

VISUALIZANDO EL CONFLICTO DISCURSIVO A TRAVÉS DE LA EXPRESIÓN FÓNICA: UN ESTUDIO A PARTIR DE DOS CONVERSACIONES

VISUALIZING DISCURSIVE CONFLICT THROUGH PHONETIC EXPRESSION: A STUDY BASED ON TWO CONVERSATIONS

Maria Estellés Arguedas

Universitat de València

Resumen

El presente análisis estudia la variación fónica en conversaciones conflictivas y, a partir de dos conversaciones espontáneas del corpus ESPRINT_VLC, y utilizando el pitch y la dirección de los intercambios como indicadores, presenta dos métodos de visualización. Los gráficos de dispersión con Z-Scores permiten rastrear anomalías fónicas en el tiempo, identificando acumulaciones de tonos altos o bajos y correlacionándolos con eventos conversacionales significativos. Los diagramas de red ilustran la interacción entre hablantes, incluyendo la tendencia de cada participante a modificar su pitch y la duración de sus turnos, proporcionando una "radiografía" de las conversaciones. Así, la metodología propuesta tiene aplicaciones en terapia de conflictos y comunicación intercultural, puesto que permite detectar patrones conversacionales e interactivos, como el de 'perseguidor' y 'evitador', que se relacionan con dinámicas interaccionales negativas.

Palabras Clave: expresiones fónicas, conflicto discursivo, tono, diagrama de red, interacciones discursivas, perseguidor, evitador.

Abstract

The current analysis examines phonetic variation in conflictive conversations and, drawing from two spontaneous conversations from the ESPRINT_VLC corpus, utilizing pitch and the direction of exchanges as indicators, presents two visualization methods. Scatter plots with Z-Scores allow for tracking phonetic anomalies over time, identifying clusters of high or low tones and correlating them with significant conversational events. Network diagrams illustrate the interaction among speakers, including the tendency of each participant to modify their pitch and the duration of their turns, providing a "radiography" of the conversations. Thus, the proposed methodology has applications in conflict therapy and intercultural communication, as it enables the detection of conversational and interactive patterns, such as 'pursuer' and 'withdrawer', which are related to negative interactional dynamics.

Keywords: phonetic expressions, discursive conflict, voice pitch, network diagram, discursive interactions, pursuer, withdrawer.

1 INTRODUCCIÓN

En este artículo de investigación se presenta una primera aproximación al análisis fónico del conflicto discursivo mediante el uso de un corpus de conversaciones coloquiales, el Corpus ESPRINT-conversación¹. Este corpus, compuesto por conversaciones coloquiales espontáneas en español, se convierte en la base fundamental para examinar cómo ciertas características fónicas, como el tono, la intensidad, la duración y la velocidad de habla, y otras interaccionales, como la dinámica de toma de turnos, se relacionan con el conflicto verbal y pueden ayudar a crear una radiografía de cada interacción que dé cuenta, a un simple golpe de vista, de la dinámica exhibida por los participantes en dicha interacción o en una serie de interacciones.

El enfoque metodológico adoptado en este trabajo plantea un abordaje cuantitativo y cualitativo. Por un lado, los fragmentos conflictivos seleccionados se han obtenido gracias a la percepción de los participantes en la interacción, puesto que una de las participantes fue también la encargada de transcribir y etiquetar los fragmentos con conflicto, y las percepciones de la persona implicada se han completado con las de un analista (cf. Estellés, en prensa). Por otro, el material transcrito y etiquetado se ha sometido a un análisis cuantitativo, basado en datos obtenidos de manera automática, que permite identificar patrones y tendencias que podrían no ser evidentes de manera intuitiva.

Los datos procedentes del análisis cuantitativo se han procesado mediante dos herramientas de visualización: gráficos lineales de avance prosódico (Cabedo, 2022) y diagramas de redes, para representar gráficamente la dinámica del conflicto discursivo en el espacio fónico. Al integrar métodos cualitativos y cuantitativos gracias a las herramientas de visualización, no solo se contribuye a la comprensión teórica de la comunicación conflictiva y a comprender la complejidad inherente a la intersección entre el conflicto discursivo y la expresión fónica (Couper-Kuhlen & Selting, 1996; Zellers & Ogden, 2014), sino que se apuntan aplicaciones prácticas en campos como la terapia de conflictos y la comunicación intercultural (Greenberg & Johnson, 1988; Johnson & Whiffen, 1999). La prosodia se convierte en un reflejo sonoro de las dinámicas comunicativas en situaciones de conflicto y se rastrea la huella fónica de los hablantes como indicador gráfico de las dinámicas interaccionales que, más allá de su interés lingüístico, pueden ayudar a pacientes y psicólogos en la identificación de conductas interactivas negativas y en su abordaje terapéutico.

Previamente a explorar la manifestación fónica del conflicto en los hablantes en conversación coloquial, el apartado 2 procede a apuntar brevemente, en primer lugar, las bases psicológicas del conflicto entre íntimos (como en el caso que nos ocupa, la de una pareja y sus hijas), para lo cual se partirá de la explicación proporcionada por la Emotionally Focused Therapy (Johnson, 2019); tras ello, en segundo lugar, se ahondará en los aspectos pragmático-discursivos del conflicto, centrándose en la reacción a lo largo de los turnos conversacionales y en la modulación fónica del desacuerdo.

¹ Esta investigación ha sido posible gracias a la ayuda recibida por el Ministerio de Ciencia e Innovación para el proyecto ESPRINT, "Estrategias pragmático-retóricas en la interacción conversacional conflictiva entre íntimos y conocidos: intensificación, atenuación y gestión interaccional" (ref. PID2020-114805GB-I00).

2 CONFLICTO: PROSODIA EN DISCURSO

El estudio del conflicto ha despertado el interés de diversas disciplinas sociales y humanísticas, desde la comunicación y sociología hasta la psicología, ciencias políticas, historia o antropología (Stewart & Maxwell, 2010: 11-12), y, más allá de la propia definición del término (cf. Sifianou 2019, Estellés, en prensa), una de las grandes preguntas que surgen es precisamente el origen del conflicto.

En el ámbito en que se enfoca este artículo, el de las relaciones familiares e íntimas, Johnson (2008) señala que la raíz de la mayor parte de los conflictos radica en la desconexión emocional y la inseguridad en el vínculo afectivo entre las parejas. La idea subyacente es que las emociones que habitualmente se manifiestan en los conflictos, como el enojo, la frustración o la crítica, son en realidad emociones subsidiarias que esconden otras emociones más básicas o primarias, como el miedo, la inseguridad, o la necesidad básica de conexión. Para Johnson, tras la mayoría de los conflictos hay un deseo profundo de conexión emocional y un miedo al rechazo o abandono. Las personas buscan seguridad emocional, y cuando sienten que su vínculo está amenazado, pueden surgir conflictos que, a menudo, involucran por parte de los participantes la asunción de dos patrones de interacción: el de perseguidor, que es quien busca más cercanía y comunicación, y el de evitador, que tiende a retirarse ante el conflicto. Estos dos roles, como veremos, tienen un claro correlato en los comportamientos lingüísticos de los participantes, entre los cuales se halla el comportamiento fónico.

En los últimos veinte años, la investigación sobre la relación entre conflicto y prosodia ha proporcionado valiosas perspectivas sobre cómo los elementos fónicos del habla influyen y reflejan las interacciones conflictivas (Roth & Tobin, 2010; Szczepek Reed, 2010; Wittfoth et al., 2010). La prosodia, que abarca aspectos acústicos como el tono, la intensidad o el ritmo, desempeña un papel fundamental en la expresión de emociones y actitudes durante el discurso (Ang et al., 2018; Bänziger & Scherer, 2005; Garrido Almiñana & Chica Sabariego, 2018), y emerge como una herramienta comunicativa polifuncional que va más allá de ser un simple acompañante en la transmisión de la información verbal y revela dimensiones cruciales en la dinámica del discurso.

The point that we wish to make here is that prosody can be seen as one of the orderly 'details' of interaction, a resource which interlocutors rely on to accomplish social action and as a means of steering inferential processes. Prosodic features, we suggest, can be reconstructed as members' devices, designed for the organization and management of talk in social interaction. They can be shown to function as part of a signalling system which - together with syntax, lexico-semantics, kinesics and other contextualization cues - is used to construct and interpret turn-constructural units and turns-at-talk (Couper-Kuhlen & Selting, 1996: 37)

En esta línea, algunas investigaciones han destacado la importancia del componente fónico en la expresión de la intensidad emocional durante el conflicto verbal (Kim et al., 2012; Roth & Tobin, 2010; Szczepek, 2010; Wittfoth et al., 2010; véase también Briz, en prensa, para el español). La entonación, el ritmo y otros aspectos fónicos del habla, por tanto, se han presentado como indicadores cruciales de la intensidad emocional (aunque no siempre de manera específica para cada emoción; véase Padilla García, 2023), la dirección del conflicto y los momentos de resolución.

En el análisis de la manifestación del conflicto, se ha observado una doble tendencia. Por un lado, se tiende a modular el desacuerdo para mitigar posibles consecuencias negativas en la interacción social (Clancy, 2018; Pomerantz, 1984). No obstante, junto a esta

tendencia existe un desacuerdo intensificado, presente en el habla conflictiva familiar, donde las estrategias mitigadoras son menos frecuentes debido a la dinámica conversacional (Kotthoff, 1993). Esta manifestación diferencial de las dos caras del conflicto también se ha descrito fónicamente, tanto en los momentos de escalada de la tensión discursiva, donde suele ir acompañada de un aumento general de los parámetros fónicos, como en los intentos por desescalarla mediante una reducción de estos:

When speakers increase their pitch level, speech intensity, and speech rates over the previous speaker, then they “heat up” the situation and “up the ante,” literally trump the commitment made before. On the other hand, speakers calm the situation when their conversational contribution is produced with lower speech volume, pitch, and speech rates (Roth & Tobin, 2006: 29)

La intensificación de la prosodia, caracterizada por un tono más alto, mayor intensidad y velocidad del habla, contribuye a caldear la situación, término usado por Roth y Tobin (2006: 40), proporcionando recursos para expresar emociones. En el contexto de conflictos verbales, la prosodia revela aún más sobre las dinámicas de poder. En este sentido, la prosodia no se limita a ser un mero componente del discurso; más bien, Couper-Kuhlen y Selting (1996) la conceptualizan como un elemento organizador clave en la interacción social. Entendido el conflicto como la existencia de un desacuerdo mantenido a lo largo de varios turnos conversacionales (Muntigl y Turnbull, 1998; Clancy 2018; Briz, en prensa 2023; Estellés, en prensa), algunos autores han manifestado también el papel del alineamiento o desalineamiento prosódico en relación con la prosodia manifestada en los turnos de habla anteriores (véanse Szczepek Reed, 2010; Wichmann, 2012, entre otros). La teoría de análisis conversacional sugiere que las características prosódicas, junto con elementos contextuales como sintaxis, semántica léxica y kinésica, contribuyen a la construcción e interpretación de unidades de construcción de turno y turnos de habla. Wichmann (2011) destaca que, en discusiones entre participantes con relaciones de poder simétricas, el aumento del tono puede coexistir con una base de cooperación subyacente. Sin embargo, en relaciones de poder asimétricas, la prosodia puede indicar un cambio significativo, donde el participante más poderoso habla en un modo de subordinación frente a otro interlocutor (Wichmann, 2012: 343).

