

Articles

Acoustics of persuasion: a systematic review of prosodic features on the communicative attitudes of confidence

Acústica da persuasão: uma revisão sistemática de aspectos prosódicos na atitude comunicativa de confiança

Dianete Ângela do Valle Gomes¹

Rodrigo Dornelas²

Livia Maria Santos de Souza³

Yonara Caetano de Santana Strauss⁴

Eduardo Magalhães da Silva⁵

Letícia Corrêa Celeste⁶

1. Universidade de Brasília – UNB. Distrito Federal – Brasil. <https://orcid.org/0000-0002-3481-6927>. E-mail: dianete.vg@gmail.com

2. Universidade Federal do Rio de Janeiro – UFRJ. Rio de Janeiro – Brasil. <https://orcid.org/0000-0002-9710-5751>. E-mail: rdgdornelas@gmail.com

3. Universidade de Brasília – UNB. Distrito Federal – Brasil. <https://orcid.org/0000-0001-9491-6156>. E-mail: livia_mss@hotmail.com

4. Universidade de Brasília – UNB. Distrito Federal – Brasil. <https://orcid.org/0000-0002-6012-8757>. E-mail: ycaetano@hotmail.com

5. Universidade de Brasília – UNB. Distrito Federal – Brasil. <https://orcid.org/0000-0001-5132-8370>. E-mail: edmagalhaes@gmail.com

6. Universidade de Brasília – UNB. Distrito Federal – Brasil. <https://orcid.org/0000-0002-2384-3989>. E-mail: leticiacceleste@gmail.com



This content is licensed under a Creative Commons Attribution License, which permits unrestricted use and distribution, provided the original author and source are credited.

ABSTRACT

Purpose: To review the literature regarding prosodic acoustic features found in communicative attitudes related to confidence, certainty, and persuasion. Method: A systematic review was carried out in the databases VHL, Web of Science, Science Direct, SciELO, and SCOPUS with no temporal restriction. Data Extraction: The data from each article was extracted based on the STROBE Statement checklist. To analyze the prosodic variables, the data were subdivided according to Couper-Kuhlen's (1986) theoretical assumptions, and the variables were grouped into "temporal organization of speech," "intensity," and "pitch". Conclusion: The data suggested that there are relevant but not consensual variables to characterize a persuasive speech and some variables can be considered as positive, negative, or neutral in different language contexts. The variables that stand out as relevant and are characteristic of the persuasive or confident speech were faster speech speed and higher intensity. The only negative variable that stood out regarding persuasion was the increase in pitch.

Keywords: *persuasion; prosody; communication; acoustics; speech.*

RESUMO

Objetivo: revisar a literatura sobre características acústicas prosódicas encontradas em atitudes comunicativas relacionadas à confiança, certeza ou persuasão. Método: Foi realizada uma revisão sistemática nas bases de dados BVS, Web of Science, Science Direct, SciELO e SCOPUS sem restrição temporal. Extração de dados: os dados de cada artigo foram extraídos com base na lista de verificação da instrução STROBE. Para analisar as variáveis prosódicas, os dados foram subdivididos a partir do pressuposto teórico de Couper-Kuhlen (1986) e as variáveis foram agrupadas em "organização temporal da fala", "intensidade" e "variação melódica". Conclusão: Os dados sugerem que existem variáveis relevantes, mas não consensuais, para caracterizar um discurso persuasivo e algumas variáveis podem ser consideradas positivas, negativas ou neutras em diferentes contextos de linguagem. As variáveis que se destacam como relevantes e características do discurso persuasivo ou confiante são maior velocidade de fala e maior intensidade. A única variável negativa que se destacou em relação à persuasão foi o aumento do pitch.

Palavras-chave: *persuasão; prosódia; comunicação; acústica; discurso.*

1. Introduction

Humans communicate using a complex system that involves non-verbal and verbal features (Chomsky, 1975; Couper-Kuhlen, 1986; Mozziconacci, 2001; Hardcastle et al., 2012). Gestures and face expressions are part of the non-verbal communication elements (Hardcastle et al., 2012; Pouw et al., 2017). Additionally, dress style, or stance may influence how the speaker is perceived by others (Mehrabian, 1969; Barnard, 2013; 2014), so we could also view such aspects as a form of communication.

Segmental and suprasegmental phonation aspects are a part of verbal communication (Couper-Kuhlen, 1986; Hardcastle et al., 2012). In the field of Speech-Language Pathology, the improvement of communicative performance makes use of linguistic techniques related to prosody and oratory with the purpose of developing skills related to communicative functionality. Our study focuses on the communicative aspects surrounding prosody, particularly how attitudinal prosody is conveyed.

Prosodic features can alter linguistic interpretation of the same segmental text (Fónagy, 2003), and this can be done voluntarily or not (Petty & Cacioppo, 1986). From the perspective of Sperber and Wilson's relevance theory (Petty & Cacioppo, 1986), prosody provides information unintentionally and intentionally, and the latter may happen covertly or overtly (Wilson & Wharton, 2006). Intentional covert prosodic effects are the ones in which the speaker intentionally lets a message be conveyed to the receptor, but not directly. On the other hand, intentional overt prosodic effects are direct and openly intentional; they convey exactly what the speaker wants to say without ambiguity (Petty & Cacioppo, 1986; Petty, 2018).

Attitudes are cognitively monitored internal states, with a communicative purpose that only exists and is expressed in the listener's presence. This occurs because attitudes implicitly involve an action towards someone else, that is, the speaker is always expressing himself/herself towards the listener in a condescending, friendly, etc. way. (Zhao et al., 2018). An attitude is a controlled and conscious determined behavior consisting of intellectual and moral components.

Emotions, on the other hand, escape from speakers' control – they are individual psychic manifestations. In this sense, irony, doubt, certainty, and approval are examples of attitudes, while joy, sadness, and anguish are seen as emotions (Jiang & Pell, 2017).

It is known that prosody effects are highly dependent on the context (Zhao et al., 2018). They are influenced by the context and interact on the message interpretation level. The same prosodic input has various effects depending on how it is used. To minimize such trends, attitudinal prosody studies have been trying to eliminate context bias by providing a single text to all participants in experiments (Jiang & Pell, 2016). This kind of research allows us to better understand the organization of prosodic features in attitudes by interpreting their communicative mechanisms and their relevance to human performance.

A key issue in studying persuasion is the vast quantity of theories in social sciences developed to try to understand how people are influenced by someone or something (Petty & Cacioppo, 1986; Petty, 2018). Although it is an extremely interesting and relevant point, this paper will focus on speakers' attitudes conceptually related to persuasion, namely certainty and feeling of knowledge.

Prosody can be studied from a phonetic or a phonological perspective. From a phonetic point of view, prosody is composed of three main features: pitch (fundamental frequency), intensity, and speech temporal organization (temporal measures) (Blakemore, 1987). Considering the relevance of prosody to pragmatics, especially when it correlates with attitudes, the aim of this study is to examine the following question: what are the acoustic prosodic characteristics found in persuasive communicative attitudes in adults?

2. Method

Search strategy. Relevant studies were identified in the electronic literature databases Virtual Health Library (VHS), Web of Science, Science Direct, SciELO, and SCOPUS. No time restrictions were considered. The search terms were previously discussed by the group and classified using the DeCS-Health Sciences Descriptors section of the VHL. Then, to identify eligible publications, the following

combination of keywords and Mesh Terms was entered in the topic/subject fields of the databases: (prosody OR “speech rate” OR pitch OR intonation) AND (confidence OR persuasion OR certainty OR doubt OR uncertainty OR knowledge). The last search was performed in April 2018.

Selection of publications. After the search, the references were imported to Mendeley to remove duplicates and their titles and abstracts were reviewed to determine the full-text reading selection. Full-text reading of selected articles was conducted by two members of the research team to decide upon eligibility for inclusion. In case of disagreement about the inclusion of a certain paper, an independent third member also assessed the paper, and the final decision was based on a majority vote.

An adaptation of the PICO framework was used in this review to define the population, context and/or problem situation (P); the variables (V), and the desired outcomes (O).

Inclusion criteria. Studies were included if they were (1) dealing with the relationship between prosodic acoustic characteristics in communicative attitudes and persuasion, (2) written in English, Portuguese or Spanish, and (3) were available in full text at *Portal de Periódicos CAPES*, a Brazilian online research library.

Exclusion criteria. Papers were excluded if they (1) included participants under 18 years old, (2) reported on single-case studies, reviews, books, commentaries, letters to the editor, bibliographic review, and unpublished dissertations, or (3) were written in languages other than the selected ones.

Data analysis. Data from each publication was collected and summarized in a standardized table based on the STROBE Statement checklist (Malta et al., 2010) namely, title, year of publication, journal, authors, and country of origin, identification of the background, objectives, and speech corpus data measurement. The variables were grouped into temporal organization of speech, intensity, and pitch. The prosodic variables were then subdivided (Couper-Kuhlen, 1986). Within the temporal organization of the discourse, the following variables were listed: speech rate, pauses, and hesitations (Celeste &

Dianete Ângela do Valle Gomes, Rodrigo Dornelas, Livia Maria Santos de Souza, Yonara Caetano de Santana Strauss, Eduardo Magalhães da Silva, Letícia Corrêa Celeste

Reis, 2013). The prosodic subdivisions (Couper-Kuhlen, 1986) had to be later read and subdivided into more specific variables, since some subdivisions appeared only after analysis (Table 1).

Table 1 – Specific variables of the prosodic feature analyzed from the theoretical assumption of Couper-Kuhlen (1986).

