JoSS Journal of Speech Sciences

DOI: 10.20396/joss.v14i00.20385

THE AFFECTIVE FOUNDATION OF SPEECH. AFFECT, PROSODY, ILLOCUTION

CRESTI, Emanuela[1*]

[1]University of Florence - ORCID: https://orcid.org/0000-0002-3018-6120

Abstract: *According to Language into Act Theory, the identification of illocutionary types in speech is based on two premises: a) the speaker's affect is the activating force of illocution; b) prosodic contours convey the illocutionary force. This paper aims to clarify both assumptions by presenting arguments that link affect to prosody and prosody to pragmatic activities. The first part is dedicated to sketching the theoretical framework (Human Birth Theory), grounding the concept of affect, which must be distinguished from emotion and attitude. It* ***focuses*** *on its application to the domain of language performance and the relation between affect and prosody. The second part presents empirical arguments supporting the idea that prosodic contours guide the interpretation of illocution. The "Taken-for-granted assertions," "Evident assertions," "Requests for confirmation," and "Open questions," illocutionary types not identified in previous corpus-based taxonomies, are shaped by dedicated prosodic contours and defined according to pragmatic parameters. Parameters are scalar and connected to the speaker's affective force: a) the speaker's commitment to the truth, the relevance of the content, and the speaker's affective involvement for assertions; b) the speaker's estimation of the addressee's knowledge, the type of linguistic behavior requested, and the speaker's affective involvement for questions. The distinction between prosodic contours bearing illocutionary value and the prosodic properties that are a function of the speaker's attitude is faced, showing that the latter does not change the prosodic model. The comparison between "Taken for granted assertion" and "Open question" shows how subtle prosodic cues can change the pragmatic interpretation.*

Keywords: Language into Act Theory; Affect; Prosody; Illocution; Corpus-based research

[*] Corresponding author: emanuela.cresti@unifi.it

1 Introduction

1.1 Premises

Within the framework of Language into Act Theory (L-AcT; Cresti 2000), the identification of illocutionary types that occur in speech performance is based on two premises: the first is that the speaker's affect is the activating force of illocution and the second is that the illocutionary force is conveyed by modeled prosodic contours. Therefore, our interpretation in assigning illocutionary values to the utterance is guided by modeled prosodic contours. This paper aims to clarify both assumptions by presenting empirical arguments that link affect to prosody and prosody to pragmatic activities.

The concept of affect is rarely considered in linguistic studies.[1] The first part of this paper is dedicated to outlining the theoretical framework that underpins the concept of affect and its application to the domain of language performance. Relying on the Human Birth Theory (HBT-Fagioli 1972[1], Eng translation 2019a), it is argued that affect is the deep psychic spring to speech activation, depending on pulsion, which must be distinguished from emotion, depending on instinct.

The second part of the paper presents empirical arguments. They support our interpretation in assigning illocutionary values guided by prosodic contours rather than logical or lexical ones, as assumed in classical and recent corpus-based taxonomies, which crucially do not consider prosody. We will illustrate the relation between prosodic contours and specific illocutionary values within the general illocutionary assertion class and the directive subclass of questions.

We will make explicit the criteria for identifying, within these classes, new types that are not considered in other corpus-based taxonomies. We will deal with "Taken-for-granted assertion" and "Evident assertion" types for the assertive class and "Request for confirmation" and "Open question" types for the question sub-class within the directive class. Considering the prosodic contours bearing illocutionary value of those types, the distinction with the prosodic properties that are a function of the speaker's attitude will be faced. Finally, based on a paradigmatic case, we will experimentally test how prosodic contours convey distinctive pragmatic values.

The paper is structured as follows. After these premises, the exemplification of a laboratory case is given in 1.2. 2 clarifies the relationship between affect and speech: in 2.1, we will briefly present the theoretical framework of Human Birth Theory grounding the affective foundation of speech; in 2.2, we will distinguish affect and emotion; in 2.3, we will sketch the relation between affect and prosody; and in 2.4, we will consider how attitudinal connotations are seen in our perspective.

Paragraph 3. is devoted to framing the L-AcT approach within the tradition of corpus-based studies of illocution. In Section 3.1, we will sketch the main corpus-based illocutionary taxonomies, which pave the way for discovering, through corpus analysis, illocutionary types not considered in classical works based on performatives. A Table of L-AcT's illocutionary taxonomy is added. In 3.2, a brief description of the methodology used in the Language into Act Theory will be presented to connect prosodic properties and illocutionary values in corpus data.

Paragraph 4 is devoted to the assertive class, presenting an overall introduction to assertive illocution. Then, the specific illocutionary types that are the object of this research are introduced in 4.1 and 4.2. Their prosodic properties and pragmatic definition will be proposed.

[1] The term '*affect*' is commonly employed as adjective (*affective*) and as noun with a technical meaning in the research on emotions (Scherer 1986; Magno-Caldognetto 2002).

5 is dedicated to an overall introduction to the directive class, followed by that of the subclass of questions. In Section 5.1, the Request for Confirmation type is described, and in Section 5.2, the Open question type is discussed. Then, in 5.3, a comparison with Kohler's attitudinal hypothesis is considered. Finally, in Section 6, a prosodic comparison is carried out between the 'Taken-for-granted' assertion and the 'Open question', which reveals similar prosodic contours. In 7, a brief conclusion ends the work.

1.2 Nuances of meaning or illocutionary types in the assertive class?

The empirical issue of the paper is introduced by showing two assertive utterances. Other examples of the directive class of questions will be presented throughout the paper. (1) and (2) correspond to different prosodic performances of the duplicate lexical content taken from an utterance of the Italian corpus, which is taken from a dialogue between two friends who are preparing a typical Italian dessert (Tiramisù) (*Ci sbatto la chiara* 'I beat in the egg white'). A female speaker performs the two utterances in a laboratory setting, and each is realized through a different prosodic contour conventionally labeled TG in Figure 1 and EA in Figure 2. Each utterance replicates the prosodic contour of assertive utterances found in our corpus, accomplishing comparable pragmatic activities.[2]

The Figures display the spectrograms of the productions, where the f0 track has been overlaid on the first harmonica to highlight the differential contour. [3] Let us see examples (1) and (2) and Figures 1 and 2.

In the eliciting context of (1), the speaker presents well-known information on how to make that cake. The information is low-relevant for both the speaker and the addressee since it may be expected to be based on shared knowledge.

(1) *ABC: ci sbatto la chiara //
'I beat the egg white'

[2] The research is part of ongoing work on a set of assertive illocutionary types carried out with Massimo Moneglia (Cresti & Moneglia forthcoming).

[3] The reader can download the audio files referred to in the paper from https://doi.org/10.5281/zenodo.17079732. The spectrograms of the examples are calculated through WinPitch (Martin 2004). The stylization of f0 tracks identifies movements over the glissando threshold. Glissando is the rate of f0 change above which a melodic change is supposed to be perceived and determines the perceptual boundary between a static pitch and a melodic variation. If the variation is less than the threshold, the perception will correspond to a static tone; if it is higher, it will be perceived as a melodic variation (Rossi 1971; Rossi 1978; 't Hart 1976; Martin 2015). The colors of the stylization lines are automatically assigned by Winpitch for the sole purpose of visual emphasis but have no functional equivalent.

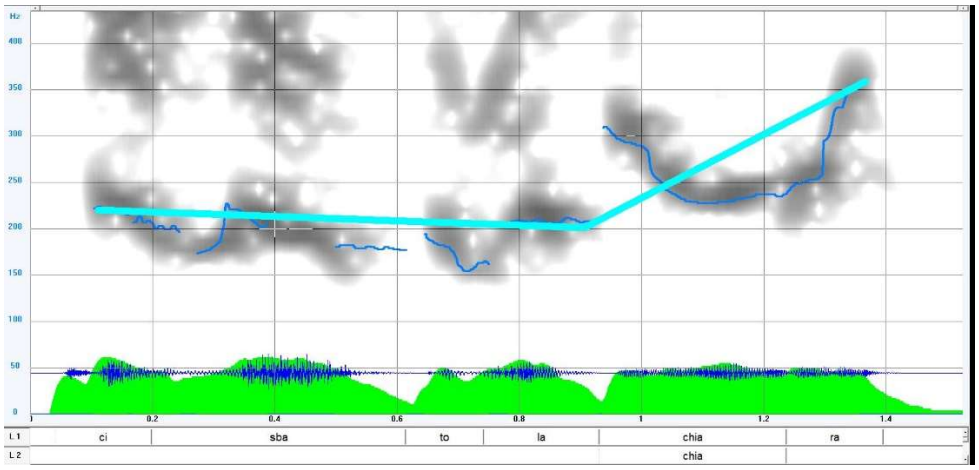

Figure 1: The stylized contour is conventionally tagged as TG for Taken-for-granted assertion

The eliciting context of (2) foresees, on the contrary, that the speaker, who is not acquainted with making cakes, presents a personal evaluation of the prosecution of the recipe, like proposing an attempt or offering an unexpected judgment.

(2) *ABC: ci sbatto la chiara //
'I beat the egg white'

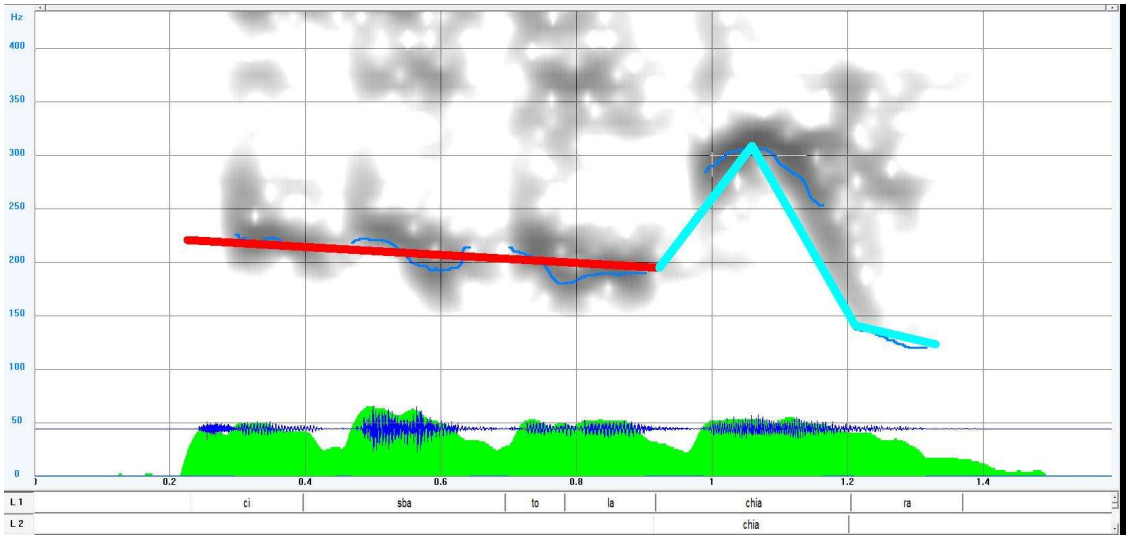

Figure 2: The stylized contour is conventionally tagged as EA for Evident assertion

When presented to a listening group, utterances (1) and (2) were recognized as assertive, and although subjects could not assign a precise value to their difference, they were still distinguished. The question arises whether the different prosodic realizations of the exact wording bear different "nuances" of meaning, expressing attitudinal stances, or whether they correspond to distinct illocutionary types within the assertive class.

The question has its root in the assumption that illocution is based on affect and prosodic contours manifest illocutionary types. In our view, while an affect moves an illocutionary type as a deep psychic involvement conveyed by a dedicated prosodic contour (Cresti 2020), a nuance of meaning retraces to what in the literature is referred to as the broad domain of stance, encompassing epistemic and attitudinal devices (Biber et al. 1999). We consider that nuances are connotations that should be primarily attributed to the lexical and semantic level of the

utterance rather than to the illocutionary foundation of the act. Nuances can be expressed through grammatical devices and operate on the f0 range variations of prosodic contours and voice quality. Still, they do not alter the distinctive movements that define the contour's identity (4.3). We will present arguments supporting the hypothesis that the different prosodic contours, such as those in Figures 1 and 2, convey specific forces moved by affects.

2 Affect and speech

2.1 The Human Birth Theory

The Human Birth Theory (HBT), briefly presented here, is the reference framework for L-AcT. I must cite "Istinto di morte e conoscenza" (Death instinct and knowledge) by Fagioli (1972[1]; 2017[14]), which contains the basic principles of HTB.[4]

The definition of affect can be framed within HBT. Affects are neither generic, sentimental manifestations nor determined by the interactive context, but are free, libidinal human responses linked to mental images and intentionally directed toward the addressee.[5] The affective foundation of speech, as foreseen by L-AcT, is based on the distinction between the concepts of instinct and pulsion, a core notion of HBT. In turn, emotion and affect, which may be inputs to initiate communication, respectively derive from instinct and pulsion.

Conti (2019) reports and discusses exhaustive literature on instinct. The center of his exposition deals with the comparison between instinct and pulsion, as proposed by Fagioli. Generally, we can refer to instincts as a set of adaptive behaviors that all living organisms share in response to external and internal stimuli. They include incorporation, rejection, protection, destruction, reproduction, reintegration, orientation, and exploration (Plutchik 1980). We can say that while instincts are dominant in determining the lives of all other animals (Colgan 1989),[6] according to HBT, human behavior, especially in terms of thought and speech, is a consequence of pulsion.