La interpretación de la prosodia presenta desafíos adicionales. Szczepek (2010) advierte de que la atribución de emociones basada en la prosodia puede ser engañosa, ya que la interpretación de los participantes no siempre refleja con precisión las emociones transmitidas, un punto de vista que se retomará más abajo.

3 CUESTIONES METODOLÓGICAS DEL CORPUS

Este estudio supone la exploración de la primera muestra conversacional del corpus ESPRINT, coordinado por Marta Albelda y Maria Estellés, centrado en interacciones conflictivas entre íntimos y conocidos en diversas ciudades españolas. Específicamente, el subcorpus empleado es el ESPRINT-conversación VLC (Valencia), que presenta la singularidad de haber sido grabado, transcrito y etiquetado por la misma persona (la hablante C en las transcripciones), por lo que se proporciona información de primera mano sobre los segmentos percibidos como conflictivos por los participantes.

Tanto este subcorpus como otros en el corpus ESPRINT-conversación documentan interacciones entre íntimos y conocidos seleccionadas porque los implicados reportan tener frecuentes conflictos. Estos conflictos, categorizados como intratables (Coleman, 2000) se caracterizan por ser intensos, persistentes en el tiempo y centrados en

necesidades o valores fundamentales, afectando múltiples aspectos de la vida de los contendientes. Coleman (2000) identifica que estas situaciones de conflicto suelen originarse en diferencias morales irreconciliables, problemas de distribución de alto riesgo y relaciones de poder asimétricas, aspectos todos ellos presentes en el subcorpus estudiado, cuyo foco conflictivo gira fundamentalmente en torno a la relación entre A y B. Es importante señalar que, debido a las limitaciones en la representatividad de la muestra del subcorpus ESPRINT-conversación VLC, tanto en términos de cantidad de hablantes (solo 4) como de diversidad de sexo (solo un hombre y tres mujeres) y franjas de edad (dos franjas etarias), este estudio adopta una perspectiva que combina métodos de visualización exploratoria con técnicas de estadística multivariante, aunque estos solo tienen validez sobre la muestra recogida en el corpus ESPRINT-conversación VLC (Valencia). Se trata, por tanto, de un enfoque de aproximación al objeto de estudio, aunque los resultados que aquí se detectan podrán posteriormente certificarse en otros datos de distinta procedencia a medida que se procesen los datos del corpus recogidos en otras ciudades. A pesar de estas limitaciones, el enfoque metodológico propuesto proporciona un método de análisis diseñado y pensado para comprender las dinámicas comunicativas del grupo de hablantes específicos del corpus ESPRINT-conversación VLC. Junto a la exploración cualitativa de los datos recopilados, se presenta también una manera de extrapolar y adaptar este estudio a muestras más amplias y diversificadas en futuras investigaciones que enriquezcan así la comprensión del conflicto como fenómeno global dentro de la comunidad de hablantes española.

3.1 *Muestra de análisis y hablantes*

El subcorpus de análisis, ESPRINT-conversación VLC, comprende originalmente trece conversaciones con un total de 221 minutos, es decir, algo más de tres horas y media. Los hablantes que participan en las grabaciones son miembros de una misma familia, seleccionada por reportar la presencia de un conflicto intratable y cuyos miembros aceptaron voluntariamente ser grabados. La familia está compuesta por cuatro hablantes (A, B, C y D). La madre, A, una mujer de 61 años, con nivel de instrucción alto y bilingüe en castellano y valenciano. El hablante B, padre, es un varón de 61 años con nivel de instrucción alto y bilingüe pasivo. Por otro lado, hay dos hijas: D, mujer de 25 años de nivel de instrucción alto, y C, mujer de 22 años, con nivel de instrucción alto, que ejerce además como observadora participante, grabadora y etiquetadora de la muestra.

Como parte del procedimiento legal, se ha solicitado el consentimiento formal de los participantes en tres fases (Uclés y Carcelén, 2019; Carcelén 2023; Carcelén, en prensa): en primer lugar, los sujetos aceptan previamente ser grabados; se les detalla la metodología propuesta y el interés lingüístico del proyecto (no se detalla el interés particular en el conflicto para no sesgar las interacciones), y se les explica que uno de los miembros de la familia procederá a grabar de manera discreta situaciones conflictivas en el hogar mediante dispositivos móviles o grabadoras espía, siempre en momentos en los que no sean conscientes de ello. Así pues, los hablantes aceptan ser grabados, pero sin saber el momento concreto en el que acontecerá la grabación. En segundo lugar, una vez realizadas las grabaciones, se solicita la segunda firma; por medio de esta, autorizan la utilización de los audios para fines investigadores, previa anonimización del audio y de los datos personales sensibles (nombres, calles, menciones explícitas a la profesión, etc.). Con esta segunda firma, se ofrece a los sujetos la posibilidad de escuchar las grabaciones y, si lo consideran oportuno, retirar todas o parte de ellas en cualquier momento (en este caso,

no se vetó ninguna). Por último, se firma un tercer consentimiento sobre la política europea de protección de datos, de acuerdo con la normativa vigente.

De las conversaciones originales, se han procesado para el análisis fónico las conversaciones c1, c4, c6, c10, c11, c12 y c13, que presentan la mayor calidad de audio y suman un total de 56 minutos de grabación. De estas, para la visualización, se han seleccionado c4 y c13 por dos razones fundamentales: son relativamente extensas (suman un total de 28 minutos de grabación) y representan dos situaciones discursivas diferentes, una en la que participa casi exclusivamente el matrimonio y otra en la que participan los cuatro hablantes. Todos los audios del corpus fueron divididos en grupos entonativos, transcritos y anotados mediante el programa ELAN (Max Planck Institute, 2022). Aunque la división en grupos entonativos no la llevó a cabo automáticamente PRAAT, sino que los delimitó manualmente la persona transcriptor, esta recibió la instrucción de seleccionar cada grupo de habla ubicado entre pausas o entre inflexiones marcadas de la frecuencia fundamental (Quilis et al., 1993).

3.2 *Técnica del observador participante*

Como se detalla más arriba, en el contexto específico abordado, la persona encargada de la grabación, la hablante C, estuvo presente durante los intercambios, seleccionó las grabaciones, las transcribió y etiquetó mediante ELAN los fragmentos que percibió como conflictivos.

La utilización de la técnica del observador participante reduce la interpretación por parte de investigadores ajenos a la interacción original. Este enfoque, alineado con las propuestas, entre otros, de Watts (2003), Haugh (2007) y Zhu y Boxer (2013), resulta especialmente valioso en el análisis de secuencias conflictivas. No obstante, este acercamiento no carece de inconvenientes; como se señalaba arriba, Szczepek (2010) señala lo confuso que puede ser evaluar las emociones experimentadas por otros interlocutores, incluso en casos en los que analista e interlocutores hayan compartido espacio discursivo inmediato:

even when participants themselves tell each other that they heard a previous prosody as displaying a certain emotion ('sadness'), we cannot automatically assume that therefore the prosody did indeed display that emotion, let alone make any assumptions regarding the speaker's inner state. Both aspects—participants' use of displays of emotions, and their use of displayed interpretations of the emotional states of others—are reminders to be careful when identifying emotions in talk-in-interaction. (Szczepek: 2010)

En cualquier caso, en el corpus analizado, la hablante no solo comparte espacio con los hablantes, sino que se ve involucrada como parte del conflicto en varias ocasiones. Un estudio detallado de las partes en que la participante etiqueta el conflicto proporciona información valiosa sobre la propia naturaleza del conflicto y desvela (Estellés, en prensa) que una estructura en la que se encuentran turnos despreferidos que se suceden a lo largo de varios turnos conversacionales, incluso con marcas lingüísticas emocionales, puede no considerarse un conflicto a ojos de un participante dependiendo de quiénes son los implicados o de cuál es el foco del conflicto en un momento dado. Así, exacta o no, la identificación del conflicto se basa en un testimonio de primera mano que incorpora su visión y que será, en este caso, la utilizada para delimitar los fragmentos conflictivos del estudio.

Si bien únicamente se ha tenido en cuenta el etiquetado manual de los fragmentos conflictivos, y de acuerdo con las diferencias entre el etiquetado de la participante y del analista (véase Estellés, en prensa), se acudirá a la interpretación de este último a la hora de explicar algunos fenómenos fónicos revelados en las visualizaciones derivados de los problemas a la hora de delimitar el conflicto o de 'errores' en el etiquetado de la transcriptor.

3.3 Variables de análisis y categorías

En los datos proporcionados, se presenta información sobre ciertos aspectos relacionados con situaciones de conflicto y no conflicto, organizados en diferentes categorías. El conjunto de datos se organiza en torno a la presencia o no de conflicto y la clasificación en diferentes hablantes o tiers, identificados como A, B, C y D. En todo el corpus ESPRINT_conversación VLC, de acuerdo con la catalogación de la participante, se registran 799 instancias de conflicto y 359 instancias de no conflicto. Como se ha explicado anteriormente, las situaciones de conflicto y no conflicto de este trabajo se han obtenido de dos fuentes o conversaciones, representadas por las categorías c13, en la que se registran 308 grupos entonativos, y c4, con 259 grupos entonativos.