Prosodic feature	Specific variables
Speech temporal organization	Higher/faster speech rate Lower speech rate Pauses Short pauses Long pauses Disfluencies (including filled pauses and stuttering-like disfluencies)
Intensity	Moderate intensity Increased mean intensity Lower mean amplitude Increased variation in intensity
Pitch	Rising pitch/intonation High Mean F0 Increased F0 range/increased variation in pitch Decreased F0 range/unmarked intonation Lower mean F0 Wide falling tone (pitch range) Falling-rising tone (pitch range)

Geographic and temporal data analyses of the articles were also carried out. The temporal analysis was quantitative and considered the year of publication and, as for the geographic analysis, the country of origin or the university of the main author were included for consideration. *Statistics*. The Shapiro-Wilk normality test was used to determine the data's normal distribution. For comparison between non-parametric groups, the Mann-Whitney U test was applied.

3. Results

The initial search yielded 684 articles and 65 duplicates were removed after being imported to Mendeley. After revision of the titles and abstracts, 466 articles were excluded. Full-text access was evaluated, and another 102 articles were excluded. The remaining 51

papers were read, and 30 texts were excluded for dealing with other aspects of the communicative performance or not presenting acoustic features. Ultimately, 21 articles were included in the review.

The studies were conducted in nine different countries and were published between 1973 and 2018. The countries that produced the most articles were Canada (n=6) and the United States (n=5), followed by Brazil (n=3). In the 21st century, other countries contributed one publication and the number of publications increased (n=18) (Table 2). Figure 1 shows the geospatial and temporal surface view of the distribution of the publications by decades.

Table 2 – General data of the 21 analyzed papers

N	Author	Year	Journal	Title	Country of origin
1	Klofstad & Anderson	2018	Evolution and Human Behaviour	Voice pitch predicts electability, but does not signal leadership ability	USA
2	Zhao et al.	2018	Journal of Pragmatics	From pitch to purpose: The prosodic pragmatic mapping of [+verb] belief constructions in English and Mandarin	UK
3	Jiang & Pell	2017	Speech Communication	The sound of confidence and doubt	Canada
4	Jiang & Pell	2016	Neuropsychology	Neural responses towards a speaker's feeling of unknowing	Canada
5	Chen et al.	2016	PLOS ONE	Perceived masculinity predicts U.S. Supreme Court Outcomes	France
6	Jiang & Pell	2015	Cortex	On how the brain decodes vocal cues about speaker confidence	Canada
7	Elbert & Dijkstra	2014	Psychology & Health	An experimental test of the relationship between voice intonation and persuasion in the domain of health	Netherlands
8	Azevedo et al.	2014	Arq. Neuro-Psiquiatria	Effect of Speech Therapy and pharmacological treatment in prosody of parkinsonians	Brazil
9	Celeste & Reis	2013	Alfa	Attitudes expression on stuttering speech: perception of fluent speakers	Brazil
10	Celeste & Reis	2013	Ver. CEFAC	Expression of certainty and doubt in stuttering: study of temporal aspects of speech	Brazil
11	Yokoyama & Daibo	2012	Psychological Reports	Effects of gaze and speech rate on receivers, evaluations of persuasive speech	Japan
12	Tomlinson & Tree	2011	Cognition	Listeners' comprehension of uptalk inspontaneous speech	USA
13	Jones et al.	2007	Speech and Computer Language	Synthesized speech intelligibility and persuasion: Speech rate and non-native listeners.	Australia

Dianete Ângela do Valle Gomes, Rodrigo Dornelas, Livia Maria Santos de Souza, Yonara Caetano de Santana Strauss, Eduardo Magalhães da Silva, Letícia Corrêa Celeste

N	Author	Year	Journal	Title	Country of origin
14	Chebat et al.	2007	Perceptual and motorskills	Voice and persuasion in a banking telemarketing context	Canada
15	Pell MD	2006	Brain and Language	Reduced sensitivity to prosodic attitudes in adults with focal right hemisphere brain damage	Canada
16	Verdugo DR	2005	Journal of Pragmatics	The nature and patterning of native and non-native intonation in the expression of certainty and uncertainty: Pragmatic effects	Spain
17	Krahmer & Swerts	2005	Language and Speech	How children and adults produce and perceive uncertainty in audiovisual speech	Netherlands
18	Barr, DJ	2003	Psychonomic Bulletin & Review	Paralinguistic correlates of conceptual structure	USA
19	Gélinas-Chebat et al.	1996	Perceptual and motorskills	Voice and advertising: Effects of intonation and intensity of voice on source credibility, attitudes toward the advertised service and the intent to buy	Canada
20	Erickson et al.	1978	Journal of experimental social psychology	Speech style and impression formation in a court setting: The effects of "powerful" and "powerless" speech.	USA

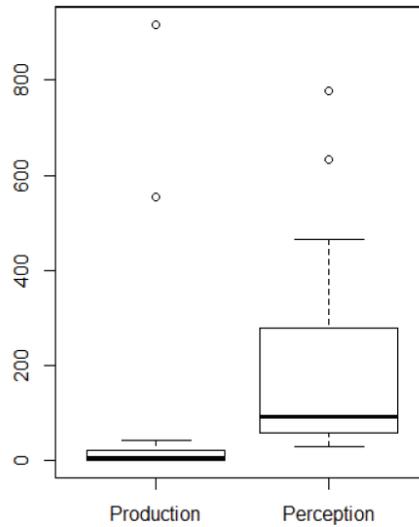
Figure 1 – Geospatial and temporal surface view



(A) 1970s, USA (n=2), (B) 1990s, Canada (n=1), (C) 2000s, Canada (n=2), USA (n=1), Netherlands (n=1), Australia (n=1), Spain (n=1), and (D) 2010s, Canada (n=3), USA (n=2), Brazil (n=3), Netherlands (n=1), UK (n=1), Japan (n=1), France (n=1).

Among the 21 remaining papers, 12 were about speech perception and two about speech production. Seven of them dealt with both speech perception and production. The distribution/number of participants in the groups was different ($U=330.5$; $p=0.0024$) and more prominent in the perception ($CI=-145.9$) than in the production ($CI=-29.1$) side. It is noteworthy that in perception studies, the samples were bigger and had greater variability (Figure 2).

Figure 2 – Distribution of the participants in production and perception studies. The perception studies showed a bigger sample and a broader distribution of participants.



These numbers were not under a normal distribution for either production ($w=0.39459$ and $p=0.00$) or perception ($w=0.7801$ and $p=0.00$) studies. Participants came from different contexts (college students, members of Congress, stutterers), were native or non-native speakers, and the evaluated samples of speech production or perception were uniform or varied by participant (a specific utterance or connected speech) (Table 3).

Dianete Ângela do Valle Gomes, Rodrigo Dornelas, Livia Maria Santos de Souza, Yonara Caetano de Santana Strauss, Eduardo Magalhães da Silva, Letícia Corrêa Celeste

Table 3 – Methodological details of the selected studies

Number 1 Objective: To assess whether voice pitch is a reliable signal of leadership ability, if individuals with lower pitched voices are better leaders, individuals with lower pitched voices should be more persuasive when making policy appeals. Global design: Production and perception		
Corpus (speech production)	Perceptual evaluation (perception)	Participants
Study 1. Observational study. Recordings of each Member of Congress. Voice sources from YouTube and CSPAN online archive videos.	Study 1. Leadership ability is measured based on a power ranking of the Members of the 109th US Congress created by Knowlegis (cqrollcall.com/knowlegis) and compared to results of acoustic analysis by PRAAT used to measure the mean F0 of each recording.	Study 1. 536 Members of Congress
Study 2. Policy persuasion experiment. Four male and four female native English speakers were recorded speaking the following six policy advocacy statements: "You should support same sex marriage"	Study 2. The experiment required subjects to listen to recordings of policy advocacy statements delivered by speakers with differently pitched voices to test whether speakers with lower voices are more persuasive.	Study 2. 777 participants. 344 men and 433 women. Cooperative; Congressional.

Number 2 Objective: Study the "prosodic variation as a cue to pragmatic interpretation in a specific linguistic construction namely [+verb]". Global design: Production and perception		
Corpus (speech production)	Perceptual evaluation (perception)	Participants
An online language questionnaire adapted from the questionnaire Zaho's (2013) was used and online perception tasks created via Qualtrics (Qualtrics Lab, INC, 2015) and validated by pilot test. The tasks combine the studies of Meyer and Mleineks (2006) that use electronic research to present audio recordings for the perception and categorization of tasks involving the prosodic interpretation of issues/utterances in Russian; and the studies of Kurumada et al. (2012) with electronic human intelligence (artificial intelligence) in tasks of weekly interpretations and prosodic variation of the utterance: "It looks like", and that similarly, analyzed the effect of prosodic variation in the semantic interpretation.	"Think" ('juede' in Mandarin), the most used belief verb, was used across trials. All target nouns were imaginable, common words and were chosen for their neutrality and contextual plausibility. (...) These were audio recorded with one of the six prosodic variations and presented for the listener to click and hear. Each recording was followed by a question asking the participant to choose whether the speaker in the recording sounds, "Deliberate," "Uncertain," or "Reluctant", with a short definition following each term" (...) "This task included 24 target nouns with each of the six prosodic variations appearing four times".	59 individuals consisting of 30 in the US/English group and 29 in the China/Mandarin Group.