According to HBT, pulsion is a reaction, the movement of biological reality, that emerges at birth and triggers a unique achievement of our species, called the *disappearance fantasy.* [7] The *disappearance fantasy* is the human response to the photon hitting the newborn's retina,

[4] The scientific work of Fagioli is vast, I can only mention the main volumes through which the Human Birth Theory was developed during the '70: *Istinto di morte e conoscenza* 'Death instinct and knowledge'(Fagioli 1972[1)], *La marionetta e il burattino* 'The Marionette and the Puppet' (Fagioli 1974[1)], *Teoria della nascita e castrazione umana* 'Theory of Human Birth and Castration'(Fagioli 1975[1]), *Bambino, donna e trasformazione dell'uomo* 'Child, woman and man transformation'(Fagioli 1980[1]), *Istinto di morte e conoscenza* has been translated into German, English, French, Portuguese, and Spanish. Until Fagioli died in 2017, this core work was accompanied by books, articles, and public interventions. Among this production, Fagioli's lectures at the University of Chieti and the weekly articles written for the LEFT magazine from 2006-2017, published in annual collections, must be remembered. Given the complexity of the subject, I take full responsibility. Interpretations and errors are mine alone.

[5] Concerning intentionality, the term must be seen in the tradition of Husserl's phenomenology (Husserl 2015; 2021) as the act of perceiving something that exists independently of the percipient as an unconscious direction toward an object. According to the emergence of pulsion, Fagioli developed the concept of intention as the newborn libidinal tendency, which depends on desire and investment (Fagioli 2024:91).

[6] This assumption is connected to a long philosophical tradition that regards Herder as a well-known forerunner. In his Essay on the Origin of Language (Herder 1772), he claims that humans, as regards the strength and certainty of their instincts, are inferior to animals, since it is an established fact that they do not possess at all innate aptitudes and instincts.

[7] The foundation of the theory is presented in the second chapter of "Death instinct and knowledge" Fagioli (2017: 47-105).

after which the brain's activation starts.[8] Due to a "scarcity of vitality," [9] the newborn closes his eyes with a virtual annulment of the surrounding hostile world because of the shock of first immersion in light, noise, and a cold environment. After a short moment (so-called 20 seconds), in the instant of the union of light (inanimate reality) with the biological reality of the brain, when the pulsion emerges, vitality takes on strength.

Thus, the virtual annulment of the world is countered by vitality. The newborn attempts to return to the previous state by creating a *memory fantasy* of the sensory experience that the fetus had in the homeostasis with the amniotic fluid. The fetal experience initially is tactile, but in the last month of pregnancy, it also becomes auditory. It constitutes a neuronal memory trace (mnemic trace), in terms of spontaneous electrical activity, sensory-driven, and represents the physiological maturation of the fetal auditory system. Then, during the early months of life, it facilitates the newborn's ability to decode acoustic input (Perani et al. 2010) and justifies the newborn's clear preference for human spoken productions (see paragraph 2.3).

Conversely, memory fantasy is a newborn creation, constituting the first image and the start of thought. Images are unconscious, not material reality, which contains the libidinal possibility of object relationship, as the fetus has with the amniotic fluid (Fagioli 2024:89). [10] Thus, the memory fantasy that is consequent to the pulsion becomes the "certainty that a breast exists," pushing the newborn to search the human world. Memory fantasy distinguishes our species from other animals, grounding our capacity to imagine and our affective nature.

Then, the internal libidinal situation meets the external datum. The newborn gives origin to unconscious images through continuous and ongoing events of disappearance fantasy within his relationship with the breast, the "physical object".[11] Images do not correspond to the exact mental representation of objects or events from which the newborn separates. Still, they are the transformations of sensory and perceptual activity producing mental entities that are totally new. Images are colored by light and shadow, rich in auditory, tactile, olfactory, taste, and only later, visual experience, and are endowed by the newborn's libidinal relationship with humans.[12] Images are indefinite representations without linear boundaries. Saussure's notion of 'amorphous thought' (Saussure 1967) offers a useful conceptual parallel to what Fagioli describes as the newborn's indefinite representations.

[8] Fagioli's discovery of the brain's activation at birth, triggered by photons hitting the retina, was made on a deductive basis and built upon extensive experience in treating mental illness (Calesini 2017). It has been confirmed over the years through neurobiological research (Long et al. 2005; Polli et al. 2010; Tinsley et al. 2016; Polese et al. 2022). See an interesting review on the impact of light on cognitive mechanisms (Mahagoney & Schimdt 2024).

[9] Fagioli (2019:190). "Lack of vitality (carenza di vitalità), which allows the pulsion of annulment to make the world non-existent, is necessary at birth to create the memory which is the certainty that a breast exists." [my Eng. translation].

[10] Fagioli (2024), "At birth, the death instinct, as a fantasy of the non-existence of the neonatal situation, makes the intrauterine self, that is, the self in relation to an object, exist in the memory trace." […] "What previously outside (amniotic fluid) comes to be inside the newborn as a memory trace of the lost object, a memory trace that includes the image of the object and the possibility of an object relationship" […]"Instead of a projection of image and recollection, a possibility of object relation and a libidinal investment of the object" [my Eng. translation].

[11] It is not the aim of this article to address the complex story that characterizes the newborn's relationship with the breast until separation from the physical object, and the image of the whole psychic object is formed. (Fagioli 2017:147-152). (See below 2. 2).

[12] See Marcella Fagioli (2025) for an exhaustive description of internal images, always accompanied by affective "movements", and their distinction from the concept of figure, which in the HBT stands for the mental reproduction of the exact perception of things in the world.

The unconscious, not material reality, accompanies us from birth to death, as dreams reveal. While instinct characterizes the behavior of all other animals, only the pulsion from which originates the capacity to imagine characterizes human beings in a peculiar way.

Concerning speech, at the threshold of the first year, it is generally assumed that the infant reaches the stage of symbolization (Bates et al., 1979). Fagioli claims that the infant passes from images, as an unconscious, indefinite reality, to verbal images. The infant realizes again what he has done at birth with the creation of the image, since he "abandons" even the image, making a disappearance fantasy of it in favor of the creation of the verbal image (symbol).

It is usually assumed that within the symbolization process, a thing, an entity, or a percept *means* something different to the subject, distinct from its perceptual processing. For instance, the tricolor flag means 'nation' to Italians, and the smell of madeleines means '*le temps perdu*' to Proust (1987). With respect to the subject (C), the symbolic meaning (A) stands for the object (B).

The symbolization has been connected to the infant's self-recognition in the mirror (Lacan 1949). The visual perception in the mirror (B) would stand for self-recognition (A) to the infant (C), thus laying the basis for the symbolization process. Fagioli observes that first, the infant cannot *recognize* himself; he has never seen himself before, and the only visive perception of his face cannot stand for the meaning of self. [13]

But when and how does a process of symbolization occur? Not every perception of a piece of colored fabric waving in the air or every cooking smell is a symbol. We hypothesize that the subject must *give a shape* to something for it to become a symbol. The assignment of a shape allows us first to distinguish an entity from all the others of the same type, and at the same time to give it a value.

Fagioli assumes that the transition to the symbol requires an additional step by the newborn, namely, the creation of the line (Fagioli 2019). [14] The line does not exist in nature, and Fagioli highlights that it is not created by the newborn when he realizes the *disappearance fantasy* at birth, with the onset of unconscious, indefinite images. The line is neither an image nor a recollection of a conscious perception. The line is not even a memory-fantasy of lived experience. Given that the line does not exist in the world, it means that at the threshold of the first year, the infant creates the line.

The infant, looking at himself in the mirror, mentally "draws" a line around the conscious perception of the face that becomes his image and consciousness of self. [15] The infant gives *shape* to his face, defining its form with the realization of his own personal identity distinct from that of others, who are nevertheless similar in their being human. It must be considered that self-knowledge is achieved in the evolution and maturation of the infant who, upon weaning, separates from the mother and a partial image of her dependent on a lustful or even desirous libido. The containment of libido enables the infant to form a complete and well-defined image of the mother, endowed with her human qualities.

The infant creates the line in the moment he gives shape to the conscious perception of his face and assigns it the *sense* of his identity ("I am"), thus reaching a new dimension of creativity. [16] The assignment of meaning to visual perception, defined by the line, establishes

[13] Fagioli taking part in the meetings of the Aula Magna of the University "La Sapienza" of Rome, held in the autumn of 2015, made this fundamental theoretical distinction.
[14] The collection of articles "Left 2016-17" explores and deepens the theme of the line that was proposed for the first time by Fagioli in 1999 in public intervention on collective analysis in Palau, whose partial selection is published as "Negazione e rifiuto" (nagation and refusal) in Sogno della farfalla 2000.
[15] The line gives a sense to the perception (Fagioli 2019:120) [my Eng. translation].
[16] See Bruner (2018) for the visuo-spatial integration of the body and self-noesis.

the capacity to create symbols. From the ability to assign meaning to visual conscious percepts, the infant can move on to giving shape to acoustic sequences and assigning meaning to them.

We assume that after the infant has reached the stage of symbolization, he can mentally give shape to a sequence of syllables (two) with an accent structure. [17] In this regard, it is necessary to cite Saussure, who defines the linguistic sign (word) as a psychic entity that exists in the speaker's mind as a *significant* (acoustic image) associated with a *signifié* (meaning). We can recall that the phonic substance is indefinite, like the amorphous thought, and it acquires definition only within the linguistic relationship with the meaning, becoming the acoustic image of the sign. The word's concept intrinsically implies its acoustic image endowed by articulatory features.

This capacity is consequent to the long infant's auditory experience, and the production of syllables derives from the hard training within babbling (D'Odorico 2005) (see below in 2.3). Shaping the syllabic sequence achieves a fusion between articulation and thought in a way that allows the infant to assign the sense of his experience to the syllabic unit: mommy, dog, toy, ball, music. The indefinite, unconscious image disappears, and the definite, conscious representation of the object appears in the symbolic form of the word (Fagioli 2024).

Although the sense of the word remains internal, the infant can communicate it to others through the physical articulation of the word. However, it is essential to remember that spoken communication emerges after weaning, as the child's libido, which was once greedy or desirous to suck, becomes receptive and investing. This allows the mouth to be freed up to speak, investing his achievements of knowledge in another human being.

In the early acquisition, the infant creates verbal symbols. Then, in life, the human subject proceeds with a largely conscious realization (verbal thought) composed of words taken from his speaker community and provided with conventional meaning. Verbal thoughts must be pragmatically performed as speech acts (Austin 1962), moved by affects.

2.2 Emotion and Affect

Emotion and affect, respectively deriving from instinct and pulsion, are often confused in the linguistic terminology, favoring the exclusive use of the former with a general characterization of spontaneous speech as emotional (Poggi & Magno-Caldognetto 2004; De Marco & Paone 2016; Orrico & Schettino 2022).

We would like to briefly outline a definition of affect according to the HBT by referring to De Simone (1994). The author proposes that affect occurs when the pulsion is linked to the image. That is, humans behave affectively, starting with images. The image maintains a directionality in the relationship with reality, and the affect activates a charge in relation to it.

Affects can have qualities and intensity depending on the dynamics of the psychic relationship with the object. Even if we cannot delve into their complexity, we can mention their general explanation by Fagioli (2017). In short, the subject's relationship with the object (partner) generates an image; if it follows from blind avidity, it matches with an angry affect (Fagioli 2017: 86). If it follows from envy, seeing and denying, it matches with an affect of hatred. (Fagioli 2017: 165). If it follows from listening and availability, it matches with an affect of receptivity (Fagioli 2017:152). If it follows from a vision and interest, it matches with an affect of libidinal investment in the object (Fagioli 2017: 249). In any case, regardless of their

[17] The syllable represents the minimum acoustic unit of all human languages and is missing in the communication systems of other animals. It is composed of a vocalic nucleus in whose resonance box a consonantal sound can be articulated, constituting the prosodic basis.

quality, affects are connected to an image depending on the specific relationship with the object/partner.

The literature on emotions is impressive: it is possible to start with Aristotle (1996), pass through Descartes (2003), and arrive in more recent times at James (1884), and finally at various scientific theories of emotions (Ekman 1982, 1992; Damasio 1994). According to most parts of the literature, a minimal set of basic/universal emotions is identified (fear, aggression, depressive tone, elation, disgust, joy). They are non-intentional physiological reactions triggered by stimuli, non-voluntary, resulting from the automatic processing of the stimulus, characterized by the rapidity with which they activate and the short duration of their manifestation. They are also possessed by lower-ranking animals and consequently regulated by instincts specific to each species. They are the trigger of a large part of animal communication (Colgan 1989).

In classic studies (Arnold 1960; Weiner 1985; Ellsworth 1991; Scherer 2014), two aspects of emotions are distinguished: *arousal*, which refers to the activation of an organism by a suitable emotional stimulus, and appraisal, which involves the cognitive elaboration of emotional states by the subject and encompasses their evaluation. From the perspective of appraisal relevance, Scherer, the most authoritative scholar in the study of voice expression of emotions, proposes the Component Process Model (Scherer 2009). The author disagrees with considering a limited number of basic emotions, as he foresees the occurrence of motivational, neurophysiological, and expressive-motor components, all coordinated by a subjective component, which, according to a defined scheme, contributes to the emotional process. Thus, the Model is centred on the concept of appraisal, which originates from a complex set of emotional families. However, the author feels obliged to justify the basis of appraisal not as a cognitive calculus but as an unconscious modality (Scherer 2009).

We would like to note that every kind of evaluation and even ethical values become "tacked" onto appraisal; however, the evaluative qualities attributed to appraisal ultimately depend on thought. Given that affects are pulsional responses linked to mental images, appraisal seems to belong to an affective rather than an emotional dimension. It is therefore no coincidence that even Scherer, a staunch supporter of appraisal, when noting its unconscious aspect, moves to consider emotions not different from affects in substance.

