A Cognitive View on Prosodic Relations

Doina Jitcă

Correspondence: Doina Jitcă, Institute of Computer Science of Romanian Academy-Iassy Branch, Romania.

Received: May 24, 2024 Accepted: June 26, 2024 Online Published: July 9, 2024

doi:10.5430/elr.v13n2p1 URL: https://doi.org/10.5430/elr.v13n2p1

Abstract

The paper presents a new F0 contour partitioning approach by using the category of prosodic relations also used by Ladd (2008) for improving the F0 contour descriptions based on phonological categories. The new approach involves a cognitive view on the low-high and high-low 'metrical' structures of prosodic relations, by relating them to the structures of cognitive relations generated during speech object representations at the cortical level. In section 2, the paper presents the information structure model by defining the cognitive categories that describes prosodic relations of F0 contours involving their two overlapped structures and nuclear positions. CU_predicate-CU_argument and CU_theme-CU_rheme are the two structural levels of prosodic relations. The model proposes a binary-tree hierarchy to describe the articulation of prosodic relations within utterances. Two rules are formulated for the identification of nuclear constituents of prosodic relations. The utterances analysed in section 3 illustrate how prosodic and prominence relations can be identified by analysing acoustic cues of their F0 contours. Utterances correspond to English borad focus and narrow focus statements. Focus positions are deduced by only using the rule of the cognitive model.

Keywords. prosodic relation, cognitive structure, information structure, intonational contour

1. Introduction

This paper investigates the role of prosody in relation to the cognitive processes underlying cortical speech object representations. We hypothesize that neuronal activation patterns during word evocation modulate the fundamental frequency (F0) and generate the F0 contour of speech, thus imparting cognitive meaning to prosodic elements. According to our interpretation, prosodic features of speech, such as prosodic words, reflect the cognitive functions associated with speech objects as represented within cortical utterance frameworks.

In the cognitive view, information structure (IS) of utterances is the structure of speech objects within their cortical representations and this explains why we describe IS in terms of cognitive categories and not in terms of semantic categories; i.e focus-background in Halliday (1967) or Steedman (2000). All linguistic models consider prosody as the pragmatic support of the semantic IS functions of utterance constituents but semantic categories cannot generate in all cases different IS descriptions to different intonational contours. Describing prosody by using cognitive categories, the resulted structures exactly reflect Information Structure applied by the cortical representations to speech objects and then to utterance constituents of speech outputs.

Ladd (2008) attempts to improve the phonological description of intonational contours by adding 'metrical' features to other phonological categories and introduces the notion of prosodic relations. He argues for two structural variants of prosodic relations, high-low and low-high, and to a 'metric' of prosodic relations in addition to the 'metric' of prominence relations, the latter ones having weak-strong structure: "Just as two phonological constituents or domains can be related as either weak-strong or strong-weak, so I proposed that at least certain kinds of prosodic constituents can also be related as either low-high (an accent and a following non-downstepped one) or high-low (an accent and a following downstepped one)" Ladd (2008: 304-306).

In Ladd (2008), a 'metrical' feature named 'downstep' feature is proposed to describe the structure of prosodic relations. Ladd illustrates the presence of prosodic relations by presenting the F0 contour of the sentence (1.b) where the 'downstep' feature is used to solve the problem with the relationship between the target tones of the syntactic constituents *Anna* and *Manny* in the context of the focus event on *Manny*.

(1.a) Q: What about Anna? Who did she come with?

(1.b) A: [[Anna /came][with/ **Manny**]].

About the contour of the answer (1.b) where *Manny* bears new information and the focus function, Ladd says that "the ratio between the heights of the two peaks of the contour is not consistent with the pragmatic structure of the contour" because the two peaks of *Anna* and *Manny* are approximately equal Ladd (2008: 68-69). He presents the (Liberman & Pierrehumbert 1987)'s opinion related to this problem: "the second peak is a bit higher than the first but is lowered by final lowering and they have almost the same height". In other words, they consider an emphasis exists on *Mann*y but its high target level is decreased by final lowering.

In Ladd's view, the target tones of *Anna* and *Manny* being approximately equal is not due to the final lowering effect, but rather the consequence of a prosodic relation involving the "downstep" feature of their prosodic relation. Ladd argues that a primary accent and focus function are attributed to the low target tone constituent of a prosodic relation with the "downstep" feature. He asserts that *Manny* is the constituent on the 'downstep' tendency, and its focus function arises from this feature rather than from emphasis generated by its target tone.

In the cognitive view, prosodic relations at any utterance tree level convey how speech objects are merged into cognitive relations at the cortical level. In other words, the paper gives a cognitive meaning to prosodic relations and presents how to decode their structures by interpreting the phonological/phonetic cues of related prosodic words. In this view, all constituents of utterances are merged into prosodic relations, not only those which correspond to syntagmatic syntactic constituents, as in the *Anna / Manny* syntactic context presented in Ladd (2008).

In the cognitive perspective, prosodic relations across all levels of utterance trees elucidate how speech objects are integrated into cognitive relations at the cortical level. Put differently, the paper assigns a cognitive significance to prosodic relations and outlines how their structures can be deciphered by interpreting the phonological and phonetic cues of associated prosodic words. According to this viewpoint, all components of utterances are merged into prosodic relations, not solely those corresponding to syntagmatic syntactic constituents, as illustrated in the context of *Anna* and *Manny* presented in (Ladd 2008).

In other words, the cognitive perspective integrates prominence relations with prosodic relations within utterances, attributing cognitive significance to both. In phonology, prominence relations are described by weak-strong structure after identifying phrases within the utterance tree. According to Ladd (2008: 261-262), the sentence "*A friend of mine works for NASA*" is described as shown in (2). This sentence represents a broad focus statement with the focus on the last word. Therefore, the word "NASA" is annotated with the label S-Strong.