Acoustics of persuasion

<p>Number 3 Objective: To shed light on how speakers and listeners use systematic variations in verbal and vocal cues to communicate confidence along a continuum of expressed meanings (confident, close-to-confident, unconfident) and in the context of utterances serving different communicative functions (facts, judgments, intentions). Globaldesign: Production and perception</p>		
Corpus (speech production)	Perceptual evaluation (perception)	Participants
<p>Experiment 1. Sentences were produced by native Canadian English speakers. Speakers were encouraged to produce statements with a certain level of confidence by responding to a question from a female examiner in a mini dialogue format (e.g., “What will happen?” – Target: “Maybe, we will run out of gas”).</p>	<p>Listeners were asked to make two consecutive judgments. They were first asked to judge whether the speaker intended to convey some level of confidence by clicking on a YES.</p>	<p>Experiment 1. Six native Canadian English speakers (mean age of 22.8 years, three female and three male). Experiment 2. A total of 72 native speakers of Canadian English were recruited from McGill University and consented to serve as listeners in the study (37 females and 35 males, mean age of 24.8 years - from 18 to 35 YO; mean education of 16.34 years, from 13 to 21 YO).</p>
<p>Number 4 Objective: To Investigate the time course and neural responses underlying a listener’s ability to evaluate speaker confidence from combined verbal and vocal cues. Globaldesign: Perception</p>		
Corpus (speech production)	Perceptual evaluation (perception)	Participants
<p>Recordings were selected for four different actors (two females and two males, aged from 19-25, 24 triplets/ speaker); these speakers produced the most perceptually salient distinctions in the low to high levels of expressed confidence based on ratings from 60 listeners in a validation study.</p>	<p>96 stimulus triplets were selected from a database of vocal confidence recordings (Jiang and Pell, 2014). A stimulus was composed of a lexical phrase communicating one of three distinct levels of confidence (confident, close-to-confident, unconfident), followed by the main utterance that was linguistically identical across confidence conditions, but produced in a tone of voice that was congruent with the lexical phrase.</p>	<p>30 university students consented to participate in the electroencephalogram (EEG) experiment (15 female/15 male, mean age of 22.6 years, range from 18 to 30 YO). All were right-handed, native Canadian English speakers who took part in a companion study focusing on vocal confidence perception.</p>
<p>Number 5 Objective: To test whether people’s subjective voice-based trait judgments are predictive of the Supreme Court of the US outcomes. Globaldesign: Perception</p>		
Corpus (speech production)	Perceptual evaluation (perception)	Participants
<p>1634 oral arguments made by 916 distinct male advocates. The stimuli for this study were drawn from oral arguments made in the Supreme Court of the US between 1998 and 2012. Use of identical two to three seconds of content delivered at the outset of each argument: “Mr. Chief Justice, (and) may it please the Court”.</p>	<p>Participants from Amazon Mechanical Turk (AMT) rated the voice clips of the Supreme Court advocates. About half (321) of the 634 distinct participants who completed our survey were female. Participants were asked to rate the voice clips of Supreme Court advocates on a scale of 1 to 7 in terms of aggressiveness, attractiveness, confidence, intelligence, masculinity, and trustworthiness. Each participant rated 66 voice recordings.</p>	<p>916 male advocates; 634 raters (321 female).</p>

Dianete Ângela do Valle Gomes, Rodrigo Dornelas, Livia Maria Santos de Souza, Yonara Caetano de Santana Strauss, Eduardo Magalhães da Silva, Letícia Corrêa Celeste

Number 6
Objective: Examine the temporal neural dynamics underlying a listener's ability to infer speaker confidence from vocal cues during speech processing and investigate how and when the brain responds to expressed confidence in the vocal channel of speech, to provide new data on how listeners infer key aspects of a speaker's mental state (i.e., "feeling of knowing").
Global design: Perception

Corpus (speech production)	Perceptual evaluation (perception)	Participants
The EEG study used recordings from two male and two female speakers who encoded the most salient distinctions in the low to high levels of confidence when judged by native listeners in the validation study. Altogether, 384 recordings (96 statements and four confidence levels) were chosen as the experimental stimuli. Another 480 utterances that contained an initial lexical phrase for inferring confidence (e.g., "I'm positive...", "Maybe...") were also presented for the purpose of a companion study investigating the effects of verbal cues on vocal confidence perception (for these items, sentences were identical to the critical stimuli but included an initial lexical phrase that was not removed prior to the EEG experiment).	96 stimulus "quartets" were selected from a database of vocal expressions of confidence. Within each quartet, identical statements expressed either a confident, close-to-confident, unconfident, or neutral intending message based only on changes in tone of voice. The 96 statements varied in communicative function: 32 were descriptions of fact (e.g., "She has access to the building"), 32 were statements of judgment (e.g., "He's too old to split the wood"), and 32 were statements of intention (e.g., "We'll help them with it").	To production - used recordings by two male and two female speakers. To perception - 30 university students (15 female/15 male, mean age 22.6 years, range from 18 to 30 YO), native Canadian English speakers, right-handed. None had suffered from major psychiatric or neurological illness, had speech or hearing problems, or had participated in our previous studies.

Number 7
Objective: To test to what extent different levels of intonation are related to persuasion and whether for some recipients the threat posed by the message information might become too strong to face.
Global design: Perception

Corpus (speech production)	Perceptual evaluation (perception)	Participants
An actress with sufficient control over her voice was selected and directed by the first author to produce three versions of the health message which varied on the dimension of intonation (low, moderate, and low).	<p>Study 1. 130 respondents listened to a health message with either a low, moderate, or high level of intonation.</p> <p>Study 2. 143 respondents. The same manipulations of intonation were applied, but half of the respondents were affirmed before they listened to the persuasive message. Intention to increase fruit and vegetable intake was used as a dependent variable.</p>	<p>Study 1. 130 Dutch student participants (76.2% females), ranging from 17 to 32 YO, randomly distributed over the following conditions: low intonation condition; moderate intonation condition; high intonation condition.</p> <p>Study 2: 143 participants (72.7% females), who varied in age from 17 to 31 YO; self-affirmation conditions: low, moderate, high level of intonation; no self-affirmation conditions: low, moderate, high level of intonation.</p>

Acoustics of persuasion

<p>Number 8 Objective: To verify the influence of levodopa and of the adapted Lee Silverman Vocal Treatment Method on prosodic parameters employed by parkinsonian patients. Globaldesign: Production</p>		
Corpus (speech production)	Perceptual evaluation (perception)	Participants
Recording of utterances produced in four stages: expressing attitudes of certainty and doubt and declarative and interrogative modalities. The sentences were recorded under the effect of levodopa (on), without the effect of levodopa (off), and before and after speech therapy during the on and off periods.	No perceptual evaluation done.	Ten patients with idiopathic Parkinson's disease using levodopa, consisting of five men ranging from 59 to 88 YO and five women, ranging from 59 to 75 YO.
<p>Number 9 Objective: To verify how fluent speakers of Brazilian Portuguese perceive the expression of certainty and doubt attitudes on stutterers' speech. Globaldesign: Perception</p>		
Corpus (speech production)	Perceptual evaluation (perception)	Participants
12 stutterers (EG) and 12 non-stutterers (CG) recorded two utterances in each of the three forms studied (neutral, certainty, and doubt)	60 judges have participated: 30 judged CG's utterances and 30 judged EG's utterances.	60 male judges ranging from 20 to 40 YO. 24 participants, 12 with stuttering and 12 without stuttering for speech data.
<p>Number 10 Objective: To examine the role of speech temporal organization on the expression for attitudes of certainty and doubt in a group of adults who stutter, comparing such analysis with a group of speech-fluent adults. Globaldesign: Production</p>		
Corpus (speech production)	Perceptual evaluation (perception)	Participants
10 key phrases that were produced in the neutral form, expressing doubt and certainty, totaling 840 utterances.	No perceptual evaluation done.	24 individuals, 12 who stuttered (EG1 and EG2) and 12 non-stuttering (CG).
<p>Number 11 Objective: This study examined how gaze and speech rate affect the perceptions of a speaker. Globaldesign: Production and perception</p>		
Corpus (speech production)	Perceptual evaluation (perception)	Participants
The speaker was a 23 YO female graduate student. The speaker wore a black suit and glasses, showing her shoulders and head. Two gazes were directed at the camera: 8% of the total visual recording time (low gaze) and 83% of total visual recording time (high gaze). A group of recordings with a slow speech rate (330 syllables/min) and another group with a fast speech rate (510 syllables/min.), based on the average speech rate of 420 words/min. Pitch was set constant during the editing process. This produced four stimulus types: low gaze + fast speech, low gaze + slow speech, high gaze + fast speech, and high gaze + slow speech.	Participants were divided into four groups according to stimulus gaze (high gaze + low gaze groups) and stimulus speech rate (fast speech + slow speech groups). Participants were initially told that the aim of the study was to examine participants' comprehension of a provided communication. Participants were shown a static image of the speaker for five seconds, so that they could become accustomed to the study environment. Participants were asked to listen to and observe the speaker from the perspective of a person who is attending the conference.	466 students with a mean age of 19.6 years.