While the arousal of emotions remains within the limits of physiology and serves as an adequate input for the manifestations of animal communication, appraisal seems to be correctly hypothesized only for humans. Thus, when dealing with emotions, we consider only the arousal of basic emotions, which are also active in human beings, although they are reduced to a marginal sphere of human behavior, which, on the contrary, is led by thought.

Crucially, while the emotion triggered by the stimulus mainly determines animal communication through different modalities suitable to the specific endowment of the species, linguistic behavior is not determined by the stimulus. Chomsky authoritatively argued that language is free (not determined by the stimulus) and creative (albeit with limited creativity in syntax) (Chomsky 1966, 1986). Language possesses a vast lexicon, structural complexity (including double articulation, morphology, hierarchical, and recursive syntax), and, specifically, symbolic meaning, all of which are not found in animal communication systems.

Language is free because it does not respond directly to stimuli but to its psychic processing, following the creation of images. Language is activated by affect, which, unlike emotion, is connected to mental images and is intentionally directed at a human being. Affects

are free libidinal human responses, and what is relevant here is that they are actual input for speech.[18]

When the subject responds to relevant input, he usually intervenes with speech. In so doing, he abandons the unconscious image. However, the affective charge of the image remains, and the speaker directs the speech act to the addressee following this affect.[19] There is no language without thought, but the physical manifestation of speech needs an "embodiment," which is turned on by the affect.

Moreover, to be understood and shared, linguistic communication must fulfill social conditions through historical systems (phonological, morphological, lexical, semantic, syntactic) unique to each language (locution), and conventional speech acts (illocution) as assertion, question, order, instruction, deixis, irony, excuse, thanks, greetings, etc. (see Table 1 below). Unlike linguistic systems, the essential repertoire of speech acts is shared across diverse human societies.

2.3 Affect and prosody

The L-AcT framework extends Austin's speech act theory by positing that prosody serves as the necessary interface between the linguistic fulfillment (locution) and the illocution of the speech act. Prosody is the physical means that expresses the affect, moving the communication activity. We might wonder why affect should be conveyed specifically by prosody.

This long story begins with the auditory sensations of the fetus, which, in the last month of pregnancy, are stimulated by sequences of sounds from the maternal voice filtered through the uterine wall. Only the selection of waves with a fundamental frequency (f0) below 600 Hertz can cross the wall. It is worth noting that these sequences replicate the rhythmic alternations of f0 peaks and valleys in the maternal sound, which are the core characteristics of human prosodic patterns (Ghio et al. 2022). As anticipated, unlike what is assumed in the literature (Hewes 1996; Gervain & Mehler 2010; Gervain 2015; Mariani et al. 2023), we are not compelled to hypothesize mental memory in the fetus.

According to HBT, the fetus cannot mentally process sensory stimuli (Fagioli 2019), and the neuronal memory trace only participates in the physiological maturation of the fetal auditory system. At birth, the newborn cannot form a "cognitive memory" of the auditory sensations it experienced in the womb. However, the auditory system of the newborn is almost complete, adequate for the task, and preferentially directed towards human voices. The fact that the fetal auditory system has been physiologically trained by prosodic patterns of the human voice in the womb seems sufficient to explain at birth its adequacy and preference for speech without the necessity of a cognitive recollection.

Thus, concerning perception, the stimulation of human voices shaped by prosody is pervasive and accompanies the response of care that keeps the newborn alive. Voices are intrinsically endowed with the affective value of human relationships.

Regarding newborn vocal productions, they must be considered within a prelinguistic communicative system, which can be briefly distinguished into three developmental stages: crying and vocalization sounds with control of the phonatory system, vocal games, and only at a later stage, babbling (D'Odorico 2005).

[18] "The word must therefore be understood with the pulsion story of those who use it to bring closer or push away those who receive it" (Fagioli 2017) [My Eng. translation].

[19] "The image is transformed into a signal, and the libido, freed from the internal image, invests the object" (Fagioli 2024) [My Eng. translation].

The birth cry is socially expected and conventionally considered the beginning of every human being's life, whereas the birth of other mammals is silent and has no social significance. Fagioli (2009:49) writes that the "word of the first cry says, in the universal language, 'I am a human being'" [my Eng. translation].

Crying vocalizations are the first newborn's forms of expression. They are simple vocalic sounds, like screaming, repeated according to different rhythms. They seem uncontrolled, as their fundamental frequency is high, they exhibit few variations in height, and they progress in a crescendo. Even if they mainly depend on the stimuli of hunger, huff, and malaise, they often also express contact requests, distinguishing between annoyance and discomfort, depending on the psychic relationships with the breast and the affective images developed by the newborn. However, soon, during the exploration of the phonatory system, the newborn begins to vary his productions, which is a typical human behavior. Each newborn exhibits his personal "style", ritualizing the expression of states of mind, especially well-being (D'Odorico & Franco 1991). Their quality is up to prosody. Overall, a newborn's vocalization usually receives attention from the caregiver, who interprets its internal affective value and hopefully satisfies the baby. These initial vocal productions, with their fundamental prosodic distinctions, are intentionally directed at humans and work from a communicative perspective (Flax et al. 1991; Papaleiu & Trevarthen 2006; Balog & Brentary 2008; Esteve-Gibert & Prieto 2012).

Then, auditory feedback maturation enables the infant to become aware of his productions, allowing him to shape and distinguish his vocalizations in vocal games that correspond to prosodic patterns systematically conveying the newborn's state of mind. Only later does the child engage in the articulation of syllabic sounds (babbling), which acquisition and control are very challenging. They, unlike prosodic patterns, have no evident communicative value, specifically serving as training for the formation of the syllables of a particular language. The production of the child's first verbal symbols occurs through a creative act of giving shape (in a linear form) to a syllabic sequence, thereby creating the signal of meaning.

In parallel, the implementation of prosodic patterns continues throughout the entire early stage of language acquisition, gradually conforming to those of the language in which the infant lives (Furrow 1984; Avesani et al. 2022).

The choice of prosodic patterns as the first means of internal affective expression and intentional communication does not depend on the complex articulatory programs required to produce syllables. Prosody represents an "easy" means of expression and communication before the symbolic capacity. Thus, prosody, affects, and human relationships are linked from the very beginning of human life.

2.4 Emotions and attitudes

Of course, humans, like all other animals, activate emotions. Non-voluntary physiological changes in humans as a response (arousal) to inputs triggering emotions include a rapid heartbeat, paleness or rush of blood, slowing or speeding up of breathing, movement blockages or urges to distance oneself, skin effects, and changes in body hair, which are characterized by quick activation and a short duration. In humans, emotions remain as sensory experiences confined to the subject, and their external manifestations are not intentionally directed at another, even if they transmit information to people who interpret them.

Studies are specifically dedicated to the effects on the speaker's phonation when subjected to emotional stimuli. They focus on the narrowing or widening of the trachea and oral cavity and the changes in thickness and tension of the vocal cords (Anolli & Ciceri 1999; Johnston & Scherer 1999; Magno-Caldognetto 2002). The result is an overall characterization

in voice (breathy, laryngealized, whispered), which does not concern the carrying out of any speech act.

It must be highlighted that the analysis of spontaneous speech corpora shows that in regular interactive exchanges, the isolated realization of "pure" emotion is rare (screams of fright, exclamations of joy, shouts of anger) (Poggi & Magno-Caldognetto 2004; Moneglia 2008).

Emotional connotations of speech acts (such as sadness and joy) are frequently considered together within stance and attitudinal meanings (including uncertainty, haste, seductiveness, and mock) and social habits (such as politeness and professional roles) (Mello et al. 2012; Moraes 2012; Moraes & Rillard 2014).

The distinction between emotion and attitudes on one side and affect on the other is crucial from our perspective.[20] Emotions cannot properly constitute the foundation of a speech act, lacking in origin the thought content of the image itself that is intentionally directed at the addressee. An image with its affective charge must already exist; only on this basis can an emotional or social attitude connotation take place. We will address the issue of prosodic connotations associated with an illocutionary type by highlighting that they are, in some sense, overlaid on its prosodic performance, observing how they can operate on range variations of the prosodic contour but not on its form (4.3).

3 L-AcT and the corpus-based study of Illocution

3.1 Corpus-based Illocutionary Taxonomies

Proposing new illocutionary types based on an affective criterion, as anticipated in the introduction, may be surprising. Our objective must be considered within a recently redefined landscape. Since the 2000s, corpus-based observations have led to the emergence of speech act taxonomies that have transformed the illocution classification as it had been defined up to that point according to classical logical taxonomies (Austin 1962; Searle 1979; Searle & Vanderveken 1985; Sbisà 2023). Briefly, the main corpus-based taxonomies are the following:

- Tokyo University for Foreign Studies (TUF) (Kawaguchi et al. 2006)
- Dialogue Annotation and Research Tool (DART) (Weisser 2018)
- Dialogue Act Markup using Several Layers (DAMSL)
- Taxonomy of Dialogue Acts, Annotation Scheme**,** (DIT++) (Bunt et al. 2017)

These taxonomies have identified illocutionary types that were not previously foreseen and are based on lexical and deductive criteria. Furthermore, corpus-based taxonomies document a vast repertoire of illocutionary types (TUF approximately 60; DART approximately 120; DIT++ approximately 100), suggesting the possibility of making discoveries while exploring new data.

DIT++ and DART can be considered the most significant outcomes of this tradition. For assertion, within DIT++, the illocution is introduced in the following terms: the speaker provides the addressee with certain information that he believes the addressee is not aware of, and which he assumes to be correct. Assertive speech acts are identified as the speaker's reasons for providing the information: Inform, Answer, Confirm, Disconfirm. The four subtypes have a logical definition based on what the speaker thinks the addressee believes.

[20] In the L-AcT framework, most of the emotional connotations and the expression of stance are classified within the illocutionary class of expression (see Table1 below) according to a set of conventional speech acts correlating with dedicated prosodic contours (Cresti 2020).

The DART taxonomy, on the contrary, is highly granular. Each assertive illocutionary type depends on the interpretation of the linguistic content of the utterance in the context and lacks a logical definition. Around 65 main assertive illocutionary types are foreseen. For instance, the act of referring varies according to what is referred to (number, condition, date, day, reason, thing…). This approach suggests that expanding the dataset will lead to a random extension of illocutionary types.

However, none of the previous taxonomies can offer us useful tools to analyze and pragmatically distinguish examples (1) and (2): according to DITT++ they are both cases of Information production and according to DART, due to the lack of a performance explaining the speaker's action, can be brought back to a generic assertion used for Referring. Moreover, neither of the two approaches considers prosodic realization as a distinctive characteristic of the illocutionary type.[21]

3.2 The L-AcT approach to illocution

The Language into Act Theory (Cresti 2000), has generated a corpus-based taxonomy of around a hundred illocutionary types (Cresti 2020) (see Table 1). The taxonomy derives from the analysis of spoken Romance corpora (Italian, French, Spanish, European, and Brazilian Portuguese), collected according to a comparable corpus design), known as the C-ORAL family (Cresti & Moneglia 2005; Raso & Mello 2012; Nicolas 2012.

Table 1: L-AcT illocutionary Taxonomy (from Cresti 2020 with modifications)

Assertive		**Directive**	
WEAK	Self-conclusion On-going comment Confirmation Neutral assertion / Explanation Assertion taken for granted Literal citation	**APPEARANCE COMMUNICATIVE INVOLVEMENT**	Distal recall (non-visible addressee) Distal recall (visible addressee) Proximal recall Functional recall (CMM)
		CHANGE OF ATTENTION	Distal deixis (moving object) Distal deixis (still object) Proximal deixis Prompt Event presentation Mental deixis
		MENTAL TRANSFORMATION	Instruction Introduction of a person Correction Reported speech Warning Alert
STRONG	Answer Ascertainment Evident assertion Hypothesis Conclusion Assent/Dissent	**LINGUISTIC BEHAVIOR**	Partial question Open questions Request for confirmation/ Challenging questions Alternative question (Positive /Negative) Double questions Tag question
		BEHAVIOR	Order Request Prohibition Invite

[21] Important research on the prosodic characters of attitudes and illocution has been developed by Raso and his team (LEEL- Laboratory at the Federal University of Belo Horizonte) (Mello et al. 2012; Rocha 2016; Luma et al. 2020).

			Offer
		ENDORSEMENT	Commitment (Bet, Promise) Proposal Authorization

Expressive		**Ritual**		**Refusal**
BELIEF	Contrast Softening Expression of Obviousness Irony Disbelief /Doubt Admission Waiver /Renouncement Rhetorical question	**COURTESY RITES (SOCIAL FIELD: EDUCATION AND CIVIC LIFE)**	Thanks Greetings Welcome Excuses Wishes Congratulations Condolences Compliments	
FEELINGS **MOODS** **STATE OF MIND**	Protest Complaint Grumbling Imprecation Surprise/Wonder Wish Easing	**BOND RITES (SOCIAL FIELD: LOW, RELIGION, INSTITUTIONS)**	Legal declarations Convictions Penalties Dedications Religious rites	
SPEAKER/ADDRESSEE RELATION	Approval/Disapproval Derision Defiance Reproach Mockery Allusion /Concession	**DIALOGIC MOVES**	Back Channel Repetition request Request to stop Request to wait	

The methodology was developed through leading collaboration exchanges, such as those with Blanche-Benveniste (Blanche-Benveniste 1997) and the Laboratório de Estudos Empíricos e Experimentais da Linguagem (LEEL) (Raso & Mello 2012; 2014).

