

Invading the hearer's auditory space. A prosodic analysis of Italian insults

Patrizia Sorianello

University of Bari, Italy

patrizia.sorianello@uniba.it, ORCID iD: <https://orcid.org/0000-0002-6632-0555>

Giovanni Vinciguerra

University of Bari, Italy

giovanni.vinciguerra@uniba.it, ORCID iD: <https://orcid.org/0009-0000-9473-5566>

Submitted: 02/07/2024; Accepted: 07/01/2025; Published online: 09/05/2025

Citation: Sorianello, P. & Vinciguerra, G. (2025). Invading the hearer's auditory space. A prosodic analysis of Italian insults. *Loquens*, 12, e118. <https://doi.org/10.3989/loquens.2025.e118>.

ABSTRACT: Impoliteness is a broad field of research that also encompasses insults, which are forms of verbal attack that involve a degree of injury and social censure. From a linguistic standpoint, insults are characterised by distinctive verbal and non-verbal cues, including specific phonetic and prosodic alterations. The aim of this research was to analyse, for the first time, the prosody of Italian insults using a spectro-acoustic approach. The study utilised data from 15 oral role plays performed by six professional theatre actors. The objective of the study was to investigate the extent to which duration, intensity, and pitch range are employed to convey the specific meaning of an invasion of the listener's auditory space. The analysis demonstrated that insults exhibited longer nuclear vowels and higher intensity values than the control speech sample. Conversely, the pitch range of the target utterances was narrowed, with the majority of f_0 values situated within the high-frequency range.

Keywords: impoliteness; insults; prosody; pitch range; Italian language.

RESUMEN: *Invadir el espacio auditivo del oyente. Un análisis prosódico de los insultos en italiano.* La descortesía es un campo de investigación amplio que también abarca los insultos, que son formas de violencia verbal que implican cierto grado de injuria y censura social. Los insultos nunca son expresiones neutrales y siempre muestran alteraciones fonéticas y prosódicas. El objetivo de esta investigación fue analizar por primera vez la prosodia de los insultos italianos utilizando un enfoque espectro-acústico. El estudio utilizó datos de 15 juegos de rol orales realizados por seis actores de teatro profesionales. El objetivo del estudio fue investigar en qué medida se emplean la duración, la intensidad y el campo tonal para transmitir el significado específico de una invasión del espacio auditivo del oyente. El análisis demostró que los insultos exhibieron duraciones más largas y valores de intensidad más altos que la muestra de habla de control. Por el contrario, el campo tonal de las expresiones objetivo se redujo, con la mayoría de los valores de f_0 situados dentro del rango de alta frecuencia.

Palabras clave: descortesía; insultos; prosodia; campo tonal; lengua italiana.

Copyright: © 2025 CSIC. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License.

1. INTRODUCTION

The complexity and diversity of insults make them a challenging subject to analyse from multiple disciplinary perspectives. This is due to the fact that insults encompass social, psychological, legal, linguistic, and other aspects of human life. Over the years, and with particular reference to communication approaches, insults have been a significant area of interest from a pragmatic standpoint.

In the domain of speech impoliteness, various forms of pejorative language coexist, including insults, accusations, criticisms, threats, slurs, swear, and taboo words. Distinguishing between these speech acts can prove challenging, as their boundaries are often gradational and overlapping, and the theoretical frameworks used to delineate them exhibit considerable variability. Within the wide context of impolite speech acts, insults represent merely one of several negative acts aimed at demeaning the target. They represent a form of aggressive verbal communication, including a broad spectrum of expressions intended to cause emotional and psychological distress, and to damage the reputation of a given target (either an individual or a group), often leading to social conflict and physical confrontation. Insults are a manifestation of linguistic impoliteness or verbal aggression, as they always involve a high degree of social transgression and censorship (Moïse, 2006). The issue becomes more complex when it is recognised that insults can be considered as both a form of impolite speech and, in particular contexts, as a form of hate speech. This phenomenon occurs when insults are directed towards a specific social group on the basis of race, religion, sexual or political orientation or other identitarian factors. Such insults are frequently associated with expressions of prejudice, negative stereotyping and incitement to violence or hate (Culpeper, 2021; Guillén-Nieto, 2023). Consequently, insults can be distinguished by their level of social acceptance or violation, as well as by their degree of semantic offensiveness.

Over time, scholars have proposed various taxonomies of insults. For instance, Jucker & Taavitsainen (2000, p. 74) distinguished insults based on the following criteria: formal degree, context-dependence, semantics, speaker-attitude and reaction. Similarly, Mateo & Yus (2013) provided a detailed categorisation of insults, encompassing three key aspects: intention, interpretation, and reaction.

According to Culpeper (1996, p. 357), insults can be regarded as a form of positive impoliteness output strategies, with the purpose of threatening the positive face of the speaker. In the taxonomy of impoliteness formulae proposed by Culpeper (2011a, pp. 135-136; 2021) and van Olmen *et al.* (2023, pp. 24-25), insults can be categorised as relating to quality, relational and social identity face, highlighting their social nature.

The etymology of the word already suggests a negative content. The term derives from the Latin verb IN+SULTĀRE, which literally means "to jump on," and is metaphorically used to convey the act of attacking. Consequently, insults are often associated with words that allude to physical violence, such as strokes, punches, slaps, or bullets.

In the field of pragmatics, insults present a number of interesting aspects that have been the focal point of extensive research (see, *inter alia*, Jucker & Taavitsainen, 2000; Mateo & Yus, 2013; Alfonzetti, 2017). Studies have identified a number of subtypes of insults, including pure unfactual insults, calumnies, and factual put-downs (Feinberg, 1985).