(2) [A friend of mine^{W} /[works ^{W}/ for NASA^{S}]].

The issue with (2) lies in the partitioning of the utterance, which appears to be driven more by syntactic considerations rather than prosodic ones as inferred from its F0 contour. In the cognitive framework, prominence relations and prosodic relations operate within the same hierarchy, which reflects the hierarchy of cognitive relationships formed during utterence production at the cortical level. Here, constituents are merged into binary units we named cognitive relations.

Regarding the associations between semantic information structure (IS) functions and intonational forms of prosodic constituents, Ladd (2008: 277) proposes the metrical structure as "a better way to discuss the universal feature of the expression of focus," avoiding the problematic association of focus with specific types of phonological or phonetic features. He illustrates this proposal using the sentence "*Dogs must be carried*" in two different pragmatic contexts. According to Gussenhoven (1983), the two neutral intonational variants of the sentence correspond to the 'contingency' reading - (3.a) and the 'eventive' reading - (3.b). Ladd distinguishes between the stress patterns of the two variants by assigning the former to a description where there is a single phrase with focus on the verb "carried," and the latter to another description with two prosodic phrases. In the latter case, the prosodic phrases are involved in a Strong-Weak prominence relation because the global focus is marked on the word "dogs," and the second phrase contains the remainder of the sentence as the weak part.

(3.a) [Dogs must be **carried**$_{F}$].

(3.b) [**Dogs**$_{\mathbf{F}}$] [must be carried].

The sentence-initial position of focus in the sentence (3.b) is not in agreement with its type because it is not a narrow focus statement and in the broad focus statement the nucleus cannot be in the sentence-initial position. Thus, we propose to reanalyse the utterance of the sentence (3.b) on a cognitive basis in order to identify the hierarchy of its prosodic and prominence relations and after that, the local and global focus positions result.

In section 2, the cognitive model of Information Structure proposes a set of functional categories to describe the structure of cognitive relations. It defines the roles of nuclear elements within cognitive units and formulates two rules for identifying the nucleus within prosodic relations. The remainder of the paper analyses the intonational contours of few English sentences to illustrate how to interpret the prosodic and prominence relations of utterances. The partitioning and stress patterns of prominence are described using the categories from the cognitive model of Information Structure.

2. The Cognitive Model of Information Structure

Section 2.1 presents the functional categories that describe the two structural levels of prosodic relations, generated by mapping cognitive relations from the cortical level into the tonal space of F0 frequency. In section 2.2, the aspect of prominence within prosodic relations is discussed, and the Nucleus Identification Rules (NIR) are formulated. From this perspective, prosodic relations carry the same structural meaning as their corresponding cognitive units (CUs), and utterance trees are essentially hierarchies of CUs.

2.1 The Two Structural Levels

In the theory presented, the predicate-argument structure forms the foundational structural level of cognitive units (CU) or cognitive relations that underpin cortical representations of speech objects. According to this framework, the brain's ability to track and enumerate objects involves a conceptualized propositional structure of predicate-argument type. Here, the pre-linguistic predicate (PP) is the element that attributes a feature (content) to another object, which is conceptualized as the pre-linguistic argument (PA).

Based on Quilty-Dunn's (2020), we know that Perceptual Object Representations (PORs), in our discussion we refer to speech object, consist of distinct components for individuals (corresponding to PA constituents) and properties (corresponding to PP constituents). Quilty-Dunn argues that without the ability to bind objects into argument-predicate relations, there is a risk of losing featural information while retaining index-like elements. He supports this argument by referencing experiments involving six-month-old infants, where it was observed their “working memory supports an object representation that is featureless”.

Hurford (2003) is cited for his insight into predicate-argument structures within mental representations of objects, which he views as fundamental to both phylogenetic and ontogenetic stages of mental development before language. He posits that these structures serve as the core of primitive (pre-linguistic) mental representations. Additionally, Hurford suggests that the structures found in modern natural languages can be mapped onto these primitive representations. In this paper, we adopt this perspective by decomposing utterances with various syntactic structures into hierarchies of binary prosodic relations aligned with predicate-argument structures.

Gabelentz's model, as summarized in von Heusinger (2002), marks a significant period in the history of Information Structure modeling, particularly within psychological frameworks, aligning closely with our cognitive approach to Information Structure (IS). Gabelentz presents the argument-predicate structure in terms of a psychological subject (PS), which relates to "that about which the hearer should think" (an indexical reference), and a psychological predicate (PP), which relates to "what he should think about it" (a feature or attribute). This formulation leads to a PS-PP structure, reflecting the foundational elements of cognitive processes involved in structuring information within mental representations.

Quilty-Dunn (2020) defines the structure of units within perceptual object representations at the cognitive level, using the framework of predicate and argument relations. In language modelling, perceptual objects are speech objects and their cognitive relations attribute features (predicates) to arguments.

In a visual perspective, Zacks (2020) treat perceptual object as events unfolding over time and suggests that these event representations exhibit common structural elements across perception and memory. At the first structural level, he describes cognitive relations between visual objects by part-subpart structures, which articulate the relationship between a visual object (the "part") and one of its components (the "subpart"). This structure allows the “subpart” to function as a feature or attribute (analogous to a predicate) of the larger visual object, the “part”, which serves as the argument in this visual POR context.