L-AcT's pragmatic approach is speaker-oriented. Prosody is considered the necessary interface between locution and illocution. The boundaries of each utterance are identified by terminal prosodic breaks (Swerts 1997; Barbosa & Raso 2018), and utterances are aligned to the acoustic source accordingly (Moneglia 2005; Raso & Mello 2012; Izre'el et al. 2020). Utterances are segmented into Information units corresponding to prosodic units marked by non-terminal breaks (Moneglia 2011; Moneglia & Raso 2014).

We only recall that the information unit of Comment is considered the necessary and sufficient information unit of the utterance since it conveys illocutionary force, allowing for its pragmatic interpretability. The Comment is performed through dedicated prosodic units of the *root* type ('t Hart et al. 1990), which contours correlate with illocutionary variations (Cresti 2000; Firenzuoli 2003). Therefore, the pragmatic and prosodic analyses in the article will be conducted only on the Comment units of assertive and interrogative utterances.[22]

4 Assertive illocutionary types

[22] The article does not address the affective implications of the stanza, considered the second reference unit of speech in L-AcT, and those that regard the information structure of the utterance. In the stanza, an affective constancy toward the addressee is required, with the possibility of varying degrees of involvement. Various considerations regarding the Topic, the Parenthetic and the Allocutive information units have been proposed in Moneglia (2021), Cresti and Moneglia (2022; forthcoming), and Cresti (forthcoming).

Keeping in mind the above theoretical background, this second part of the article examines assertive and question illocutionary types. The research is run on real examples taken from the LABLITA corpus. We introduce the assertive illocution class, explaining its pragmatic definition based on specific parameters that reflect the speaker's affective involvement. We will show that the pragmatic interpretation of spoken data is guided by the modeled prosodic contours grounded on the speaker's affect. Prosodic cues are necessary for assigning an illocutionary value to the utterance.

The pragmatic intentionality of assertion, as is generally assumed, is such that the speaker assures the addressee of his commitment to the truth of what he states (Searle 1979; Sbisà 2023). According to L-AcT, we emphasize that this commitment is explicitly realized as a manifestation of the speaker's identity. Only because of his own identity originating at birth, does the speaker develop a thought of unconscious affective images on which the verbal thought is based. Thus, even the conscious verbal thought derives its perspective from the unconscious reality of the speaker's identity in relation to the world and, above all, to other human beings.

The speaker intentionally presents descriptions, representations, conclusions, judgments, discoveries, reflections, and evaluations to the addressee as realizations of his verbal thought. Through the assertion, the speaker shows his self-confidence by bringing his identity into play with the addressee.

The speaker's commitment to the truth, however, can also vary depending on the relevance of the semantic content of the assertion. This is not measured by its objective relevance (for instance, Einstein's relativity) but rather by the speaker's interest in its value for the ongoing situation and the human relationship.

Affective involvement varies in quality and intensity (2.2). Even if we cannot determine from corpus data its libidinal nature (we should be immersed in the situation, sharing the relationship between the participants), we can at least evaluate its affective charge through voice qualities.

Accordingly, the distinctive pragmatic features defining the assertive class depend on:

- The strength of the speaker's commitment to the truth of the content
- The relevance of the semantic content of the utterance for the speaker
- The affective involvement/interest of the speaker towards the addressee

The previous distinctive features must be considered 'scalar'. According to their respective grades and combinations, assertive illocutionary types are grouped into two subclasses: "strong" and "weak" assertions. A low affective force substantially unifies the weak assertive subclass, while the strong subclass shows a high affective force.

Both the weak and strong assertive subclasses include several illocutionary types (see Table 1). In 4.1 and 4.2, we will deal with two assertive types: Evident assertion of the strong subclass and Taken-for-granted assertion of the weak subclass. We will show that prosodic performance, encompassing intensity, speed, phonetic accuracy, and voice quality, contributes to assigning their respective illocutionary value. Moreover, we will see that each type may be optionally connotated by a speaker's attitude, working as a mitigation or intensification.

4.1 Evident Assertion

Let's look at a prototypical example of Evident assertion according to the scalar features listed above. The Evident assertion presents a free judgment on an aspect of the context that the speaker finds evident. [23]

Example (3) is taken from a poker game between young friends. The players are offered diluted whisky, and after modifying the name of a well-known brand to suggest it has "knock-out" properties, the speaker ironically asserts that the drink's quality is quite the opposite. Figure 3 shows the spectrogram and stylization of the F0 contour of (3).

(3) *LAK: lo Sballantines //
'the Sballantines'
*CIC: c' è un po' d' acqua eh // ma l' è meglio //
'there's a bit of water, huh. But it's better.
*LAK: mh // **fa digerire //**
'mh. it helps digestion'

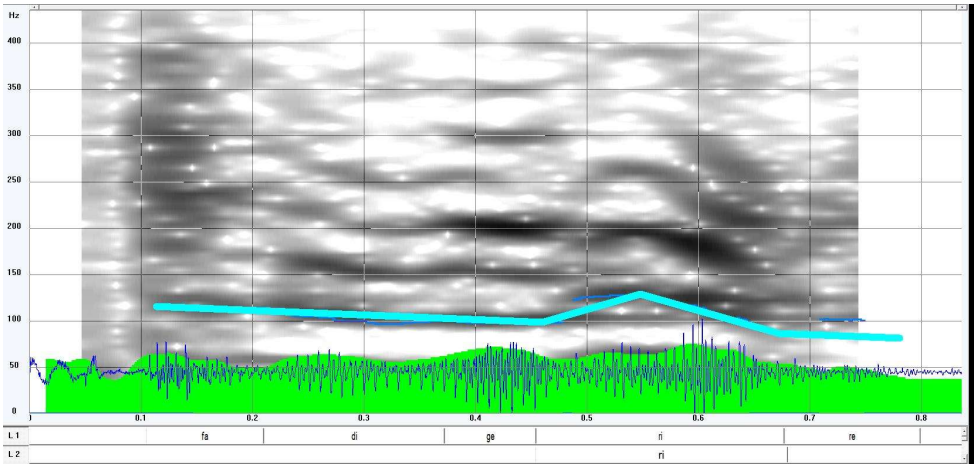

Figure 3: Evident Assertion Type (Strong Sub-Class)

The prosodic contour of an Evident assertion consists of three movements:

- A plateau at medium f0 values
- A brief rising movement
- A falling movement aligned with the tonic syllable of the final semantic word in the Comment unit

Each of the movements composing the contour is over the glissando threshold. An impressionistic observation shows that intensity is high, and the speed is medium to slow. The phonetic realization is precise, and the voice is clear with some syllable-by-syllable performance. The contour of (3), together with the other acoustic features, is comparable to that of example (2), independently of their different wording.

The corpus-based research on spoken Italian, more generally, confirms that the prosodic contours of utterances conveying a similar pragmatic value are comparable (Cresti & Moneglia

[23] The term *evident* records an ambiguity that derives from the meaning of the term *evidence*, which can refer to both material proofs, as in a judicial process, and perceptual and mental evidence implying supposition. We use the term in this second meaning.

forthcoming). All contexts found in corpora matching the above contour highlight the speaker's assessment of an aspect of the context, that is, the addressee's behavior or a linguistic act by the addressee, and, in the case of (3), the quality of an object. The speaker freely offers his unexpected judgment as if it were an evident fact and aims at the addressee to accept his point of view without further proof. However, knowing that this is just his personal view, the speaker often softens the claim by giving it an ironic connotation, as in example (3).

The parameters of this type are as follows:

- The strength of commitment to truth is medium, as the speaker is aware he is presenting a personal judgment.
- The relevance of semantic content for the speaker is high since it represents a personal way of characterizing the interaction.
- The affective involvement is high, as the speaker aims to gain agreement on his point of view, which he knows is not substantiated.

The pragmatic definition of an Evident assertion can be summarized as follows:

The speaker asserts a free judgment of an aspect of the context, thus offering an unexpected implementation of the exchange. The commitment to truth has a mid-value and is often mitigated by ironic connotations. However, the content is relevant for the speaker because it is a function of his creativity. Moreover, given that the speaker is aiming for the addressee's agreement, this implies a high involvement with the latter.

4.2 Assertion Taken-for-Granted

The Taken-for-granted assertion belongs to the weak assertive subclass. It frequently appears in the flow of speech as a transition between arguments. With a Taken-for-granted assertion, the speaker provides background information, adding little novelty to the conversation and little push to the interaction.

In (4), still from the poker game, a player fails to realize it is his turn to bet, and the speaker sarcastically points out indirectly the trivial fact that turns in poker are fixed (*I'm the one after you*). Figure 4 shows the spectrogram of an example of a Taken-for-granted assertion.

(4) *CIC: le metti / o no ?
'are you betting, or not?'
*LAK: le metto sì // ma perché / t' aspettai me ?
'yes, I am. But why, were you waiting for me?'
*CIC: **son dopo di te** //
'I'm the one after you'

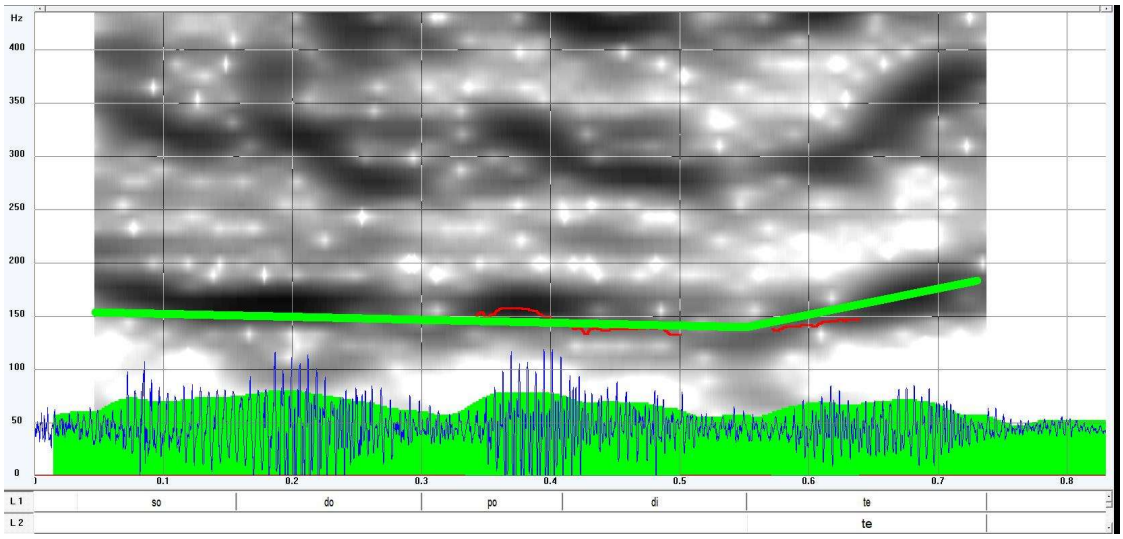

Figure 4: Taken-for-granted Assertion (Weak Sub-Class)

The prosodic contour of a Taken-for-granted assertion consists of two movements:

- A gently descending plateau at mid-values of f0
- A rise aligned to the final tonic syllable, which is also lengthened and continues into the post-tonic.

The contour of (4) is comparable to that of example (1), independently of their wording. An impressionistic observation shows that intensity is medium, speed is slowed, the phonetic realization is precise, and the voice is clear.

All contexts found in corpora matching the above contour highlight that the speaker provides information that is not particularly significant or that the speaker assumes might already be known to the addressee or expected. The speaker does not aim to expand the addressee's knowledge and often downplays the content, suggesting it is trivial.

The parameters of this type are as follows:

- The speaker is stating information that should be available to the addressee; thus, his strength of commitment to truth is low
- The speaker knows that he is giving little impulse to the exchange; therefore, the relevance of the semantic content is low
- The speaker is not very interested in sharing his point of view with the addressee, only expecting the latter's registration, and his affective involvement is low

Contrary to the assertive types of the strong subclass, all parameters of the Taken-for-granted assertion are set to a low grade. The pragmatic definition of a Taken-for-granted assertion within the weak assertive subclass can be summarized as follows:

The speaker asserts that information is low-relevant since it should have already been shared or expected by the addressee. The speaker gives little force to his commitment to the truth, adding a frequent diminishing connotation to the content. Given that he waits for the addressee to register the assertion, he demonstrates a scarce investment in the latter.

Compared with other corpus-based taxonomies for the classification of new illocutionary types within the assertion, Evident assertion and Taken-for-granted assertion can be identified as intentional conditions on the speaker's affective involvement toward the addressee outside of any logical principles. Such conditions can be appreciated only when considering prosodic performance.

Regarding the connotations of illocutionary types (2.4), in previous examples, the Evident assertion's connotation is ironic, while that of the Taken-for-granted assertion is sarcastic. We noticed that connotations are overlaid on prosodic performance, observing that they can operate on f0 range variations of the prosodic contour and voice quality, but do not modify the distinctive movements that constitute the contour's identity.

For instance, if we compare the contours of the two Taken-for-granted types in (1) and (4), the contour is the same, but the rise in height of (1) is steeper and higher than that of (4). In (1), the speaker assumes that the addressee already knows the content and consequently performs a taken-for-granted version without any attitude. In contrast, in (4), where the addressee appears to lack a clear understanding of the game, the speaker lowers their profile due to sarcasm. Although both semantic contents remain of scarce relevance, justifying the Taken-for-granted illocutionary type, their connotations can be appreciated through the prosodic variants.

