A Cognitive View on Prominence Relations

Doina Jitcă

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

Received: December 8, 2024 Accepted: January 10, 2025 Online Published: January 14, 2025

doi:10.5430/elr.v14n1p8 URL: https://doi.org/10.5430/elr.v14n1p8

Abstract

The present paper proposes a cognitive view on prominence relations of utterances by treating them as prosodic relations with weak-strong metrical structure. Considering this aspect, prosodic relations have a cognitive structure and a prominence structure, both of them reflecting the structures of cognitive relations generated at cortical level during the evocation of auditory objects corresponding to speech constituents. Ladd (2008) introduces two types of relations in the F0 contour description, prosodic and prominence relations, but we suggest a single type of relation between speech constituents, named prosodic relation, and consider it is in the same time a prominence relation due to its weak-strong metrical structure. In section 2, the paper presents the cognitive model of information structure by defining the categories used for the description of cognitive and prominence structures of prosodic relations. In section 3, the cognitive and prominence structures of four utterances of the word *permit* as noun and verb, in descending and ascending F0 contours, are compared with those of four utterances of certain small statements and questions having words as constituents. All the contours mentioned are useful for understanding how utterances can be decomposed into binary hierarchies of prosodic relations with local or global prominent (strong) constituent, aiming to motivate researchers to take into account the cognitive interpretation of F0 contours.

Keywords: information structure, prosodic relation, prominence structure, cognitive structure

1. Introduction

The Autosegmental-Metrical (AM) model of intonation introduces the notion of pitch accent as an ensemble of perceptual cues that generates events within F0 contours, related to different linguistic constituents. The model distinguishes between pitch accent and stress, and relates pitch accents to both stressed and non-stressed syllables. Ladd (2008: 48-50) accepts that pitch accents correspond to those local pitch contour features including pitch change or not. However, he further claims pitch accents have to mark the prominence of one syllable against other syllables of the same prosodic word. Thus, pitch accent has to be linked to prominent syllables and Ladd suggests using linguistic information of stressed syllable positions in order to identify the positions of pitch accents within F0 contours. But prominences are also accepted on non-stressed syllables, and we must have reasons for identifying them in all pragmatic contexts; for example, when words are uttered by two speech constituents and each of them have their own pitch accent. This paper suggests a cognitive view on intonational contours in order to explain what makes a syllable to be prominent, without using linguistic information. This leads to identify pitch accents in both cases of stressed and non-stressed syllables.

There are cases of intonational contours where it is very easy to identify pitch accents and their related stressed syllables. Ladd (2008: 48:50) exemplifies these cases by using two different pitch patterns of the word *permit* as noun and verb. They are reproduced in Figure 1.a-b where we can observe that the two patterns have in common a high pitch accent "despite other differences referring the pitch rise on the stress syllables and the pitch fall from the top of the peak". In these particular cases, we can relate the stressed syllables to the acoustical prominences of the two high pitch accents. In Figure 1.a the prominence is related to the first syllable (the noun case) and in Figure 1.b, to the second syllable (the verb case).

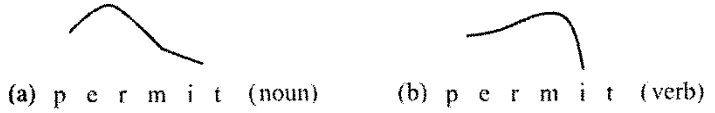

Figure 1. Descending F0 contours of the word *permit* uttered as noun (a) and verb (b)

In Fry's view, stress is a complex amalgam of F0 features, duration and intensity where the presence of pitch accent is the most important. Ladd (2008: 48:50) accepts Fry's definition of pitch accent in the cases of the two contours in Figure 1, but he aims to find a more general prosodic characterization of pitch accents, in order to be available in the case of pitch prominences not marked by high pitch accents. He exemplifies with the patterns of the word *permit* as noun and verb within utterances with ascending phrase-final contours (e.g. in questions). The contours are reproduced in Fig 2

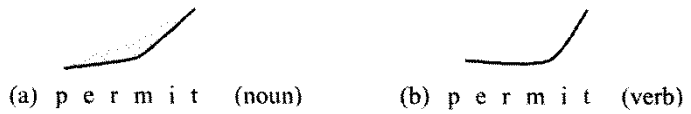

Figure 2. Ascending F0 contours of the word *permit* uttered as (a) noun and verb (b)

Ladd concludes that a complete AM approach of pitch patterns related to prominence event has to incorporate a theory of stress or linguistic prominence and introduces prominence relations with their weak-strong structure in phonology, as a solution to achieve this goal. The notion of prominence relation is preceded by Liberman's notion of linguistic prominence that involves a relation between nodes in a binary-branching tree structure associated to any utterance, Liberman (1975) and Liberman & Price (1977). According to Liberman and Price, in any such relation, one node is strong and the other one is weak. In the case of ascending contour utterances related to Figure 2, Ladd proposes to use the linguistic information in order to correctly describe the prominence relations between their syllables and (1.a)-(1.b) descriptions results in the two cases of the word *permit* as noun and verb. Thus, the pitch accents will be related on the contours to the syllable annotated by S.

(1.a) Per $_{S}$-mit $_{W}$

(1.b) Per $_{W}$-mit $_{S}$

In the present paper, prominence relations with weak-strong metrical structure are treated as prosodic relations with non-nuclear--nuclear element structure, as they are defined in Jitcă (2024). From this point of view, the weak-strong prominence structure can be considered as a structural level of prosodic relations. The cognitive perspective on information structure (IS) of utterances, where cognitive relations generated at cortical level between auditory objects are reflected at intonational level by prosodic relations, gives us a cognitive basis for the identification of prominence events within F0 contours without using linguistic information. The notion of prosodic relation was also introduced by Ladd (2008) at phonological level, generating High-Low or Low-High 'metrical' structures but he does not define a related weak-strong metrical structure. The cognitive model of IS defines prosodic relations between speech constituents with cognitive functions, one of constituents bearing the nuclear function.