In line with Zacks’s view presented in Zacks (2020), the part-subpart structure is determined by the spatial delimitation of the two related events in the representational space of their features. The second structural level proposed by Zack is determined by the existence of the temporal delimitation between events. One constituent of cognitive relations is the “cause” event and the other event of the respective relation, is the “effect” event. In the case of the language, the “cause” and “effect” events of cognitive units may be viewed as “theme” and “rheme” speech objects because their relation involves a causality relation. “Theme” and “rheme” are viewed as cognitive categories in this paper but we can deduce their involvement in supporting the realizations of the theme and rheme semantic categories. These concepts are important in understanding how information is organized and communicated both cognitively and linguistically.

The cognitive model introduces the CU_argument and CU_predicate categories to describe the argument-predicate structure of prosodic (cognitive) relations and the CU_rheme and CU_theme categories to describe the “theme”-“rheme” structure. At the prosodic level, the CU_argument corresponds to *higher target tone constituent* and CU_predicate, to the *lower target tone constituent*. In other words, the predicate-argument and argument-predicate structures give a cognitive meaning to Ladd’s low-high and high-low structures.

The tonal difference between the constituents at the CU_predicate-CU_argument structure level can be translated at the neural level by the difference between the activation levels of the neurons related to the two auditory items merged into a relation by a delta wave rhythm. Delta wave rhythm has the capability to modulate the firing level of the neurons that process auditory items by relating them to its different phases -see Oblesser et al. (2019) about delta wave phases. We formulate the hypothesis that different firing levels of the neurons involved in the perception of one utterance are translated within speech output by different tonal levels of the F0 contour segments synchronized with the respective neurons. The parallel between Linguistic/speech, cognitive and neurobiological primitives have to be the subjects of the future researches that will give the arguments required for a complete justification of the new prosodic phrasing approach. The aim of the present paper is to formulate these hypotheses and to use them and to illustrate how they improve the intonational contour understanding concerning prosodic phrases, nuclear accent, and phonological representations.

The CU_theme- CU_rheme structure is marked at the prosodic level by different temporal features/shapes of pitch movement during the corresponding prosodic words; e. g. CU_rheme is usually marked by *slow pitch variation* and the CU_theme is marked by *abrupt pitch movements or constant level pitch movements*. This characterization of CU_theme and CU_rheme marks are in agreement with the two types of intonational forms assigned in Steedman (2000) to the semantic theme and rheme constituents: with H* pitch accent for elements in a former category, and with L+H* pitch accent for those in the latter category.

We conclude that the cognitive model defines two structural levels of prosodic relations. Structure overlapping is possible because pitch events have two independent features: the level of the target tone that encode the first structural level and the time pitch variation within the pitch movements of prosodic words that encode the second structural level.

In the cognitive model presented in this paper, utterance structures are described using CU hierarchies with a set of functional labels. The labels P and A are used to annotate CU_Predicate and CU_Argument, while R and T are used for CU_rheme and CU_theme annotation. In this proposed annotation system, each element of the partition is annotated with two labels because it serves functions at two structural levels. These labels are connected by "+" and enclosed in round parentheses. CUs (Cognitive Units) are described using sequences of two round parentheses separated by a slash, corresponding to the two constituents of the CU. In (4.a)-(4.d), all four possible sequence variants are presented.

(4.a) (A+R)/(P+T)

(4.b) (A+T)/(P+R)

(4.c) (P+R)/(A+T)

(4.d) (P+T)/(A+R)

At the higher level, any CU serves as a functional constituent that carries the two cognitive functions of its core element. Describing the entire utterance tree requires identifying all local nuclear elements, which underscores our interest in establishing rules for identifying the nuclear elements of prosodic relations.

2.2 Nucleus Identification Rules

In the cognitive model view, at each level of the utterance tree, a prosodic relation with a nuclear element can be identified. The nuclear element of the highest-level relation carries the sentence accent, and all local nuclear elements project their cognitive functions to the entire unit they belong to. Therefore, each relation is represented at the next level by its nuclear constituent.

The hierarchical perspective on the articulation of prosodic relations is based on the idea that there is competition at the cortical level among speech objects, supported by neurons that evoke the corresponding speech elements during the information packaging process. Under the influence of delta waves, speech objects are merged, and the non-nuclear object is excluded from the evocation space in the high-gamma domain. Neurons associated with the nuclear element remain active in this domain, continuing to compete for higher-level nuclear functions.

This is in line with Nelson's observation from 2017: "each merge is reflected by a sudden decrease in high gamma activity in language areas." This means that the phonetic characteristics of non-nuclear constituents are no longer encoded by high gamma waves, which act on the synaptic inputs of neurons involved in higher-level competitions for nuclear functions.

The research on F0 patterns of pitch accents that mark constituents according to cognitive functions leads us to conclude that nuclear elements are not always acoustically prominent (emphasized). Emphasis can occur on constituents with CU_argument function that reach the top level of the unit during the tonic syllable. After reaching the top level of the emphasized element, a falling pitch variation follows on the same accented syllable or on the next syllable of the same word or the next word. In essence, after reaching the high target tone, there must be a subsequent falling pitch movement down to lower levels. If this condition is not met, the high tone does not emphasize the CU_argument constituent.

Therefore, we formulate two rules for nucleus identification (NIRs) related to relations with emphasized and non-emphasized CU_argument constituent. The first rule, NIR_E (Emphasized nucleus), is formulated in (5), and the second one, NIR_NE (Non-Emphasized nucleus), is presented in (6).

(5) NIR_E (Emphasized nucleus): If the CU_argument of a relation is marked for emphasis within its prosodic word, then it assumes the nuclear function in that relation. Additionally, if the CU_argument of the current relation is involved in a lower-level relation where it is marked for emphasis (as a local nucleus) due to a falling pitch movement during its locally paired constituent, and this local group is followed by a CU_predicate in the current relation, then the group must prosodically subordinate the CU_predicate. Specifically, the CU_predicate must have tones below the lowest tone of the group, ensuring that the group also assumes the nuclear function at the current relation level.