Dianete Ângela do Valle Gomes, Rodrigo Dornelas, Livia Maria Santos de Souza, Yonara Caetano de Santana Strauss, Eduardo Magalhães da Silva, Letícia Corrêa Celeste

Number 12		
<p>Objective: To test the hypothesis that the extensions at the end of the declarative utterances, or uptalk, can differentiate the conflicting functions of the increase of the tone. To verify whether the listeners assign different levels of certainty to the increase of tone and the prolongations.</p> <p>Global design: Perception</p>		
Corpus (speech production)	Perceptual evaluation (perception)	Participants
<p>For each utterance with a rising contour, a matched utterance was digitally stylized with a falling contour using the PSOLA function on PRAAT. Likewise, for each utterance with a falling contour, a matched utterance was digitally created with a rising contour. The beginning of the rise or fall was kept the same. The upward and downward of the contour was multiplied or divided by a factor of 1.5 and then further stylized to smooth out the contour. In this way, the same item could be heard with and without uptalk.</p>	<p>Experiment 1. 48 utterance pairs were selected from a specially compiled corpus of spontaneous speech. Speakers created spontaneous sentence frames to convey celebrity facts to an addressee who attempted to select the celebrity out of an array. For example, upon reading place of birth: Brooklyn, the speaker might say to the addressee, "This actor was born in Brooklyn, New York." The utterance pairs consisted of two sentences spontaneously produced by the same speaker; for example, "I have two children. I was the princess of Wales." The first utterance of the pair had either rising prolonged pitch (n=12), rising non-prolonged pitch (n=12), falling prolonged pitch (n=12), or falling non-prolonged pitch (n=12).</p> <p>Experiment 2 and 3. The same 48 stimuli were used as in Experiment 1. In addition, 48 filler stimuli were created. The filler stimuli were of two types. The first had the target word in the first utterance. The second did not contain the target word. When the target word was presented in the stimulus, there was a 1000ms pause until the next trial started.</p>	<p>All 26 participants were native English speakers, students from The University of California Santa Cruz.</p> <p>Experiment 2. 21 students participated. Experiment 3. 48 students participated.</p>

Acoustics of persuasion

<p>Number 13 Objective: To assess the effect of variation in speech rate on comprehension and persuasiveness of a message presented in text-to-speech (TTS) synthesis to native and non-native listeners. Global design: Perception</p>		
Corpus (speech production)	Perceptual evaluation (perception)	Participants
<p>Eight banking product descriptions under one of four conditions: normal rate (155 words/minute) with no background noise, normal rate with multi-talker background noise (+6 dB SNR), fast-normal (178 words/minute) with no background noise, and fast-normal with multi-talker background noise.</p>	<p>Participants rated each product on a scale from 1-5 (where 1 = not useful at all, and 5 = extremely useful), and answered two true/false questions about the product description they had just heard.</p>	<p>80 non-native English speakers and 80 native Australian English speakers.</p>
<p>Number 14 Objective: To examine the influence of voice and sex on the credibility of the voice source in a banking telemarketing context as well as with regards to the attitude toward the advertisement and subjects' behavioral intention. Global design: Perception</p>		
Corpus (speech production)	Perceptual evaluation (perception)	Participants
<p>Two professional actors (a woman and a man) manipulated the intensity and the intonation of the voice and the speech rate for each condition. The professionals, supervised by a phonetician, modified their voices to have a low, moderate, or high intensity, a marked, moderate, or unmarked intonation, and a slow, moderate, or fast speech rate. The message consisted of an advertisement for an ATM (Automatic Teller Machine) card offered by a Canadian bank.</p>	<p>Source credibility was evaluated with a semantic differential scale containing nine items. This scale ranged from -3 to +3 with zero as a midpoint. Participants were asked if the source of the voice who conveyed the message was competent or incompetent in the financial service field (V1), inspired confidence or distrust (V2), inspired honesty or dishonesty (V3), whether the listener believed or did not believe everything that was said (V4), had or did not have the power of persuasion (V5), was or was not prestigious (V6), had or did not have a similar cultural background (V7), was pleasant or unpleasant (V8), and was attractive or repulsive (V9). Participants completed questionnaires for attitude towards the advertisement and behavioral intention.</p>	<p>399 native speakers of Australian English.</p>

Dianete Ângela do Valle Gomes, Rodrigo Dornelas, Lívia Maria Santos de Souza, Yonara Caetano de Santana Strauss, Eduardo Magalhães da Silva, Letícia Corrêa Celeste

<p>Number 15 Objective: Study the relationship of the right hemisphere with the ability to classify the emotional attitudes of confidence and politeness of the speaker, based on prosody, through the comparison between individuals with and without impairment of the right hemisphere. Globaldesign: Production and perception</p>		
Corpus (speech production)	Perceptual evaluation (perception)	Participants
<p>Two experiments were carried out. The first one had prosody as a clue to judge confidence level and the second had prosody as a clue to judge the speaker's politeness. Sentences from six to 11 syllables, recorded by 4 male actors, were used. Half of the utterances were semantically congruent, and half were pseudo utterances only with acoustic clues, both with high, medium, or low confidence levels. 10 male and 10 female judges rated the recordings performed by actors on a scale from 1 to 5, resulting in 45 selected stimulators.</p> <p>In the second experiment, eight simple commands were used in sentences of three to five syllables recorded by two female actors, in two tones, and two auditory forms: with ascending prosody and descending prosody, and with four different linguistic beginnings: declaratory, direct, indirect, rather indirect and with the word please. The linguistic and prosodic utterances were judged by eight adults. They resulted in 80 linguistic and 32 prosodic-only utterances, totaling 112 selected statements.</p>	<p>Experiment 1. Evaluation of statements expressing confidence in two situations: (1) linguistic context provided by an informative semantical content; the sentences began with linguistic phrases indicating high, moderate, or low level of confidence; (2) pseudo utterance of exclusively auditory characteristics, produced with high, moderate, and low degrees of confidence.</p> <p>Experiment 2. Evaluation of utterances that express levels of politeness based on linguistic situations expressed by (1) semantic content and prosody and (2) prosody alone.</p>	<p>Experiment 1. Nine individuals with involvement of the right hemisphere and 11 individuals without such involvement. All English speakers with an average age of 63 and an average formal education time of 12 years.</p> <p>Experiment 2. Six individuals with an impairment and 11 without and impairment participated. Their mean age was 65 and they had an average of 12 years of formal education. There was an equal number of men and women.</p>

<p>Number 16 Objective: To investigate whether the interlanguage intonation produced by non-native speakers of English might lead to pragmatic differences that could affect their spoken discourse in the expression of certainty and uncertainty. Globaldesign: Production and perception</p>		
Corpus (speech production)	Perceptual evaluation (perception)	Participants
<p>A cross-linguistic computerized corpus has been compiled to study Spanish learners of English and English native speakers' intonation. The speech of both groups of speakers was digitally recorded while they performed two tasks: reading aloud and interpreting short English conversations. The conversations highlight the contrast between different degrees of certainty and uncertainty.</p>	<p>The data collected in the corpus, totaling over 3 million words, were analyzed acoustically, so that comparable qualitative and quantitative information on the prosodic characteristics produced by the two language user groups could be obtained.</p>	<p>Two homogeneous groups of undergraduate students: a group of 10 Spanish learners of English and a group of 10 native speakers of English.</p>

Acoustics of persuasion

<p>Number 17 Objective: The objective of this study is twofold: (1) To assess whether children’s feelings of perceived judgment are similar to those of adults and (2) to assess whether children use fewer audiovisual cues for uncertainty in a less systematic way. Global design: Production and perception</p>		
Corpus (speech production)	Perceptual evaluation (perception)	Participants
<p>Experiment 1. First, participants were asked a series of questions by the researcher and their responses were filmed using a digital camera. This part of the experiment was set up in such a way that participants could not see the researcher. Second, the same sequence of questions was asked again, for participants to indicate how sure they were that they would recognize the correct answer if they had to find it in a multiple-choice test. Third, the same sequence of questions was asked once more in a multiple-choice paper-and-pencil test, in which the correct answer was mixed with three plausible alternatives.</p> <p>Experiment 2. Detecting uncertainty. 30 adult utterances and 30 child utterances were selected from the corpus of answers collected in the first study. Stimuli were presented on a screen where the judges first saw the stimulus ID (1 through 30) and then the actual stimulus.</p>	<p>Experiment 1. Certain and uncertain speaker utterances from adults and children were collected by asking them a series of factual questions (40 for adults, 30 for children). For adults, the questions came from two sources. First, we selected questions with a single word answer (e.g., “Who wrote Faust?”), and then we added a supplementary list from the Dutch version of the game Trivial Pursuit. For children, we selected questions with a single word answer (e.g., “How much is a dozen?”) only. For both groups questions were selected in a way that different types of responses could be elicited, namely answers, non-answers, certain, and uncertain responses.</p> <p>Experiment 2. Judges were instructed to estimate whether speakers were uncertain about their answers or not. Adult judges scored on a seven-point Likert scale, child judges on a five-point Likert scale (using the same facial representation as above)</p>	<p>Experiment 1. Twenty adults and 21 children participated as speakers. The adults (11 males, 9 females) were colleagues and students from Tilburg University, between 20 and 50 YO.</p> <p>Experiment 2. 80 native speakers of Dutch participated as judges, 40 adults and 40 children (20 male and 20 females per group), and all different from the speakers that participated in the production experiments.</p>

<p>Number 18 Objective: Because atypical instances of a category are more difficult to classify than typical instances, when speakers refer to these instances their lack of confidence will be demonstrated in a paralinguistic way, that is, in the form of hesitations, filled pauses, or rising prosody. These features can help listeners learn by enabling them to differentiate good from bad examples of a category. Global design: Perception</p>		
Corpus (speech production)	Perceptual evaluation (perception)	Participants
<p>Participants were asked to imagine that they were new employees in the stockroom of an Internet clothing store, where their job was to fill orders. A male “expert teacher” trained the participants on the store color scheme. The voice used for the teacher’s sound recordings was that of a male confederate, who was instructed to speak each color name five times with confidence that varied from very unsure to very sure. 95% of the least confident utterances began with a filled pause (e.g., “uh”, “um”, “mm”), whereas none of the most confident ones did. 92% of the least confident utterances also included rising intonation, whereas none of the most confident utterances did.</p>	<p>Listeners learned a set of novel color categories from an expert teacher. For half of the categories, the paralinguistic characteristics of the teacher’s utterances were consistent with the category structure: difficult-to-classify instances (i.e., atypical) were marked by unconfident-sounding utterances, and easy-to-classify instances (i.e., typical), by confident-sounding ones. For the other half, the characteristics were inconsistent with the category structure.</p>	<p>36 undergraduates, 22 females and 14 males.</p>