We suggest, in this paper, a discussion about several prominence patterns between syllables of prosodic words, after discussing the same prominence patterns between prosodic phrases having words as constituents. In both cases we will decompose utterances into binary-branching tree structures where weak and strong constituents will be syllables, syllables group, words or words groups. We aim to identify prominence positions of prosodic relations at any utterance tree level by only using tonal information deduced from their pitch features. The prominent syllable of prosodic words usually corresponds to the stressed syllable of one word and the prominent word of word groups, to their focus word. Phonological events of pitch accents can be associated to prominent syllables, no matter they are stressed or non-stressed syllables at linguistic level.

In section 2, the cognitive model of information structure is presented where functional categories and rules for the prominence identification are defined in order to deduce weak-strong prominence structures at any level of utterance hierarchies. In section 3, the rules will be applied for deducing the stressed syllables of the word *permit* in the cases presented in Figures 1-2 and another case where the word *permit* is uttered in a longer sentence which contour is represented in Figure 3.

Figure 3. F0 contours of the word permit uttered as verb on the non-focal part of the two sentences.

The prominence patterns of the four utterances of the word *permit* as noun and verb are compared with equivalent prominence patterns of the utterances of certain small statements and questions having words as constituents. All presented contours are useful to decompose utterances into binary hierarchies of prosodic relations with local prominence event leading to a better understanding of supra-segmental features of speech signals.

2. The Cognitive Model of Information Structure

The paragraph summarizes the main aspects of the cognitive model of information structure by defining the categories that describe the cognitive and prominence structures of prosodic relations. Taking this into account, prosodic relations observed within the F0 contours reflect the binary cognitive units (CUs) generated during the evocation of auditory objects related to speech constituents of corresponding utterances.

2.1 The Two Cognitive Structure Levels of Prosodic Relations

The cognitive model introduces the predicate-argument structure as the first structural level of cognitive units (CU) or cognitive relations on which Perceptual Object Representations (POR) are based during the cortical process of speech processing (the word evocation). In the light of Quilty-Dunn (2020)'s POR theory, the brain's capacities to track and enumerate objects employ conceptualized propositional structure of predicate-argument type within which the pre-linguistic predicate (PP) is the object which adds a feature (content) to another object considered as pre-linguistic argument (PA).

Hurford (2003) also considers that primitive (pre-linguistic) mental representations contain units with predicate-argument structure. He claims that "structures of modern natural languages can be mapped onto these primitive representations" and this is exactly what we do in this paper by relating utterances with different syntactic structures to hierarchies of binary prosodic relations with predicate-argument structures.

Gabelentz' model summarized in von Heusinger (2002) discusses the information structure of sentences in terms of a psychological subject (PS) related to "that about which the hearer should think", and a psychological predicate (PP) related to "what he should think about it" (feature). Thus, we can understand that PS-PP structure of prosodic phrases has the same meaning as predicate-argument structure of cognitive units that have generated the respective speech output.

Zacks (2020) discusses the structure of perceptual object representations in more general terms, referring to visual objects. He treats them as events because they have temporal unfolding. Discussing the common structures of event representations at both perception and memory levels, Zacks (2020) presents their first structural level as part-subpart structures. That means that part-subpart structure refers to a relation between a perceptual object defining a "part" category and another one belonging to a "subpart" subcategory of its 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 cognitive categories supporting the realizations of the theme and rheme semantic events.

The cognitive model of information structure introduces the CU_argument and CU_predicate categories to describe the argument-predicate structure of prosodic relations and the CU_rheme and CU_theme categories to describe the "theme"- "rheme" structure. At prosodic level, the CU_argument is marked by the higher target tone of prosodic unit and the CU_predicate, by the lower target tone. While prosodic relations defined by Ladd (2008) in phonological terms, have only a low-high 'metrical' structure, prosodic relations with CU_predicate-CU_argument structure have cognitive meaning.

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; for example, CU_rheme is usually marked by slow pitch variation and the CU_theme is marked by abrupt 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 semantic theme and rheme constituents: forms with H* pitch accent for elements in a former category, and forms with L+H* pitch accent for those in the latter category.

In the cognitive model view presented in this paper, utterance structures can be described by CU hierarchies by using a set of functional labels. P and A labels were introduced for the CU_Predicate and CU_Argument annotation, and T and R labels for the CU_theme and CU_rheme annotation. In the proposed description system, two labels are used for annotating one element of partition because it has functions at the two structural levels. Labels are linked by "+" and

enclosed between round parentheses.

CUs are described by sequences of two round parentheses separated by slash corresponding to the two CU constituents. In (2.a)-(2.d) all four possible sequence variants are presented.

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

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

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

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

At the next higher level, any CU is a functional constituent which bears the two cognitive functions of its prominent element. The description of the whole utterance tree involves the identification of all local prominent (nuclear) elements and this explains our interest to find rules for the prominence identification within prosodic relations.

2.2 The Prominence Structure of Prosodic Relations

In the cognitive model view, at each level of utterance tree, a prosodic relation with prominent element can be identified. We name it nuclear element. The nuclear element of the top-level relation of phrases/sentence bears the phrase/sentence accent and all local nuclear elements project their cognitive functions to the whole units which they belong to. Thus, each relation is represented at the next level by its nuclear constituent.

This hierarchy of prosodic relations within utterances and the existence of prominent constituent, leads us to the idea that a competition exists at the cortical level between auditory objects and the nuclear object corresponds to the winner element. The competition is supported by the neurons that evoke those elements during the information packaging process. After auditory objects are merged, the non-nuclear object is discarded from the evocation space in high-gamma domain and the neuron related to the nuclear element continues to be active in this domain remaining in the competition for the higher-level nuclear functions. This is in agreement with (Nelson 2017)'s observation: „each merge is reflected by a sudden decrease of high gamma activity in language areas".