From a more general pragmatic perspective, doing away with possible subtyping, insults can be considered as a hybrid phenomenon, combining the illocutionary force of a verdict and that of an expressive act (Jucker & Taavitsainen, 2000; Moïse, 2006; Alfonzetti, 2017). This hybrid nature is attributed to the fact that insults convey a negative evaluation of an individual or objects. Moreover, they are known to elicit a perlocutionary effect that exploits a speech face-attack to immediately produce a form of violence and denigrate the offended person. The act of insult is defined as a speech act in which both the speaker (the insulter) and the addressee (the insultee) assume active roles. Insulting actions are typically deliberate and targeted towards a specific individual by the speaker¹. In contrast, the addressee of the insult is not merely the recipient of the negative assessment; they also play an active role in evaluating the offensiveness of the act by implementing various strategies, ranging from no reaction to the insult (silent and ignoring strategies) to counter-insult. This latter reaction frequently gives rise to a verbal escalation².

Furthermore, in the absence of a verb that has an explicit performative function (e.g. **I insult you*), insults are typically expressed through the use of the so called pejorative language. This is defined as the use of lexemes with a negative semantic value, which are drawn from semantic fields that are often considered taboo (e.g. sex, bodily fluids, genitals, etc.). However, recent studies have shown that any word can be employed with a dysphemic value, when the speaker uses specific facial expressions, aggressive gestures, body postures, and certain prosodic and paralinguistic signals. In spoken language, a range of prosodic strategies, in conjunction with specific lexical choices, intensifying modifiers and morpho-syntactic structures, are used to increase the intensity of insults (Culpeper, 2011b). In contrast, in written and computer-mediated discourse (CMD), graphic devices, such as emojis, emoticons and the inappropriate use of capital letters, or punctuation marks, are frequently employed in conjunction with or as substitutes for text as symbols of affective states³.

¹ The reference to intentionality is indeed a crucial aspect. Culpeper (2021) points out that not all forms of speech impoliteness are intentional.

² For further insight into this topic, see Jucker & Taavitsainen (2000) and Ilie (2001).

³ These symbols perform a range of functions such as the expression of emphasis and emotion, the emulation of direct and spoken speech and they also serve as prosodic modifiers, indicating the intention and the emotion conveyed (for Italian language, see Vinciguerra, 2023a).

2. PROSODIC STUDIES

The majority of studies on impolite speech acts have focused on pragmatic strategies, neglecting the analysis of prosody for many years. Due to their peculiar pragmatic nature, insults are never neutral speech acts, since their expressive and derogatory force is conveyed through both form and content. In spoken language, insults are characterised by a substantial alteration of all acoustic parameters. Nevertheless, there is a surprising lack of studies on the prosody of insults, probably due to the difficulty of collecting them in spontaneous contexts. However, extensive research on linguistic (im)politeness and on the acoustic correlates of emotions offers valuable insights.

In the field of the impoliteness theory, Culpeper (1996; 2011a; 2011b) and Culpeper *et al.* (2003) conducted pioneering research on impolite prosody, focusing on exploitative TV shows and quizzes or intercepted phone-calls. The researchers demonstrated the feasibility of employing impoliteness strategies, whether negative or positive, through the utilisation of prosody alone. They identified several strategies, the first of which is employed to hinder the linguistic process, as it serves to interrupt the interlocutor's turn in a dialogue, conveying a sense of speech closure and dissociation. In terms of intonation, such utterances are characterised by high f_0 onsets and falling intonation contours.

The second strategy involves invading the auditory space of the listener by increasing the volume and the pitch range of the voice. It is important to note that speakers are sensitive to the perception of appropriate levels of voice in everyday interactions. In interpersonal communication, the production of a speech that exceeds the necessary level of loudness, given the physical distance between interlocutors, is often perceived negatively and may be interpreted as an auditory invasion. The use of a high-pitched voice in communication can be considered an attempt to openly express one's emotional state to another individual during the course of an interaction. This strategy can be employed as standalone act of rudeness or, when combined with other strategies such as insults, as a means of increasing the offensiveness of an utterance⁴.

Nevertheless, Culpeper *et al.* (2003) posit an observation pertaining to the predictability of prosodic cues. They further note that the association between these indices and pragmatic meanings is frequently contradictory. For instance, in some languages, certain phonetic cues are employed to convey polite nuances, while in others, these cues are utilised to convey impolite meanings⁵. With regard to this aspect, research has shown that the pitch range is one of the most problematic acoustic parameters when comparing the prosodic cues of polite and impolite utterances (Brown & Prieto, 2017; Caballero *et al.*, 2018), as well as distinguishing between genuine and mock impoliteness (McKinnon & Prieto, 2014; Andreeva *et al.*, 2016). This can be attributed to the gradient and expressive nature of tonal excursion (both in level and span), as well as its direct involvement in the experience of emotions.

In some languages, a wider pitch range may be perceived as more polite to the listener, as is the case in Dutch, English (Chen *et al.*, 2004), Canadian English (Caballero *et al.*, 2018), and Japanese (Ohara, 2001). Conversely, a narrower pitch extension tends to convey rudeness (Uldall, 1960). In contrast to this trend, Stadler (2007) demonstrated that high pitch can convey aggressive attitudes in German and New Zealand English. Furthermore, Winter & Grawunder (2012) found a correlation between a narrow pitch range and the use of a polite register in Korean speakers. Finally, Andreeva *et al.* (2016) found no significant differences in pitch range when comparing genuine and mock utterances. The researchers employed both production and perception texts to compare four utterances with an offensive or supportive attitude produced in German and Polish. The two languages seem to select different prosodic indices, e.g. mean intensity is used to distinguish between the two attitudes in Polish, but not in German⁶. Conversely, the opposite is true for speech rate, with German speakers exhibiting a higher rate of speech in a supportive context. In contrast, the nuclear pitch accent in Polish serves a differentiating function, exhibiting a falling intonation in instances of genuine impoliteness and a rising intonation in instances of mock attitude. However, in German, no clear differentiation is exhibited.

In a similar way, duration and speech rate show conflicting trends. The use of polite speech is often slowed down by the inclusion of pauses and fillers, which serve to convey a greater sense of insecurity and the conveyance of politeness (Winter & Grawunder, 2012). In Japanese, polite speech is characterised by slower speech rate (Ofuka *et al.*, 2000); however Álvarez *et al.* (2011) found for Venezuelan Spanish that impolite utterances showed greater temporal variability and longer syllables than polite or neutral ones. In parallel, a study focusing on the prosodic distinction between polite and impolite requests elicited by controlled tasks in Canadian English found slower speech rates and longer durations in rude utterances (Caballero *et al.*, 2018). These results suggest a complex interpretation of linguistic impoliteness, in which the role of phonological and phonetic indices (local and global) is combined with both lexical and morpho-syntactic indices and information derived from the pragmatic context, leading to a dynamic scenario.