Junto a estos datos y metadatos, proporcionados por la persona encargada de transcribir y etiquetar, se consideran los datos procedentes del análisis fónico automático, realizado mediante la herramienta Oralstats (Cabedo, 2022). Para un análisis descriptivo, se muestra el histograma del subcorpus completo en lo que respecta a tono, inflexión tonal, intensidad, duración y velocidad de habla (Sección 4). Sin embargo, a efectos ilustrativos de las posibilidades de representación gráfica, para las herramientas de visualización se ha considerado únicamente la variable pitch (tono) en las dos conversaciones mencionadas, c4 y c13. Esta variable pitch refleja el tono del discurso y se mide en semitonos para salvar las diferencias entre las voces femeninas y masculinas (para un estudio completo con las demás variables fónicas, cf. Cabedo, en prensa). En última instancia, se mide el Z-Score del pitch, dado que esta medición permite observar cuán lejos está cada registro en relación con la media de pitch correspondiente del hablante y, de nuevo, este estudio podría extenderse al de la intensidad, la duración o la velocidad de habla, por ejemplo. A continuación, se ofrece con mayor detalle una descripción de las pruebas estadísticas y los mecanismos de visualización empleados.

3.4 Técnicas estadísticas y de visualización aplicadas

En esta sección, se presentan las pruebas estadísticas empleadas para describir los datos, y se explican las técnicas de visualización que desempeñan un papel relevante en el análisis de las intervenciones del corpus ESPRINT_Conversación VLC. En concreto, se examinan detenidamente dos enfoques: Z-Score, visualizado con línea temporal (3.4.1) y diagrama de red (3.4.2).

3.4.1 Z-Score y línea temporal

El Z-score, o puntuación estándar, es una medida estadística que cuantifica la posición relativa de un dato dentro de un conjunto de datos, expresada en términos de desviaciones estándar desde la media (Field et al., 2012). En otras palabras, el Z-score indica cuántas desviaciones estándar tiene un dato particular por encima o por debajo de la media de una variable. El Z-score es útil para comparar y analizar datos que se encuentran en diferentes

escalas o distribuciones. Un Z-score positivo indica que el valor está por encima de la media, mientras que un Z-score negativo indica que está por debajo de la media.

Para visualizar la variación de los Z-Scores de las variables fónicas, por ejemplo, de pitch a lo largo del tiempo, considerando la interacción entre el conflicto y el hablante, se emplea ggplot2 (Wickham, 2016) en R (R Core Team, 2023). Se calculan los Z-Scores para normalizar los datos y luego se crea un gráfico de dispersión en el que el eje Y representa los Z-Scores de pitch, el eje X refleja el tiempo, y el color de los puntos indica la interacción entre conflicto o no conflicto y los distintos hablantes.

En esta investigación, este enfoque permite visualizar de manera efectiva la distribución temporal de los Z-Scores de pitch en una conversación dada (recuérdese que se ejemplificará con c4 y c13) y cómo esta variación se relaciona con la interacción entre el conflicto y el hablante, es decir, la línea temporal permite observar cómo grupos entonativos superiores o inferiores a la media del hablante se distribuyen a lo largo del tiempo y, según el color, se permite detectar en qué zonas de la conversación ha habido más conflicto (o menos) y cuál de los hablantes ha sido su autor. Puede observarse en la Sección 5.2.

3.4.2 Diagrama de red

La utilización de diagramas de redes en el análisis de conversaciones proporciona una perspectiva novedosa para observar la complejidad de las interacciones entre hablantes. Este enfoque de análisis ofrece una representación visual global de las conexiones y patrones emergentes en la red de interacciones verbales.

La ventaja de emplear diagramas de redes radica en su capacidad de visualizar relaciones complejas. Las conversaciones a menudo involucran múltiples participantes que interactúan simultáneamente. Los diagramas de redes permiten visualizar de manera clara y comprensible las conexiones entre todos los hablantes, revelando la estructura completa de la interacción.

At an abstract level, a network refers to various structures comprising variables, which are represented by nodes, and the relationships (formally called edges) between these nodes. For example, from the Foresight Report the variables such as stress, peer pressure, functional fitness, nutritional quality of food and drink represent nodes in the network, and the positive and negative relationships between those nodes are edges. There are some differences in nomenclature in the network literature: nodes are sometimes referred to as vertices, edges are sometimes referred to as links, and networks are also called graphs (Hevey, 2018: 304-305)

Esta representación facilita la identificación de roles específicos y la comprensión de la posición relativa de cada hablante en la red. La centralidad de un nodo en la red puede indicar su importancia y grado de influencia en la conversación. Los nodos más centrales pueden ser aquellos que reciben y emiten más interacciones, de manera que se plasma gráficamente su papel crucial en la dinámica discursiva.

Por ello, los diagramas de redes pueden ofrecer una herramienta visual que facilita el análisis de la particular complejidad de las interacciones conflictivas en conversación coloquial. Desde la identificación de roles hasta la visualización de patrones emergentes, este enfoque proporciona perspectivas útiles para entender la estructura y dinámica de las interacciones comunicativas.

4 RESULTADOS. DISTRIBUCIÓN DE LAS VARIABLES

De acuerdo con lo planteado en 3.3, esta sección presenta la exploración de las variables de estudio, tanto nominales como numéricas, y su distribución en las siete conversaciones

del corpus ESPRINT_conversación VLC seleccionadas por sus características acústicas y estructurales (c1, c4, c6, c10, c11, c12 y c13). Cada subepígrafe se dedicará a analizar estos dos tipos de variables. En las nominales, se desglosan patrones específicos de manera general y se exploran posibles relaciones entre categorías. En las numéricas, por su parte, se comenta la distribución de cada variable a partir de histogramas y de sus valores de asimetría y curtosis.

4.1 Variables nominales

El gráfico de la Figura 2 presenta gráficos de barras con las frecuencias absolutas de las variables nominales para este estudio, conflicto, hablante y conversación, en las

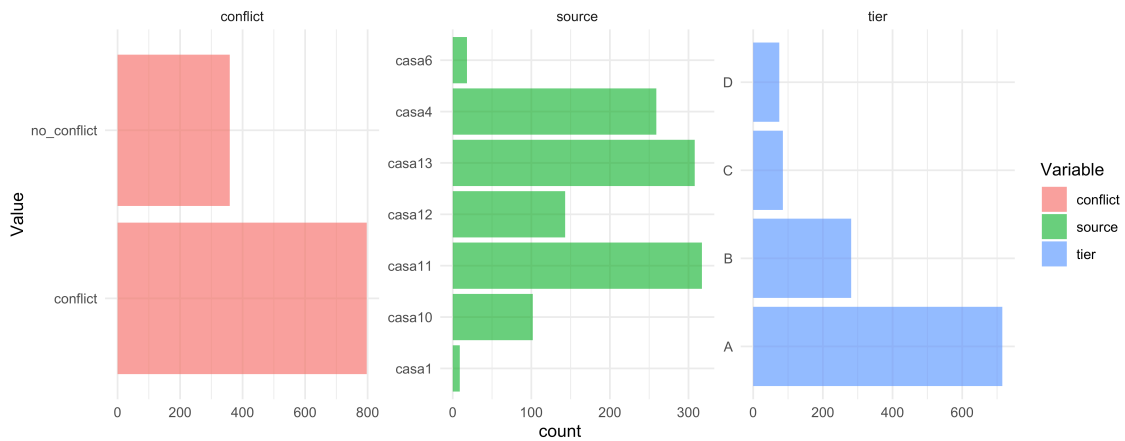


Figura 2. Distribución de las variables categóricas analizadas

El análisis de las variables nominales revela información importante sobre la distribución de las categorías en el conjunto de datos. En cuanto a la conflictividad de los enunciados en la muestra, siempre con el etiquetado original de la participante, se observa que cerca de 800 enunciados son clasificados como conflictivos, frente a 359 enunciados no conflictivos. En cuanto a la variable conversación, los registros provienen principalmente de c11 (317 observaciones), c13 (308 observaciones), y c4 (259 observaciones), mientras que otras presentan menor representación, como la c6 con solo 19 observaciones, datos esperables teniendo en cuenta que, como se explicaba más arriba, se trata de las conversaciones más extensas.

La diferencia más notable se produce en la variable hablante, donde se observa cómo A es la interlocutora que ocupa mayor tiempo de posesión de la palabra, con más de 600 grupos entonativos, seguida de lejos por B, con unos 250 grupos entonativos. La participación de C y D es bastante similar, con unos 100 grupos.

4.2 Variables fónicas

El gráfico de la Figura 1 presenta el histograma de las cinco variables numéricas, de base fónica, que se consideran en este estudio: tono, inflexión tonal, intensidad, duración y velocidad de habla.

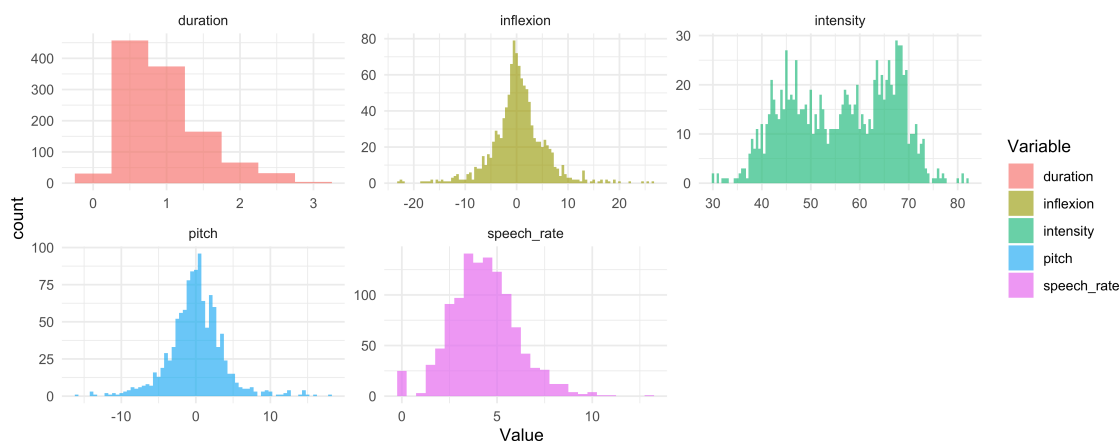


Figura 1. Distribución de las variables numéricas analizadas

Las distribuciones de las variables analizadas revelan características distintivas. En términos de tono de voz (pitch), se observa una asimetría cercana a cero (0.22), indicando relativa simetría, mientras que la curtosis positiva (3.08) sugiere colas más densas y una distribución más puntiaguda, con valores más cercanos a la media y con desviaciones estándares menos amplias. Para la intensidad, la asimetría negativa (-0.58) señala una inclinación hacia valores más bajos, mientras que la curtosis cercana a cero (0.58) indica una forma similar a la distribución normal. En relación con la duración del discurso, la asimetría positiva (1.44) indica una inclinación hacia valores más altos, y la curtosis positiva (3.01) sugiere colas más pesadas y una distribución más puntiaguda, similar a la variable tono. Por último, la velocidad de habla exhibe una asimetría positiva (0.69) que indica una inclinación hacia valores más altos, junto con una curtosis positiva (3.15). Generalmente, a excepción de las variables intensidad y duración, que presentan una variación mayor, se trata de variables con una distribución más o menos normal.