The research on F0 patterns of pitch accents that marks the constituents for cognitive functions leads us to conclude that nuclear elements are not in all cases acoustically prominent elements (e.g. emphasized elements). The presence of emphasis on constituents is an important mark of their nuclear position, and we have to recognize emphasis in all pitch pattern contexts. Generally, emphasis is identified on constituents that reach a local or global maximum tone at the unit/utterance level during its accented syllable. The maximum tone has to be followed by a falling pitch variation on the same accented syllable, when emphasis is more prominent, or on the next syllable of the same word, or of the next word, when it is less prominent (in neutral utterances). If this condition is not fulfilled, the maximum tone does not generate emphasis. Thus, we formulate two nucleus identification rules (NIRs) used in the cases of relations with emphasized and non-emphasized CU_argument constituents. For the first category of relations, the NIR_E (Emphasized CU_argument) is formulated in (3), and for the second one, the NIR_NE (Non- Emphasized CU_argument) is presented in (4).

(3) NIR_E: If the CU_argument of a relation is marked for emphasis and prosodically subordinates its paired CU_predicate, then it assumes the nuclear function in that relation. Additionally, if the CU_argument of the current relation was involved in a lower-level relation where it was marked for emphasis (as a local nucleus) due to a falling pitch movement during the following constituent(s), then it bears the nuclear function at the current relation only if it prosodically subordinates the CU_predicate that follows it within the current relation. Specifically, the CU_predicate must have tones below the lowest tone of the group.

(4) NIR_NE: If the CU_argument is not marked for emphasis, then the CU_predicate assumes the nuclear function in that relation. Another case is that of the CU_argument in the current relation which is involved in a lower-level relation where it is marked for emphasis due to a falling pitch movement during the following constituent(s). If 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 in the current relation.

2.3 Prosodic Phrases in the Cognitive Model Perspective

The syllables of one utterance generate a related number of speech objects during the cortical evocation. Objects are in essence neurons, to which their phonetic features are applied. Neuronal activity is also influenced by other inhibitory input signals that modulate the output level of neurons. One of these inhibitory inputs is dictated by the phase of delta oscillation to which the respective auditory object (syllable) is assigned - see Boucher et al. (2019) about the role of delta-band oscillations. Delta oscillations generate time frames within which corresponding auditory objects compete

for the nuclear function on different levels of cognitive relation hierarchy. At prosodic level, auditory objects linked to one delta oscillation correspond to speech constituents of one prosodic phrase. The nucleus of the top-level relation of the hierarchy related to one delta oscillation correspond to focus constituent of the respective phrase. The delta oscillations which process one utterance, assign all auditory objects to different delta wave phases. The inhibition of delta oscillation on neurons has different levels corresponding to different positions of the related syllables within the time frame generated by each oscillation - see Obleser et al. (2019) about phases of delta–band oscillations.

After all objects are processed by merge operations which generate cognitive relations between objects assigned to one oscillation, a single neuron remains active and it corresponds to the local nuclear auditory element of the respective delta oscillation. A higher-level oscillation will include it through the auditory objects that compete for the next higher-level nuclear function. The neuronal activity level during the evocation of one speech output is reflected by the prosodic phrases of the resulted utterance. The prosodic relations of prosodic phrases correspond to the cortical merge operations that have occurred between the auditory objects of the corresponding delta oscillations. This paper explains how the hierarchy of prosodic relations can be identified within utterances by analyzing the tonal features of their corresponding F0 contours.

3. Utterance Descriptions by Prosodic Relation Hierarchies

3.1 Noun Permit Uttered with Descending F0 Contour

The utterance of the noun *permit* represented in Figure 4 corresponds to the generic contour presented in Figure 1. The syllable *per-* has long duration at the highest tonal level. In the last part of the syllable, the falling pitch movement begins. It continues on the second syllable, which is tonally subordinated to the syllable *per-*.

In the cognitive interpretation, the syllables *per-* and *-mit* are the two constituents of the prosodic relation, where the first one is the CU_argument (high target tone) and the second one, the CU_predicate (low target tone). On the second structural level, the syllable *per-* has the CU_theme function (constant level pitch movement) and the syllable *–mit*, the CU_rheme function (slow falling pitch movement).

The F0 contour pattern in Figure 4 generates emphasis on the first syllable due to the top-level target tone followed by the falling pitch variation on *-mit*, and this leads to its nuclear function (NIR_E). We claim that the stressed syllable of words corresponds to the nuclear constituent generated by utterances of the respective words.

We now suggest a recognition in Figure 5 of the same F0 contour pattern as that in Figure 4, but presented in a case of a prosodic relation having the words *five* and *francs* as constituents, corresponding to the syntactic group *five francs* of the clause *I gave him (five francs)* within the sentence *I didn't give him three francs, I gave him (five francs)*$_{F.}$ The sentence is presented in Ladd (2008) and also in Jitcă et al. (2023).