⁴ In addition, strategies related to denying common ground or distancing oneself from others are also formulated.

⁵ In this regard, a recent study conducted on several languages has demonstrated that the correlation between polite communication and rising pitch contours is not always observed, resulting in a re-evaluation of the Frequency Code (Winter *et al.*, 2021).

⁶ In particular, it can be observed that Polish utterances produced with a genuine impoliteness are louder than those produced with a supportive attitude.

In the domain of hate speech studies, a novel research approach has been developed which analyses the correlation between the prosody of impolite speech acts and their level of social acceptability. In Niebuhr & Neitsch (2020), 28 participants were presented with offensive stimuli in German that were intended to convey irony or genuine insults towards the Holocaust. The participants were asked to assess the personal acceptability and degree of social violation of the stimuli. The results indicated that stimuli with lower acoustic indices, including mean f_0 and speech rate, were evaluated as less acceptable and more socially censored.

The quality of the voice and the use of non-verbal cues, such as gestures, facial expressions, gaze and bodily movements, also constitute a significant factor in the performance of insults in spoken language. The research on multimodal aspects of impolite speech revealed the pivotal role of non-verbal cues, including gestural, physical proximity towards the addressee, hand and arm gestures, many of which have obscene connotations, and overt alterations of facial expression, such as frowning eyebrows (McKinnon & Prieto, 2014; Brown & Prieto, 2017). The gestural component is frequently accompanied by a hoarse, rough, and nasalised voice which is typically indicative of the expression of anger and aggression⁷.

As regards Italian language, Vinciguerra (2023b) presented the first experimental findings concerning the speech rate in insults. The insults, comprising personalised negative vocatives as the incipit of impolite sentences, exhibited no differences in speech rate when compared to unmarked utterances. Nevertheless, the analysis demonstrated that the epithet was produced by the speakers at a substantially slower rate than the “impolite coda”. The author postulated that the impolite utterances may have undergone an overall slowing down due to the insulting epithet being placed in the initial position.

3. THE STUDY

3.1. Research questions

To date, the majority of studies on Italian insults have been developed from a pragmatic perspective (Alfonzetti, 2017; Bazzanella, 2020; Bianchi, 2021), with a particular focus on slurs (Tenchini & Frigerio, 2020). However, there have been no studies on the prosody of insults in Italian, which highlights an important gap in the existing research. To this end, we present the initial findings of a broader study that focuses on the prosodic production of Italian insults. The aim of this research is to explore certain prosodic aspects of a corpus of insults elicited through oral role plays using spectro-acoustic analysis.

Given the frequent association of insults with hot anger and aggression, it can be assumed that the acoustic patterns used in insulting expressions are amplified with respect to non-aggressive forms. The initial hypothesis is that insults are employed as a strategy of impoliteness which threatens the face of the addressee, thereby reducing the pragmatic space of conversation and cooperative agreement. The reduction of conversational space is reflected in the phonetic domain by a widening of the auditory space, which the objective of increasing the audibility of the offensive utterance and ensuring the success of the verbal attack. Based on the aforementioned observations, three distinct research hypotheses have been formulated. Due to their expressive status, insults are expected to have (i) greater intensity, (ii) faster speech rate and (iii) wider pitch range than assertive sentences produced without offensive meaning.

The increase in mean intensity and pitch range is consistent with the speaker's intention to dominate the listener verbally. Furthermore, the use of a fast speech rate is also consistent with the metaphorical content of 'throwing insults', as is often found in the extensive and interdisciplinary literature on the subject.

3.2 Materials and methods

The linguistic materials were collected using audio recordings of mimetic-pretending role play, a method which is commonly used in the field of (im)politeness research⁸.

For the purposes of this research, the focus has been directed towards one of the most typical forms of insult, which involves the utilisation of an epithet with a direct address to the listener. This is also designated as a 'personalised negative vocative' (Culpeper, 2011a). The participants were provided with 15 scripts, each comprising a narrative extract that contextualises the communicative exchange and a target sentence. An example of role play is presented in (A) in both the Italian and the English versions. The words included in each target sentence were selected on the basis of their frequency of use in Italian and their offensive connotations⁹. The participants were instructed to read the content of the role play aloud in order to enhance their identification with the context. They were then required to produce the sentence containing the insult. In each scenario, the focus of the analysis was on the target sentence containing the insult. The statements were compared

⁷ Hot anger is characterised by a greater mean intensity and mean f_0 , higher pitch range and pitch variability; among others, Anolli & Ciceri (1997); Kerr & Schneider (2008), Johnstone & Scherer (2000); Scherer *et al.* (2003).

⁸ For a review of methods in pragmatics, see, among others, Kasper (2008) and Landone (2022).

⁹ A pre-test was administered in order to select only those words with a similar level of offensiveness. To gain further insight, a second assessment is currently being conducted.

with neutral sentences produced by the same subjects, recorded using the same methodology. These were designated as the control group, and were found to be free of offensive content or emotional bias.

(A) *Vito accompagna Annamaria al lavoro. Nel tragitto, Annamaria scorge un rossetto che non le appartiene nel portaoggetti dell'auto, così uscendo inferocita dall'auto gli urla contro:*

'Stronzo, a casa facciamo i conti'

(A) *A Vito is driving Annamaria to work. On the way, Annamaria discovers a lipstick in the glove compartment that does not belong to her, so she gets out of the car and angrily yells at him:*

'Asshole, I'll deal with you at home'

The acoustic recordings were conducted in silent environments with the sole presence of the researcher. A professional recorder was utilised for this purpose (Zoom H5, Wav, Hz 44.100, 16 bit)¹⁰. The study analysed 90 insulting utterances and 90 control utterances, employing the PRAAT software (Boersma & Weenink, 2024, vers. 6.4.13). For each sentence, we measured the mean and the standard deviation of the following acoustic parameters: median, maximum and minimum of f_0 , pitch range (semitones, ST), mean intensity (dB), duration of nuclear vowels (ms), speech rate (syll/sec). The pitch range is equivalent to the pitch span, which is the distance in semitones calculated by determining the range of fundamental frequency (f_0) values between $f_{0\max}$ and $f_{0\min}$.