5 MÉTODO DE ANÁLISIS Y VISUALIZACIÓN DE LOS RESULTADOS: LA RADIOGRAFÍA DE UNA CONVERSACIÓN CONFLICTIVA

Tal como se ha planteado en los apartados anteriores, a continuación, se presentan dos métodos de visualización que, tomando como base los datos cuantitativos del pitch y los cualitativos de hablante, conversación y presencia/ausencia de conflicto, proporcionan una radiografía de la interacción en la que se pueden observar las dinámicas fónicas y de toma de turno. Esta visualización proporciona información (a) acerca de los picos fónicos de la conversación, que pueden cotejarse con la transcripción para ver qué sucede exactamente en el fragmento, y (b) acerca de los roles adoptados por los participantes y el modo en que estos cambian tanto en conflicto vs. no conflicto como dependiendo del interlocutor al que se dirijan.

Para obtener la información de (a), el apartado 5.1 presenta el sistema de visualización del desvío prosódico por conversación, mientras que 5.2 contribuye al objetivo de (b) mediante la presentación de la visualización por diagramas de red. Dado que se trata de fotografiar e interpretar una conversación concreta, como se ha detallado arriba, en estos apartados se emplearán las conversaciones c4 y c13.

5.1 Visualización del desvío prosódico por conversación

El primer sistema de visualización empleado establece la relación temporal entre variables tonales y las interacciones entre conflictos y niveles de hablante en un conjunto de datos. La estética del gráfico se define con ejes X e Y que representan, respectivamente, time_start (tiempo de inicio en el que suceden los grupos entonativos) y pitch, y la forma de puntos y segmentos determinada por la interacción de conflict y tier. Los segmentos representan la relación entre time_start y time_end respecto a la media de pitch. Los puntos, con tamaño y color según la interacción entre conflict y tier, permiten visualizar cómo estas variables afectan los valores de pitch. Este gráfico detallado facilita la identificación de patrones temporales y variaciones tonales en la investigación, de modo que permite ver en qué momentos de la conversación se están produciendo comportamientos prosódicos alejados de la tónica general de los hablantes.

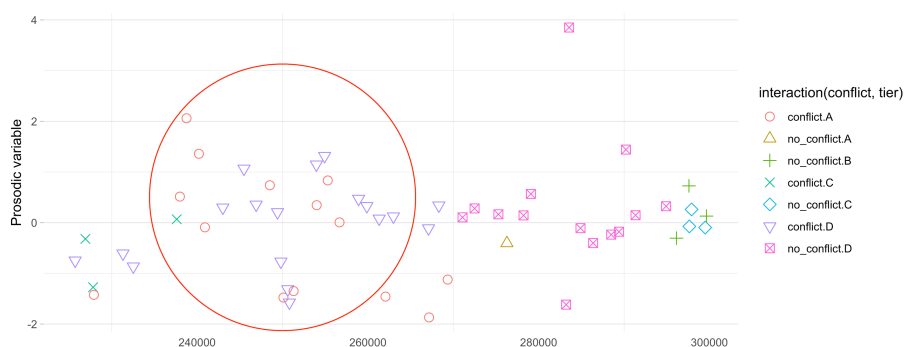


Figura 8. Gráfico de los valores de tono de los grupos entonativos dispuestos a lo largo del eje temporal en la conversación 13. Un círculo se ha sobrepuesto para remarcar una zona de conflicto y desvío tonal

tier	text	pitch	pitch_z_score	conflict
D	crujientísima<alargamiento/>	-2.06	-0.75	conflict
C	sí sí>	-0.75	-0.32	conflict
C	en [su punto]	-3.68	-1.27	conflict
A	[porque no se ha] hecho como lo he dicho yo	-4.13	-1.42	conflict
D	pues la próxima vez lo haces tú	-1.64	-0.61	conflict
D	si sabes hacerlo	-2.42	-0.86	conflict
C	[<risas/>]	0.43	0.07	conflict
A	[estoy bastante]	1.80	0.51	conflict
A	[<entre risas>har]ta<alargamiento/></entre risas>	6.54	2.06	conflict
A	harta	4.39	1.36	conflict
A	de que me contestes como me contestas	-0.06	-0.09	conflict
D	no pero es que <fsr t="amá">mamá<alargamiento/></fsr> tú tampoco te das cuenta de cómo hablas	1.14	0.30	conflict
D	y estás menospreciando<alargamiento/>	3.49	1.07	conflict
D	que una persona te ha hecho la cena	1.32	0.36	conflict
A	yo no estoy menosprecian[do<alargamiento/>]	2.49	0.74	conflict
D	[sí<alargamiento/>	0.86	0.21	conflict
D	sí<alargamiento/>	-2.13	-0.77	conflict
A	no<alargamiento/>	-4.30	-1.48	conflict
D	¡sí!	-3.80	-1.31	conflict
D	vale	-4.60	-1.58	conflict
A	no	-3.91	-1.35	conflict
D	[<ininteligible/>]	3.74	1.15	conflict
A	[si a] mí me preguntan	1.28	0.35	conflict

D	hay que poner [(cada cosa en su <ininteligible/>)]	4.26	1.32	conflict
A	[de cómo lo hago<alargamiento/>]	2.78	0.83	conflict
A	lo que no puedes hacer es como te da la gana	0.24	0.01	conflict
D	jno! como te da la gana no	1.67	0.47	conflict
D	como una persona sabe	1.23	0.33	conflict
D	que es diferente	0.47	0.08	conflict
A	vale	-4.24	-1.46	conflict
D	porque una persona intencionadamente no te está haciendo las cosas porque le da la gana mal	0.60	0.12	conflict
D	que [no están] como tú quieres	-0.12	-0.11	conflict
A	[<susurro>vale<alargamiento/></susurro>]	-5.50	-1.87	conflict
D	es algo diferente<alargamiento/>	1.26	0.34	conflict
A	vale vale	-3.21	-1.12	conflict
D	si tú quieres que esté más crujiente	0.55	0.10	no_conflict
D	pues la próxima vez que tú sabes hacerlo hazlo tú<alargamiento/>	1.09	0.28	no_conflict
D	es lo que te he dicho	0.73	0.16	no_conflict
A	uhum	-1.01	-0.40	no_conflict
D	si tú lo sabes hacer mejor	0.67	0.14	no_conflict
D	o más crujiente pues hazlo tú	1.96	0.57	no_conflict

Tabla 6. Selección de grupos entonativos de la conversación 13 junto con los valores de tono y su correspondiente Z-Score

En este fragmento de la c13 se observan datos que pueden estar relacionados con la evolución de una situación de conflicto, sobre todo entre A (madre) y D (hija). Toda la familia se está preparando para cenar sardinas rebozadas, que ha preparado B. Las sardinas resultan estar mal rebozadas, poco crujientes y demasiado saladas para el gusto de A, que le recrimina a B que no haya seguido sus instrucciones. D, la hija, afea abiertamente a A su manera de hablar y su modo de recriminar las acciones de quienes, con su mejor intención, le han preparado la cena.

De acuerdo con la tabla 6, los grupos entonativos con valores más altos de tono corresponden a secuencias paralingüísticas (risas) o a solapamientos, por lo que, al menos las primeras, no pueden tenerse en consideración desde un punto de vista acústico. No obstante, dos de los enunciados con valores altos muestran formas de conflicto directas, expresiones del estado de ánimo y menciones metapragmáticas al modo en que se han percibido las intervenciones de la otra: el segundo caso de harta (eliminado el primer harta que se produce en un solapamiento con unas risas) pronunciado por A, y y estás menospreciando<alargamiento/>, pronunciado por D.

Por su parte, también son significativas, y más abundantes, las secuencias breves de respuesta con un tono muy bajo, inferiores a -3 semitonos. Principalmente las emite la madre y son los casos de vale, vale vale, o no. A ellos se une una secuencia en forma de queja, [porque no se ha] hecho como lo he dicho yo, que se ve parcialmente afectada por un solapamiento. Finalmente, el segmento conflictivo termina con una acumulación de secuencias de D, continuadas por C y mínimamente por B, en las que no interviene A. Son secuencias etiquetadas como no conflictivas, dado que no hay respuesta despreferida de ninguno de los interlocutores y, aparentemente, la hablante A, que planteaba la situación de intercambio conflictivo, decide callarse durante esta porción temporal.