Dianete Ângela do Valle Gomes, Rodrigo Dornelas, Livia Maria Santos de Souza, Yonara Caetano de Santana Strauss, Eduardo Magalhães da Silva, Letícia Corrêa Celeste

<p>Number 19 Objective: In an experiment with 2x2x2 factorial design, several hypotheses derived from the Elaboration Likelihood Model and from phonetics literature were tested. Global design: Perception</p>		
Corpus (speech production)	Perceptual evaluation (perception)	Participants
<p>Two linguistically similar advertising messages of financial services of high (student loan) versus low (Automatic Teller Machine cards) involvement were recorded by a professional actor using four types of voice (two levels of intonation of voice x two levels of intensity).</p>	<p>Once exposed to the experimental advertisement, subjects were asked to complete a questionnaire. Credibility was operationalized using two sets of four items each, one measuring “internalization” and the other “identification.” “Internalization” was measured by four 7-point bipolar scales of “trustworthy”, “honest”, “competent”, and “believes in what he says”. Identification was measured by four 7-point bipolar scales of “prestigious”, “same culture as me”, “pleasant”, and “attractive”.</p>	<p>229 university students.</p>

<p>Number 20 Objective: Assess the effects of powerful and powerless speech on the impressions of the speaker and on persuasion. Global design: Perception</p>		
Corpus (speech production)	Perceptual evaluation (perception)	Participants
<p>To develop the verbal stimuli, we examined the tapes of actual court trials and chose a trial in which a female witness had given her testimony in a style characterized by frequent use of powerless forms (i.e., hedges, intensifiers, rising intonations, especially formal grammar, and polite forms). Actors then reproduced this original testimony (changing only names, dates, places, and editing certain legal technicalities). A female witness powerful-style tape was also made (using the same actors), in which most of the powerless features were omitted, but the substance of the testimony remained unchanged. The “powerless” style is characterized by the frequent use of linguistic features such as intensifiers, hedges, hesitation forms, and questioning intonations. The “powerful” style is marked by less frequent use of these features.</p>	<p>Subjects were provided either with earphones through which they could hear the stimulus tape (in the oral-presentation conditions) or with envelopes containing one of the four experimental transcripts (in the written-presentation conditions). Subjects rated the witness on a series of 11-point semantic differential-type rating scales with the endpoints: powerful-powerless, competent-incompetent, masculine-feminine, trustworthy not trustworthy, likely able-unlikely able, strong-weak, intelligent-not intelligent, and active-passive. Subjects also used 11-point scales to rate how much they believed the witness, how convincing the witness was, how sympathetic they felt toward the witness, how similar the witness was to them, and how qualified they felt the witness was to testify. Additional questions inquired how responsible and how negligent the defendants were for the victim’s death and how much the defendants should pay the plaintiffs in damages.</p>	<p>152 undergraduate students (73 male and 79 female). Five additional subjects did not fill in the whole questionnaire and were omitted from the analysis.</p>

Acoustics of persuasion

<p>Number 21 Objective: Study the paralinguistic cues in the speaker when coding confidence and doubt in speeches well as the attributes observed in the speaker when exposed to these cues. Globaldesign: Perception</p>		
Corpus (speech production)	Perceptual evaluation (perception)	Participants
<p>The study was conducted in two parts, a recording of linguistically confident texts and another of linguistically doubtful texts. This process generated four texts with two disparate channels and two congruent channels.</p>	<p>Evaluation of the audio recordings to classify the speakers as trustworthy, experts, and legally competent. Textual characteristics were rated on a scale from zero to 100 and speech features were rated on a scale from zero to nine simultaneously, to establish the pairs of effectiveness of manipulation and to evaluate about 23 personality characteristics and features of speech. The recording was classified as “text and voice confident”, “confident text and voice doubtful”, “doubtful text and confident voice”, “doubtful text with doubtful voice”. About 23 personality characteristics and seven features of speech on a scale from zero to nine were also evaluated.</p>	<p>To assess trust or doubt: 10 women college students, paid volunteers, were tested one at a time, using peer comparison techniques. They were asked about which pair expressed greater confidence. The order and the peers were randomized. To assess personality and attributes of voice 47 women from a female school, who were paid volunteers, in 4 sessions, were asked to evaluate the ability of law students to practice.</p>

The variables analyzed in the 21 remaining papers were divided into three settings: “speech temporal organization,” “intensity,” and “pitch”. In all settings there were positive, negative, and neutral effects on persuasion. The most associated variable with a positive effect on persuasion was “increased mean intensity” (n=5), followed by “higher speech rate” or “faster speech rate” and “increased F0 range” or “increased variation in pitch” (n=4 each). The negative effect on persuasion was represented by “high mean F0” (n=6), “lower speech rate,” and “disfluencies” (n=4, each), followed by “rising pitch” or “rising intonation” (n=3). The main findings and conclusions of the studies are summarized in Tables 4 and 5.

Dianete Ângela do Valle Gomes, Rodrigo Dornelas, Lívia Maria Santos de Souza, Yonara Caetano de Santana Strauss, Eduardo Magalhães da Silva, Letícia Corrêa Celeste

Table 4 – Results and conclusions of the 21 remaining papers about prosodic variables of theattitudinal expressions related to persuasion.

N	Results	Conclusion
1	The observational study shows that Members of Congress with lower voices are not more effective leaders, and the experiment shows that speakers with lower voices are not more persuasive policy advocates. Taken together, the two studies suggest that voice pitch is not a reliable signal of leadership ability.	Voice pitch does not correlate with leadership ability. The data presented here supports the hypothesis that voice pitch contains information about leadership ability. Members of Congress with lower voices are not more effective legislators, and speakers with lower voices are not more persuasive when making statements about governmental policies.
2	The study showed that prosody with contrasting foci seem to be performed universally as a pragmatic tool to present new information. In fact, there is congruent evidence in the use of pitch to signal the contrast of focus strategy that is used both in English and Mandarin. In the intonation (emphasis) the focus is marked by the accentuation of the pitch. However, the prosodic structure within the lexical component, makes lexical access easier and determines the lexical identity, rather than the pragmatical interpretation of the statement.	The findings strengthen the notion that both pitch and pitch range function successfully as a cue of prominence. The [+verb] perception task discussed in this paper revealed that perceived confidence of [+verb] constructs correspond to alternative semantic interpretations created by prosodic variation of focus-marking, with direct statements appearing most confident, followed by the main clause, epistemic marker, and finally discourse marker [+verb] variations.
3	Confident voices tend to display increased intensity, and increased variation in both intensity and pitch, whereas doubtful voices are marked by a higher pitch, more pauses, and decreased rate of speech. Neutral statements were significantly lower in pitch and weakest in intensity, varied the least in pitch and intensity, and were spoken faster than all the other expressions that were seen as communicating some level of confidence. Intonation contour (upward versus downward) can play an important role in how expressed confidence is communicated.	The findings provide new information on how metacognitive states such as confidence and doubt are communicated by vocal and linguistic cues which permit listeners to arrive at graded impressions of a speaker's feeling of (un)knowing.
4	For the lexical phrase, the mean F0 was highest when a speaker was unconfident, lower in the close-to-confident condition, and lowest in confident expressions. Mean speaker amplitude was highest in confident expressions and lowest in unconfident expressions. Speech rate was notably slower when speakers were unconfident and was slower in close-to-confident than in confident expressions. The analysis of F0 and amplitude range did not yield significant differences across conditions, although unconfident expressions tended to display the largest F0 range.	The results show the mechanisms of cognitive processing and the individual factors that govern how we infer the speaker's mental state from the speech signal.
5	The study found significant correlation between vocal characteristics and court outcomes and the correlation is specific to perceived masculinity, even when judgement of masculinity is based only on less than three seconds of exposure to a lawyer's speech sample. Specifically, male advocates are more likely to win when they are perceived as less masculine. No other personality dimension predicts court outcomes.	While the study does not aim to establish any casual connections, the findings suggest that vocal characteristics may be relevant in even as solemn a setting as the Supreme Court of the United States.