(6) NIR_NE (Non-Emphasized nucleus): If the CU_argument of a prosodic relation is not marked for emphasis, then the CU_predicate assumes the nuclear function in that relation. However, if the CU_argument of the current relation is involved in a lower-level relation where it is marked for emphasis due to a falling pitch movement during its locally paired constituent, and this local group is followed by a prosodically non-subordinated CU_predicate in the current relation (where the CU_predicate has tones above the lowest tone of the group), then the CU_predicate assumes the nuclear function at the current relation level.

The first rule differentiates between the cases of nuclear CU_argument marked for emphasis within its prosodic word and that of its marking for emphasis within lower-level relation. If the local group of the CU_argument does not acoustically subordinate the following current-level CU_predicate, then CU_argument does not bear emphasis at the current level and the CU_predicate wins the nuclear function at this level (NIR_NE).

We hypothesize that neural mechanisms exist to aid the CU_predicate element in winning the competition in the NIR_NE case. One such mechanism could be inferred from Meyer et al. (2012), whose neural investigations during sentence processing with argument-verb syntactic structures revealed sustained oscillations at 10 Hz (alpha band) "during the storage phase of the argument (about 2 seconds after argument presentation), with a peak just before memory retrieval of the main verb in sentence-final position."

Based on this, we propose the hypothesis that the argument-verb syntactic structure is implemented at the cognitive level by two entities linked through the CU_argument-CU_predicate structure. Alpha oscillations exert inhibitory effects during the CU_argument element, which may facilitate the CU_predicate element in assuming the nuclear function.

Now we exemplify how NIR_E and NIE_NE rules (for emphasized and non-emphasized partitions) can be useful for differentiate between two types of wh-question contours of the same language. Two contours extracted from the (Frotta & Prieto 2015)'s database, are analyzed. They are related to two types of Friulian wh-question interrogations: information-seeking (I-S) and echo types. Roseano et al. (2015) describe them by using the same sentence stress pattern with the nuclear configuration on the verb in the sentence-final position. They identify only one phonological difference between the two contours consisting in the pitch accent type of the nuclear accents. We are interested to find cognitive differences in order to relate them to the linguistic differences (I-S vs. echo wh-questions).

The first contour is that of (Negrons) Friulian I-S WHQ, represented in Figure 1 and described in (7). It is presented in Roseano et al. (2015:126) as a contour with the nuclear element in the sentence-final position marked by the nuclear configuration H*+L L%. Following the wh-word "quant," the contour exhibits a tonal step up to the top level, where the target tone is reached on the stressed syllable of the verb.

In Figure 1 the wh-word is annotated by a L* pitch accent and the verb, by H* pitch accent with very high target tone which does not generate emphasis because the falling part of the peak has no pitch variations. It is only a step down to the low boundary tone. NIR_NE rule says that if the CU_argument is not emphasized, the CU_predicate bears the nuclear function. This leads to the nuclear function of the wh-word as it is described in (7). NIR_NE rule generalizes Ladd's rule that applies the primary accent on the lower-target tone constituent, only in the case of prosodic relations with the 'downtep' feature. But the nuclear function must be related to the CU_predicate function in all non-emphasized relations.

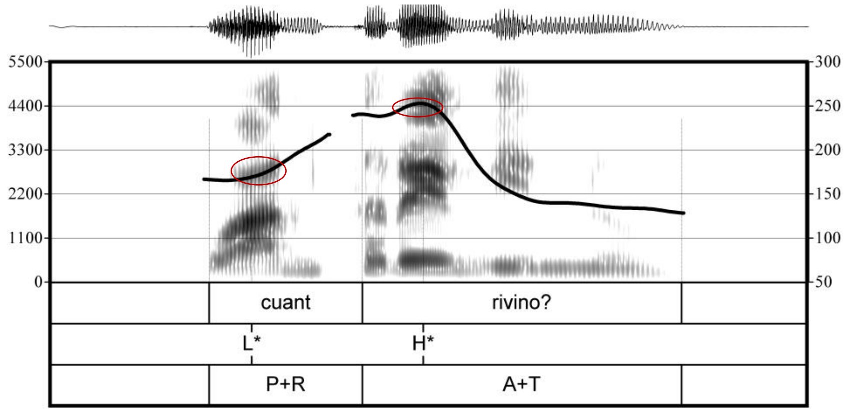

Figure 1. The F0 contour and the spectrogram of the (Negrons) Northern Friulian I-S WHQ Cuant RIvino? 'When will they arrive?'

(7) Cuant $_{N}$ $^{P+R}$/ RIvino^{A+T}?

Now we can compare the contour of (Negrons) Friulian I-S WHQ with the (Negrons) Friulian echo WHQ contour, the latter one being represented in Figure 2 and described in (8). In Figure 2 we observe that the highest target tone is synchronized with the middle of the accented syllable and the falling pitch variations on this syllable generate emphasis on the last word. Thus, the second pitch accent marks the verb for the nuclear function (NIR_E).