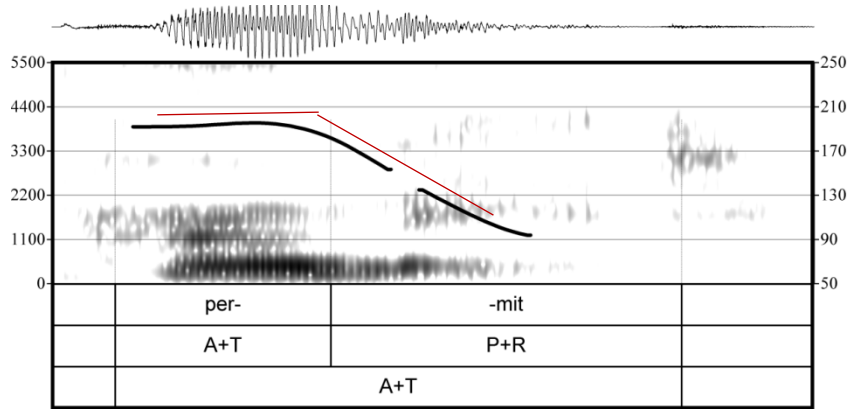

Figure 4. The F0 contour and the spectrogram of the utterance of the noun permit in descending F0 contour

The F0 contour pattern already presented in Figure 4 applies the CU_argument and CU_theme functions to the word *five* and CU_predicate and CU_rheme function to the word *francs* within the local cognitive relation. The meaning of the sentence requires a narrow focus on *five* but the utterance has a neutral intonation with the focus on the group *five francs* at the second clause level. Within the group, *five* has a top-level target tone followed by pitch falling movement on *francs*. Thus, the former element *five* bears local nuclear function, similarly to the syllable *per-* in the utterance of the noun *permit*.

The intonational phrase of the second clause is divided by two prosodic phrases as it is described in (5): one of the

verbal phrase and the other, of the noun phrase. Prosodic phrases are separated in Figure 5 by the red vertical line. At cortical level, they correspond to the two low-level oscillations that assign to their two phases, the syllable *gave* and *him*, and *five* and *francs*, respectively.

Two nested prosodic relations, ((I/gave)/him), that structure the verbal phrase *I gave him* have the verb as low-level nucleus (local emphasis) and the pronoun *him* as the nucleus of the whole prosodic phrase because the pronoun is not acoustically subordinated to the group *I gave*.

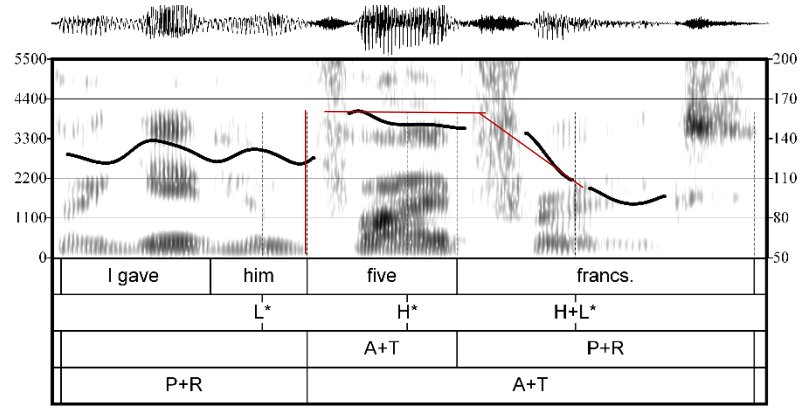

Figure 5. The contour of the clause I gave him five francs in the context of the narrow focus statement *I didn't give him three francs, I gave him (five francs)*$_F$

(5) I didn't give him three francs, (I gave him) $^{P+R}$ / [five $_N$ $^{A+T}$ / francs $^{P+R}$] $_{A+T}$

The second prosodic phrase related to the group *five francs* has to be interpreted as in Figure 4, and a local emphasis results on *five*. At intonational phrase level, the word *five* also bears emphasis because it is not followed by a non-subordinated phrase (NIR_E). It is preceded by the non-subordinated phrase *I gave him*.

3.2 Verb permit Uttered with Descending F0 Contour

The utterance of the verb *permit*, with the F0 contour represented in Figure 6, shows a peak reaching a maximum tonal level on the second syllable. This F0 contour pattern generates emphasis on the syllable *–mit* due to the top-level target tone followed by falling pitch movement on the vowel /i/. Thus, the stress position shifts from the first syllable in the noun case, to the second syllable in the verb case. The stressed syllable *–mit* corresponds to the nuclear constituent within the utterance of the verb *permit*.

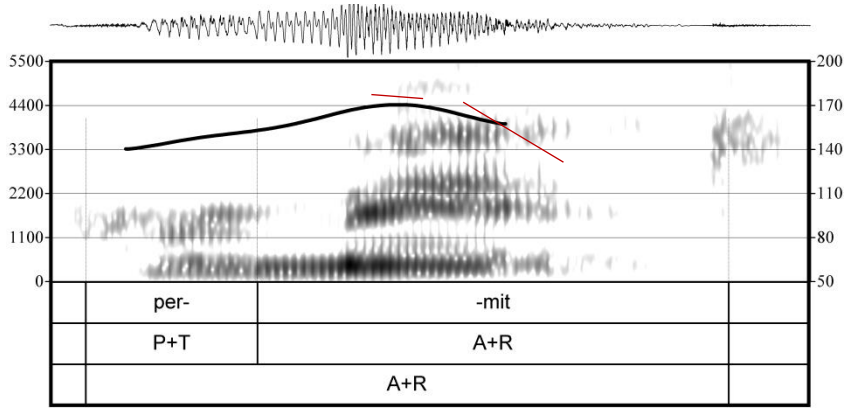

Figure 6. The F0 contour and the spectrogram of the utterance of the verb *permit*

The F0 contour pattern discussed in Figure 6 can be observed in Figure 7 between the word constituents of the elliptic sentence *My mother's diaries*: *My mother's* and *diaries.* The sentence accent pattern is also presented in Ladd (2008). In the relation between the syllable of the word *mother's*, the syllable *mo-* is the nuclear syllable because it bears emphasis due to the falling pitch variation on the last syllable *-ther's*. The syllable *mo-* is also the nuclear syllable of the noun phrase *My mother's* because the noun is the CU_argument with local emphasis. At cortical level, the three constituents of the noun phrase *My mother's* are processed by one delta oscillation.