In order to quantify intonation variation, the median f_0 value ($f_0\tilde{x}$) was employed as the neutral reference point for each speaker. This approach was based on the mean pitch values that were measured for each speaker. The rationale for utilising the median f_0 rather than the mean f_0 is due to the former's greater resistance to outliers (Sicoli *et al.*, 2015). The median f_0 was assumed as the pitch level of reference for each speaker. The durations were normalised in relation to the speaking rate calculated for each speaker (duration/speaker speech rate). The statistical significance of the results was evaluated using One-Way ANOVA with post-hoc Tukey¹¹, Pearson's correlation coefficient with IBM SPSS software.

3.3. Participants

The participants selected for the oral productions were all professional actors. This approach mitigated the potential for emotional bias to influence the performance of non-professional speakers. Furthermore, it permitted the acquisition of natural productions without the potential risk of inhibition due to verbal self-censorship. The study involved six theatre actors, three males and three females, aged between 21 and 28 years old (mean age $\pm$ 2.48), all from the Apulia region, which is situated in the south-east of Italy. The actors had an average of nine years of experience in acting and had attended advanced-level theatre courses. The participants were recorded individually in a quiet room. Prior to the recording, all subjects involved in the study provided with written consent. The participants were not informed of the specific aims of the research.

4. THE RESULTS

The results indicated that insults have distinctive acoustic properties when compared to the utterances employed as control. The study revealed that the mean intensity of offensive sentences was higher (+6 dB) (see Figure 1) and statistically significant, as indicated by the one-way Anova [$F(3, 180) = 6.02, p = .001$] and Tukey post-hoc test. In detail, $INS_{\text{male}}/CON_{\text{male}}$ ($p = .037$); $INS_{\text{female}}/CON_{\text{female}}$ ($p = .020$); the comparisons between $INS_{\text{male}}/INS_{\text{female}}$ ($p = .755$) and $CON_{\text{male}}/CON_{\text{female}}$ ($p = .801$) do not reach statistical difference.

The first hypothesis was supported by the results; the actors employed a louder volume of voice when uttering insults, in comparison to the control forms. In addition to being more intense, the insulting portion tends to be expressed with a slower speech rate (-1 syll/sec), with an average rate of 5 syll/sec, as shown in Figure 2.

The mean speech rate in insults was observed to be lower than that calculated for the control sentences, despite the high level of expressiveness and emotional value of the former. However, this difference was not found to be statistically significant [$F(3, 180) = .856, p = .465$]. Therefore, the second hypothesis was not confirmed.

In contrast with the initial hypothesis, the production of insults was not found to be significantly slower. However, it should be noted that there is a certain degree of variability in the way speakers perform these speech acts. While some speakers showed no notable difference between insults and control sentences, two female subjects (F_1 and F_3) employed contrasting strategies for insults. Specifically, subject F_1 used a faster speaking rate (6.8 syll/sec), while subject F_3 used a slower rate (4.5 syll/sec).

¹⁰ These are standard parameters that are conventionally used to ensure the high quality of the audio signal.

¹¹ One-way ANOVA is a common statistical test used to compare the means of more than two experimental variables or groups. The post-hoc Tukey test is used for equal sample sizes and after the ANOVA test to determine which groups are statistically different.

Despite the absence of statistical significance with regard to the difference in speech rate, a clear discrepancy was observed in the duration of the nuclear vowel in the intonation unit containing the insult when compared with that of the control sentences. In order to mitigate the potential for phonetic distortions, all vowel durations were normalised based on the speaking rate of each participant. Normalisation made it possible to reduce the amount of temporal conditioning between fast and slow speakers. As illustrated in Figure 3, the duration of the vowels on which the nuclear pitch accent falls underwent a process of lengthening in insults, resulting in a temporal prominence that was also discernible through perceptual feedback. The statistical comparisons conducted using One-way Anova yield significant results [$F(3, 180) = 10.35, p = .000$], as did comparisons between subgroups; INS_{male}/CON_{male} ($p = .001$); $INS_{female}/CON_{female}$ ($p = .000$). The post-hoc analysis between INS_{male}/INS_{female} did not reach statistical results ($p = .996$), as well as CON_{male}/CON_{female} (.939); this suggest that the groups are homogeneous, regardless of the gender of the subjects.

The Pearson correlation coefficient indicates that there is a linear negative relationship between the speech rate of the sentences and the duration of nuclear stressed vowels. An increase in the duration of nuclear stressed vowels (NV) is associated with a decrease in the speech rate of sentences; female groups= $r(45) = -.85, p=.000$); male groups= $r(45) = -.75, p=.000$.

Figure 1: Mean values for intensity (dB) (I: insults, C: controls) (M: males, F: females)

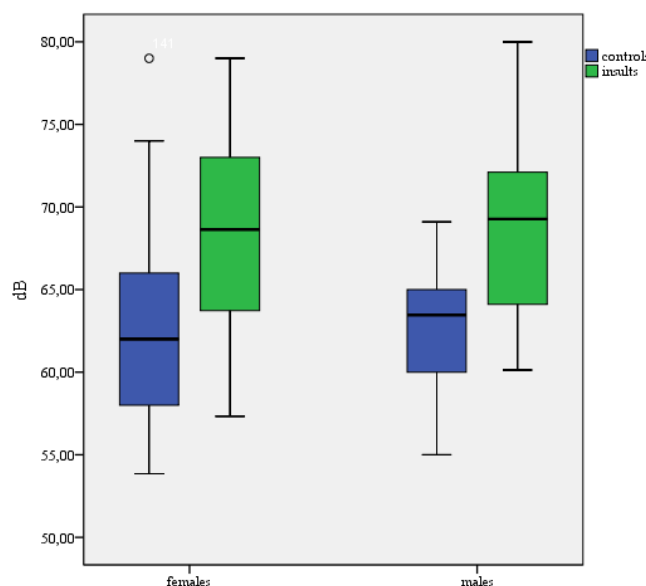


Figure 2: Mean values for speech rate (I: insults, C: controls) (M: males, F: females); error bar indicates the standard deviation