5 Directive illocutions: Questions

Within L-AcT, the directive illocution is based on the speaker's affect corresponding to desire and satisfaction from the addressee (Cresti 2017). It implies that the speaker considers the addressee's knowledge, capacities, and willingness to collaborate, waiting for the latter to provide some modification of the world. The affective basis of direction is well differentiated from that of assertion, which is centred instead on the speaker's self-confidence and the expectation of acceptance by the addressee.

In the Direction class, the addressee should transform the world, his attentional horizon, behavior, or knowledge. Accordingly, the repertory records various sub-classes, among which we mention Calling, Deixis, Presentation, Request for action, Request for mental change, and Request for linguistic behavior (see Table 1).

Within L-AcT, questions are more generally described as activities directed to *the addressee's linguistic behavior* rather than seeking information acts (Braun et al. 2019), as they are also defined in corpus-based taxonomies (Bunt et al. 2017; Weisser 2016). In other words, we consider that what is prompted by a question is not exclusively limited to getting a piece of information (Cresti 2020; Cresti & Moneglia 2023).

The sub-class of Request for linguistic behavior includes: a) Total questions (Request for confirmation and Challenging questions); b) Partial questions, extended to Open questions; c) Illocutionary patterns of questions (Affirmative and Negative Alternative questions, Double questions and Tag-questions).

The salient features of the pragmatic definition of the Request for linguistic behavior are still scalar but different from those of the assertive type, depending on:

- The speaker's assumptions about the background knowledge of the addressee
- The type of linguistic behavior the addressee is requested
- The speaker's affective involvement toward the addressee

Each illocutionary type of question is characterized by the previous distinctive parameters, which change in quality and grade. We present two examples taken from the LABLITA corpus, belonging to the Question sub-class: Request for confirmation and Open question types.

5.1 Request for confirmation

Requests for confirmation are Total questions satisfied by a yes/no answer. In the Request for confirmation, the speaker positively represents an eventuality. In so doing, he should at least assume its possible occurrence as a *hypothesis*. This seems like a minimal intentional preparatory condition to keep the question as a meaningful speech act. Under this assumption, the linguistic behavior enacted by the speaker should not be defined as a *seeking for information* act since it is instead a *request for the addressee's agreement on the speaker's hypothesis*. The condition is independent of the speaker's presupposition of the truth of his hypothesis and must be held in all circumstances (*biased* or *unbiased*). [24]

It is worth noting that the Request for confirmation is the most frequent type among Total questions in our Italian corpus, at 64%. Let's examine example (5) from a dialogue between two women. The speaker wonders where the village (*Castiglion de' Pepoli*) mentioned by the addressee is, and she is uncertain about the place suggested by the latter (*Sasso Marconi*). Thus, she hypothesizes that the village is near Bologna, asking for confirmation of her suggestion. Figure 5 shows the spectrogram of this example of a Request for confirmation.

(5) *ELA: […] ndo' l' è / Castiglion de' Pepoli ?
'where is Castiglion de' Pepoli?'
*LIA: dunque / avanti di arrivare a Sasso Marconi // sull' autostrada //
'well, before reaching Sasso Marconi. On the highway.'
*ELA: […] **vicino Bologna** ?
'near Bologna?**'**

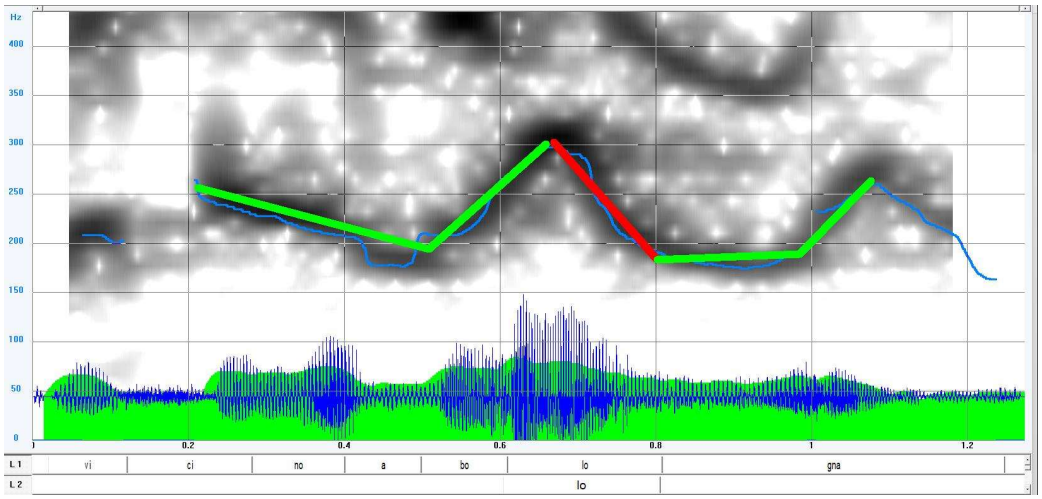

Figure 5: Request for confirmation (Request for linguistic behavior sub-class)

The prosodic contour of Figure 5 corresponds to a *valley* which is composed of three movements in the nucleus of the root unit (Cresti & Moneglia 2023):

- A falling movement on the tonic syllable of the last word of the Comment
- A holding movement (under the glissando)
- A rising movement on the post-tonic

[24] Being a state of mind, the presupposition of the truth of the content remains underdetermined. Conversely, the consistency of the hypothesis condition has been verified in the Italian corpus (Cresti & Moneglia 2023). Moreover, our corpus data does not verify the variation of a contour in correspondence with either a positive or negative presupposition (*bias*) (Savino 2012; Vanrell et al. 2013).

The nucleus, which is responsible for the illocutionary force, is often preceded by some optional syllables of preparation on the pretonic syllables (*vicino*), as can be seen in Figure 5.

The pragmatic parameters of the Request for confirmation are as follows:

- The speaker provides a hypothesis without any necessary assumptions on its truth.
- The behavior expected by the addressee is confirmation or disconfirmation of the speaker's hypothesis.
- The speaker's affective involvement toward the addressee is mid

When the speaker makes a Request for confirmation, he only hopes that the addressee is conciliatory with his hypothesis, and to this end, the speaker often mitigates his request with a polite attitude (Leech 2014). The pragmatic definition of a Request for confirmation can be summarized as follows:

The speaker makes a hypothesis and asks the addressee to confirm it. The addressee is expected to answer while the consideration of his actual knowledge and opinion remains in the background. The speaker's affective involvement toward the addressee is mid. The speaker may mitigate his request with a polite attitude.

We can compare the speaker's behavior in the case of a Request for confirmation and when he accomplishes the other type of Total question, which we call Challenging questions. They are realized by a dedicated prosodic contour, consisting of a rising movement followed by a rapid falling one on the tonic syllable (Cresti & Moneglia 2023). These Total question types are known in other Romance languages as *declarative questions* (Martin 2015). They are activities challenging the addressee to *validate* the speaker's linguistic behavior as an intersubjectively shared assumption. They are injunctive and fully fit with a directive request, so their affective involvement is also strong, unlike the mild one of a Request for confirmation.

Let's see example (6):

(6) *LID: […] gli facevo / Pallino / Pallino // avanti ha fatto / fr fr ha girato / va' // poi s' è
messo / a schiena 'n su / e come &s / si gongolava //
'I called him: Pallino, Pallino. Before he did: fr fr. He turned, finally. Then he put himself to his back 'n' up and gloated'
*ELA: [<] <cioè> /a pancia all' aria ?
'you mean: belly in the air?
*LID: a pancia all'aria//
'belly in the air'

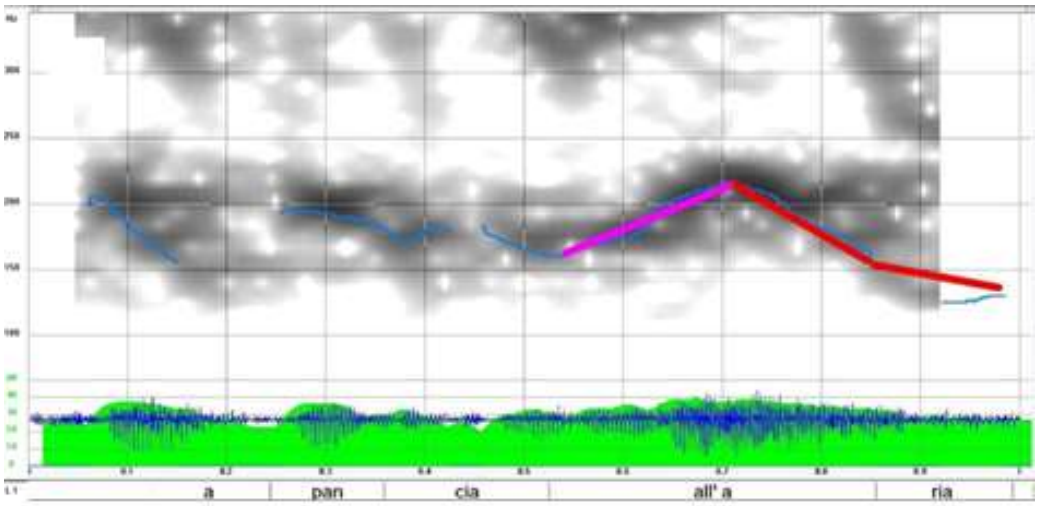

Figure 7: Challenging question (Request for linguistic behaviour subclass)

In the above contexts, the information questioned —the description of the purring cat —is already at the disposal of the addressee. Crucially, the question is not supported by any speaker's *hypothesis*. In no circumstance can the above context be paraphrased as *the speaker makes the hypothesis of the occurrence of the given eventuality, asking the addressee for its confirmation*. The act may be peremptory in connection with the greater or lesser speaker's interest in validating the ongoing linguistic exchange.

We have seen that an ironic attitude may mitigate the Evident assertions, and a sarcastic attitude may diminish the value of the content of the Taken-for-granted assertions. Mitigation may also apply to questions whose prosodic contour can be overlaid by a *courtesy* attitude (Moraes & Rillard 2014). This is what happens in (7), where the speaker's Request for confirmation (*before Christmas*) contains information already presented by the addressee (*on December 23rd or 24th*). Therefore, the Request for confirmation is only a courtesy move, and the latter is not foreseen to offer any disagreement. Figure 7 shows the spectrogram of the example of Request for confirmation realized to a lower f0 range than example (5).

(7) *CLA: mh / va a sciare //
'he goes skiing]
*EST: sì ? quando parte ?
'really? when is he leaving?
*CLA: il / ventitré o ventiquattro //
'on December 23rd or 24th'
*EST: **ah / prima di Natale** ?
'so / before Christmas ?

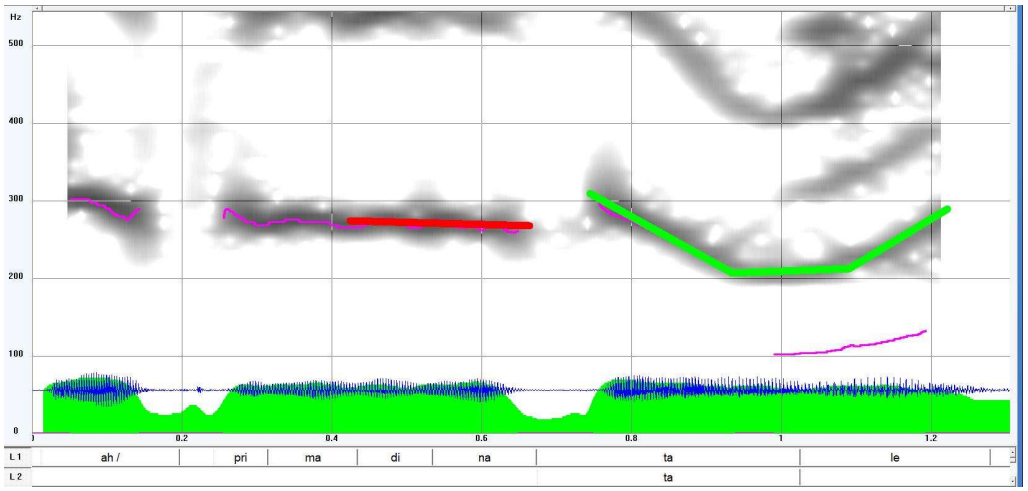

Figure 7: Request for confirmation (Request for linguistic behavior sub-class)

5.2 Partial Questions and the Open type

Partial questions are properly seeking information acts and can be defined as speech acts in which the speaker asks for information, whose focus is a variable to be satisfied in the answer. According to Cresti & Moneglia (2023), they represent nearly 38% of questions in Italian spontaneous speech. Their syntactic structures typically correspond to verbal phrases preceded by a *wh*-interrogative morpheme, which can occur alone or at the beginning of the Comment unit. The prosody of Partial questions is a falling contour as in (8).

(8) *EST: **quando parte**?
'When does she leave?'