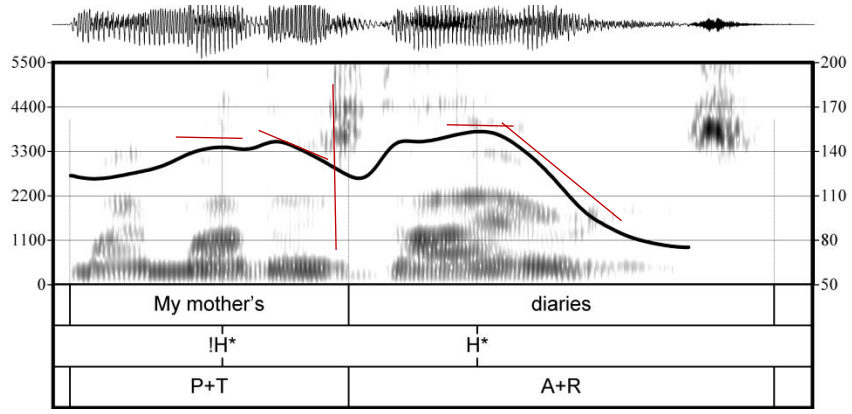

Figure 7. F0 contour and spectrogram of the statement *My mother's diaries*$_F$.

The second delta oscillation assigns the syllables of the word *diaries* to its two phases. The high part of the oscillation is synchronized with the first syllable *dia-*, and the F0 contour reaches the top-level target tone. After that, the contour falls to low levels on the last non-accented syllable generating emphasis and nuclear function on the first syllable *dia-*. At intonational phrase level, the syllable *dia-* also bears emphasis leading to the nuclear function of the word *diaries* (NIR_E) because it is not followed by a non-subordinated CU_predicate. *Diaries* is the element with CU_argument and CU_rheme functions while the constituent *my mother's* is the CU_predicate and the CU_thematic element. The pattern presented in Figures 6-7 is a prosodic solution for applying the CU_predicate-CU_argument structure to the sentence/phrase and for marking the prominence of the last constituent by emphasis within the utterances with descending F0 contours.

3.3 Noun Permit Uttered with Ascending F0 Contour

The word *permit* in the generic contours presented in Figures 2.a-b has the syllable *per-* at low tone and the syllable *-mit* reaches the highest tone by a rising pitch movement. We have to demonstrate that the contour in Figure 2.a applies the nucleus on the first syllable in agreement with the linguistic information referring the stressed syllable position of the noun *permit*.

We suggest an analysis of the prominence pattern that applies nucleus on the first syllable in Figure 2.a by firstly understanding how the interrogative contour of the (Lille) French echo WHQ (wh-question), presented in Delais-Roussarie et al. (2015), applies the nucleus on the wh-word. The contour is represented in Figure 8. It has two prosodic relations. *Où* and *je* are the constituents of the low-level relation where the wh-word *Où* is the CU_predicate and CU_theme (constant level pitch pattern) and the pronoun *je* is the CU_argument and CU_rheme, as it is described in (6). The wh-word is the nucleus because the pronoun does not generate emphasis at this local level.

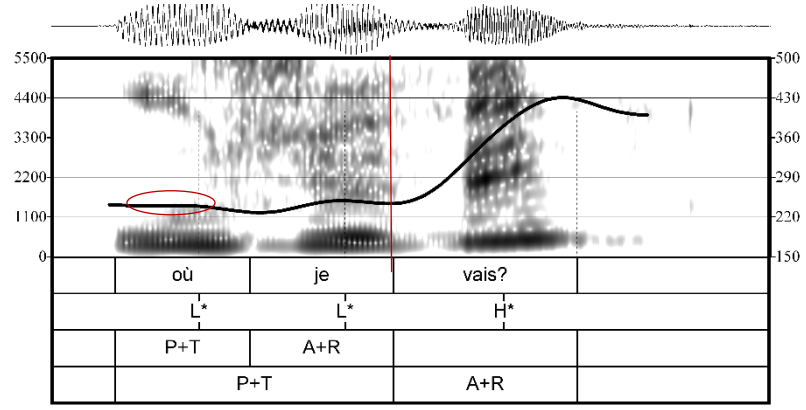

Figure 8. The F0 contour and the spectrogram of the (Lille) French echo WHQ *Où je vais?* 'Where I'm going?'

(6) (Où $_N$ $^{P+T}$ / je^{A+R}) $_{P+T}$ /vais $^{A+R}$?

The high-level relation pairs the group syllable *Où je* with the verb *vais*, where the latter one is the global CU_argument with no emphasis having only rising pitch movement. The contour pattern with low part followed by a rising accentual pitch movement which does not generate emphasis, applies the nuclear function on the wh-word with the CU_predicate

function and NIR_NE decide its nuclear function. This result is not in agreement with the claim about the nuclear pitch accent, presented in Delais-Roussarie et al. (2015). From a phonological point of view, the nuclear pitch accent of the contour in Figure 8 corresponds to the last syllable.

Now we analyse the utterance of the noun *permit* uttered with ascending contour, by using the F0 contour of (Ladd 2008)'s database, represented in Figure 9. The syllable *per-* is the CU_predicate and CU_theme and the syllable *-mit* is the CU_argument and CU_rheme. Thus, the first syllable is nuclearized and marked as stressed syllable because the rising pitch movement does not generate emphasis (has only rising pitch movement) and NIR_NE rule gives the nuclear function to the constituent with CU_predicate function.