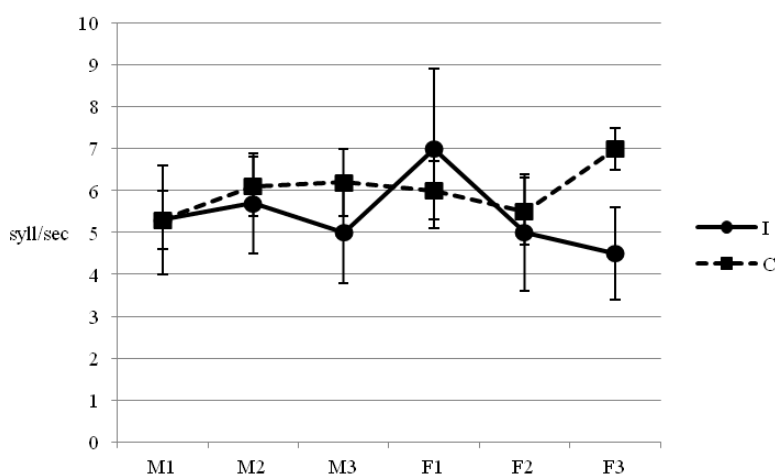
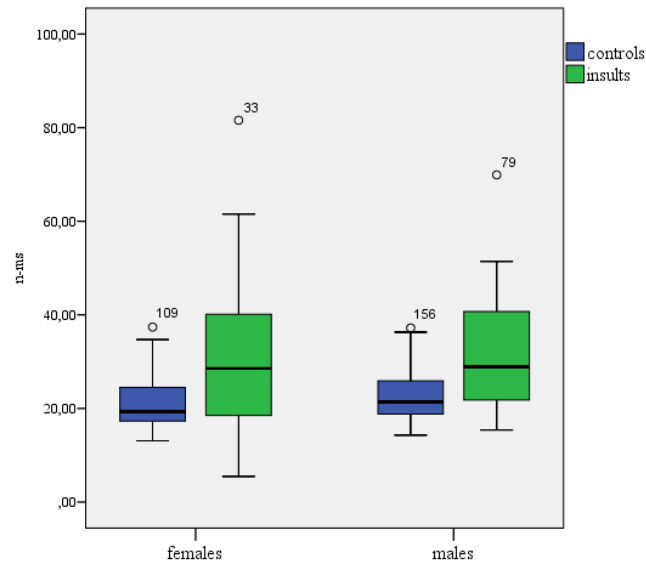


Figure 3: Normalised mean values for nuclear vowel duration (ms)
(I: insults, C: controls), (M: males, F: females)



The data suggests that insults are generally associated with a temporal decrease in speaking rate, rather than an increase as originally postulated. This is supported by the observed behaviour of nuclear stressed vowels. In the control groups (both male and female groups), the normalised duration of the nuclear vowel shows a modest degree of dispersion, as evidenced by the relatively low height of the boxplot. In the control group, 50% of the values were observed to fall between 29 and 32 ms (normalised durations), whereas the time range for the insult group exhibited a wider distribution.

With regard to intonation, it was observed that insulting expressions were produced as distinct tonal units, separated from the rest of the utterance by a pause or a boundary tone. Insults were typically identified by an intonation contour that, after a high f_0 onset, remained at a high level without reaching a true f_0 minimum. The remainder of the statement showed a falling intonation pattern, which was consistent with the conveyed offensive meaning.

For each speaker, the mean and median fundamental frequency (f_0) were higher in the insults than in the control sentences. This had an impact on both the pitch level and pitch span. The behaviour of the pitch range is unclear as it was unexpectedly lower (-5 ST) in the insulting sentences compared to the control group (Figure 4). This differs from the findings of Culpeper *et al.* (2003) and the hypothesis proposed in our research¹². Nevertheless, despite the presence of some degree of internal variability (± 3 ST), a one-way ANOVA between groups yielded a statistically significant result [$F(3, 180) = 31.62, p = .000$], as did comparisons between subgroups; INS_{male}/CON_{male} ($p = .000$); $INS_{female}/CON_{female}$ ($p = .000$). The post-hoc analysis between INS_{male}/INS_{female} did not reach statistical significance. ($p = .095$).

A review of the data indicates that the highest normalised duration of nuclear vowels is observed in the utterances where the highest pitch range values are also reached. In order to assess the relationship between these two parameters, a Pearson correlation coefficient was calculated between the duration of the nuclear vowel (NV) and the pitch range value in semitones within the same sentence. The participants were divided into two groups according to gender. The results demonstrated that in both groups, the two parameters exhibited a linear positive correlation, with an increase in one parameter (duration) resulting in a increase in the other (pitch range); female groups= $r(45) = .44, p = .001$; male groups= $r(45) = .39, p = .004$.

The data indicates that the pitch range in insults was narrower. This is due to the fact that, in many cases, the f_0 minimum was not reached at the end of the sentence. The difference between insults and control sentences in terms of semitones is evident for all participants except for speaker F_2 (see Figure 5).

To shed light on this unexpected outcome, an attempt was made to more accurately quantify the intonation variation of the sentences by employing the median f_0 as the neutral frequency point to represent the pitch level. The distances in semitones between the median fundamental frequency ($f_0\tilde{x}$) of each sentence and the maximum and minimum values of

¹² There are several reasons for this difference. Culpeper *et al.* (2003) concentrated on impolite sentences extracted from a BBC television programme, which were not exclusively intended as insults. Furthermore, the prosodic analysis was mainly conducted at the auditory level, with only limited instrumental control of intonation contours.