Nótese, sin embargo, que los límites del conflicto y no conflicto, tal como aparecen etiquetados por la participante, son confusos: no hay diferencias estructurales ni temáticas entre los pares D1-A1 y D2-A2, marcadas como conflicto, y D3-A3 y D4, no marcadas como tal:

(1)

D1: porque una persona intencionadamente no te está haciendo las cosas porque le da la gana mal/
que [no están] como tú quieres=
A1: [ˈvaleeː]
D2:= es algo diferentee
A2: vale vale
D3: **si tú quieres que esté más crujiente/ pues la próxima vez que tú sabes hacerlo hazlo túu/ es
lo que te he dicho**
A3: uhum
D4: **si tú lo sabes hacer mejor/ o más crujiente pues hazlo tú/ es que/ no sée/ parece que estás
esperando a ver que falle/ paraa recriminar // si sabes hacerlo/ pos ¡ayuda! /yo no sé
hacerlo/ a mí no me lo- no me lo- no me lo mandes hacer si no sé hacerlo/ me va a salir
peor que eso**

Justamente, el fragmento que aparece resaltado en negrita no fue marcado como conflicto pero responde exactamente a la misma estructura que los anteriores: reproches de D, seguidos por un ‘vale’ en tono bajo de A. Por tanto, desde un punto de vista externo podrían considerarse como un conflicto. La visualización de la figura 8, recuperada aquí como 9, permite localizar este fragmento concreto en la línea temporal y comprobar cómo el fragmento en negrita (marcado con un círculo con trazo intermitente) tampoco difiere fónicamente en exceso del marcado como conflicto; es más, aquí aparece uno de los picos más altos de pitch de toda la conversación, marcado con una flecha, que corresponde con el grupo entonativo *¡yo no sé hacerlo!* emitido por D.

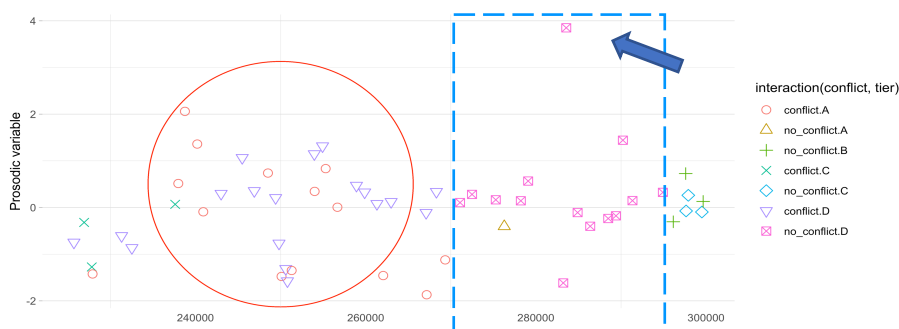


Figura 9. Gráfico de los valores de tono de los grupos entonativos dispuestos a lo largo del eje temporal en la conversación 13. Un círculo marca la zona de conflicto y desvío tonal, y un rectángulo discontinuo marca el fragmento supuestamente conflictivo pero no etiquetado.

Compárese ese comportamiento fónico con el de la cola del gráfico, marcada con un triángulo, en que B y C departen de manera más distendida sobre la sal de la comida y B admite su error sin problemas:

(2)

B: mucha sal ¿eeh/[<ininteligible/>]?
C: apá/ t'has'pasao en sal en todo/ [en esto y en eso]
B: [ahí noo]/pero en esto sí/ [no sé por quée]

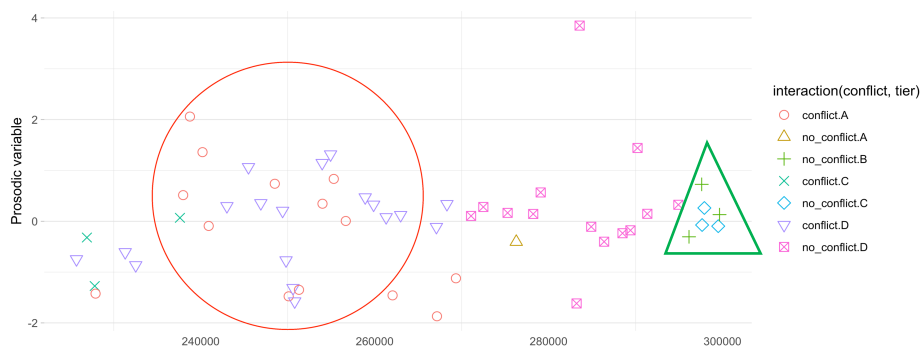


Figura 10. Gráfico de los valores de tono de los grupos entonativos dispuestos a lo largo del eje temporal en la conversación 13. Un círculo marca la zona de conflicto y desvío tonal, y un triángulo marca el fragmento no conflictivo.

En la conversación c4, que sucede fundamentalmente entre A y B, puede observarse cómo el gráfico de valores muestra diferencias en los fragmentos conflictivos y los no conflictivos:

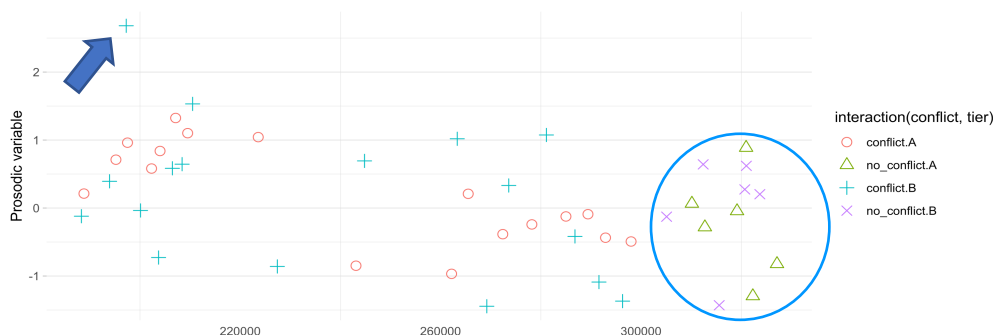


Figura 11. Gráfico de los valores de tono de los grupos entonativos dispuestos a lo largo del eje temporal en la conversación 4. Un círculo marca la zona de no conflicto.

En el primer fragmento de la figura, entre los 180 000 y los 240 000 ms, se produce una acumulación de valores de *pitch* altos, entre ellos el pico indicado con una flecha, en que B pronuncia el grupo entonativo marcado en negrita, en una secuencia estructuralmente muy marcada, con muchas repeticiones (*¡y ya está! / ¡ya está! / ¡y ya está!; te estoy diciendoo ¡y ya está!; ¡pues ya está!*), insultos (un poco *maleduca*), etc. En el fragmento, A trata de dar una información a B sobre los festivos de ese año, pero B responde con displicencia que ya lo sabe y A se ofende porque B nunca la escucha ni da fiabilidad a sus palabras:

(3)

- A: yo digo simplemente que te lo mandé [para que]=
 B: [¡ya estamos!]
 A:= vieras que era todo el [mundo]
 B: [¡hombre! ¿te djas cuenta que no contestoo?/ pues me parece perfectoo/me lo han mandado
 A: [pues sí] no contestas es que eres un poco=
 B: [¡y ya está! / ¡ya está! / **¡y ya es[ta]!**
 A:= [maleduca / ¡po]rque te estoy hablan[do]

B: [te es]toy diciendoo/ ¡y ya
estáa// ¿quieres que te vuelva a repetir lo mismoo?
A: noo
B: ¡pues ya estáa!

En cambio, el fragmento final, señalado con un triángulo, los valores de ambos hablantes se encuentran más centrados (solo dos grupos entonativos quedan por debajo). Esa configuración de *pitch* corresponde con la parte en que A y B abandonan progresivamente el tema que ha provocado el conflictivo: dejan de hablar de quién lleva la razón respecto a los días festivos y pasan a lamentarse de que han cogido días libres y no van a poder viajar por las restricciones de movilidad del COVID; A sugiere a B que viajen por la Comunidad Valenciana, donde sí está permitido.

(4)

B: de poco ha servidoo
A: ¿el quéé?
B: que de poco ha servido/ pedimos ell/pedirme yo el tres y el cuatro
A: yo el tres no me lo he pedidoo
B: ¡yo síi!
A: ¡hombree!/ siempre se puede viajar por la comunidad ¿noo?
B: yaa/ síi
A: Valencia Castellón y [Alicante sí que te puedes moveer]
B [has]ta el puente de Contreras se puede viajar
(...)
A: y Alicante hasta abajo/ también te puedes mover hasta Murcia ¿no?
B: °<ininteligible/>°
A: [o Albace]te
B: [°<ininteligible>°]
A: po' ya está
B: mm
A: habrá cosas que no conozcaas/ y que sean interesantes de ver

El método de visualización con eje temporal, en síntesis, permite localizar puntos calientes en lo que respecta a la manifestación de un parámetro, en este caso el *pitch*. No obstante, atender solamente al *pitch* puede conducir a interpretaciones erróneas o incompletas, puesto que no todos los picos de *pitch* corresponden con conflicto. En el caso de B, por ejemplo, se registra *pitch* muy elevado en los fragmentos joviales, jocosos, en las despedidas antes de irse a dormir, etc. Esta capa de *pitch* se puede superponer a otras de intensidad, velocidad, etc., de modo que se puedan delimitar cada vez con más precisión las zonas de acumulación de recursos fónicos prominentes y relacionarlos con la parte de la conversación que está teniendo lugar en ese momento.

5.2 Configuración prosódica y dirección de hablante

La relación entre conflicto y prosodia puede observarse en ocasiones en desalineaciones fónicas, es decir, casos en los que hay intercambios discursivos donde el hablante previo utiliza un valor fónico concreto (tono alto, por ejemplo) y el siguiente hablante, al tomar el turno de habla, utiliza un valor fónico distinto, no necesariamente contrario, pero sí modificado sustancialmente (por ejemplo, un tono menor al anterior Roth & Tobin, 2010; Szczepek Reed, 2010; Wichmann, 2012).

Las conversaciones seleccionadas para la visualización se caracterizan, como se ha señalado arriba, por la distinta participación de los hablantes. En c4, salvo en un fragmento breve al final, la conversación se desarrolla exclusivamente entre A y B, y en c13 participan los cuatro hablantes de manera regular en toda la interacción. En el análisis de los grupos entonativos emitidos por los participantes, se observa cómo estos pueden entenderse a

nivel monológico y dialógico. A nivel monológico, en ocasiones un mismo hablante mantiene el turno durante amplios periodos de tiempo, lo que implica la consecución de largos fragmentos monológicos con varios grupos entonativos pertenecientes al mismo hablante. Como se apuntaba en el análisis cuantitativo previo, la tendencia del corpus general es la predominancia discursiva de la hablante A, que se confirma en c4 y c13. En la tabla 7 de frecuencias y medias de tono, que refleja los resultados de la conversación 13, puede observarse que las secuencias de mayor continuidad del turno son las de A, acompañadas de una media tonal neutra de 0.54 semitonos. El resto de hablantes, a excepción de D con solo cuatro grupos entonativos consecutivos, no mantiene el turno de manera continuada y simplemente responde a turnos anteriores con intervenciones breves de uno o dos grupos entonativos.