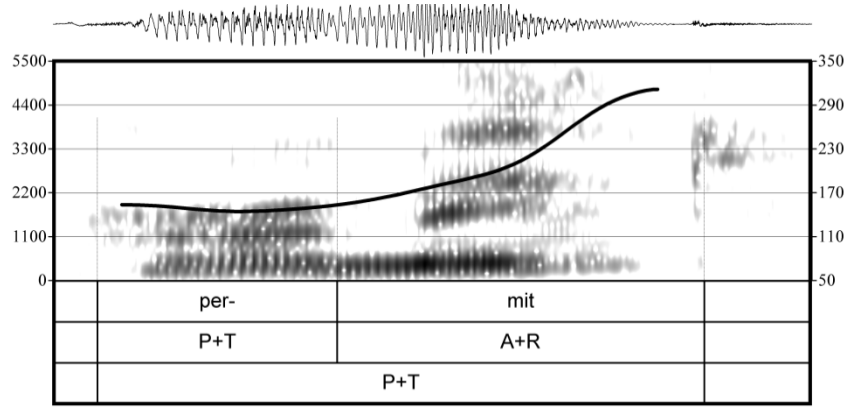

Figure 9. The F0 contour and the spectrogram of the utterance of the noun *permit* in ascending F0 contours

This result is in agreement with the linguistic information about the stressed syllable *per-* of the noun *permit*. If we accept the arguments for the nuclear function of the low tonal level constituent in the case of the contour in Figure 9, then we also have to accept them for the nuclear pitch accent position in the contour in Figure 8.

3.4 Verb Permit Uttered with Ascending F0 Contour

Peninsular Spanish I-S WHQ contour is firstly analysed in order to understand how the second syllable of the verb *permit* can be nuclearized by an ascending F0 contour. Spanish I-S WHQ contour is presented in Hualde et al. (2015) and the paper reproduced it in Figure 10, where we can observe that both target tones related to wh-word and the verb have low tonal levels, but the last one reaches the lowest level of the tonal space. After the lowest tonal level is reached, the F0 contour has a rising pitch movement and it ends in the high boundary tone.

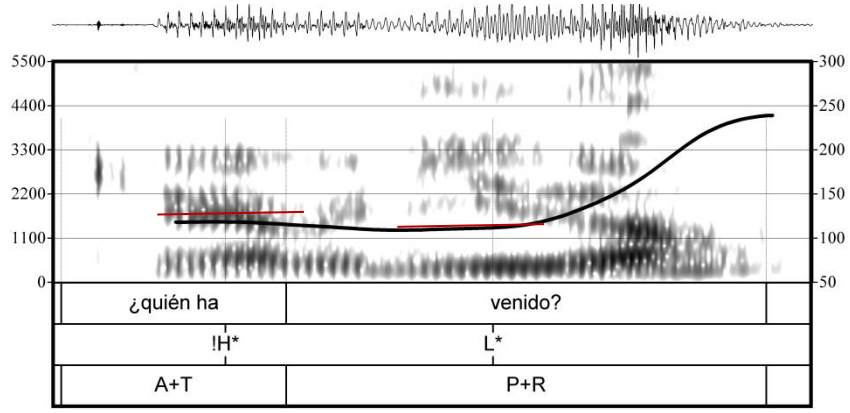

Figure 10. The F0 contour and the spectrogram of the Peninsular Spanish I-S WHQ *¿Quién ha venido?* 'Who came?'

The Spanish Information-Seeking WHQ contour has a small difference between the target tones of the two constituents, but that of the wh-group *Quién ha* is higher and this leads to the CU_argument function of the wh-word and the CU_predicate function of the verb. In the description in (7), the wh-group is annotated as the CU_theme (constant level pitch accent) and the verb as the CU_rheme (slow rising pitch movement). The CU_argument does not generate emphasis and the nucleus is on the CU_predicate *venido*. This is in agreement with the nuclear pitch accent position presented in Hualde et al. (2015).

(7) ¿Quién ha $^{A+T}$/ venido $_{N}{}^{P+R}$?

With this F0 contour pattern in mind we can understand how the contour of Ladd (2008)'s database, represented in Figure 11, accentuates the verb *permit* in interrogative utterances with ascending contour. After the first low target tone on the syllable *per-*, the contour holds low tonal levels on the consonant /m/ and on the first part of the vowel /i/.

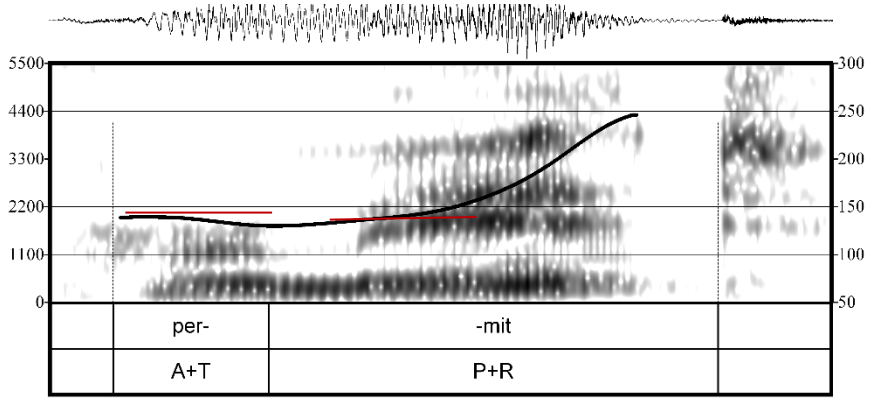

Figure 11. The F0 contour and the spectrogram of the utterance of the verb *permit* in ascending F0 contours

This contour has to be thought as having a CU_argument-CU_predicate structure by taking into account that it holds low levels on the first half of the second syllable that marks the last syllable *per-* for the CU_predicate function. Thus, we can consider it has the same cognitive structure as that represented in Figure 10. The late rising pitch movement on the second syllable gives to this syllable the CU_predicate and nuclear functions (NIR_NE).

3.5 Word Permit in Long Utterance with Descending Contour

Ladd (2008) considers the identification of the stress syllable position within the verb *permit* as an uncomfortable task when the verb *permit* occurs within the compressed part of the contour represented in Figure 12. The contour is extracted from an utterance of the sentence *I told you they permit him to retire*.