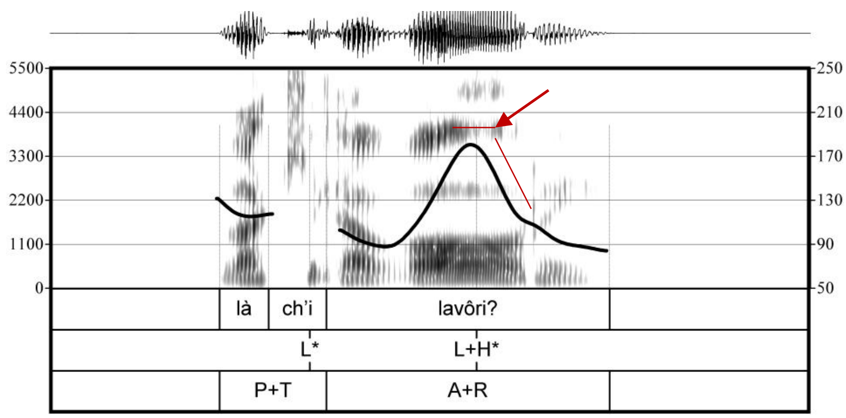

Figure 2. The F0 contour and the spectrogram of the (Negrons) Northern Friulian echo WHQ *Là ch'i laVOri?* 'Where do I work?'

(8) Là ch'i $^{P+T}$/ laVOri $_{N}$$^{A+R}$?

We conclude that linguistic differences between the information-seeking and echo WHQ contours are reflected by the cognitive differences involving the nuclear position and the second structural level that changes from the CU_rheme-CU_theme structure in the former contour, into the CU_theme-CU_rheme structure in the latter case, leading to the CU_rheme function of the nuclear element in both cases.

3. Utterances Viewed as Prosodic Relations Hierarchies

In section 3, several F0 contours with emphasized and non-emphasized partitions are presented and NIR_E and NIR_NE rules are applied to deduce the nuclear events of their prosodic relations. F0 contours are related to broad focus and narrow focus statements.

3.1 Broad Focus Statement with Neutral Intonation

One utterance of the sentence *A friend of mine works for NASA* is analyzed from a cognitive perspective by interpreting its F0 contour, represented in Figure 3. The contour is described by Ladd (2008: 261-262) with structure (2), but from a cognitive perspective, description (9) results. The sentence is a broad focus statement where the first constituent has a target tone near the top level, though the top level is not reached during the vowel of the tonic syllable. The top level is reached on the final consonant part /nd/ of the word *friend.* This is followed by a falling pitch movement during the next constituent "of mine works," creating a local emphasis on the noun "friend" within the embedded phrase *A friend of mine works*. The peak of the F0 contour indicates the lower-level relation between the constituents *A friend* and *of mine works*, where the word *friend* bears local emphasis and nuclear function, as described in (9). At the global level, the CU_predicate for NASA is not acoustically subordinated by the group *A friend and of mine works* because its beginning tone is above the last tone of the group. Thus, the CU_predicate *for NASA* bears the global nuclear function (NIR_NE) and it supports the semantic focus function of the broad focus statement.

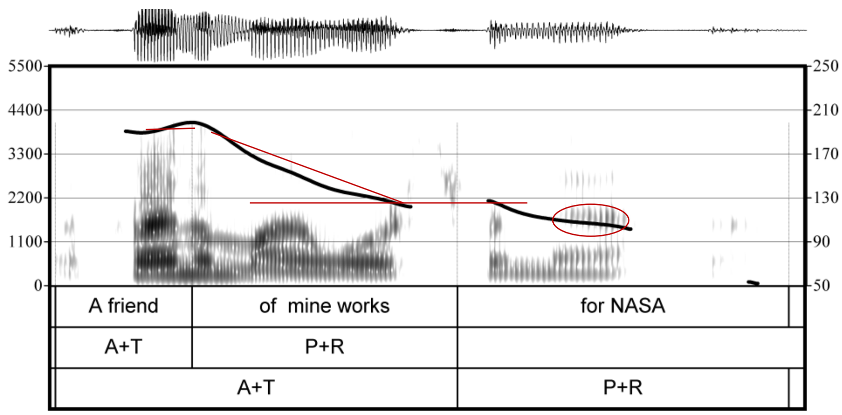

Figure 3. F0 contour of the statement *A friend of mine works for NASA*$_F$

(9) [A friend $_n$ of $^{A+T}$/ mine works $^{P+R}$]$_{A+T}$ / for NASA$_N$$^{P+R}$

The cognitive perspective leads to a different utterance tree structure than that described in (2). The local nuclear element corresponds to the topic element *a friend* and the global nuclear element is related to the global rhematic element *for NASA*.

3.2 Broad Focus Statement with Acoustical Prominent Topic

The sentence *Dogs must be carried* is a broad focus statement with acoustical prominent topic. In Ladd's interpretation, a global focus function is applied to the topic as it is described in (3.b). The F0 contour is represented in Figure 4 and its cognitive description is presented in (10) by using a two-levels hierarchical structure.

At the lower level, the subject dogs is the CU_argument and the partial verbal group *must be* is the CU_predicate. In Figure 4 we observe the peak pattern with the top-level tone on the accented syllable of *dogs* and the falling pitch movement on verbal group *must be*. This F0 patterns marks the relation between the two constituents and generates local emphasis and nuclear function on the subject dogs (NIR_E). The emphasis is more prominent than in the contour represented in Figure 3 because the word dogs holds its tones near the top level for a longer time. At the second structural level, the CU_argument dogs is the CU_theme element of the relation (rising pitch movement with narrow pitch range) and the CU_predicate *must be*, the CU_rheme (slow falling pitch movement).

At the global level, the global CU_ predicate carried is also prominent, because it is not subordinated by the preceding group *Dogs must be*. Thus, NIR_NE rule gives the global nuclear function to the CU_predicate *carried* annotated by N in (10).