A nivel dialógico, los grupos entonativos se agrupan en turnos que se alternan dinámicamente, de modo que un hablante cede o pierde su turno de habla en relación con otros interlocutores. Para las conversaciones con más de dos hablantes, es interesante observar la dirección de esos intercambios y observar si existen casos más frecuentes que otros. Es decir, es interesante determinar cuáles son los hablantes que interactúan más frecuentemente entre sí en una conversación y cuáles son las características fónicas de estas interacciones, para ver si son diferentes de los comportamientos fónicos desplegados con los demás interlocutores. Por ejemplo, en los fragmentos en que la hablante A emite un grupo entonativo y el hablante B emite el siguiente, pueden observarse las características fónicas de estas unidades de habla (en este caso, el *pitch*) y ponerse en relación con el hablante anterior o con el hablante posterior, de manera que nos permite descubrir alineamiento o desalineamiento prosódico en este sentido.

from	tierto	frecuencia	pitch
B	A B	1	8.78
C	B D	1	5.72
A	C A	2	4.41
B	B A	2	3.92
D	D D	4	3.84
D	C B	1	2.38
A	D A	10	1.27
C	B C	2	1.25
D	A A	10	1.13
A	A D	10	0.84
A	A A	83	0.54
D	C A	1	0.43
B	A A	6	0.35
B	C A	1	0.31
B	B C	5	0.30
A	A C	2	0.02
A	A B	7	-0.11
C	B B	3	-0.16
B	C C	1	-0.17

from	tierto	frecuencia	pitch
	A A	1	-0.28
A	D D	13	-0.47
A	B A	5	-0.53
C	C C	1	-0.59
C	D A	1	-0.71
D	C C	1	-0.75
D	D A	11	-0.98
C	A C	1	-1.25
A	B B	3	-1.26
C	A D	2	-1.26
D	A D	12	-1.74
C	C A	2	-1.82
C	A A	3	-1.85
D	D C	2	-2.24
A	C D	1	-2.49
B	D C	1	-2.92
B	C B	5	-3.36
B	B B	2	-4.43

Tabla 7. Combinaciones de hablante anterior, hablante y hablante posterior y su correspondiente media de tono (semitonos) en un fragmento de c13.

La tabla 7, por ejemplo, muestra las secuencias de dos grupos entonativos consecutivos del mismo hablante seguidos de un grupo entonativo de otro hablante en un fragmento de

c13. En esta selección, por ejemplo, hay diez casos en los que A, después de emitir dos grupos entonativos, es seguida de D; y similares resultados en la dirección inversa, es decir, once situaciones de intercambio en las que dos grupos de D son seguidos por un grupo de A. Curiosamente, los intercambios similares entre B y A no son demasiado frecuentes, aunque llaman la atención los 6 intercambios en los que B emite un grupo y A emite los dos siguientes y 7 casos en los que B sigue a dos grupos entonativos de A. En ambos casos los valores tonales son neutros y apenas perceptibles. En esta última situación comunicativa, son llamativos los intercambios de A con D, con doce intercambios, y D con A, con diez.

La extracción de la dirección de los intercambios es automática: es la sucesión de hablantes la que determina que uno responde al otro. Por ejemplo, si se encuentra una secuencia A-A-A-B, se entiende que el turno compuesto por tres grupos entonativos va dirigido de A a B y que la de B es una respuesta a A. Por supuesto, este análisis automático no es infalible, ya que puede darse el caso de que, por ejemplo, la secuencia continúe como A-A-A-B-C, con lo cual, independientemente de las intenciones de A, el turno de B podría estar dirigido a C, etc. Sin embargo, las conversaciones c4 y c13 fueron sometidas a un análisis manual de la dirección de los intercambios y los resultados de la anotación automática fueron muy satisfactorios.

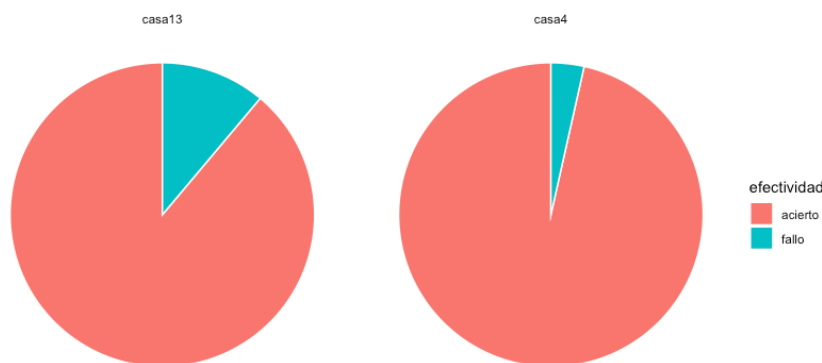


Figura 12. Porcentaje de aciertos y fallos del etiquetador automático en la c13 (izquierda) y c4 (derecha)

Por ejemplo, c13 cuenta con un total de 136 intervenciones conversacionales, de las que el sistema automático falló en 17, y c4 falló solo 4 de las 110, todas ellas en la parte final, en que participan los cuatro hablantes. Lógicamente, el acierto del sistema es completo para los intercambios entre dos hablantes, lo que lo hace especialmente apto para el análisis de conflictos de pareja.

De vuelta a la Tabla 7, en cuanto al nivel tonal asociado a los intercambios de c13, los valores más altos, superiores a 1.5 semitonos, son puntuales y no superan nunca los 5 casos. Los picos corresponden a un intercambio de A con B y a un intercambio de B con sus hijas, C y D. También C sube el tono cuando habla con su madre. En cuanto a los valores bajos, la tendencia es similar, aunque los intercambios de C con B son habitualmente los que manifiestan un tono muy bajo. Los datos de la Tabla 7 permiten establecer aproximaciones descriptivas a los datos: en primer lugar, la hablante A es una figura conversacional céntrica, dado que participa en gran parte de los intercambios discursivos con los otros hablantes y también usa amplios periodos de continuidad

discursiva, prácticamente secuencias monológicas, como se señalaba arriba. En c13, la hablante que tiene una participación importante es D, normalmente en interacciones que responden a A o que incluso toman y se apropian del turno con posterioridad. Como se ha comprobado en el apartado anterior, el foco del conflicto en c13 es precisamente la recriminación de D a A por haber menospreciado la cena preparada por B.

A continuación, la Figura 9 y la Figura 10 proporcionan una visualización de las relaciones entre los hablantes. Cada línea corresponde a un grupo entonativo y las flechas indican la dirección del intercambio, mientras que el color señala el nivel tonal utilizado (rojo, nivel alto, superior a 3 semitonos; azul, nivel bajo, inferior a -3 semitonos; verde, neutro). Para la catalogación entre niveles tonales altos y bajos se han calculado los semitonos observando la relación de los valores en Hz de cada grupo entonativo en relación con la mediana de cada hablante; al mismo tiempo, se ha considerado el doble de semitonos que según la bibliografía (Pamies Bertrán et al., 2002), son perceptibles por el oído humano y que se sitúan en 1.5 semitonos.

En la Figura 9, el diagrama de red correspondería a los casos de alineación o desalineación discursiva con los turnos anteriores, dado que son los casos que la bibliografía ha determinado como potenciales marcas del conflicto discursivo (Roth & Tobin, 2010; Szczepek Reed, 2010; Wichmann, 2012).

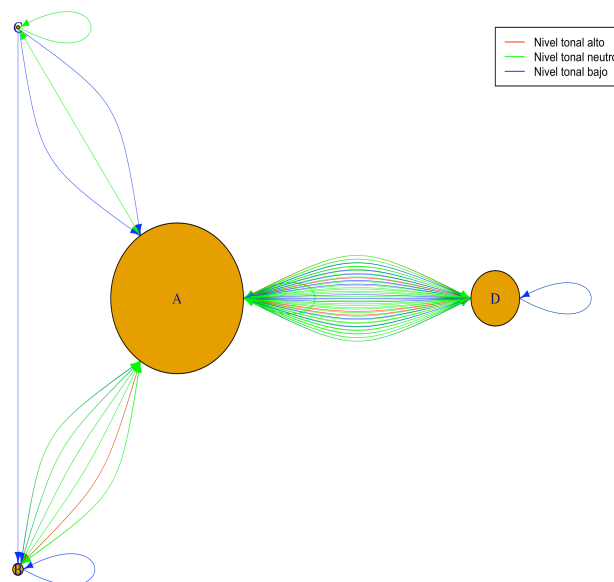


Figura 13. Diagrama de red que muestra quién habla en el turno precedente y el comportamiento tonal de los hablantes en fragmentos de conflicto en c13

Los nodos de A y D, es decir, los círculos que los rodean manifiestan el número de grupos entonativos, que son claramente superiores en el caso de A y que configuran a esta hablante como el centro de los intercambios discursivos: es la hablante que más habla cuando hay conflicto. Sin embargo, los intercambios que presentan una mayor frecuencia, los que acontecen entre A y D, son mayoritariamente neutros, aunque también se aprecian zonas de mayor o menor tonalidad. Existen además algunas flechas que salen y vuelven sobre el mismo hablante, que corresponden con los casos en los que hay continuidad discursiva, es decir, en los que el hablante del grupo entonativo anterior es el mismo.

En el gráfico, en cuanto a A, esa línea autodirigida se ve sobrepuesta por las líneas de los intercambios con D. Cuando B tiene a C como hablante anterior usa un tono más bajo, y no hay intercambios conflictivos entre D y B. Por su parte, en la relación discursiva de B con A hay proporcionalmente un número mayor de tonos bajos.

Compárese la figura 13 con la 14, que muestra el diagrama de red de los fragmentos no conflictivos de c13. El esquema ilustrado por esta figura, equilibrado en tono y en participación, serviría como ejemplo del desarrollo ordinario de una conversación a cuatro: se observa claramente como el papel preponderante de A ha desaparecido y cómo el reparto de tiempo de habla es notablemente más equitativo; de hecho, en este caso es C, el gancho, quien más participa y A es quien menos lo hace. El análisis del contenido conversacional parece sugerir que A se encuentra molesta por el conflicto con D, que le ha recriminado su comportamiento, y hasta cierto punto se retira o autoexcluye de las conversaciones triviales de los demás. El tono predominante es neutro (verde) y hay apenas 6 intercambios con tono alto, distribuidos entre todos los hablantes salvo B con D.