The first prosodic phrase corresponding to the main clause *I told you* applies emphasis on the verb which is synchronized with the top level of the contour and it is followed by a falling pitch movement until the end of the word *you*. The verb is the CU_argument and the CU_theme element as it is described in (8). Its nuclear function is deduced by NIR_E rule. Thus, the main clause is the thematic part of the sentence.

The falling pitch movement continues after the pronoun *you*. The question is: why does the first phrase not end after the syllable *they'd* when the minimum tonal level is reached? The local emphasized verb *told* becomes a global emphasized element only if the following part of the utterance remains under the minimum tonal level of the first phrase *I told you*. The minimum tonal level is marked by the horizontal red line in Figure 12. If we place the vertical red line after the syllable *they'd*, the compressed part of the F0 contour is not under the new horizontal line (the dotted line) and another interpretation of the utterance results. In the second variant, the verb has only a local emphasis and the global nucleus in the following prosodic phrase of the compressed part of the F0 contour.

In the second phrase, in the compressed part of the contour, the first nested relations related to the group *they /(permit / (him to))* have to be analysed. The nuclear syllable of the prosodic relations *him to* is that of the preposition *to*, because *him* has no local emphasis (the contour falls on consonant /m/ not on the last part of the vowel /i/). The higher-level relation *permit /(him to)* has the verb as A+R element due to its nuclear syllable *–mit* that has a higher tonal level than that of the preposition *to*. Within the verb *permit*, the syllable *per-* does not generate local emphasis (the duration of the vowel /e/ is too short and the pitch falls on the consonant /m/ not on the vowel /i/). Thus, the syllable *–mit* is the local CU_predicate and the nuclear syllable of the word *permit* (NIR_NE).

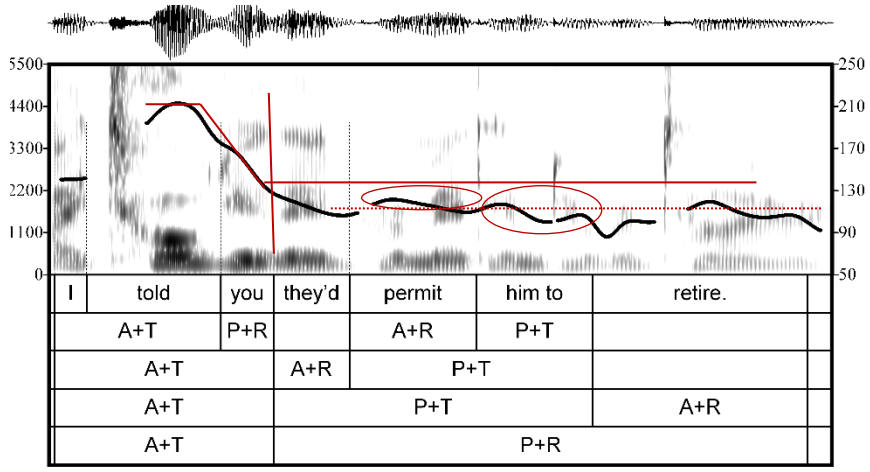

Figure 12. The F0 contour and the spectrogram of the utterance of the sentence I told you they permit him to retire

(8) [I told $_{N}$ $^{A+T}$/ you $^{P+R}$] $_{A+T}$/ [[they'd $^{A+R}$ /[permit $^{A+R}$ / him to $^{P+T}$]$_{P+T}$]$_{P+T}$ / retire $_{n}$ $^{A+R}$]$_{P+R}$

In the higher-level relation of the group *they permit him to*, the syllable *they'd* is the A+R element without emphasis and the preposition *to* is the local nuclear element of the group having the CU_predicate function (NIR_NE).

In the next higher-level group, *they permit him to retire*, the verb *retire* is the CU_argument and CU_rheme (slow falling pitch movement) and the subgroup *they permit him to* is the CU_predicate and CU_theme. Within the word *retire*, the highest target tone is reached on its accented syllable *–ti-* where a local peak with emphasis is generated. Thus, the verb *retire* is the CU_argument, CU_rheme and nuclear element in the compressed F0 contour part of the utterance. It is marked by label *n* in (8).

At intonational phrase level, the first part *I told you* is the CU_argument and CU_theme and the second part *they'd permit him to retire* is the CU_predicate and CU_rheme. The former constituent prosodically subordinates the latter one and its emphasis is a global one leading to the sentence accent function of the verb *told*.

3.6 Prosodic Phrase with Pitch Accent on Non-stressed Syllable

We exemplify how the cognitive perspective helps us to improve our pitch accent identification by analyzing the contour related to information-seeking (I-S) French yes-no question corresponding to one Swiss speaker from Geneva. Delais-Roussarie et al. (2015: 84-87) describe the question by the sequence H* L H* L%. In this paper, the contour is described at a cognitive level in (9). In Figure 13 the contour of one utterance of the French Yes-No question *Vous avez des mandarines? 'Do you have tangerines?'* is illustrated.

The first accentual rising movement is generated within the verbal phrase on the auxiliary verb *avez* (H* pitch accent) marking it as the CU_argument and CU_theme element. The syllable *-vez* is the nuclear syllable within the verbal group because it is a CU_argument which carries emphasis (NIR_E). The local emphasis is generated by the high target tone followed by the falling pitch variation on the article *des* and the following syllables *manda-* of the noun *mandarines*. Within the low-level prosodic phrase (low-level delta oscillation, at cortical level), the highest relation has the verbal group as the CU_argument with local emphasis and the group of syllables *des manda-* as the CU_predicate. The verbal phrase has the pitch accent on the nuclear syllable *-vez* and the CU_predicate has its nucleus on the syllable *-da-* where the second pitch accent is annotated on the F0 contour. On the second structural level, the verb is the CU_theme and the CU_predicate is the CU_rheme marked by slow falling pitch movement.