Figure 4: Mean values of pitch range (ST) grouped for gender (males, females) and attitude (insults, controls)

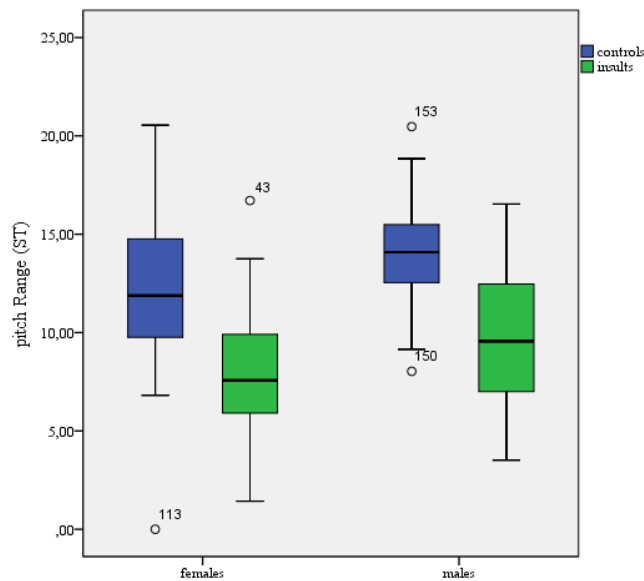
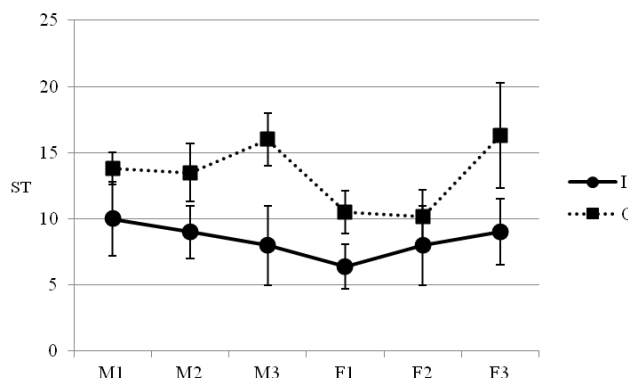


Figure 5: Mean values of pitch range (ST) (I: insults, C: controls) and speakers (M: males, F: females); error bar indicates the standard deviation



fundamental frequency ($f_{0\max}$ and $f_{0\min}$) were calculated for each speaker. In the sample of insults, the median f_0 is higher compared to that of the control sentences (M: $p = .03$; F: $p = .04$).

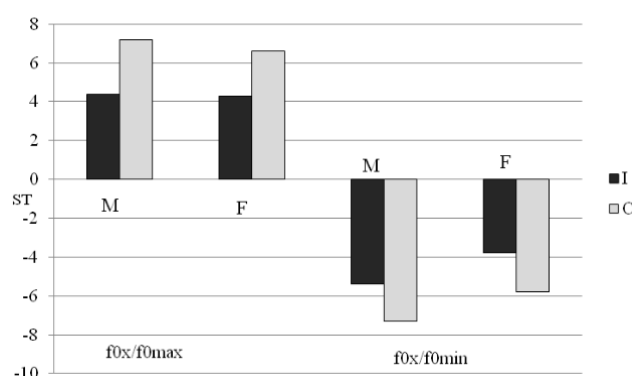
As shown in Figure 6, regardless of gender, the $f_{0x}/f_{0\min}$ distance remained closer to the reference point, graphically represented by the zero value, while the $f_{0x}/f_{0\max}$ distance was smaller. It should be noted that the average value of $f_{0\min}$ is higher in the insults than in the control sentences (males: .005; females: .006). Diversely, the value of $f_{0\max}$ may undergo a substantial increase, as this parameter determines the degree of expressiveness and prominence.

Thus, in insults, the pitch level is higher than in control speech, while the pitch span is reduced. The pitch range of sentences containing insults was found to be contracted, due to the failure to reach a minimum frequency value, resulting in a smaller range than that observed in the control speech. This trend aligns with the use of the vocative function conveyed by the nominal phrase employed during the construction of the target sentences.

5. DISCUSSION

The study has produced preliminary findings on the prosody of Italian insulting expressions, with significant trends being observed in the data. However, it is important to note that the results only pertain to linguistic forms elicited by professional theatre actors in predetermined contexts. Although the speech sample was produced with a high degree of naturalness and without inhibition, it cannot be considered representative of spontaneous speech due to its controlled nature. To address this limitation, a new corpus of insults produced by naive speakers in less controlled contexts is currently being collected. This will provide a basis for comparison with the existing data in future research.

Figure 6: Mean differences in ST between $f_0\tilde{x}$ and $f_0\max$ and $f_0\min$, respectively, by gender (M, F) and subgroup (I, C)



The analysis showed that insults have a specific prosodic realisation that leads to the activation of numerous acoustic indices. The offensive nature of insults is conveyed not solely through lexical selections and the conveyance of a negative polarity evaluation, but also by means of phonetic components. It is noteworthy that insults share diverse aspects with highly activated negative emotions, such as hot anger, as they always contain an aggressive component achieved through a gradual increase in voice volume and tonal extension.

In their research, Culpeper (2011a; 2011b) and Culpeper *et al.* (2003) postulated a broadening of the pitch range of impolite sentences, but our findings did not fully support this. Instead, it was observed that the pitch range shifted upwards while remaining compressed, creating a perceptual effect of a dominant voice. Furthermore, the pitch span has been found to correlate with the duration of the nuclear vowels, with a longer duration corresponding to a wider pitch range.

The effect was further enhanced by the use of a scratchy voice quality, resulting in a verbal attack that sounded aggressive. This overpowered the auditory space of the addressee, who is denigrated and, at the same time, dominated. This is in line with the fact that all types of violent speech, including insults, along a semantic continuum, are overt strategies of denigration and clear face-attacks.