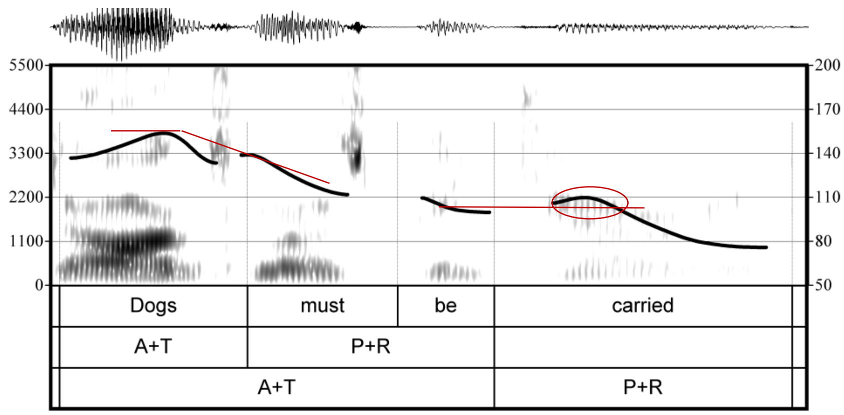

Figure 4. F0 contour and spectrogram of the statement *Dogs*$_{CT}$ *must be carried*$_{F}$

(10) [Dogs $_{n}^{\ A+T}$ / must be $^{P+R}$]$_{A+T}$ /carried $_{N}^{\ P+R}$

If we compare the emphasis on subject in Figure 3 with that on subject in Figure 4 we conclude that local emphasis in the latter case is acoustically more prominent but they are both broad focus statements having the same prosodic (cognitive) structure that applies the focus function on the last word with global CU_predicate and CU_rheme functions at the cognitive level.

3.3 Narrow Focus Statements

The sentence *Anna came with Manny* is a narrow focus statement where both referents *Anna* and *Manny* bear local emphasis but the latter one bears the global focus or sentence accent. The contour corresponding to the utterance with the focus on *Manny* is illustrated in Figure 5 and described in (11).

Ladd (2008: 153) describes the contour by introducing the prosodic relation between *Anna* and *Manny* but we have to also observe the lower-level relations between *Anna* and the verb *came*, and between the preposition *with* and *Manny*. In the lower-level prosodic relations, *Anna* and *Manny* are CU_arguments with emphasis and bear the local nuclear functions (NIR_E).

At the second structural level, the former lower-level relation has the CU_theme-CU_rheme structure and the latter one, the CU_theme-CU_rheme structure. Thus, the global relation has *Anna* as the global CU_theme constituent and *Manny* as the global CU_rheme element as it also described in (11).

We observe in the contour that the subject *Anna* and the complement *Manny* have both very high target tones but the last target is a little higher than the first target tone because *Anna* reaches the maximum peak's height during the consonant /*n/* of the last non-accented syllable and not on the first vowel /a/. This explains why *Manny* is the global CU_argument and *Anna* is the global CU_predicate.

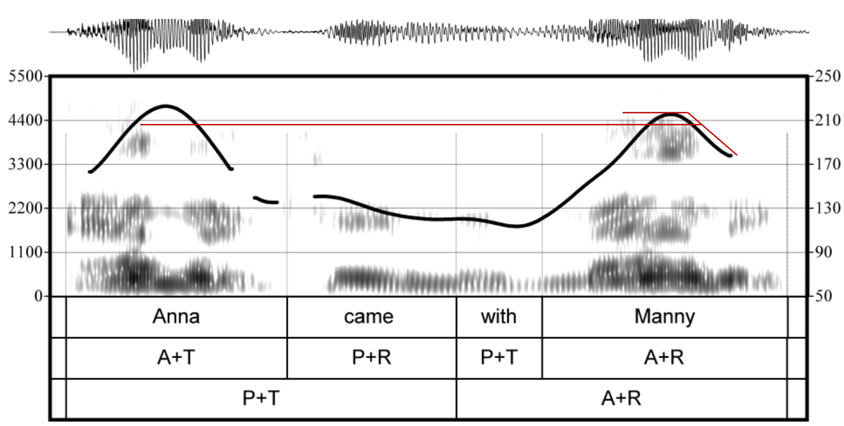

Figure 5. The F0 contours of the utterances of the sentence *Anna came with Manny* in the Background/Answer semantic context

(11) [Anna $_{n}^{\ A+T}$/ came $^{P+R}$]$_{P+T}$ / [with $^{P+T}$/Manny $_{N}^{\ A+R}$]$_{A+R}$

We conclude that *Manny* bears the global nuclear function bearing the CU_ argument function and emphasis marked by the higher target tone followed by a falling pitch movement (NIR_E). *Manny* and *Anna* are annotated in (11) by N and n labels, respectively. The global emphasis is prominent because the top level is reached in the middle of the vowel of the accented syllable and it is followed by falling pitch variation.

The global prosodic relation has the CU_predicate-CU_argument structure because no matter which is the high level of the target tone of the word *Anna*, the target tone of the word *Manny* is a little higher but enough to mark it for global emphasis and global nuclear function. This tonal relationship was observed by the linguists who previously have analyzed utterances of the sentence with different degree of emphasis. This corresponds in the phonological Ladd's view, to a Low-High structure of the global relation and the 'downstep' feature cannot characterize the relation between *Anna* and *Manny* as it is suggested in (Ladd 2008:153).

Another contour where the target tone of the focus constituent is a little higher than the other element of the global prosodic relation corresponds to the sentence (12) and its F0 contour is represented in Figure 6.

(12) I did it for **him**$_F$.

The contour in Figure 6 has a lower-level relation between the preposition *for* and the pronoun *him*. The contour is described in (13) by a two-level hierarchy.