The low target tone level on the syllable *–da* is annotated by L in Delais-Roussarie et al. (2015: 84), but it has to be annotated as L* pitch accent marking the local nucleus of the CU_predicate *des manda-* of the first prosodic phrase where the verbal phrase *vous avez* has the CU_argument and nuclear functions.

In the second phrase (to the right of the red line) corresponding to the second delta oscillation at cortical level, the syllable *—ri-* is marked as the CU_predicate and local nuclear element and the syllable *-nes*, as the CU_argument constituent.

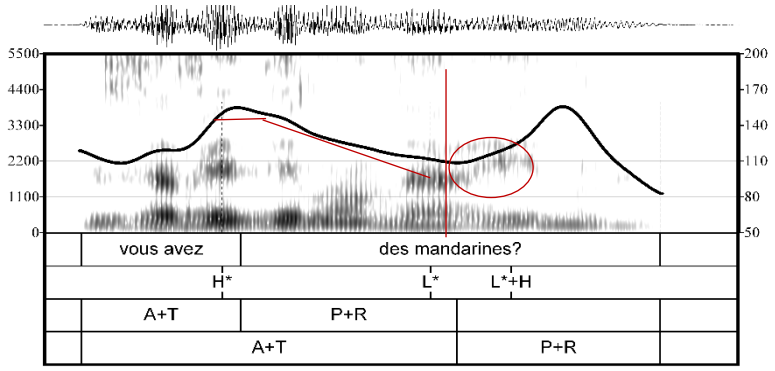

Figure 13. The contour of (Geneva) French I-S YNQ *Vous avez des mandarines?* 'Do you have tangerines?'

(9) [[Vous avez $^{A+T}$/ des manda $^{P+R}$]$_{A+T}$ /-rines $_{N}$ $^{P+R}$]$_{P+R}$

At intonational phrase level, the syllable *-ri-* is the global nucleus because the tones of the second phrase does not hold their levels under the target tone level of the L* pitch accent. Thus, the verb cannot be globally emphasized and the CU_predicate *-rines* bears the global nuclear function (NIR_NE).

The interpretation of this interrogative contour demonstrate that the cognitive perspective influences the identification of pitch of pitch accent positions and also the pitch accent type. Thus, we can understand why the last pitch accent has low target tone and it is followed by the HL% boundary tone.

4. Conclusions

In the cognitive perspective, utterances are viewed as hierarchies of prosodic relations where constituents (syllables, part of words, words, or words groups) are linked into cognitive and prominence structures. At cortical level, constituents of utterances are auditory objects supported by neuronal elements which compete for nuclear functions into a hierarchical manner.

The paper presents how we have to interpret F0 contours by using a minimal set of rules in order to deduce the prosodic relations hierarchy of utterances. Prosodic relations within prosodic phrases on the compressed parts of F0 contours can be identified in the same manner as those corresponding to non-compressed parts of F0 contours. Further, we have explained the reason for the amplitude limitation of delta oscillations processing constituents that follows a local emphasized constituent, in order to mark it as a global emphasized constituent.

We conclude that it is crucial to correctly appreciate the presence or absence of emphasis within prosodic relations, in order to correctly identify their nuclear elements. If prosodic relations and their cognitive and prominence structures are correctly deduced within prosodic phrases, we can compare the resulted nuclear events with the linguistic prominence events of stressed syllables and focus words, and we will conclude about their correspondence.

This paper suggests a cognitive view on intonational contours in order to explain what makes prominent constituents of low and high-level prosodic relations, without using linguistic information. The prominence identification is important in semantics for understanding local and global focus positions of utterances. This is also important in the phonological analysis of F0 contours for deducing the pitch accent positions in both cases of stressed and non-stressed syllables. The cognitive IS model is useful in speech processing because it leads to a better understanding of supra-segmental features of speech signals.

By understanding how speech objects are structured at cortical level, researchers will be motivated for taking into account the interpretation of F0 contours by using the cognitive model of information structure.

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

Boucher, V. J., Gilbert, A. C., & Jemel, B. (2019). The Role of Low-frequency Neural Oscillations in Speech Processing: Revisiting Delta Entrainment. *Journal of Cognitive Neuroscience*, *31*(8), 1205-1215. https://doi.org/10.1162/jocn_a_01410

Delais-Roussarie, E., Post, B., Avanzi, M., Buthke, C., Di Cristo, A., Feldhausen, I., … & Hiyon Yoo. (2015). Intonational phonology of French: Developing a ToBI system for French. In Frota, S. and Prieto, P. (Eds), Intonation in Romance, 63-100. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199685332.003.0011

Hualde, J. I., & Prieto, P. (2015). Intonational variation in Spanish European and American varieties. In Frota, S. and Prieto, P. (Eds), Intonation in Romance, 350-391. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199685332.003.0011

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. (2025). A cognitive view on pprosodic relations. *English Linguistics Research, 13*(2). https://doi.org/10.5430/elr.v13n2p1

Jitcă, D., & Maruşca, S. (2023). A common view on broad and contrastive focus events. *Bucharest Working Papers in Linguistics, XXV*(2). https://doi.org/10.31178/BWPL.25.2.2

Ladd, D. R. (2008). *Intonational phonology.* Cambridge University Press, pp: 68, 152-153. https://doi.org/0.4236/ce.2012.326134.

Liberman, A. M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry, 8*, 249-336. http://www.jstor.org/stable/4177987

Liberman, M. (1975). The Intonational System of English, unpublished Doctoral dissertation, MIT, Cambridge, Massachusetts.

Nelson, M. J., El Karoui, I., Giber, K., 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. Epub2017 Apr 17.

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

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

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 meaning. *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. https://doi.org/10.1146/annurev-psych-010419-051101. PMID: 31905113; PMCID: PMC8679009.