There is a belief that insults are a form of invasion into the space of the listener. The speaker adopts a centric-position, appearing dominant and tending to invade the personal and auditory space of the listener. This acoustic strategy is the result of several processes. The most noticeable of these strategies involves increasing the volume of the voice, making the insult perfectly audible and effectively offensive¹³. The maintenance of a high pitch level serves to emphasise insults and ensure the transmission of their hurtful content, a behaviour consistent with the expression of intense anger. The findings of the research indicate that the duration of nuclear vowels has also a significant impact on the production of insults.

It is our opinion that the invasion of auditory space occurs along two dimensions: vertical and horizontal. On the vertical axis, it results in an increase in voice volume and pitch level, a voice strategy which serves to ensure audibility and give prominence to the insult. On the horizontal axis, the insult causes the nuclear vowel to be lengthened, making it more prominent, not only in terms of loudness and f_0 frequency, but also in terms of duration. The use of an insulting word has been shown to take up more time, thereby creating a specific effect of dominance and hostility. At the same time it denies the offended person the opportunity to take the turn in the dialogue, which can be seen as a form to prevent a reaction. Verification of this behaviour should be carried out in natural conversations.

There is still much work to be done in this area. One noteworthy finding from our analysis is that during the production of target sentences, the actors frequently hurled insults at an unknown listener, resulting in a performance that closely resembled a spontaneous interaction. However, the speaker occasionally opted to deliver the insult in a hushed tone, almost whispering it, which can be viewed as an alternative but equally offensive tactic. It seems reasonable to posit that this behaviour is responsible for a certain degree of cross-speakers variability, and thus represents a worthy subject for future investigation. This data appear to suggest that the average speech rate and intensity may be influenced by the speaker's position on the dominance and power continuum as well as the nature of the interaction. In this regard, a perception test is being administered to assess the speakers' degree of offensiveness and dominance.

Moreover, a comprehensive examination of prosodic indices in a less controlled speech sample, encompassing aspects such as voice quality, is essential to enhance our understanding of the role of parameters in conveying offensive meaning.

¹³ Intensity also seems to play a key role in the perception of (im) politeness nuances, as observed by Idemaru *et al.* (2020) for the Korean language.

ACKNOWLEDGMENTS

The authors would like to thank all the participants who agreed to take part in this study, and the anonymous reviewers for their careful reading of the manuscript.

DECLARATION OF COMPETING INTEREST

The authors of this article declare that they have no financial, professional or personal conflicts of interest that could have inappropriately influenced this work.

FUNDING SOURCES

This Publication was produced with the co-funding of the European Union-Next Generation EU: NRRP Initiative, Mission 4, Component 2, Investment 1.3 - Partnerships extended to universities, research centres, companies and research D.D. MUR n. 341 del 15.03.2022 – Next Generation EU (PE0000014 - "SEcurity and Rights In the CyberSpace - SERICS" - CUP: H93C22000620001).

AUTHORSHIP CONTRIBUTION STATEMENT

This article is the result of an extended collaboration between the authors. However, for academic purposes: Patrizia Sorianello: Conceptualisation, Formal analysis, Methodology, Writing- original draft, Writing-review & editing, Funding acquisition.

Giovanni Vinciguerra: Conceptualisation, Data Curation, Methodology.

REFERENCES

- Alfonzetti, G. (2017). *Questioni di (s)cortesia: Complimenti e insulti*. Edizioni Sinestesie.
- Álvarez, A., Blondet, M. A. and Darcy, R. (2011). Sobre (des)cortesia y prosodia: una relación necesaria. *Oralia*, (14), 437-450. <https://doi.org/10.25115/oralia.v14i1.8206>
- Andreeva, B., Bonacchi S. and Barry W. (2016). Prosodic cues of genuine and mock impoliteness in German and Polish, *Proceedings of the International Conference on Speech Prosody* (pp. 999-1003). <https://doi.org/10.21437/SpeechProsody.2016-205>
- Anolli, L. and Ciceri, R. (1997). *La voce delle emozioni: verso una semiosi della comunicazione vocale non-verbale delle emozioni*. Franco Angeli.
- Bazzanella, C. (2020). Insulti e pragmatica: complessità, contesto, intensità. *Quaderns d'Italia*, (25), 11-26.
- Bianchi, C. (2021). *Hate speech. Il lato oscuro del linguaggio*. Laterza.
- Boersma, P. and Weenink, D. (2024). *PRAAT: doing phonetics by computer* [computer program]. Version 6.4.13, retrieved from <http://www.praat.org/>.
- Brown, L. and Prieto, P. (2017). (Im)politeness: prosody and gesture. In J. Culpeper, M. Haugh & D. Kádár (Eds.), *The Palgrave Handbook of Linguistic (Im)politeness* (pp. 323-355). Palgrave.
- Caballero, J.A., Vergis, N., Jiang, X. and Pell, M. D. (2018). The sound of im/politeness. *Speech Communication*, (102), 39-53. <https://doi.org/10.1016/j.specom.2018.06.004>
- Chen, A., Gussenhoven, C. and Rietveld, T. (2004). Language-specificity in the perception of paralinguistic intonational meaning. *Language and Speech*, 47(4), 311-349. <https://doi.org/10.1177/00238309040470040101>
- Culpeper, J. (1996). Towards an anatomy of impoliteness. *Journal of Pragmatics*, 25(3), 257-286. [https://doi.org/10.1016/0378-2166\(95\)00014-3](https://doi.org/10.1016/0378-2166(95)00014-3)
- Culpeper, J. (2011a). *Impoliteness: Using language to cause offence*. Cambridge University Press.
- Culpeper, J. (2011b). It's not what you said, it's how you said it: Prosody and impoliteness. In Linguistic Politeness Research Group (Ed.), *Discursive approaches to politeness* (pp. 57-83). (Mouton Series in Pragmatics [MSP]; Vol. 8). Mouton de Gruyter.
- Culpeper, J. (2021). Impoliteness and hate speech: Compare and contrast. *Journal of Pragmatics*, 179(2), 4-11. <https://doi.org/10.1016/j.pragma.2021.04.019>
- Culpeper, J., Bousfield, D. and Wichmann, A. (2003). Impoliteness revisited: with special reference to dynamic and prosodic aspects. *Journal of Pragmatics*, (35), 1546-1579. [https://doi.org/10.1016/S0378-2166\(02\)00118-2](https://doi.org/10.1016/S0378-2166(02)00118-2)
- Feinberg, J., (1985). *Offense to others*, vol. 2. Oxford University Press.
- Guillén-Nieto, V. (2023). *Hate Speech: Linguistic Perspectives*. De Gruyter Mouton.
- Idemaru, K., Winter, B., Brown, L. and Oh, G. E. (2020). Loudness Trumps pitch in politeness judgments: evidence from Korean deferential speech. *Language and Speech*, 63(1), 123-148. <https://doi.org/10.1177/0023830918824344>
- Ilie, C., (2001). Unparliamentary language: Insults as cognitive forms of ideological confrontation. In R. Dirven, R. Frank & C. Ilie (Eds.), *Language and ideology* (pp. 255-264). *Volume II: Descriptive cognitive approaches*. John Benjamins.
- Johnstone, T. and Scherer, K. R. (2000). Vocal communication of emotion. In M. Lewis & J. Haviland (Eds.), *Handbook of emotion* (2nd ed., pp. 220-235). Guilford Press.