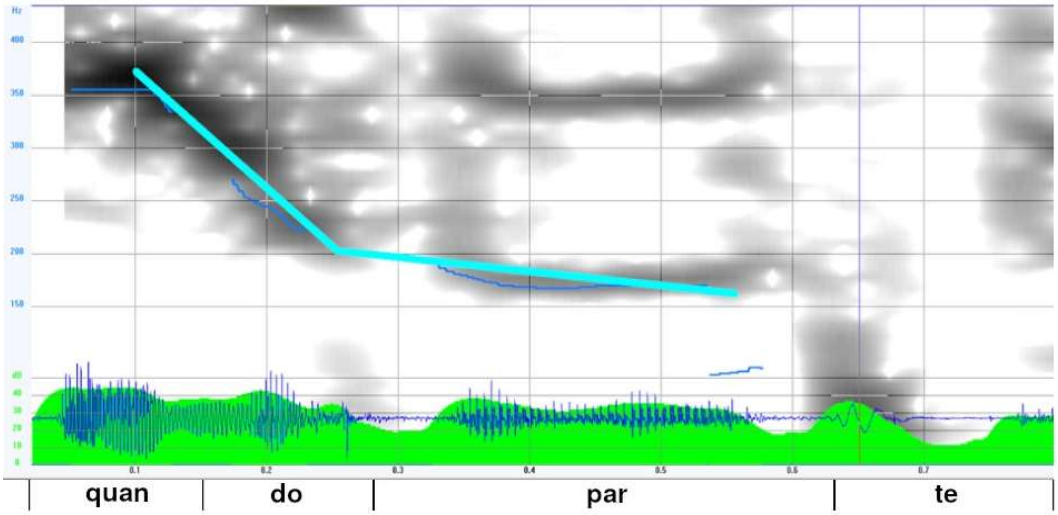

Figure 8: Partial Question

The falling contour of Partial questions occurs on the tonic vowel of the variable, which a rapid declination can characterize depending on the height at the start.

However, it is worth noticing the idiosyncratic prosodic performance of questions filled by *perché (why*). This *wh*-morpheme is often performed through a rising contour in isolation as in (9).

(9) *ELA **perché**?
' Why?'

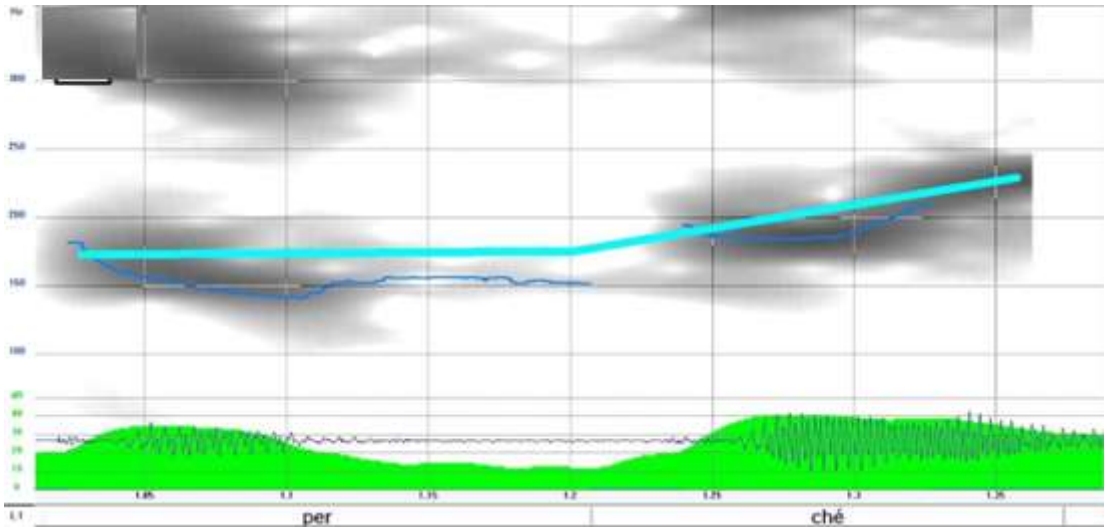

Figure 9: Partial Question (Why)

Partial questions introduced by *perché* are realized through a unique movement *slightly rising* (over the glissando), mostly performed in isolation or followed by an optional descending plateau (*tail*) on the rest of the Comment.

Moreover, among Partial questions, we must face a specific type called Open question. This represents a peculiar case characterized by a rising movement comparable to that of questions introduced by *perché*.

Let's examine example (10), taken from a dialogue between an esthetician and a friend during a depilation session. They are chatting about holidays. Figure 10 shows the f0 track of the Open question.

(10) *PAO: perché la domenica / tu se' qui / a <negozio> //
'because on Sunday, you are here, at the store'
*FRA: da urlo //
'screaming'
*PAO: ma / &he / **pe' quest' estate** ? come / tu se' messa ?
'but, and for this summer? how do you manage it?

In Open questions, the syntactic structure is relevant since it corresponds to a verb-less phrase (noun phrase, adverbial phrase, adjectival phrase, verb nominalization).

Even if Open questions are requests without a *wh*-morpheme, they are also seeking information questions, and as Partial questions, they ask for a full semantic answer. Their phrasal structure invites completion. For instance, the Open question '*and for this summer*?' asks for a free explanation by the addressee to give sense to the incomplete prepositional phrase constituting the question.

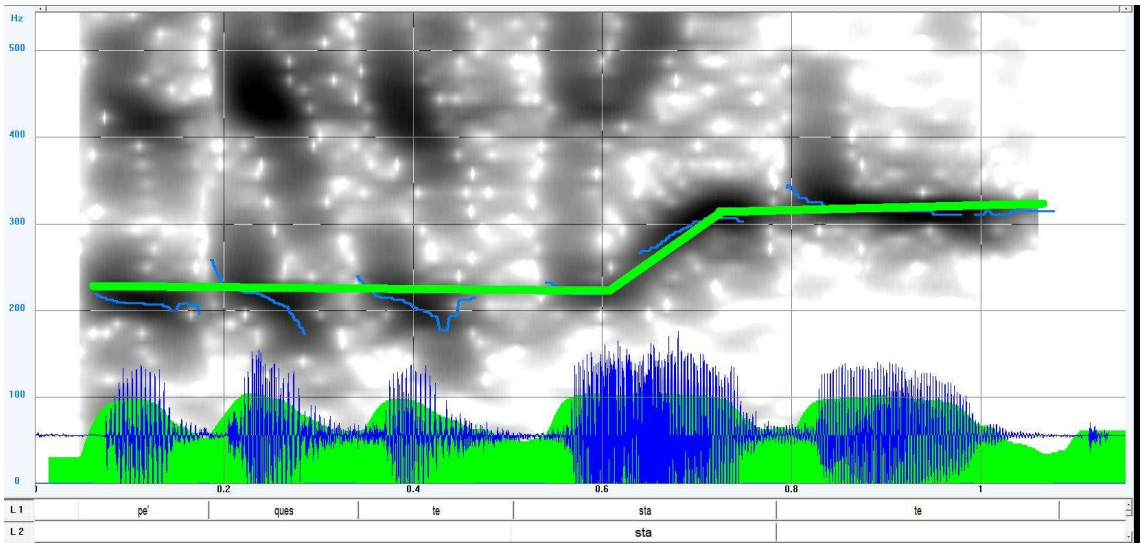

Figure 10: Open question

Of course, Open questions cannot be answered by a y/n utterance just as Total questions can. However, from a syntactic point of view, they show a syntax that can be equivalent to that of a Total question. For instance, the Request for Confirmation (5) (*near Bologna?*) and the Open question (10) (*for this summer?)* do not offer any syntactic criterion of distinction since they are both prepositional phrases. Only prosody distinguishes the pragmatic value of Open questions.

The prosodic contour of Open questions differs from the falling one of Partial Questions, as shown in (8), and the valley of the Request for Confirmation, as seen in (5) and (7). The Open's contour is composed of three movements: [25]

- an initial flat plateau at the mid-value of f0
- a rising movement on the tonic syllable reaching a high f0 value
- a gently rising plateau/holding on the post-tonic syllable

At first glance, the intensity is high, the speed is fast, and the voice is clear with a final modulation.

When the speaker accomplishes an Open question, he cannot accept a confirmation nor intend to obtain a predetermined completion of a variable from the addressee, like in Partial questions. The speaker waits for the addressee to complete the subject. [26] The pragmatic parameters of Open questions are as follows:

- The speaker does not provide a hypothesis but presents an incomplete argument as the subject of the exchange, positively estimating the addressee's knowledge.
- The speaker expects the addressee to participate in the exchange freely, completing the argument according to his knowledge.

[25] The contour of the Open question has been roughly described in Cresti & Moneglia (2023), as a *continuous rising movement*, failing to provide a more realistic description. The reason has been to distinguish the Open question from the falling contour of the other Partial question types, highlighting its rising form.

[26] Open questions are often followed by a speaker's Partial question, limiting the addressee's freedom to determine the answer. In this case, the content of the Open question, as in example (7) '*for this summer'*, could be interpreted as the Topic of the Comment that should accomplish the Partial question '*How do you manage it*?'. However, this is not the case since the prepositional phrase is performed through a dedicated contour bearing an illocutionary force, as shown in Figure 10. Thus, '*for this summer*' also results in a Comment that can be interpreted as an independent question in isolation.

- The speaker's affective involvement toward the addressee is high.

The pragmatic definition of an Open question can be summarized as follows:

The speaker presents to the addressee an incomplete argument as the subject of the spoken exchange, positively estimating the addressee's knowledge. The speaker aims for the addressee to complete the argument freely. The speaker's affective involvement toward the addressee is high.

Despite their peculiar prosodic realization, partial questions introduced by *why* and open questions without an introducer, from a pragmatic point of view belong to partial questions because they are acts asking for information and are focused on a morphological variable to be saturated, explicit in partial questions and as an implicit constituent in open questions (*how is it possible? how do you do it*?). Open questions, even if they are rarely considered in the literature (Weisser 2018), are not marginal in the language usage, representing a significant percentage of Italian Partial questions (13%).

5.3 Question variations and Kohler's attitudinal hypothesis

The explanation of the different question types according to our affective-illocutionary hypothesis can be compared to the proposal by Kohler (2004), who foresees that the prosodic variations of questions found in the Kiel corpus of spoken German depend on attitudinal meanings.

Questions - that are syntactically classified by Kohler - find *default* correlations between, on one side, the rising prosodic contour and the syntactic structure marked by the verb in the first position in German (Total questions), and on the other, the *falling* contour and the syntactic structure marked by *wh*- variables (*Partial questions*). He claims that the two prosodic contours bear different *attitudinal meanings*: the rising one (Total) expresses a positive orientation toward the addressee with interest and openness towards him, while the falling (Partial) is oriented toward the speaker who needs a response.

However, he frequently found unexpected prosodic variations of syntactically marked questions in his corpus. Then Kohler hypothesizes that the default link between attitudes and syntax can be restructured during communication. Each syntactic type may occur with a not-expected contour because of a change in attitudinal meaning. Therefore, according to Kohler, the prosodic contour is a function of attitudinal meanings that can contradict the one expected for a specific syntactic construction.

It should be noted, however, that if this were the case, it would be necessary to identify a) the set of attitudes that remain underdetermined and b) the correlations between those attitudes and their prosodic counterparts.

Beyond the fact that the L-AcT framework is based on the Italian corpus, which mainly consists of nominal questions without a verbal predicate, which makes the syntactic criterion of classification only partially usable, it must be considered the distinction between affects and attitudes, which Kohler does not consider.

Indeed, the default Kohler's explanation of the two prosodic contours for the Total and Partial questions, attributed to different *attitudinal meanings* (positive orientation toward the addressee vs. orientation toward the speaker), is relevant. Still, from our perspective, they should be conceived as affective parameters that ground each speech act type.

The Request for Confirmation, which syntactically is a Total question, expresses rather than an attitude of positive orientation toward the addressee, as proposed by Kohler, a specific speech act (defined in 5.2). From our perspective, the request asks the addressee to confirm the

speaker's hypothesis, characterized by mild affective involvement. A Partial question, instead of an orientation toward the speaker, is an injunctive act with a strong speaker's affective involvement.

Regarding the unexpected prosodic performance of syntactically marked questions, the problem is that Kohler takes syntax as the primary character for question type identification in German. Our empirical research on Romance corpora shows that every time a prosodic contour is repeated and is consistently found in the corpus, it corresponds to one illocutionary type with a pragmatic identity independently of its syntactic structure. For instance, the previous cases of Request for Confirmation (Total) and Open Questions (Partial) cannot be distinguished by their syntax.

The same is true within Total questions. As anticipated (5.2), they split between Requests for Confirmation and Challenging questions. Challenging questions are injunctive, and the speaker aims to have the addressee validate his assumption. Their affective involvement is strong, unlike the mild one of the Request for confirmation. Accordingly, the two pragmatic types vary in their prosodic contours: valley contour for Request for confirmation and rising movement followed by a rapid falling one on the tonic syllable for Challenging questions. Again, the two types cannot be distinguished by their syntactic form.

This observation is even more valid in the assertive field, where no syntactic distinction can be made for the different illocutionary types. In our examples (1) and (2), not only do they share the same syntactic structure of the declarative sentence, but even the same wording gives rise to two illocutionary types characterized by different affective involvement (Evident assertion and Taken-for-granted), depending on the different prosodic realizations. Thus, the prosodic contour goes hand in hand with an illocutionary type, which, in turn, has a specific affective foundation.

6 Comparison between Taken-for-granted assertion and Open question

We want to add a piece to demonstrate the distinctive value of the prosodic contour for the pragmatic interpretation of an utterance. Two prosodic contours can be well differentiated to convey distinctive illocutionary values, as in the case of Evident and Taken-for-granted assertions. However, it may be the case that their prosodic distinction is subtle.

For instance, comparing the prosodic contours of the 'Taken-for-granted' and the 'Open' questions, which are strongly distinct pragmatic activities, a rising movement characterizes both (see Figures 1, 4, and 10). We wonder how the same apparent contour might be interpreted as an assertion or a question.

It could be assumed that contextual and interactional aspects determine the interpretation of the two similar contours. This explanation will contradict our assumption concerning the necessary function of prosody in assigning illocutionary value. We dealt with the issue through an experiment.

The similarity of the above contours found in corpora has been replicated in the Lab to disentangle them. To obtain the actor's right elicitation, the same wording (*la macchina rotta* 'the broken car') was inserted into two different contexts (Context 1 and 2) supporting the respective pragmatic values.