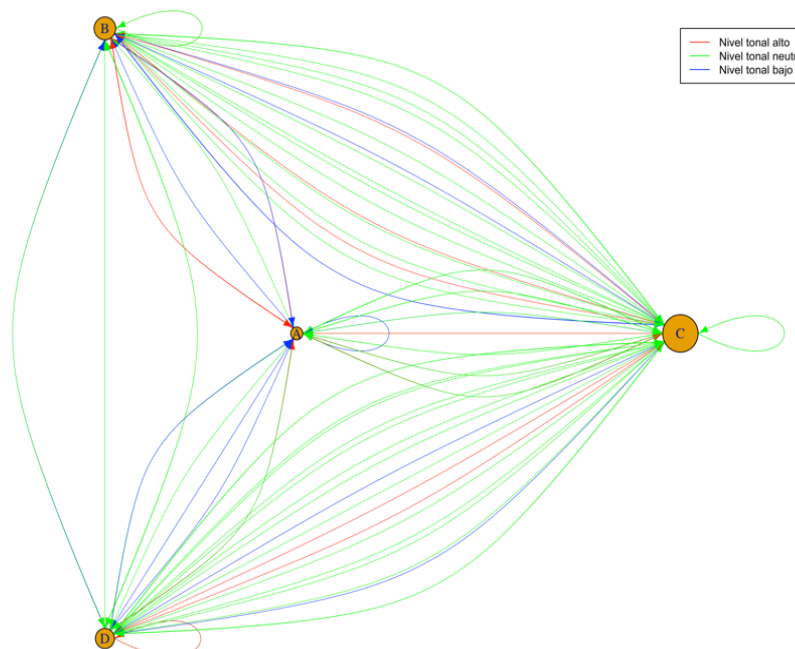


Figura 14. Diagrama de red que muestra quién habla en el turno precedente y el comportamiento tonal de los hablantes en fragmentos de **no** conflicto en c13

En la conversación 4, por su parte, se registran 124 grupos entonativos en fragmentos de conflicto y 135 de no conflicto. La Figura 15 muestra el mapa del conflicto en c4. En esta conversación, solo 10 de los 259 grupos entonativos registrados no son emitidos por A ni por B, lo cual explica que el mapa del conflicto sea una fotografía del intercambio entre estos dos hablantes:

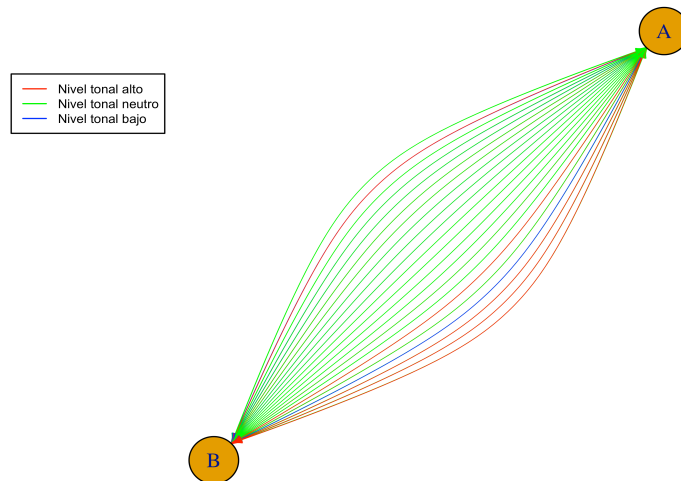


Figura 15. Diagrama de red que muestra quién habla en el turno precedente y el comportamiento tonal de los hablantes en fragmentos de conflicto en c4

De la Figura 15, llama la atención la multiplicidad de intercambios entre A y B, la mayoría con un nivel neutro, aunque se observan muchas secuencias con tono alto. Más allá de esa relación conflictiva entre estos dos hablantes, el diagrama manifiesta un dato interesante sobre B: este hablante, que a la luz de los resultados globales del corpus y de lo mostrado en la c13 utiliza pocas intervenciones, se involucra en este caso en un intercambio conflictivo equilibrado con su esposa, en términos de posesión de la palabra. Los tonos altos se dirigen fundamentalmente de A a B, pero en uno de los turnos de B a A, como se mostraba en la visualización temporal, se da el pico más alto de pitch de la conversación, concretamente cuando B pronuncia *¡y ya es[táa!]* por tercera vez. El alto número de solapamientos de este intercambio explica la dinamicidad conversacional y, por tanto, el tamaño equilibrado de los círculos de A y B.

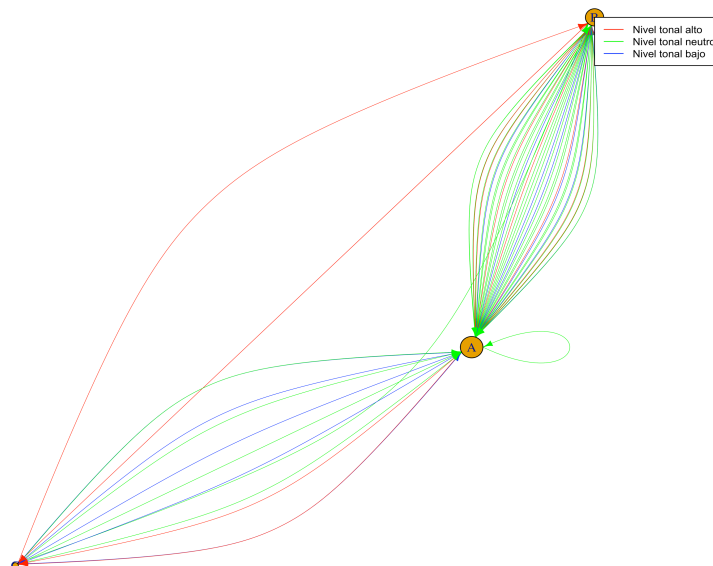


Figura 16. Diagrama de red que muestra quién habla en el turno precedente y el comportamiento tonal de los hablantes en fragmentos de **no** conflicto en c4

Si se compara la figura 15 con la 16, que representa los grupos entonativos no conflictivos, lo primero que se observa es la entrada en escena de la hablante C (D, en este caso, no está presente en la conversación). Como se mostraba en el apartado 5.1, el fragmento no conflictivo de la c4 sucede inmediatamente en el tiempo al fragmento conflictivo: se abandona parcialmente el tema de la frustración de A por no sentirse escuchada cuando comunica información sobre los días festivos y se pasa a una fase de desescalada en la que comienzan a hablar de dónde pueden ir a pasar los días festivos, dado que las medidas de protección del COVID impiden viajar fuera de la Comunidad Valenciana. Se da un predominio de enunciados neutros, con un par de excepciones y algunos grupos entonativos cuyo pitch elevado se debe a que B está ya subiendo por las escaleras para marcharse ya a la cama. En el fragmento, B sube el *pitch* como rasgo concomitante a la subida de la intensidad necesaria para que lo oigan desde la estancia familiar, hasta que finalmente decide volver a bajar y el pitch se normaliza:

(5)

A: ¿está cerrada la verjaa?
B: sí/ °<ininteligible/>°
A: ¿eh?
C: [°<ininteligible/>°/ °¿qué] le pasa a esto°°
B: [°<ininteligible/>°]
A: no te he entendido
B: ya está cerrada
A: ¿pero qué has dicho?
B: que así ya no tengo que salir/ ni pasas frío

A nivel discursivo, este comportamiento de epicentro discursivo de A, tanto en c4 como, especialmente, en c13, correspondería con su rol de perseguidora, mientras que B, por sus interacciones y su manifestación fónica, es un evitador (cf., entre otros, Greenberg & Johnson, 1988; Johnson, 2008; Gottman, 1999). Tal como se entiende en psicología y en terapia de pareja (Johnson, 2008), el perseguidor en una relación tiende a buscar cercanía y comunicación constante, un comportamiento que puede provenir de la ansiedad ante la percepción de distancia emocional por parte del otro o de una falta de comunicación. Lingüísticamente, esto se manifiesta en que suele ser quien inicia las conversaciones sobre los problemas de la relación, a menudo de manera insistente o confrontativa, lo que puede percibirse como una persecución por parte de la otra persona. El perseguidor, ante el temor a ser abandonado o rechazado, trata de mantener la conexión activamente, incluso si eso significa iniciar o perpetuar un conflicto, por ejemplo, con turnos conversacionales muy extensos que buscan constantemente la reacción del otro. Estas son exactamente las características exhibidas por A en todo el corpus, y que se perciben, en parte, en la visualización de la parte conflictiva de c13, en contraste con la no conflictiva.

Por otro lado, el evitador tiende a retirarse o distanciarse ante el conflicto o la presión emocional, de modo que minimiza los problemas, se retira del enfrentamiento y puede parecer indiferente; este comportamiento, que a menudo es en realidad un mecanismo de defensa desplegado para evitar el dolor o la incomodidad, provoca que el perseguidor se sienta aún más ansioso, al percibir la ya mencionada distancia emocional y falta de comunicación, de modo que comienza el denominado ciclo de persecución y evitación que se encuentra también en la base de muchos conflictos intratables. En el caso de la conversación 13, llama la atención que, pese al intento de A por recriminar la cocción de las sardinas, B no reacciona, usa tonos bajos o ignora el intercambio, y es D quien asume un papel más proactivo en defensa de los esfuerzos de B. En cambio, en c4, el hablante

B entra en la conversación y, cuando lo hace, despliega unas estrategias fónicas mucho más virulentas e invasivas que las de A, con numerosos picos tonales.

6 CONCLUSIÓN

El estudio que se acaba de presentar muestra cómo la variación fónica, que incluye tono, intensidad, duración y velocidad del habla, juega un papel en la manifestación de los conflictos. La influencia de estos factores se hace más patente cuando se combinan con el hablante y la conversación, de manera que las características fónicas pueden ayudar a saber cómo diferentes hablantes manejan los conflictos verbalmente. No todo el mundo se comporta fónica e interactivamente igual con todos los interlocutores, y a veces se emplean parámetros fónicos distintos con unos y con otros dependiendo de la relación mutua, de la desalineación con el contenido del turno precedente, etc.