- Jucker, A. H. and Taavitsainen, I. (2000). Diachronic speech act analysis. Insults from flyting to flaming. *Journal of Historical Pragmatics*, 1(1), 67-95. <https://doi.org/10.1075/jhp.1.1.07juc>
- Kasper G. (2008). Data collection in pragmatics research. In H. Spencer-Oatey (Ed.), *Culturally Speaking: Culture, Communication and Politeness* (pp. 279-302). Continuum.
- Kerr, M. A. and Schneider, B. H. (2008). Anger expression in children and adolescents: A review of the empirical literature. *Clinical Psychology Review*, 28(4), 559-577. <https://doi.org/10.1016/j.cpr.2007.08.001>
- Landone, E. (2022). *Methodology in Politeness Research*, Advances in (Im)politeness Studies. Springer International Publishing.
- Mateo, J. and Yus, F. (2013). Toward a cross-cultural pragmatic taxonomy of insult. *Journal of Language Aggression and Conflict*, 1(1), 87-114. <https://doi.org/10.1075/jlac.1.1.05mat>
- McKinnon, S. and Prieto, P. (2014). The role of prosody and gesture in the perception of mock impoliteness. *Journal of Politeness Research*, 10(2), 185-219. <https://doi.org/10.1515/pr-2014-0009>
- Moïse, C. (2006). Analyse de la violence verbale: quelques principes méthodologiques. In *Actes des XXV^{ème} Journées d'étude sur la parole* (pp. 103-114).
- Niebuhr, O. and Neitsch, J. (2020). On the role of prosody in the production and evaluation of German hate speech. In *Proceedings of the 11th International Conference on Speech Prosody* (pp. 710-714).
- Ofuka, E., McKeown, J.D., Waterman, M.G. and Roach, P.J. (2000). Prosodic cues for rated politeness in Japanese speech. *Speech Communication*, 32(3), 199-217. [https://doi.org/10.1016/S01676393\(00\)00009-1](https://doi.org/10.1016/S01676393(00)00009-1)
- Ohara, Y. (2001). Finding one's voice in Japanese: A study of the pitch levels of L2 users. In A. Pavlenko, A. Blackledge, I. Piller & M. Teutsch-Dwyer (Eds.), *Multilingualism, Second Language Learning, and Gender* (pp. 231-254). De Gruyter Mouton. <https://doi.org/10.1515/9783110889406.231>
- Scherer, K. R., Johnstone, T. and Klasmeyer, G. (2003). Vocal expression of emotion. In R. J. Davidson, H. Goldsmith, & K. R. Scherer (Eds.), *Handbook of the affective sciences* (pp. 433-456). Oxford University Press.
- Sicoli, M. A., Stivers, T., Enfield, N. and Levinson, S. C. (2015). Marked initial pitch in questions signals marked communicative function. *Language and Speech*, 58(2), 204-223. <https://doi.org/10.1177/0023830914529247>
- Stadler, S.A. (2007). *Multimodal (im)politeness: the verbal, prosodic and non-verbal realization of disagreement in German and New Zealand English*. Philologia. Verl. Dr. Kovač.
- Tenchini, M. P. and Frigerio, A., (2020). The impoliteness of slurs and other pejoratives in reported speech. *Corpus Pragmatics*, (4), 273-291. <https://doi.org/10.1007/s41701-019-00073-w>
- Uldall, E. (1960). Attitudinal meanings conveyed by intonation contours. *Language and Speech*, 3(4), 223-234. <https://doi.org/10.1177/002383096000300403>
- Van Olmen, D., Anderson, M. and Culpeper, J. (2023). Inherent linguistic impoliteness: The case of insulting YOU+NP in Dutch, English and Polish. *Journal of Pragmatics*, (215), 22-40. <https://doi.org/10.1016/j.pragma.2023.06.013>
- Vinciguerra, G. (2023a). I meccanismi di intensificazione dell'insulto nella comunicazione mediata dal computer. In V. Caruso, M. Maffia (Eds.), *Vecchie e nuove forme di comunicazione diseguale: canali, strutture e modelli*, *Studi AltLA* 17 (pp. 355-372). Officinaventuno.
- Vinciguerra, G. (2023b). Tutt'altro che implicito: componenti prosodiche e ritmi elocutivi degli atti linguistici scortesi. *Italiano LinguaDue*, (15), 102-115. <https://doi.org/10.54103/2037-3597/22020>
- Winter, B. and Grawunder, S. (2012). The phonetic profile of Korean formality. *Journal of Phonetics*, 40(6), 808-815. <http://dx.doi.org/10.1016/j.wocn.2012.08.006>
- Winter, B., Oh, G.E., Hübscher, I., Idemaru, K., Brown, L., Prieto, P. and Grawunder, S. (2021). Rethinking the frequency code: A meta-analytic review of the role of acoustic body size in communicative phenomena. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1840). <https://doi.org/10.1098/rstb.2020.0400>