Acoustics of persuasion

N	Results	Conclusion
6	<p>In general, mean F0 was higher when a speaker was unconfident than in all other conditions. Also, close-to-confident and confident expressions (which did not differ) displayed a higher mean F0 than neutral intending utterances. Regarding amplitude, neutral-intending and confident expressions are highest (which did not differ), and close-to-confident expressions are lowest. The confident expressions showed a larger range of variation than all other expression types. Speech-rate was significantly faster for neutral-intending expressions than all three types of confidence expressions. They were also notably slower when speakers were unconfident than confident or close-to-confident. The analysis of F0 range did not yield significant differences across conditions, although unconfident expressions tended to display the largest F0 range.</p>	<p>The results establish that a listener's brain is rapidly attuned to vocal cues that signal one's feeling of knowing and can differentiate subtle vocal variations as a function of the perceived degree of knowing.</p>
7	<p>The studies showed that high intonation led to a significant drop in intention amongst respondents who perceived their own health as good. After self-affirmation, persuasion was increased.</p>	<p>A high level of intonation seems to induce self-regulatory defenses in people who do not see the necessity to change their health behavior, whereas people with poor perceived health might perceive potential to change. The use of a normal level of intonation in auditory health messages is recommended.</p>
8	<p>Speech therapy and its association with drug treatment promoted the improvement of prosodic parameters, namely increases of F0 measures, reduction of measures of duration, and greater intensity.</p>	<p>Levodopa is effective in improving the duration parameter. As such, its association with speech therapy to improve otherprosodic parameters (F0 and intensity) and to improve communication performance in patients with Parkinson's Disease - thus to improve their quality of life - is noteworthy.</p>
9	<p>The results showed a statistically significant difference between the results of CG and EG judges: judges recognize attitudes expressed by non-stutterers better than those expressed by stutterers.</p>	<p>Fluent speakers succeed in expressing certainty and doubt, while stutterers do not.</p>
10	<p>The expression of doubt has the lowest rate of articulation in the CG, followed by neutral and certainty expressions, with statistically significant differences. Pauses were present in the CG, but disfluencies happened only in the expression of doubt. In the EG, the largest difference was found in the vowel duration of the stressed syllable.</p>	<p>CG showed a more varied temporal organization in the expression of attitudes. However, it's also possible to note a trend in the group of people who stutter. Regarding the speech rate, after removing the pauses and disfluencies, both EG1 and EG2 differentiate certainty, articulating each syllable faster.</p>
11	<p>High incidences of persuasive ability are related to groups with lower speech rates in comparison to groups with faster speech rates.</p>	<p>Analysis of speech rate, gaze, and the listener's sex revealed that when combined with a small amount of gaze, slow speech rate decreased trustworthiness when compared to a fast speech rate. For women, slow speech rate was thought to be indicative of less expertise, when compared to a fast speech rate combined with low gaze. There were no significant interactions, but there were main effects of gaze and speech rate on persuasiveness. High levels of gaze and slow speech rate each enhanced perceptions of the speaker's persuasiveness.</p>

Dianete Ângela do Valle Gomes, Rodrigo Dornelas, Livia Maria Santos de Souza, Yonara Caetano de Santana Strauss, Eduardo Magalhães da Silva, Letícia Corrêa Celeste

N	Results	Conclusion
12	<p>In Experiment 1A, listeners rated speakers as less accurate when speakers prolonged syllables the end of the first utterance in the pair. Accuracy ratings were unaffected by rising pitch at the end of the first utterance in the pair. In Experiment 1B, listeners rated speakers as less certain when speakers prolonged syllables at the end of the first utterance in the pair. Certainty ratings were unaffected by rising pitch at the end of the first utterance in the pair. In Experiment 2 words were monitored for speed when they followed sentences ending in prolonged rising pitch. Words were slower when there were sentences ending in non-prolonged rising pitch. In Experiment 3 the interaction between rising pitch and prolongations depended on the listeners beliefs about the speaker's mental state.</p>	<p>Results support the theory that temporal and situational context are important in determining intonational meaning. When listeners believed that speakers were non-experts, they interpreted rising pitch and prolongations similarly to when they were provided with no information about speakers. However, when listeners thought that speakers were experts, their word monitoring was no longer affected by rising pitch or prolongations. This is strong evidence that listeners establish relationships between linguistic form and function by first presupposing a speaker's mental state.</p>
13	<p>The study indicates that simply raising the speech rate from normal to high does not result in greater persuasiveness. Such manipulation alone brought no gain for an audience of native and non-native listeners, as raising the speech rate above 178 wpm is likely to reduce comprehension even further for all listeners. Increase in speech rate is not worth pursuing to increase persuasiveness.</p>	<p>A faster speech rate lowers comprehension for both native and non-native listeners and does not influence persuasiveness. There was a non-significant tendency, however, for the usefulness of products, where the product description was delivered at a faster speech rate. There was no difference between native and non-native listeners on the effect of speech rate on persuasion.</p>
14	<p>Analysis indicated a moderate intensity, an unmarked intonation, and a fast speech rate associated with a more credible source than other combinations. Sex was not a significant moderator in the relationship between voice characteristics and source credibility.</p>	<p>Voice characteristics significantly affected attitude towards the advertisement and behavioral intention.</p>
15	<p>High confidence was expressed with a lower mean F0 than low confidence, and both high and moderate confidence utterances displayed a reduced F0 range, when compared to low confidence utterances. High confidence was expressed with a faster speech rate than moderate and low confidence for the prosody stimuli.</p>	<p>Although emotional and attitudinal meanings of prosody seem to implicate the right hemisphere in a mandatory way, perhaps at multiple processing levels, the findings are inconclusive regarding the probable input of the left hemisphere in these processes. This bias meant that when prosodic cues were continuously encoded in speech and weighed heavily for assigning speaker attitudes (Experiment 1) or signaled a non-literal interpretation of the verbal message (Experiment2), right hemisphere damage patients failed to fully appreciate the attitude of the speaker in these situations. "When speakers used a high/rising rather than a falling intonation contour, requests were invariably perceived as more polite by all participants in the study".</p>
16	<p>Spanish learners avoid the falling-rising tone to express lack of certainty, using a falling contour in all contexts instead. The intonation of the non-native speakers does not contribute to a clear distinction between certainty and uncertainty. Their falling tones are very shallow, which does not help towards discerning any absolute certainty in their responses. Instead, the responses are realized with a narrow falling pitch range and mid-level flat contours. On the gradient from strong certainty, realized with a wide falling tone, to high uncertainty, signaled by a falling-rising tone, the tones produced by non-native speakers range from narrow falling to level, just in the middle of the intonational scale described.</p>	<p>The results demonstrate that non-native English intonation, as to nature and patterning, differs from native intonation when it comes to expressing modality. The results reveal that the Spanish speakers' choice of the English tone system may lead to pragmatic incompatibility in the expression of modality in their interactions. The intonation of the non-native speakers does not contribute to a clear distinction between certainty and uncertainty. Their falling tones are very shallow, which does not help discerning any absolute certainty in their responses.</p>

Acoustics of persuasion

N	Results	Conclusion
17	Results show that when adult speakers were uncertain, they were more likely to produce fillers, delays, high intonation, eyebrow movements, and “funny faces”.	It was found that both children and adults judges rated with more accurate scores answers from adult speakers than from child speakers. Child judges provided less accurate scores than adult judges. No meta-cognitive differences were found between adults and children. It appears that adult speakers use audiovisual prosody to rate their level of uncertainty in a more consistent and clearer way than child speakers.
18	Paralinguistic expressed confidence was consistent with the underlying category structure. When speakers showed such characteristic, learners acquired the categories more rapidly and showed better category differentiation from the earliest moments of learning. 95% of the least confident utterances began with a filled pause, whereas none of the most confident ones did. 92% of the least confident utterances also included rising intonation, whereas none of the most confident utterances did show such intonation.	Listeners’ ability to exploit paralinguistic cues of uncertainty in learning novel concepts has important implications for theories of language use, categorization, and language learning.
19	Analysis by a system of simultaneous equations indicated that the effects of voice are different under low and high involvement. Voice intensity affects credibility of the source significantly more under low than high involvement. Voice intonation affects credibility more under high than low involvement.	Voice characteristics affect attitudes toward the advertised service and the intent to buy.
20	The use of a powerful style resulted in greater attraction to the witness, regardless of his/her sex or the subject, or the mode of the testimony presentation. A powerful style also resulted in greater perceived credibility of the witness than did a powerless style. However, this effect was stronger when the subject and the witness were of the same sex than when they were of the opposite sex. For the male witness-written presentation condition, a powerful style produced more acceptance of the position advocated in the testimony than did a powerless (rise intonation) style.	The results discussed the possible relationships between speech styles and person perception and persuasion processes, regarding the social psychology of legal issues.
21	Confident voice condition may indicate that paralinguistic expression of confidence is achieved, in part, through selective stress of strategic linguistic units. In the confident voice condition, the speaker used pauses less frequently and such pauses were shorter. The means suggest that the speaker used substantially longer pauses to express doubtful voice under confident text conditions, but this interaction approaches significance only for between-sentence pauses. There were no significant differences with respect to number or duration of pauses for the text variable. Findings suggest that pauses of various sorts are indicative of cognitive tension states, relevant to them, and may indicate some of the correlates of the expressed confidence and doubt.	The study suggests that confidence is expressed in a paralinguistic way by increased voice loudness, higher pitch level (under certain conditions), shorter pauses, and a faster speech rate.

Dianete Ângela do Valle Gomes, Rodrigo Dornelas, Livia Maria Santos de Souza, Yonara Caetano de Santana Strauss, Eduardo Magalhães da Silva, Letícia Corrêa Celeste

Table 5 – Prosodic variables analyzed by the 21 remaining papers and their effects on persuasion.

Setting/Variables	Effect on persuasion		
	Positive	Negative	Neutral
<i>Speech Temporal Organization</i>			
Higher/faster speech rate	4		3
Lower speech rate	1	4	1
Pauses		2	
Short pauses	1		
Long pauses		1	
Disfluencies ¹		4	
<i>Intensity</i>			
Moderate intensity ²	1		
Lower mean amplitude ²		1	
Increased mean intensity ³	5		1
Increased variation in intensity ³	2		
<i>Pitch</i>			
Rising pitch/intonation ⁴		3	
High Mean F0 ⁴	1	6	
Increased F0 range/variation in pitch ⁴	4	2	1
Decreased F0 range/unmarked intonation ⁵	1		
Lower mean F0 ⁵	2	1	1
Wide falling tone (pitch range) ⁵	1		
Falling-rising tone (pitch range) ⁵		1	

¹Including filled-pauses and stuttering-like disfluencies; ²Decrease in intensity; ³Increase in intensity; ⁴Increase in F0; ⁵Decrease in F0.

