, J., Enríquez, E., & Pardo, J. M. (1999). *Analysis and Modelling of Emotional Speech in Spanish*. 14th International Congress of Phonetic Sciences (pp. 957-960), Universidad Nacional de Educación a Distancia.
38. Murrieta, L. (2016). Acercamiento al análisis experimental del umbral de percepción entonativa en el español del centro de México. *Estudios de Lingüística Aplicada*, 34(63), 153-166.
39. Navarro Tomás, T. (1944). *Manual de entonación española*. Nueva York Hispanic Institute.
40. Pamies, A., Fernáandez Planas, A. M., Martínez, E., Ortega, A., & Amorós, M. (2002). Umbrales tonales en español peninsular. *Actas Del II Congreso Nacional de Fonética Experimental* (pp. 272–278), Universidad de Sevilla.
41. Paredes-Gallardo, A., Innes-Brown, H., Madsen, S. M. K., Dau, T., & Marozeau, J. (2018). Auditory Stream Segregation and Selective Attention for Cochlear Implant Listeners: Evidence from Behavioral Measures and Event-Related Potentials. *Frontiers in Neuroscience*, 12. <https://doi.org/10.3389/fnins.2018.00581>
42. Pierrehumbert, J. (1980). *The Phonology and Phonetics of English Intonation*. [Doctoral thesis]. Massachusetts Institute of Technology [MIT].
43. Prieto, P., & Roseano, P. (2018). Prosody: Stress, Rhythm, and Intonation. In K. L. Geeslin (Ed.), *The Cambridge Handbook of Spanish Linguistics* (pp. 211–236). Cambridge University.
44. Quilis, A. (1997). *Principios de fonología y fonética españolas*. Arco.
45. Rodero, E. (2011). Intonation and Emotion: Influence of Pitch Levels and Contour Type on Creating Emotions. *Journal of Voice*, 25(1), e25–e34. <https://doi.org/10.1016/j.jvoice.2010.02.002>

46. Rodríguez, Á., Lázaro, P., Montoya, N., Blanco, J., Bernadas, D., Oliver, J. M., & Longhi, L. (1999). Modelización acústica de la expresión emocional en el español. *Procesamiento de Lenguaje Natural*, 25, 159-166. <http://rua.ua.es/dspace/handle/10045/1964?mode=simple>
47. RStudio Team. (2021). RStudio: Integrated Development for R. (No. 2021). RStudio, PBC. <http://www.rstudio.com/>
48. Russell, J. A. (1980). A Circumplex Model of Affect. *Journal of Personality and Social Psychology*, 39(6), 1161–1178. <https://doi.org/10.1037/h0077714>
49. Russell, J. A., & Mehrabian, A. (1977). Evidence for a Three-Factor Theory of Emotions. *Journal of Research in Personality*, 11(3), 273–294. [https://doi.org/10.1016/0092-6566\(77\)90037-X](https://doi.org/10.1016/0092-6566(77)90037-X)
50. Scherer, K. R. (2003). Vocal Communication of Emotion: A Review of Research Paradigms. *Speech Communication*, 40(1–2), 227–256. [https://doi.org/10.1016/S0167-6393\(02\)00084-5](https://doi.org/10.1016/S0167-6393(02)00084-5)
51. Scherer, K. R. (2005). What Are Emotions? And How Can They Be Measured? *Social Science Information*, 44(4), 695–729. <https://doi.org/10.1177/0539018405058216>
52. Velásquez, E. P., Garzón, O. L., & Soto, D. del C. (2018). Prosodia y actos de habla: Expresividad en el habla joven. *Ideas*, 4(4), 1–20. <https://tinyurl.com/y7lcqsvw>