Context 1: A boy tells of his incident, which left him in the countryside with his car stopped. Figure 10 shows the f0 track of (11):

(11) *ABC: mi son trovato fermo in mezzo ai campi // **la macchina rotta** // non sapevo cosa fare //

‘I found myself stopped in the countryside. The car has broken down. I did not know what to do’

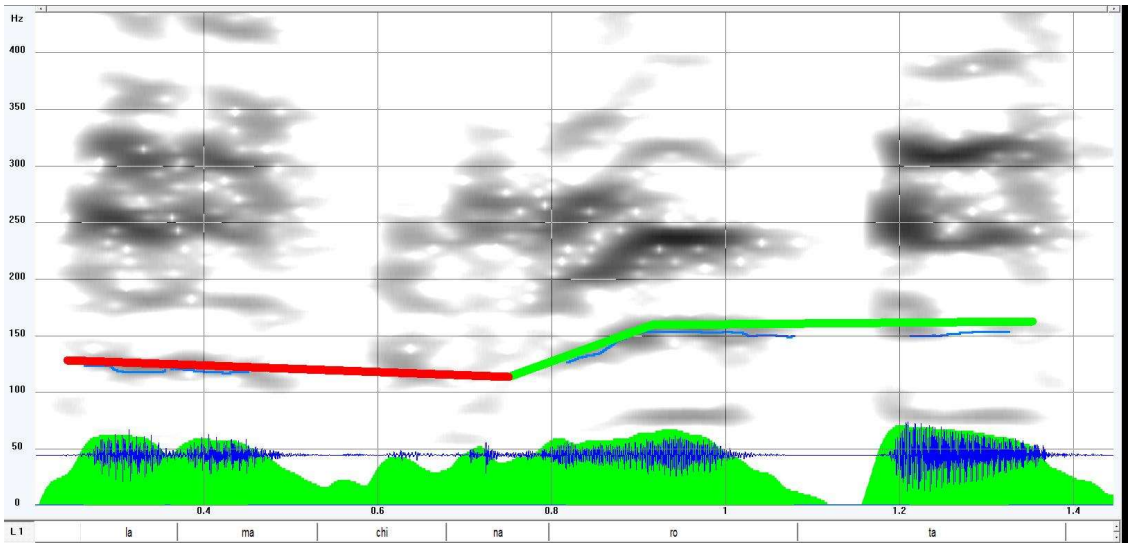

Figure 11: Taken-for-granted assertion

Context 2: A boy tells a friend about his troubles during a trip, and the other asks him how he has managed. Figure 11 shows the f0 track of (12).

(12) *ABC: mi son trovato in mezzo ai campi / la macchina era bloccata // sono venuto via a
piedi //
‘I found myself in the countryside. The car was stuck. I came back walking.’
*DEF: **e la macchina rotta ?** come hai fatto?
‘the broken car? how did you manage with it?’

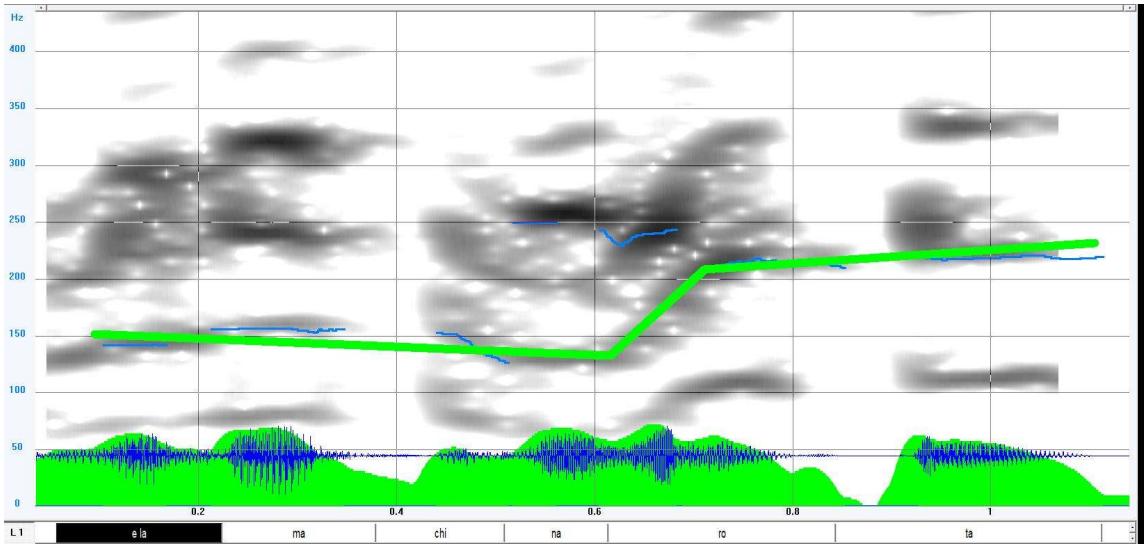

Figure 12: Open question (Request for linguistic behavior sub-class)

The prosodic performances of the elicited utterances (*the car broken down*) in (11) and (12) are found to be comparable, respectively, with the contours of the Taken-for-granted assertion in (4) and the Open question in (10).

When the two utterances are proposed in isolation to native speakers, they are distinguished. Even if subjects cannot give a precise value to their difference, a generic value of assertion is assigned to the first, and the value of the question to the second. The assertion is felt "suspended". Even if their prosodic difference is subtle, the judgments, which cannot rely on contextual features, are clear and can only be based on their prosodic differences. We can observe the following distinctive features.

The prosodic contour of the Taken for granted assertion in Figure 11 consists of two movements: a gently descending plateau at low f0 levels and a rise at medium f0 levels aligned to the final tonic syllable, which is also lengthened and continues into the post-tonic. The speed is slowed.

The prosodic contour of the Open question in Figure 12 corresponds instead to three movements: an initial flat plateau at the mid-value of f0, a steep rising movement on the tonic syllable reaching a high f0 value, and a gently rising plateau on the post-tonic syllable. Unlikely Taken-for-granted assertion, the intensity is high, and the speed is fast. The overall production of the wording is performed at a higher f0 level and is more rapid than in Figure 11.

Despite the similarity of the two contours for their prominent final rise, the above differential prosodic features and the utterance duration allow the clear perceptual distinction and the assignment of a different illocutionary value. Therefore, the context does not determine the interpretation.

7 Conclusions

Within the framework of Language into Act Theory, we have presented corpus-based and experimental arguments to identify two assertive illocutionary types (Evident assertion and Taken-for-granted assertion) and two question illocutionary types (Request for confirmation and Open question) in spoken performance, which were not previously identified in corpus-based taxonomies.

According to the Human Birth Theory, it is assumed that illocutionary acts are deeply moved by the affects, which are pulsion events linked to images and intentionally directed at humans. Affects must be distinguished from emotions, which originate from instinct, but also from attitudes considered social connotations.

We have shown the role of the speaker's affect in defining pragmatic activities and how the prosodic contours are responsible for interpreting pragmatic acts in speech. Prosody is the means of manifestation of illocutionary force that is shaped into modeled contours matching conventional pragmatic forms (speech acts). The distinction between prosodic contours bearing illocutionary value and the prosodic properties that are a function of the speaker’s attitude is faced, showing that the latter do not change the prosodic model.

The identification in the corpus of the assertive types, Evident assertion and Taken-for-granted assertion, and the question types, Request for confirmation and Open question, is based on the prosodic form of the Comment unit, whose variability is found in correlation with pragmatic parameters according to L-AcT methodology.

The pragmatic parameters of assertive types are scalar and connected to the speaker’s affective force: the degree of the speaker’s commitment to the truth of the semantic content, the relevance of the content for the continuation of the exchange, and the speaker’s affective involvement with the addressee.

In parallel, the scalar parameters of question types are the speaker’s estimation of the addressee’s knowledge, the type of linguistic behavior requested by the addressee, and the speaker’s affective involvement toward the latter.

Conditions on such parameters define the specific illocutionary value of each type. Despite variations in syntactic and semantic fulfillment, these values remain constant and are connected to a specific prosodic contour across the corpus contexts.

The differential value of prosodic contours and their capacity to determine the illocutionary interpretation have been tested experimentally by laboratory utterances with the same word sequence performed as a Taken-for-granted assertion and an Open question, both characterized by a final rise movement. Despite their similar contours, even when listened to in isolation, native speakers can distinguish the two utterances and properly assign them a generic value of assertion or question. This constitutes a test bench for the overall hypothesis that the prosodic contour precisely ensures the pragmatic interpretability of utterances beyond the contextual environment.

REFERENCES

1. Anolli L, Ciceri R. La voce delle emozioni. Verso una semiosi della comunicazione vocale non-verbale delle emozioni. Milano: Franco Angeli;1999.

2. Aristotele. Retorica. Milano: Mondadori; 1996.

3. Arnold MB. Emotion and personality. New York: Columbia University Press;1960.

4. Austin JL. How to Do Things with Words. Oxford: Oxford University Press; 1962.

5. Avesani C, Bonifacio S, Calignano G, D'Aloia V, Olivucci F, Vayra M, Zmarich C. The emergence of lexical and post-lexical prominence in Italian. A case study. In: Orrico R., Schettino L., editors. La Posizione del Parlante nell'Interazione: Atteggiamenti, Intenzioni ed Emozioni nella Comunicazione. Milano: Officina21; 2022. p. 11-34.

6. Balog HL, Brentari D. The relationship between early gestures and intonation. First Language. 2008; 28: 141–63.

7. Barbosa P, Raso T. Spontaneous speech segmentation: functional and prosodic aspects with applications for automatic segmentation. Revista de estudio do da linguagem, 2018; 26(4): 397-1433.

8. Bates E, Benigni L, Bretherton I, Camaioni L, Volterra V. The emergence of symbols: Cognition and communication in infancy. New York: Academic Press; 1979.

9. Biber D, Johansson S, Leech G, Conrad S, Finegan E. The Longman Grammar of Spoken and Written English. London: Longman; 1999.

10. Blanche-Benveniste C. Approches de la langue parlée en français. Paris: Ophrys; 1997.

11. Braun B, Dehé N, Neitsch J, Wochner D, Zahner K. The prosody of rhetorical and information-seeking questions in German. Language and Speech. 2019; 62(4): 779–807.

12. Bruner E. La mente oltre il cranio. Roma: Carocci; 2018.

13. Bunt H, Petukhova V, Traum D, Alexandersson J. Dialogue Acts Annotation with the ISO 24617-2 Standard. In: Dahl D., editor. Multimodal Interaction with W3C Standards Toward Natural User Interfaces to Everything. Berlin: Springer; 2017. p. 1-27.

14. Calesini I. Fagioli's Human Birth Theory and the possibility to cure mental illness, International Journal of Environment and Health. 2017; 8(3):185–192.

15. Chomsky N. Cartesian Linguistics. New York: Harper & Row; 1966.

16. Chomsky N. Knowledge of language: its Nature, Origin, and Use. New York: Praeger; 1986.

17. Colgan P. Animal Motivation. London: Chapman and Hall; 1989.

18. Conti EM. From animal instinct to Human Birth Theory: An entangled path, International Journal of Environment and Health. 2019; 9(4): 343–371.

19. Cresti E. Corpus di italiano parlato. Firenze: Accademia della Crusca; 2000.

20. Cresti E. The empirical foundation of illocutionary classification. In: De Meo A, a Dovetto FM., editors. Atti del convegno, La comunicazione parlata, Roma: Aracne; 2017. p. 243-264.

21. Cresti E. The pragmatic analysis of speech and its illocutionary classification according to the language into act theory. In: Izre'el S., Mello H., Panunzi A., Raso T., editors. In Search of Basic Units of Spoken Language: A Corpus-Driven Approach. Amsterdam: Benjamins; 2020. p. 181-220.

22. Cresti E. Topic vs. Allocutive in the Language into Act Theory. Corpus-based research on spoken Italian. In: Cimmino D, Ozerov P., editors., Disentangling Topicality Effects, Special issue, Linguistic Typology at the Crossroad, forthcoming.

23. Cresti E, Moneglia M., editors. C-ORAL-ROM. Integrated reference corpora for spoken romance languages. Amsterdam: Benjamins; 2005.

24, Cresti E, Moneglia M. The role of prosody for the expression of illocutionary types. The prosodic system of questions in spoken Italian and French according to Language into Act Theory. Frontiers in Communication. 2023; 8(1).

25. Cresti E, Moneglia M. The role of prosody for the expression of illocutionary force. Variation of Speech act types within the assertive class in spoken Italian. In: Atti del LVII Convegno internazionale della Società di linguistica italiana, forthcoming.

26. Damasio A. Descartes' Error: Emotion, Reason, and the Human Brain, New York: Putnam & Sons; 1994.

27. De Marco A, Paone E. Dalle emozioni alla voce, Nuovi orizzonti della comunicazione in italiano L2, Roma: Carocci; 2016.

28. Descartes R. Le passioni dell'anima. Milano: Bompiani; 2003.

29. De Simone G. Conoscenze e affetti. Il dire e il fare in psicoterapia. In: Lalli N, Cavaggioni G, Fiori-Nastro P., editors. Il processo terapeutico in psicoterapia. Roma: Edizioni Universitarie Romane; 1994. p. 158-183.

30. D'Odorico L. Lo sviluppo linguistico. Roma: Laterza; 2005.

31. D'Odorico L, Franco F. Selective production of vocalization types in different communication contexts. Journal of Child Language. 1991; 18: 475–99.