Interestingly, the effects of some variables on persuasion were presented as positive and negative (“lower speech rate” and “high mean F0”), or as positive and neutral (“higher” or “faster speech rate” and “increased mean intensity”), or as any of the options (“lower speech rate”, “increased F0 range” or “increased variation in pitch” and “lower mean F0”).

4. Discussion

Researchers have investigated the variables that transmit communicative attitudes to listeners and signal whether something is important. Voice conveys information beyond the content of the communication, since intentions are often associated with differences in prosodic and vocal characteristics (Guyer et al., 2018).

The aim of this study was to identify and analyze the acoustic prosodic features linked to persuasion in communicative attitudes. It is noteworthy that the data are controversial, which can probably be explained by social and ethnographic contexts that can affect an audience's perception. However, while the oral component in different speech settings is analyzed, vocal characteristics are not always present (Guyer et al., 2018).

It is commonly accepted that effective communication skills are the key to successful outcomes, and online papers on the subject have been available since the seventies. We found more than 600 publications that fit the search criteria, but only 21 fit the specified criteria for this study. The 21st century enlarged these numbers, but it is possible to affirm that studies began in the 1970s and the 1980s brought no contributions, even though publishing progressively increased. Despite increased access to different technologies for analysis, the numbers are not impressive until the last decade, when they double.

While Canada has had a consistent production over the years, Brazil is the country in Latin America with the largest number of publications, which clearly highlights the importance of encouraging and reinforcing studies within this scope, as in this region there are different spoken languages, different vocal psychodynamics and cultures, which should be respected in this kind of study. This is also the case in Africa, Asia, and Oceania, all of them with little published research on the subject.

The findings from this review, summarized in Tables 3, 4, and 5, suggest that the studies were conducted under different methodologies and focused on different variables. Interestingly, most of the studies (n=19) evaluated the perceptive aspects of persuasion and confidence. Despite their constant use in different studies, the methods for collecting and analyzing prosody perceptual data on the expression of attitude are very different, which may insert a bias in performing meta-analyses on the subject. Regarding production, it is possible to state that there are relevant prosodic characteristics in understanding confidence and persuasion. This review found that confidence was related to shifts in pitch, loudness, and speech rate (temporal organization).

Expressions of confidence are reflected by the relationship between increased intensity, short pauses, fast speech rate, and melodic variation.

Dianete Ângela do Valle Gomes, Rodrigo Dornelas, Livia Maria Santos de Souza, Yonara Caetano de Santana Strauss, Eduardo Magalhães da Silva, Letícia Corrêa Celeste

These indicators were first published in 1973 (Scherer et al., 1973) and are still present in current research, which leads us to conclude that the paralinguistic cues of confidence are manifested in the behavior of the speaker and are used by the listener to make inferences about the speaker's attitudinal state, from certainty to doubt. Moreover, due to the differences in methodology and speaker samples and manipulations, this does not seem to be an artifact of the study.

It is curious that the same variables can be considered as positive, negative, or neutral to characterize confidence and persuasion, namely, lower speech rate, increased F0 range or increased variation in pitch and lower mean F0. These data are independent whether the evaluation is conducted with native or non-native speakers, graduate or undergraduate students, or different professional profiles, whereas genre seems to have some effect. Despite the presence of neutrality statements for these variables, it is not clear if the temporal organization aspect of the discourse is a determinant factor for persuasion.

It is also not clear whether the psychoacoustic effects on the listener caused by an increase of energy, which is not present in the studies included in this research, can be questioned. Consequently, studies that correlate the intensity variable alone in the expression of communicative attitudes related to persuasion and to the listener's perception are necessary.

In a study with the objective of verifying the adjustments of voice quality and vocal dynamics performed by voice actors in an activity, it was noticed, through the application of a questionnaire, that the spectral inclination emerged was characterized as breathiness in the expression of persuasion and tension, in the contexts of conflict between the characters, which differs from the aspect of neutrality found in the other study cited here (Crochiquia, et al. 2020). It was shown that the listener can consider the relevance of pauses on several levels, such as silence and sound relation, elocution time, location, and form and duration of the silences, as they can be considered paralinguistic cues related to the stress strategy of the sentence related to persuasion.

A survey showed that the speech rate of Portuguese speakers differs notably from that of Brazilian speakers when both groups narrate, but not when both groups read. This was demonstrated through the linear

correlation between the range of durations delimited by smoothed punctuation points and the number of VV units in the same range (Barbosa et al., 2016).

There is a close relationship between linguistic cues, i.e., “auditory forms” and linguistic function - namely declarative, direct, indirect, and quite indirect (Pell, 2007) -, gestural cues (Yokoyama & Daibo, 2012), and place of emphasis within the sentence or word. The emphasis on the verb, for example, cannot only modify confidence in the message, but also completely modify its meaning (Tomlinson Jr. & Fox Tree, 2011; Zhao et al., 2018).

Increase in pitch was mentioned by most of the studies analyzed (n=11). However, worsening perception of speaker’s persuasion and pitch variation work as a cue to the prominence of the utterance focus (Zhao et al., 2018). From a neurophysiological perspective, the studies (Jiang & Pell, 2015; Pell, 2007) show that melodic (vocal and pitch) variation is perceived by the brain in association with other cues, leading to an inference about the status of the speaker, which would signal the level of confidence and certainty of speech. Thus, the studies indicate that there are mechanisms of cognitive processing and individual variations in the interpretation of acoustic cues related to the mental state of the speaker from the speech signal.

Despite the classification of prosodic variables as positive or negative, it is noteworthy that the communicative context and the predisposition of the speaker in the message should be considered (Elbert & Dijkstra, 2014; Verdugo, 2005), and some directions still need to be followed to better understand the complexity of the relationship between prosody and the expression of attitudes.

Image can be an important factor in the analysis of persuasion. In a survey that analyzed the perceptual audiovisual identification and the acoustic and visual production of four speech acts in Brazilian Portuguese, it was found that speech acts are identified with higher levels of identification in the audiovisual condition, and then in the condition’s audio only and video only. The automatic recovery showed that this result is a possible manifestation of attitudes that may accompany the production of the echo question, such as incredulity or

Dianete Ângela do Valle Gomes, Rodrigo Dornelas, Livia Maria Santos de Souza, Yonara Caetano de Santana Strauss, Eduardo Magalhães da Silva, Letícia Corrêa Celeste

strangeness, which it incorporated to separate such productions from those produced in a neutral way (Miranda et al., 2020).

There is no current consensus regarding the analysis and assessment of prosodic variables and their subgroups. Although several of them have been studied, just a few have emerged as relevant to persuasion. Differences of theoretical basis and methodological divergences have made a meta-analysis of our results impossible. Nevertheless, the systematization of these studies has shown the need to expand investigation in the area with greater interaction between research centers and theoretical and methodological approaches. It should be noted that no clinical or educational setting was present in the analyzed studies, probably because of the difficulty of having a controlled environment that involves other speaker's expectations.

Another limitation of the studies presented here is the intrinsic variation present in each culture, each population studied, and each methodological proposition, as they were conducted in different parts of the world and had to consider their internal references in the analysis. In addition, there were different focuses in showing and evaluating persuasion, as it could be accessed in synthesized and non-synthesized speeches. In the latter case, the speaker may be "imitating" some trained characteristics, which can trigger different results or interpretations by introducing variations in the measurement that may account for some variability across findings.

All the above mentioned is indicative of the complex and multifaceted nature of communication performance and highlights the need for future research to address some of the difficulties in defining and measuring confidence and persuasion. On the other hand, it could be noticed that all studies were structured and protocol-driven in their methodology, despite leading with subjective measures. This indicates that this is a helpful construct in the development of communicative performance, and in public speaking training, since there has not been much research into how to effectively do it (Celeste et al., 2018).

5. Conclusion

It was impossible to define a single variable that determines the level of speaker confidence. The papers accessed and evaluated suggest, though, that there is a relationship between some relevant variables, such as pitch, intensity, temporal organization, and linguistic stress. There is no consensus in defining and considering each variable alone as a prosodic indicator or characteristic by all researchers. Two variables emerged as relevant and characteristics of persuasive speech, namely faster speech rate and increased intensity. In the development of new studies on the subject, it is important to define variables and apply them in training both raters and speakers to perceive their uses in speaking.

Conflict of interests

The authors declare they have no conflict of interest.

Credit Author Statement

We, Dianete Ângela do Valle Gomes, Rodrigo Dornelas, Livia Maria Santos de Souza, Yonara Caetano de Santana Strauss, Eduardo Magalhães da Silva and Leticia Corrêa Celeste, hereby declare that we do not have any potential conflict of interest in this study. We have all participated in the study conceptualization, methodology, study design, formal data analysis, statistical data analysis, project administration, project supervision, data collection, data generation, data validation and editing. All authors approve the final version of the manuscript and are responsible for all aspects, including the guarantee of its veracity and integrity.

References

- Azevedo, I. L., Souza, I. S., Oliveira, P. M., & Cardoso, F. (2015). Effect of speech therapy and pharmacological treatment in prosody of parkinsonians. *Arquivos de Neuro-Psiquiatria*, 73(1), 30-35. <https://doi.org/10.1590/0004-282X20140193>.
- Barbosa, P. A., Camargo, Z. A., & Madureira (2016). Scripts for the Acoustic Analysis of Speech Data. In S. Madureira (Ed.), *Sonoridades [recurso eletrônico]: a expressividade na fala, no canto e na declamação [Sonorities [eletronic device]: speech, singing and reciting expressivity]*. Pontifícia Universidade Católica de São Paulo.