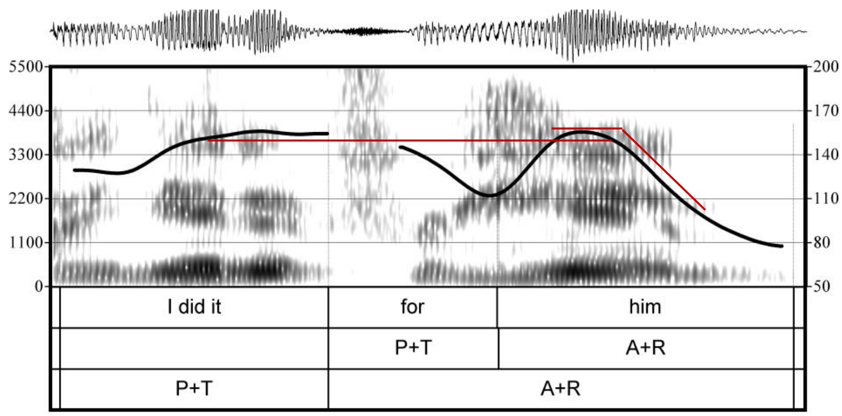

Figure 6. F0 contour and spectrogram of the statement *I did it for* ***him***$_F$

(13) I did it $^{P+T}$/ [for $^{P+T}$ /him $_N^{A+R}$]$_{A+R}$

In the lower-level prosodic relation *for* is the local CU_predicate and CU_theme element (abrupt falling pitch movement), and *him* is the CU_argument and CU_rheme element (slow falling movement). The CU_argument *him* is marked for emphasis because it reaches the top level in the middle of the syllable and the last part of it is a falling pitch movement. Thus, it bears the local nuclear function (NIR_E).

At the global level, the embedded relation is the CU_argument and CU_rheme element and the verbal phrase *I did it* is the global CU_predicate and CU_theme constituent because the first target tone is a little lower than that of the pronoun *him*. The pronoun *him* bears emphasis and the global nuclear function (NIR_E).

This intonational variant of the sentence *I did it for him* is a non-neutral one, where emphasis is placed on the pronoun *him*. This utterance corresponds to a narrow focus statement. Another intonational variant of the same sentence places focus on the entire group *for him*, and the pronoun *him* is not acoustically prominent as in the previously mentioned variant. This second variant better corresponds to the old information semantic meaning of the pronoun *him*.

In Figure 7 the contour of the second intonational variant of the sentence *I did it for him* also shows a lower-level relation between the constituents of the prepositional group *for him* where the sensitive-focus particle is the CU_argument having the target tone followed by a falling pitch movement during the pronoun. Thus, the former one bears the local nuclear function (NIR_E). It is annotated by N in (14) because the prepositional group *for him* is the global CU_argument with emphasis and bears the global nuclear function.

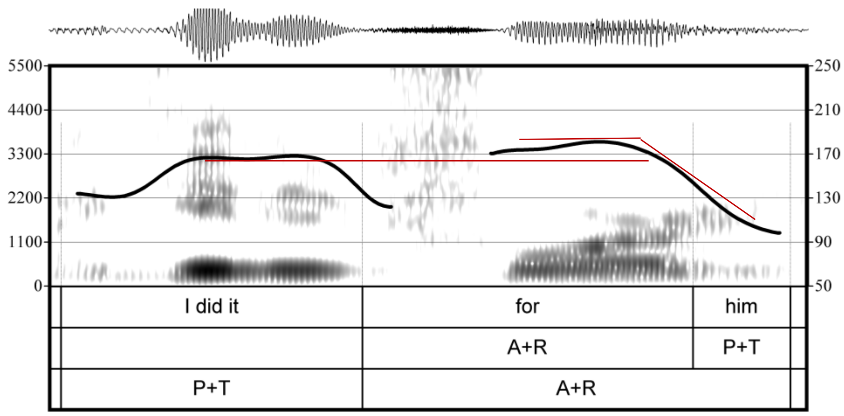

Figure 7. F0 contour and spectrogram of the statement *I did it* (***for***$_E$ *him*)$_F$

(14) I did it $^{P+T}$/ [for $_N^{A+R}$ /him $^{P+T}$]$_{A+R}$.

The verbal phrase *I did it* is the CU_predicate having lower target tone. Thus, the group *for him* bears the global nucleus and semantic focus function. The CU_predicate does not follow the local group with emphasis *for him* and it cannot take the nuclear function.

The second intonational variant places emphasis on the preposition *for* rather than on the pronoun *him*, the latter being an element of given information. The pronoun *him* contributes to the generation of the peak through its falling pitch movement and is part of the focused group.

All these analyses of the English utterances demonstrate the necessity of introducing the cognitive perspective for deducing the prosodic relations of utterances and their nuclear elements. This new approach to utterance partitioning leads to comprehensive prosodic relation hierarchies that accurately describe the cognitive meaning of utterances.

4. Conclusions

The cognitive model considers prosodic relations as prosodic correlates of cortical representations of utterances, leading to speech constituents being merged into nested binary cognitive units. The model defines a set of categories for describing the two structural levels of cognitive units and their related prosodic relations: CU_predicate, CU_argument, CU_rheme, and CU_theme.

In the cognitive view, the metrical structure of prominence relations and the 'metrical' structure of prosodic relations share the same hierarchy, which is the hierarchy of cognitive relations of speech objects. The 'strong' constituent of prominence relations is the nuclear element of the respective prosodic relations.

