

# Chapter 1

## Introducción

Desde hace años, el reconocimiento de emociones a través de la voz ha sido motivo de interés para la investigación, sin embargo siempre se ha estudiado sobre un mismo lenguaje debatiendo la habilidad de reconocer y clasificar las emociones oralmente expresadas. Esta habilidad ha sido respaldada por numerosos artículos donde se concluye que es posible distinguir e identificar entre al menos cuatro emociones básicas (felicidad, tristeza, tristeza, y enfado) a través de la voz (sin necesidad del procesamiento del lenguaje natural y por lo tanto de un contexto).

Atendiendo al estudio de las emociones expresadas según la lengua existen estudios donde se demuestra que individuos de diferentes culturas pueden reconocer emociones básicas en diferentes niveles, pero es menos abundante la evidencia de un acuerdo en cómo las emociones básicas son reconocidas desde la expresión vocal de un interlocutor [7]. Análogamente el debate del reconocimiento de emociones en un plano intercultural también se ha enfocado a través del estudio de los gestos faciales en conjunto con la expresión vocal, donde se concluye los factores sociales tienen un gran impacto, ya que la identificación de las emociones es más fácil para los miembros de la misma cultura que para los de otra distinta [7] y [6]. A pesar de ello hay una gran carencia de comparativas con respecto a la voz donde se demuestre una sólida influencia cultural, sin embargo parece claro que las dimensiones socio culturales que engloban nuestras interacciones pueden tener un gran impacto en nuestra comunicación dentro de un marco emocional.

La expresión de las emociones están íntimamente relacionadas con las propiedades fonéticas en el habla donde se observan señales y patrones para marcar contrastes lingüísticos en un idioma [4] por lo tanto, los efectos del lenguaje en la comunicación emocional son evidentes al haber sido observadas y medidas, las variaciones en el rango tonal y la frecuencia para expresarlas, cambiando no sólo el tono si no también el patrón lingüístico asociado [3]. Por ejemplo la tristeza tiende a mostrarse con un tono notablemente bajo mientras que la felicidad, la sorpresa, o el enfado son producidas con

un tono moderadamente alto. En general se esperaría que las expresiones de tristeza y enfado tiendan a marcar una mayor diferencia entre ellas y por lo tanto sean reconocidas con más precisión con independencia del lenguaje al estar situadas en los extremos (opuestos) del espectro. En literatura anterior [5] se argumenta que las 6 emociones básicas (felicidad, miedo, asco, sorpresa, enfado y tristeza) son exitosamente reconocidas desde la prosodia. Cabe destacar la diferencia entre prosodia y la calidad vocal, donde la primera se centra en características tales como la entonación, el estrés y el ritmo del habla mientras que la segunda se refiere al tono, energía y tempo [8].

Por otro lado tanto la proporción de consonantes y vocales (que hacen variar la presión de aire que se necesita) como el ratio de sílabas por palabra en cada idioma, caracterizan la expresión oral de las emociones. Existen muchos factores relacionados con el lenguaje como la morfología o la duración del estímulo que podrían ser un impacto en la decodificación de los matices en la señal vocal, tal y como explica en [2].

Existe una clasificación dependiendo de la velocidad silábica en la expresión de dichos idiomas, sin embargo poco se conoce acerca de los efectos en las medidas respiratorias en el habla. Esta observación puede llevar a que se pregunte si en lenguajes tan dispares, las emociones expresadas mediante la voz puedan ser reconocidas desde el punto de vista del otro idioma. De nuevo, en [6] se describen anómalas pseudo-manifestaciones semánticas (discurso sin sentido) que se asemejan a la lengua materna para expresar cada tipo de emoción. Aquí se evidencia claramente que la emoción de cada interlocutor puede ser reconocida con precisión desde la prosodia independientemente del contexto semántico. Además argumenta que el significado emocional en la voz se transmite por cambios en diferentes parámetros acústicos como el tono, la intensidad, la duración, el ritmo y distintos aspectos de la calidad de la voz.

Finalmente, obtener datos necesarios de la extracción de características es un problema porque apenas cubre el amplio espectro emocional que hoy conocemos. C.Bakir y M.Yuzkat [1] proponen un modelo híbrido para clasificar cinco emociones básicas en el lenguaje turco donde combinan SVM y GMM y aplicando antes MFCC y MFDWC para la extracción de características.

# Bibliography

- [1] CIGDEM BAKIR and MECIT YUZKAT. “Speech Emotion Classification and Recognition with different methods for Turkish Language”. In: *Balkan Journal of Electrical and Computer Engineering* 6.2 (2018), pp. 54–60. ISSN: 2147-284X. DOI: 10.17694/bajece.419557.
- [2] Si Chen, Yiqing Zhu, and Ratree Wayland. “Effects of stimulus duration and vowel quality in cross-linguistic categorical perception of pitch directions”. In: *PLoS ONE* 12 (July 2017), e0180656. DOI: 10.1371/journal.pone.0180656.
- [3] Assel Davletcharova et al. “Detection and Analysis of Emotion from Speech Signals”. In: *Procedia Computer Science* 58 (2015), pp. 91–96. ISSN: 18770509. DOI: 10.1016/j.procs.2015.08.032. URL: <http://dx.doi.org/10.1016/j.procs.2015.08.032>.
- [4] Marc D. Pell. “Influence of emotion and focus location on prosody in matched statements and questions”. In: *The Journal of the Acoustical Society of America* 109.4 (2001), pp. 1668–1680. ISSN: 0001-4966. DOI: 10.1121/1.1352088.
- [5] Marc D. Pell et al. “Emotional speech processing: Disentangling the effects of prosody and semantic cues”. In: *Cognition and Emotion* 25.5 (2011), pp. 834–853. ISSN: 02699931. DOI: 10.1080/02699931.2010.516915.
- [6] Marc D. Pell et al. “Factors in the recognition of vocally expressed emotions: A comparison of four languages”. In: *Journal of Phonetics* 37.4 (2009), pp. 417–435. ISSN: 00954470. DOI: 10.1016/j.wocn.2009.07.005.
- [7] Marc D. Pell et al. “Recognizing Emotions in a Foreign Language”. In: *Journal of Nonverbal Behavior* 33.2 (2009), pp. 107–120. ISSN: 01915886. DOI: 10.1007/s10919-008-0065-7.
- [8] Social Signal Processing. “Research paper Social Signal Processing !” In: (2015).