En términos prácticos, este comportamiento diferencial puede usarse como mecanismo para proporcionar una representación objetiva de las dinámicas conversacionales. En este trabajo, en concreto, a partir del corpus ESPRINT_Conversación VLC, de conversaciones conflictivas espontáneas, se han analizado el indicador fónico del *pitch* y la dirección de los intercambios como indicadores importantes en la dinámica de los conflictos verbales. Así, se han presentado dos mecanismos de visualización de las conversaciones, que se han ejemplificado con dos conversaciones en concreto, la c4 y la c13.

A través de gráficos de dispersión con Z-Scores, el analista puede observar el comportamiento fónico a lo largo de la línea de tiempo y destacar en qué momentos se produce una alteración del *pitch* medio de los hablantes y observar si, más allá de picos aislados, existe una zona de acumulación de valores notablemente altos o bajos. Esto permite al analista acudir a la transcripción y comprobar qué está sucediendo interaccionalmente en los fragmentos desviados. Por otro lado, los diagramas de red permiten observar, además, la distribución de las variables fónicas en la interacción entre hablantes, de modo que puede verse con quién tiende a elevar o reducir significativamente su *pitch* un hablante, y con quién mantiene valores neutros, así como cuánto tiempo mantiene cada participante el turno de habla. Esta visualización ofrece, así, una "radiografía" de las conversaciones conflictivas que permite visualizar los patrones de intercambio entre hablantes y revelar perfiles discursivos diferenciados. En particular, se destacan los roles psicológicos asociados a los conceptos de "perseguidor" y "evitador", los cuales tienen correlatos en la interacción verbal y el comportamiento fónico durante los conflictos.

Desde el punto de vista de la reflexión teórica, los datos obtenidos reafirman la complejidad en la interpretación de la prosodia y cómo puede ser engañosa la atribución de emociones basada únicamente en características fónicas, pero el sistema de visualización permite un análisis multicapa en el que, superponiendo otros factores fónicos además del *pitch* (como el comportamiento de la intensidad, la velocidad de habla, etc.) y otros factores pragmáticos e interactivos como la presencia de solapamientos, la extensión de los turnos conversacionales, etc., se puede lograr una delimitación cada vez más precisa de los fragmentos en que se produce un conglomerado de manifestaciones marcadas y relacionarlos con eventos conversacionales significativos.

Más allá del interés lingüístico, esta visualización de conversaciones tiene potenciales aplicaciones en campos como la terapia de conflictos y la comunicación intercultural, puesto que el seguimiento de la huella fónica puede ayudar a los pacientes en la toma de conciencia de ciertas conductas interactivas negativas, y a los terapeutas en la identificación y abordaje de estas.

7 REFERENCIAS

- Ang, N., Bein, D., Dao, D., Sanchez, L., Tran, J., & Vurdien, N. (2018). *Emotional prosody analysis on human voices*. 737-741. <https://doi.org/10.1109/CCWC.2018.8301691>
- Bänziger, T., & Scherer, K. R. (2005). The role of intonation in emotional expressions. *Speech communication*, 46(3-4), 252-267.
- Briz Gómez, Antonio. (en preparación). *Los conflictos en la conversación coloquial entre familiares, amigos o conocidos*.
- Cabedo Nebot, A. (2022). *Oralstats* (1.3) [Software]. <https://github.com/acabedo/oralstats>
- Cabedo Nebot, A. (en prensa). Análisis multicapa del habla conversacional conflictiva a partir de la herramienta computacional OralStats. *Cultura, Lenguaje y Representación*.
- Carcelén Guerrero, A. (en prensa). ¿Es posible elaborar corpus orales espontáneos y cumplir la legislación? El modelo en tres fases del corpus Ameresco. *Revista española de lingüística aplicada*.
- Carcelén Guerrero, A., & Uclés Ramada, G. (2019). Diseño y construcción de un corpus oral multidialectal. El corpus Ameresco. *Normas*, 9(1), 17. <https://doi.org/10.7203/Normas.v9i1.16007>
- Carcelén Guerrero, Andrea. (en prensa). Retos y dificultades para la construcción de un corpus espontáneo de conflicto. *Cultura, Lenguaje y Representación*.
- Clancy, B. (2018). Conflict in corpora: Investigating family conflict sequences using a corpus pragmatic approach. *Journal of Language Aggression and Conflict*, 6(2), 228-247. <https://doi.org/10.1075/jlac.00011.cla>
- Coleman, P. T. (2000). Intractable conflict. En *The handbook of conflict resolution: Theory and practice*. (pp. 428-450). Jossey-Bass/Wiley.
- Couper-Kuhlen, E., & Selting, M. (1996). Towards an interactional perspective on prosody and a prosodic perspective on interaction. *Prosody in conversation: Interactional studies*, 11.
- Estellés Arguedas, María. (en prensa). La identificación del conflicto en conversación espontánea: Participantes vs. Analistas. *Cultura, Lenguaje y Representación*.
- Field, A. P., Miles, J., & Field, Z. (2012). *Discovering Statistics using R*. SAGE.
- Garrido Almiñana, J. M., & Chica Sabariego, J. A. (2018). Pitch range and identification of emotions in Spanish speech: A perceptual study. *Estudios de fonética experimental*, 27, 13-36.
- Gottman, J. M. (1999). *Marriage clinic: A scientifically based marital therapy*. WW Norton & Company.
- Greenberg, L. S., & Johnson, S. M. (1988). *Emotionally focused therapy for couples*. Guilford Press.
- Haugh, M. (2007). The discursive challenge to politeness research: An interactional alternative. *Journal of Politeness Research. Language, Behaviour, Culture*, 3(2). <https://doi.org/10.1515/PR.2007.013>
- Hevey, D. (2018). Network analysis: A brief overview and tutorial. *Health Psychology and Behavioral Medicine*, 6(1), 301-328. <https://doi.org/10.1080/21642850.2018.1521283>
- Johnson, S. M. (2019). *The Practice of Emotionally Focused Couple Therapy: Creating Connection* (3.^a ed.). Routledge. <https://doi.org/10.4324/9781351168366>
- Johnson, S. M., & Whiffen, V. E. (1999). Made to measure: Adapting emotionally focused couple therapy to partners' attachment styles. *Clinical Psychology: Science and Practice*, 6(4), 366-381. <https://doi.org/10.1093/clipsy.6.4.366>
- Kim, S., Valente, F., & Vinciarelli, A. (2012). *Automatic detection of conflicts in spoken conversations: Ratings and analysis of broadcast political debates*. 5089-5092.
- Kotthoff, H. (1993). Disagreement and concession in disputes: On the context sensitivity of preference structures. *Language in Society*, 22(2), 193-216. <https://doi.org/10.1017/S0047404500017103>
- Max Planck Institute. (2022). *ELAN*. Max Planck Institute for Psycholinguistics. <https://archive.mpi.nl/tla/elan>
- Muntigl, P., & Turnbull, W. (1998). Conversational structure and facework in arguing. *Journal of Pragmatics*, 29(3), 225-256. [https://doi.org/10.1016/S0378-2166\(97\)00048-9](https://doi.org/10.1016/S0378-2166(97)00048-9)
- Padilla García, Xose. (2023). Cómo construimos las emociones en la entonación coloquial. *Estudios de Fonética Experimental*, 32, 155-168.
- Pamies Bertrán, A., Fernández Planas, A. M., Ortega Escandell, A., & Amorós, M. C. (2002). *Umbral tonales en español peninsular*. Universidad de Sevilla.
- Pomerantz, A. (1984). Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shaped. En Atkinson, J. M. & Heritage, J. (Eds.), *Structures of social action: Studies in conversational analysis*. Cambridge University Press.
- Quilis, A., Cantarero, M., & Esgueva, M. (1993). El grupo fónico y el grupo de entonación en español hablado. *Revista de Filología Española*, LXXIII, 55-65.
- R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. R

- Foundation for Statistical Computing.
<https://www.R-project.org/>
- Roth, W.-M., & Tobin, K. (2010). Solidarity and conflict: Aligned and misaligned prosody as a transactional resource in intra-and intercultural communication involving power differences. *Cultural Studies of Science Education*, 5, 807-847.
- Roth, W.-M., & Tobin, K. (2006). *Solidarity and conflict: Prosody as interactional resource in intra-and intercultural communication involving power differences*. International Conference on Conversation Analysis, Helsinki.
<https://api.semanticscholar.org/CorpusID:235260731>
- Sifianou, M. (2019). *Conflict, disagreement and (im)politeness*.
<https://doi.org/10.13140/RG.2.2.29715.02082>
- Stewart, K. A., & Maxwell, M. M. (2010). *Storied Conflict Talk: Narrative construction in mediation* (Vol. 12). John Benjamins Publishing Company. <https://doi.org/10.1075/sin.12>
- Szcepek Reed, B. (2010). Prosody and alignment: A sequential perspective. *Cultural Studies of Science Education*, 5, 859-867.
- Watts, R. J. (2003). *Politeness*. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511615184>
- Wichmann, A. (2011). Prosody and pragmatic effects. *Pragmatics of society*, 181-214.
- Wichmann, A. (2012). Prosody in context: The effect of sequential relationships between speaker turns. En G. Elordieta & P. Prieto (Eds.), *Prosody and Meaning* (pp. 329-348). DE GRUYTER.
<https://doi.org/10.1515/9783110261790.329>
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. <https://ggplot2.tidyverse.org>
- Wittfoth, M., Schröder, C., Schardt, D. M., Dengler, R., Heinze, H.-J., & Kotz, S. A. (2010). On emotional conflict: Interference resolution of happy and angry prosody reveals valence-specific effects. *Cerebral Cortex*, 20(2), 383-392.
- Zellers, M., & Ogden, R. (2014). Exploring interactional features with prosodic patterns. *Language and Speech*, 57(3), 285-309.
- Zhu, W., & Boxer, D. (2013). Strong disagreement in Mandarin and ELFP: Aggressive or politic? *Journal of Language Aggression and Conflict*, 1(2), 194-224.