Dianete Ângela do Valle Gomes, Rodrigo Dornelas, Livia Maria Santos de Souza, Yonara Caetano de Santana Strauss, Eduardo Magalhães da Silva, Letícia Corrêa Celeste

- Barnard, M. (2013). *Fashion as communication*, 2nd ed. Routledge. <https://doi.org/10.4324/9781315013084>.
- Barnard M. (2014). *Fashion theory: an introduction*. Loughborough: Routledge. <https://doi.org/10.4324/9780203862100>.
- Barr, D. J. (2003). Paralinguistic correlates of conceptual structure. *Psychonomic on Bulletin & Review*, 10(2), 462-467. <https://doi.org/10.3758/BF03196507>.
- Blakemore, D. (1987). *Semantic constraints on relevance*. Basil Blackwell.
- Celeste, L. C., Lima, A. M., Seixas, J. M. A., Silva, M. A., & Silva, E.M. (2018). Treinamento da performance comunicativa em universitários da área da saúde. *Audiology - Communication Research*, 23. <https://doi.org/10.1590/2317-6431-2017-1879>.
- Celeste, L. C., & Reis, C. (2013a). Expressão de atitudes na fala com gagueira: percepção de falantes fluentes. *Alfa: Revista de Linguística*, 57, 275-29. <https://doi.org/10.1590/S1981-57942013000100012>.
- Celeste, L. C., & Reis, C. (2013b). Expressão de certeza e dúvida na gagueira: estudo dos aspectos temporais da fala. *Revista CEFAC*, 15, 1609-1620. <https://doi.org/10.1590/S1516-18462012005000060>.
- Chebat, J. C., El Hedhli, K., Gelinac-Chebat, C., & Boivin, R. (2007). Voice and persuasion in a banking telemarketing context. *Perceptual Motor Skills*, 104(2), 419-437. <https://doi.org/10.2466/PMS.104.2.419-437>.
- Chen, D., Halberstam, Y., & Yu, A. C. (2016). Perceived Masculinity Predicts U.S. Supreme Court Outcomes. *PLoS One*, 11(10), e0164324. <https://doi.org/10.1371/journal.pone.0164324>.
- Chomsky, N. (1975). *Reflections on Language*. Random House.
- Couper-Kuhlen, E. (1986). *An introduction to English prosody*. Edward Arnold.
- Crochiquia, Alice, Eriksson, Anders, Fontes, Mario A. S., & Madureira, Sandra. (2020). A phonetic study of Zootopia characters' voices in Brazilian Portuguese dubbing: the role of stereotypes. *DELTA: Documentação de Estudos em Linguística Teórica e Aplicada*, 36(3), 2020360311. <https://doi.org/10.1590/1678-460X2020360311>.
- Elbert, S. P., & Dijkstra, A. (2014). An experimental test of the relationship between voice intonation and persuasion in the domain of health. *Psychology & Health*, 29(9), 1014-1031. <https://doi.org/10.1080/08870446.2014.903482>.
- Erickson, B., Lind, E. A., Johnson, B. C., & O'Barr, W. M. (1978). Speech style and impression formation in a court setting: The effects of "powerful" and "powerless" speech. *Journal of Experimental Social Psychology*, 14(3), 266-279. [https://doi.org/10.1016/0022-1031\(78\)90015-X](https://doi.org/10.1016/0022-1031(78)90015-X).
- Fónagy, I. (2003). Des fonction de l'intonation: Essai de synthèse. *Flambeau*, 29, 1-20.

- Gelinas-Chebat, C., & Chebat, J. C. (1996). Voice and advertising: effects of intonation and intensity of voice on source credibility, attitudes toward the advertised service and the intent to buy. *Perceptual and Motor Skills*, 83(1), 243-262. <https://doi.org/10.1080/08870446.2014.903482>.
- Guyer, J. J., Fabrigar, L. R., & Vaughan-Johnston, T. I. (2018). Speech Rate, Intonation, and Pitch: Investigating the Bias and Cue Effects of Vocal Confidence on Persuasion. *Personality and Social Psychology Bulletin*, 45(3). <https://doi.org/10.1177/0146167218787805>.
- Hardcastle, W. J., Laver, J., & Gibbon, F. E. (2013). *The handbook of phonetic sciences*, 2nd ed. Wiley-Blackwell.
- Jiang, X., & Pell, M. D. (2015). On how the brain decodes vocal cues about speaker confidence. *Cortex*, 66, 9-34. <https://doi.org/10.1016/j.cortex.2015.02.002>.
- Jiang, X., & Pell, M. D. (2016). Neural responses towards a speaker's feeling of (un)knowing. *Neuropsychologia*, 81, 79-93. <https://doi.org/10.1016/j.neuropsychologia.2015.12.008>.
- Jiang, X., & Pell, M. D. (2017). The sound of confidence and doubt. *Speech Communication*, 88, 106-126. <https://doi.org/10.1016/j.specom.2017.01.011>.
- Jones, C., Berry, L., & Stevens, C. (2007). Synthesized speech intelligibility and persuasion: Speech rate and non-native listeners. *Computer Speech & Language*, 21(4), 641-651. <https://doi.org/10.1016/j.csl.2007.03.001>.
- Klofstad, C. A., & Anderson, R. C. (2018). Voice pitch predicts electability but does not signal leadership ability. *Evolution and Human Behavior*, 39(3), 349-354. <https://doi.org/10.1016/j.evolhumbehav.2018.02.007>.
- Krahmer, E., & Swerts, M. (2005). How children and adults produce and perceive uncertainty in audiovisual speech. *Language and Speech*, 48(Pt 1), 29-53. <https://doi.org/10.1177/00238309050480010201>.
- Malta, M., Cardoso, L. O., Bastos, F. I., Magnanini, M. M. F., & Silva, C. M. F. P. D. (2010). STROBE initiative: guidelines on reporting observational studies. *Revista de Saúde Pública*, 44(3), 559-565. <https://doi.org/10.1590/S0034-89102010000300021>.
- Mehrabian, A. (1969). Significance of posture and position in the communication of attitude and status relationships. *Psychological Bulletin*, 71(5), 359-372. <https://doi.org/10.1037/h0027349>.
- Miranda, L. da S., de Moraes, J. A., & Rilliard, A. (2020). Percepção audiovisual da entoação modal do português do Brasil. *Gradus - Revista Brasileira De Fonologia De Laboratório*, 5(1), 47-70. <https://doi.org/10.47627/gradus.v5i1.148>.

Dianete Ângela do Valle Gomes, Rodrigo Dornelas, Lívia Maria Santos de Souza, Yonara Caetano de Santana Strauss, Eduardo Magalhães da Silva, Letícia Corrêa Celeste

- Mozziconacci, S. (2001). *Emotion and attitude conveyed in speech by means of prosody*. Sonthofen.
- Pell, M. D. (2007). Reduced sensitivity to prosodic attitudes in adults with focal right hemisphere brain damage. *Brain and Language*, 101(1), 64-79. <https://doi.org/10.1016/j.bandl.2006.10.003>.
- Petty, R., & Cacioppo, J. (1986). The elaboration likelihood model of persuasion. In *Communication and persuasion: Central and peripheral routes to attitude change*. New York: Springer-Verlag, 1-24.
- Petty, R. E. (1996). *Attitudes and persuasion: classic and contemporary approaches*. Routledge. <https://doi.org/10.4324/9780429502156>.
- Pouw, W. T. J. L., Van Gog, T., Zwaan, R. A., & Paas, F. (2017). Are gesture and speech mismatches produced by an integrated gesture-speech system? A more dynamically embodied perspective is needed for understanding gesture-related learning. *Behavioral and Brain Sciences*, 40(68). <https://doi.org/10.1017/S0140525X15003039>.
- Ramírez Verdugo, D. (2005). The nature and patterning of native and non-native intonation in the expression of certainty and uncertainty: Pragmatic effects. *Journal of Pragmatics*, 37(12), 2086-2115. <https://doi.org/10.1016/j.pragma.2005.02.012>.
- Scherer, K. R., London, H., & Wolf, J. J. (1973). The voice of confidence: Paralinguistic cues and audience evaluation. *Journal of Research in Personality*, 7(1), 31-44. [https://doi.org/10.1016/0092-6566\(73\)90030-5](https://doi.org/10.1016/0092-6566(73)90030-5).
- Tomlinson Jr., J. M., & Fox Tree, J. E. (2011) Listeners' comprehension of uptalk in spontaneous speech. *Cognition*, 119(1), 58-69. <https://doi.org/10.1016/j.cognition.2010.12.005>.
- Wilson, D., & Wharton, T. (2006). Relevance and prosody. *Journal of Pragmatics*, 38(10), 1559-1579. <https://doi.org/10.1016/j.pragma.2005.04.012>.
- Yokoyama, H., & Daibo, I. (2012). Effects of gaze and speech rate on receivers' evaluations of persuasive speech. *Psychological Reports*, 110(2), 663-676. <https://doi.org/10.2466/07.11.21.28.PR0.110.2.663-676>.
- Zhao, L., Dehé, N., & Murphy, V. A. (2018) From pitch to purpose: The prosodic-pragmatic mapping of [I+verb] belief constructions in English and Mandarin. *Journal of Pragmatics*, 123, 57-77. <https://doi.org/10.1016/j.pragma.2017.10.015>.

Recebido em: 10.06.2019

Aprovado em: 21.05.2021