The paper illustrates the application of nucleus identification rules (NIR) to generate accurate cognitive descriptions of intonational contours. For example, in broad focus statements, the cognitive analysis confirms that the nuclear constituent is typically positioned at the rightmost part of the sentence. The utterance trees of the two broad focus statements are derived from their related F0 contours, and the NIR_NE rule is applied to identify the nuclei of prosodic relations. The nuclear function of the lowest-target tone word of utterances is possible at the cortical level due to the interplay of excitation and inhibition in neuronal competition.

In narrow focus statements, identifying the global nucleus is straightforward when it is marked by the highest target tone, has significant duration, and is followed by a falling pitch movement. However, the cognitive perspective provides additional insight into cases where the constituents of global relations have very high target tones and only small differences between their target tones. In such sentences, the target tone of the narrow focus word being just slightly higher than that of the topic is sufficient for it to become the global CU_argument with emphasis. The NIR_E rule then deduces its global nuclear function.

The categories of the cognitive Information Structure model (nucleus, CU_argument, CU_predicate, CU_theme, and CU_rheme) that describe prosodic relations as cognitive relations could be taken as cognitive primitives of speech and language domains. These categories might challenge neurobiologists to find neural mechanisms that can underpin them at the cortical level. This paper aids researchers in addressing the "mapping" problem formulated in Poeppel (2012), which asks: "How to formulate the formal links between neurobiology and cognition: primitives of cognition (speech, language) and neurobiology?"

In this view, elements of cognitive descriptions based on the interpretation of F0 contours can be linked to mechanisms of representing utterances at the cortical level, where speech objects are organized into cognitive relation hierarchies. The prominence of nuclear events can be better understood when related to neuronal competition influenced by delta waves and other inhibitory rhythms that merge speech objects into cognitive relations.

Identifying the cognitive structure of speech is a complex task, requiring dedicated efforts in education and research. Future work should focus on teaching the cognitive model of information structure to young researchers in language and speech processing. This knowledge will enable them to build databases of intonational contours annotated with hierarchies of prosodic relations from a cognitive perspective. These databases can then be used to train computational modules designed to recognize prosodic relations of intonational contours, including their cognitive structures and nuclear positions.

Acknowledgments

Not applicable.

Authors contributions

Not applicable.

Funding

Not applicable.

Competing interests

Not applicable.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Sciedu Press.

The journal’s policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

References

Frotta, S., & Prieto, P. (2015). *Intonation in Romance. Frota, S. and Prieto, P.* (Eds), Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199685332.001.0001

Gussenhoven, Carlos. (1983). Focus, mode, and the nucleus. *Journal of Linguistics,* 19, 377-417. https://doi.org/10.1017/S0022226700007799

Halliday, M. A. K. (1967). *Intonation and grammar in British English*. The Hague: Mouton. https://doi.org/10.1515/9783111357447

Hurford, J. R. (2003). The neural basis of predicate-argument structure. *Behavioral and Brain Sciences, 26*(3), 261-282. https://doi.org/10.1017/S0140525X03000074

Jitcă, D., & Maruşca, S. (). *A common view on broad and contrastive focus events*, Bucharest Working Papers in Linguistics, *XXV*(2). https://doi.org/10.31178/BWPL.25.2

Ladd, D. R. (2008). *Intonational phonology.* Cambridge University Press, pp: 68, 152-153. https://doi.org/10.1017/CBO9780511808814

Liberman, M., & Pierrehumbert, J. (1984). Intonational invariance under changes in pitch range and length. In *M. Aronoff and R. Oerhle* (eds.), Language sound structure. Cambridge, MA:MIT Press, 157-233.

Meyer, L., Obleser, J., & Friederici, A. D. (2012). Left parietal alpha enhancement during working memory-intensive sentence processing. Cortex. 2013 Mar; *49*(3), 711-21. Epub 2012 Mar 20. PMID: 22513340. https://doi.org/10.1016/j.cortex.2012.03.006

Nelson, M. J., El Karoui, I., Giber, X., Yang, X. F., Cohen, L., Koopman, H., … & Dehaene, S. (2017). Neurophysiological dynamics of phrase-structure building during sentence processing. *Proceedings of the National Academy of Sciences*, May 2; *114*(18), 3669-3678. https://doi.org/10.1073/pnas.1701590114

Obleser, J., & Kayser, C. (2019). Neural entrainment and attentional selection in the listening brain. *Trends in Cognitive Sciences*, *23*(11), 913-926. https://doi.org/10.1016/j.tics.2019.08.004

Poeppel, D. (2012). The maps problem and the mapping problem: two challenges for a cognitive neuroscience of speech and language. *Cogn Neuropsychol.*, *29*(1-2), 34-55. https://doi.org/10.1080/02643294.2012.710600

Quilty-Dunn, J. (2020). Concepts and predication from perception to cognition. *Philosophical Issues*, https://doi.org/10.1111/phis.12185

Roseano, P., del Mar Vanrell, M., Prieto, P. & "Intonational phonology of Friulian". (2015). In *Frota, S. and Prieto, P.* (Eds), Intonation in Romance. Oxford: Oxford University Press, 101-139. https://doi.org/10.1093/acprof:oso/9780199685332.003.0004

Steedman, M. (2000). Information structure and the syntax-phonology interface. *Linguistic Inquiry, 34*, 649-689. https://doi.org/10.1162/002438900554505

von Heusinger, K. (2002). Information Structure and the partition of sentence. *Prague Linguistic Circle Papers: Travaux du cercle linguistique de Prague nouvelle série*, *4*.viii(376), 275-305. https://doi.org/10.1075/plcp.4.14heu

Zacks, J. M. (2020). Event Perception and Memory. Annual Review of Psychology, Jan 4; *71*, 165-191. PMID: 31905113; PMCID: PMC8679009. https://doi.org/10.1146/annurev-psych-010419-051101