32. Ekmann P. Emotion in the Human Face. Oxford: Oxford University Press; 1982.

33. Ekmann P. An argument for basic emotions. Cognition and Emotion. 1992; 6(3/4): 169–200.

34. Ellsworth PC. Some implications of cognitive appraisal theories of emotion. In: Strongman KT., editor. International review on Research on emotions. New York: Riley; 1991. p. 143-161.

35. Ellsworth HA, Smith CA. Shades of Joy: Patterns of Appraisal Differentiating Pleasant Emotions. Cognition and Emotion. 1988; 2(4): 301–331.

36. Esteve-Gibert N, Prieto P. Prosody signals the emergence of intentional communication in the first year of life: evidence from Catalan-babbling infants. Journal of Child Language. 2012; 40: 1–26.

37. Fagioli M. Istinto di morte e conoscenza: pensieri di psicoanalisi. Roma: Armando; 1972[1]; Istinto di morte e conoscenza. Roma: L'Asino d'oro; 2017[14].

38. Fagioli M. La marionetta e il burattino. Roma: Armando; 1974[1]; Roma: L'Asino d'oro edizioni; 2011.

39. Fagioli M. Teoria della nascita e castrazione umana. Roma: L'Asino d'oro edizioni; 2012 (1975[1]).

40. Fagioli M. Bambino, donna e trasformazione dell'uomo. Roma: Nuove Edizioni Romane; (1980[1]).

41. Fagioli M. Left 2006. Roma: L'Asino d'oro; 2009.

42. Fagioli M. Left 2016-2017. Roma: L'Asino d'oro; 2019.

43. Fagioli M. Negazione e rifiuto. Il Sogno della farfalla. 2000; IX(1): 9-14.

44. Fagioli M. Leggendo WR. Bion. Il Sogno della farfalle. 2024; XXXIII (2); 53-101.

45. Fagioli Marcella. La parola dell'inconscio. Roma: L'Asino d'oro edizioni; 2025.

46. Firenzuoli V. Le forme intonative di valore illocutivo dell'italiano parlato: analisi sperimentale di un corpus di parlato spontaneo (LABLITA) [PhD thesis]. Università di Firenze; 2003.

47. Flax J, Lahey M, Harris K, Boothroyd A. Relations between prosodic variables and communication function. Journal of Child Language. 1991; 18: 3–19.

48. Furrow D. Young children's use of prosody. Journal of Child Language. 1984; 11(1): 203–213.

49. Gervain J. Plasticity in early language acquisition: the effects of prenatal and early childhood experience. Current Opinion in Neurobiology. 2015; 35: 13–20.

50. Gervain J, Mehler J. Speech perception and language acquisition in the first year of life. Annual Review of Psychology. 2010; 61:191-218.

51. Ghio M, Cara C, Tettamanti M. The Prenatal brain readiness for speech processing: A review on fetal development of auditory and primordial language networks. Neuroscience and Biobehavioral Reviews. 2021; 709-719.

52. 't Hart J. Psychoacoustic backgrounds of pitch contour stylization. IPO Ann. Prog. Rep. 1976; 11: 11–19.

53. 't Hart J, Collier R, Cohen A. A Perceptual Study on Intonation. An Experimental Approach to Speech Melody. Cambridge, Cambridge University Press; 1990.

54. Herder JG. Saggio sull'origine del linguaggio (Amicone AP. editor) Parma: Nuova Pratiche Editrice; 1995.

55. Hewes GW. A History of the Study of Language Origins and the Gestural Primacy Hypothesis. In: Lock A, Peters CR., editors. Handbook of human symbolic evolution. Oxford: OUP; 1996. p. 263–269.

56. Husserl E. Ricerche logiche (G. Piana ed.) Milano: Il Saggiatore; 2015.

57. Husserl E. Fenomenologia dell'inconscio. I casi limite della coscienza. (Failla M. editor). Milano: Mimesis; 2021.

58. Izre'el S, Mello H, Panunzi A, Raso T. (2020). In Search of Basic Units of Spoken Language: A Corpus-Driven Approach. Amsterdam: Benjamins; 2020.

59. James W. What is an Emotion?. Mind Association. 1884; 9(34): 188–205.

60. Johnstone T, Scherer K. The effects of Emotions on Voice Quality. In: Ohala JJ, Hasegawa Y, Ohala M, Granvillew D, Bayley AC., editors. Proceedings of 14th International Conference of Phonetic Sciences, Berkeley: University of California; 1999. p. 2029-2032.

61. Kawaguchi Y, Zaima S, Takagaki T., editors, Spoken Language Corpus and Linguistics Informatics. Amsterdam: Benjamins; 2006.

62. Kohler K. Pragmatic and Attitudinal Meanings of Pitch Patterns in German Syntactically Marked Questions. In: Fant G, Fujisaki H, Cao J, Xu Y., editors. From Traditional Phonology to Modern Speech Processing. Bejing: Foreign Language Teaching and Research Press; 2004. p. 205–214.

63. Lacan J. Le Stad du Miroir comme formateur de la fonction de Je, telle qu'elle nous est revelee dans l'experience psychoanalytique. Revue Francaise de Psychanalyse. 1949; 13: 449–455.

64. Leech G. The Pragmatics of politeness. Oxford: Oxford University Press; 2014.

65. Long MA, Cruikshank SJ, Jutras MJ, Connors BW. Abrupt maturation of a spike-synchronizing mechanism in neocortex. Journal of Neuroscience. 2005; 25: 7309 –7316.

66 Luma M, Moraes J, Rillard A. Statistical modeling of prosodic contours of four speech acts in Brazilian Portuguese. Proceedings of 10th International Conference on Speech Prosody; 2020. p. 404–408.

67. Magno-Caldognetto E. I correlati fonetici delle emozioni. In: Bazzanella C, Kobau P., editors. Passioni, emozioni, affetti. Miano: McGraw-Hill; 2002. p. 197-213.

68. Mariani B, Nicoletti G, Barzon G, Ortiz Barajas MC, Shukla M, Guevara R, Suweis SS, Gervain J. Prenatal experience with language shapes the brain. Science advances. 2023; 9(47) Avaliable from https://pmc.ncbi.nlm.nih.gov/articles/PMC10664997/

69. Mahoney H, Schmidt T. The cognitive impact of light: illuminating ipRGC circuit mechanism. Nature Review Neuroscience. 2024; 25: 159-175.

70. Martin Ph. WinPitch Corpus, a Text to Speech Alignment Tool for Multimodal Corpora. In: Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC'04), Lisbon, Portugal. Paris: ELRA; 2004.

71. Martin Ph. The structure of Spoken language. Intonation in Romance. Cambridge: Cambridge University Press; 2015.

72. Mello H, Panunzi A, Raso T, editors. Pragmatics and Prosody. Illocution, Modality, Attitude, Information Patterning and Speech Annotation. Florence: Florence University Press; 2012.

73. Moneglia M. The C-Oral-Rom resource. In: Cresti E, Moneglia M., editors. C-ORALROM. Integrated Reference Corpora for Spoken Romance Languages. Amsterdam: Benjamins; 2005. p. 1-70.

74. Moneglia M. Le unità di riferimento per l'espressione delle emozioni: dati dal corpus C-ORAL-ROM. In: Magno Caldognetto E, Cavicchio F, Cosi P., editors. Comunicazione parlata e manifestazione delle emozioni. Atti del I convegno GSCP. Napoli: Liguori; 2008.

75. Moneglia M. Spoken corpora and Pragmatics. Revista Brasileira de Linguìstica Aplìcada. 2011; 11(2): 479-519.

76. Moneglia M. Le unità di informazione Parentetiche alla periferia destra del Comment nella Teoria della Lingua in Atto. DILEF. 2021; 1(1):88-123

77. Moneglia M, Raso T. Notes on the Language into Act Theory. In: Raso T, Mello H., editors. Spoken corpora and linguistics studies. Amsterdam: Benjamins; 2014. p. 468–494.

78. Moneglia M, Cresti E. Il Topic modale. In: Mollica A, Onesti C, editors. Studi in onore di Carla Marello. Welland (ON) Canada: éditions Soleil; 2022. p. 203-224.

79. Moraes J. From a Prosodic Point of View: Remarks on Attitudinal Meaning. In: Mello H, Panunzi A, Raso T., editors. Pragmatics and Prosody. Illocution, Modality, Attitude, Information Patterning and Speech Annotation. Florence: Florence University Press; 2012. p. 19–37.

80. Moraes J, Rillard A. Illocution, Attitudes, and Prosody: a multimodal analysis. In: Raso T, Mello H., editors. Spoken corpora and linguistics studies. Amsterdam: Benjamins; 2014. p. 233-270.

81. Nicolas C. C-or- Di Al. Corpus Oral Didactico Anotado Linguisticamente. Madrid: Liceus; 2012.

82. Orrico R, Schettino L., editors. La Posizione del Parlante nell'Interazione: Atteggiamenti, Intenzioni ed Emozioni nella Comunicazione Verbale. Milano: Officina21; 2022.

83. Perani D, Saccumanni MC, Scifo P, Spada D, Andreoli G, Rovelli R, Baldoli C, Koelsch S. Functional specializations for music processing in the human newborn brain. Proceedings of the National Academy of Sciences of the United States of America. 2010; 107 (10): 4758–4763.

84.Papaeliou CF., Trevarthen C. Prelinguistic pitch patterns expressing 'communication' and 'apprehension'. Journal of Child Language. 2006; 33: 163–78.

85. Plutchik R. Emotion: A Psychoevolutionary Synthesis. New Yor: Harper & Row; 1980.

86. Poggi I, Magno-Caldognetto E. Il parlato emotivo. Aspetti cognitivi, linguistici e fonetici. In: Albano-Leoni F, Cutugno F, Pettorino M, Savy R., editors. Atti del Convegno "Italiano Parlato. Napoli: D'Auria (CD rom) 2004.

87. Polese D, Riccio ML, Fagioli M, Mazzetta A, Fagioli F, Parisi P, Fagioli M. The Newborn's Reaction to Light as the Determinant of the Brain's Activation at Human Birth. Frontiers, Integrative Neurosciences, 2022; 16.

88. Polli D, Altoè P, Weingart O, Spillane KM, Manzoni C, Brida D, Tomasello G, Orlandi G, Kukura P, Mathies RA, Garavelli M, Cerullo G. Conical intersection dynamics of the primary photoisomerization event in vision. Nature. 2010; Sep 23. 467(7314):440-3.

89. Proust M. À la recherche du temps perdu. Paris: Gallimard; 1987.

90. Raso T, Mello H., editors. C-ORAL-BRASIL I: Corpus de referência de português brasileiro falado informal. Belo Horizont: Editora UFMA; 2012.

91. Raso T, Mello H., editors. Spoken corpora and linguistics studies. Amsterdam: Benjamins; 2014.

92. Rocha B. 2016. Uma metodologia empírica para a identificação e descrição de ilocuções e a sua aplicação para o estudo da Ordem em PB e Italiano [PhD thesis]. Belo Horizonte: UFMG; 2016.

93. Rossi M. Le seuil de glissando ou seuil de perception des variations tonales pour la parole. Phonetica. 1971; 23: 1-33.

94. Rossi M. La perception des glissando descendants dans les contours prosodiques. Phonetica. 1978; 35: 11–40.

95. de Saussure F. Corso di linguistica generale. (Tullio De Mauro editor). Roma: Laterza; 1967.

96. Savino M. The intonation of Polar Question in Italian: where is the rise? Journal of the International Phonetic Association. 2012; 42(1): 23-48.

97. Sbisà M. Essays on Speech Acts. Oxford: Oxford University Press; 2023.

98. Scherer K. Vocal affect expression: A review and a model for future research. Psychological Bulletin. 1986; 99: 143-65.

99. Scherer K. Corpus design for studying the expression of emotion in speech. In: Raso T, Mello H., editors. Spoken corpora and linguistics studies. Amsterdam: Benjamins; 2014. p. 210-32.

100. Scherer K. The Dynamic Architecture of Emotion: Evidence for the Component Process Model. Cognition and Emotion. 2009; 23(7): 1307–51.

101 Searle J. A Taxonomy of Illocutionary Acts. In: Searle J., editor. Expression and meaning: studies in the Theory of Speech Acts. Cambridge: Cambridge University Press; 1979. p. 1–29.

102. Searle J, Vanderveken D. Foundations of Illocutionary Logic. Cambridge: Cambridge University Press; 1985.

103. Swerts M. Prosodic features at discourse boundaries of different strength. The Journal of the Acoustical Society of America. 1997; 10(1): 514–21.

104. Tinsley JN, Molodtsov MI, Prevedel R, Wartmann D, Espigulé-Pons J, Lauwers M, Vaziri A. Direct detection of a single photon by humans. Nature Communications. 2016; 7: Article number: 12172.

105. Weiner B. An Attributional Theory of Achievement Motivation and Emotion. Psychological Review. 1985; 92(4): 548-94.

106. Vanrell MM, Mascaró I, Torres-Tamarit F, Prieto P. Intonation as an encoder of speaker's certainty: information and confirmation yes–no questions in Catalan. Language and Speech. 2013; 56(2): 163-90.

107. Weisser M. DART – the Dialogue Annotation and Research Tool. Corpus Linguistics and Linguistic Theory. 2016; 12 (2): 355–388.

108. Weisser M. How to Do Corpus Pragmatics on Pragmatically Annotated Data: Speech Acts and Beyond. Amsterdam, Benjamins; 2018.

Online resources:

109. DAMSL: <https://www.lancaster.ac.uk/fass/projects/eagles/restrict/manual.htm>

110. DIT++:< http://dit.uvt.nl/>

111. LABLITA CORPUS <http://corpus.lablita.it>/

112. Winpitch <https://www.winpitch.com/>