Hunyadi László: Cognitive grouping and recursion in prosody
Argumentum, 4 (2008), 231-253
Kossuth Egyetemi Kiadó (Debrecen)

Tanulmány

Hunyadi László

Cognitive grouping and recursion in prosody[1]

Abstract

This article contributes to the debate on the structure of the narrow faculty of language (FLN) as suggested in Hauser, Chomsky, and Fitch (2002), especially to the issue whether linguistic structures beyond syntax can be recursive. It argues that *a.)* speech prosody displays significant cues of recursion in the form of tonal and pausal grouping, and *b.)* recursion found in prosody is the manifestation of a more general computational mechanism.

It introduces the *principle of tonal continuity* to account for the continuous tonal phrasing of discontinuous structures with nested embedding and suggests that what underlies this cognitive computational process is the *bookmark effect*. It shows that the computational difference between nested recursion and iteration correlates with their prosodic difference, whereas the computationally indistinguishable tail recursion and iteration are similar in their prosodic realization.

Experiments are presented involving speech prosody, grouping in abstract prosodic patterns as well as grouping in abstract visual patterns demonstrating that recursive phrasing has similar properties across modalities. Their differences suggest that grouping in prosody has its cognitive basis in the grouping of less specific, more abstract, non-linguistic elements. It is concluded that recursion in prosody cannot be the effect of an interface relation between syntax and prosody, instead, it is the manifestation of a more general, more universal computational mechanism beyond linguistic structure.

1 Introduction[2]

In their article on the faculty of language Hauser, Chomsky, and Fitch make a distinction between FLN (the faculty of language in its narrow sense) and FLB (the faculty of language in its broad sense) and hypothesize that FLN only includes recursion and is the only uniquely human component of the faculty of language (Hauser, Chomsky, and Fitch 2002: 1569). According to their hypothesis the core computational mechanisms of recursion in FLN appear in narrow syntax and the mappings to the interfaces and are uniquely human. In contrast, they suggest that much of the complexity of language derives from complexity in the peripheral components of FLB that have an ancient evolutionary history. These peripheral components

[1] An earlier version of this article was presented at the conference Recursion in Human Languages in Normal, Ill., April 27-29, 2007. A revised version of this article will appear in the volume *Recursion and Human Language.* Ed. by Harry v. der Hulst and Dan Everett. Series *Syntax in Generative Grammar*. Mouton. Research this article is based on was partially supported by Grant OTKA 69042.

[2] I would like to express my thanks to the following people for their valuable comments and suggestions on previous versions of this paper or parts of it: Zoltán Bánréti, Valéria Csépe, Ray Jackendoff, Ilona Kassai, András Kertész, Katalin É. Kiss, and Shalom Lappin.

Hunyadi László: Cognitive grouping and recursion in prosody
Argumentum, 4 (2008), 231-253
Kossuth Egyetemi Kiadó (Debrecen)

of FLB include the sensory-motor system, a conceptual-intentional system and other possible systems.

In their view, FLN, the abstract linguistic computational system is independent of the other systems with which it interacts and interfaces, but, being a component of FLB, its underlying mechanisms are some subset of those underlying FLB. It is this key component of FLN that generates internal representations and maps them into the sensory-motor interface by the phonological system, and into the conceptual-intentional interface by the formal semantic system.

Apart from numerous arguments against this restricted understanding of FLN (that only recursion is special to language; cf. Jackendoff and Pinker 2005, Pinker and Jackendoff 2005), it is also arguable whether recursion is specific to narrow syntax alone. Phonology and morphology are among the possible candidates to be also included in the faculty of language as specific to humans, but if any of them shows properties of recursion, the question arises how recursion found in them is related to syntactic recursion and whether their computational mechanism is specific to humans only. The question is important since the issues raised in Hauser, Chomsky, and Fitch (2002) are directly related to the evolutionary history of human language and any findings regarding recursion in other human-specific areas of language may broaden and refine our understanding of the place of human language in the history of evolution. In this article it will be shown that a.) prosody, as part of phonology shows properties of recursion and b.) the mechanism behind prosodic recursion is related to a broader, non-linguistic cognitive capacity of abstract grouping. These findings have the potential to be seen as a challenge to the original hypothesis of Hauser, Chomsky, and Fitch (2002) regarding the structure and evolutionary status of the narrow language faculty.

2 On the syntax-prosody interface

It is not the aim of this article to give a detailed overview of the large body of work on the issue of how syntactic structure is reflected in prosodic structure. For our purpose it will suffice to mention that in a syntax-based grammar (as has been the accepted approach in much of the generative tradition) syntactic relations are phonologically and semantically interpreted through interface relations and one would expect that a.) this interpretation is as perfect as possible and b.) the phonological and the semantic sides of this interface express but relations contained in the syntactic structure proper. That is, whereas in such a view syntax is autonomous, the phonological and semantic interfaces are not. Restricting ourselves just to the interface between syntax and phonology, however, it has long been observed that the two do not fully match. It was already shown in Chomsky and Halle (1968) that there is a mismatch between syntactic and phonological (prosodic) phrase boundaries; cf. (1) – the inserted vertical bars showing where phonological phrase boundaries occur:

(1) *[|This is [the dog |that chased [the cat |that killed [the rat |that ate [the malt |that lay in [the house that Jack built.|]]]]]]*

This mismatch is considered to be the result of attempting to match one kind of structure onto another, suggesting that rules governing structures on the two sides of the interface are not necessarily identical.

Moreover, it has been found that – in addition to certain limitations of prosodic structure – the latter can, in turn, “repair” limitations of syntactic structure as well: as Price et al. (1991)

Hunyadi László: Cognitive grouping and recursion in prosody
Argumentum, 4 (2008), 231-253
Kossuth Egyetemi Kiadó (Debrecen)

showed, syntactical ambiguity (verb-preposition vs. verb-particle) is disambiguated by prosody via the differentiating use of phrase final lengthening; cf. (2) vs. (3):

(2) *[The Vikings won] % [over their enemies.] %*
(3) *[The Vikings won over] % [their enemies.] %*

It has also been shown that prosodic phrasing can disambiguate syntactic/semantic ambiguities (e.g. scope ambiguities in Hungarian resolved by prosody; cf. Hunyadi 2002); cf. (4) and (5):

(4) *JÁNOS látott mindenkit.*
John-nom. saw everyone-acc.
'It was John that, for every x, x=person, John saw x.'

(5) *JÁNOS látott MINDENKIT.*
John-nom. saw everyone-acc.
'For every x, x=person, it was John that saw x.'

In view of these and similar mismatches one may suggest that the structural/computational mechanisms behind syntax and prosody are not necessarily identical. In fact, it has been argued (cf. Nespor and Vogel 1982, Selkirk 1984) and widely accepted afterwards that mismatches like those demonstrated in (1) are accounted for by the fact that prosody is 'flatter' than syntax due to the lack of recursion in prosody. Namely, according to Selkirk's Strict Layer Hypothesis (SLH), in a prosodic tree, any domain at a given level of hierarchy consists exclusively of domains at the next lower level of the hierarchy. As a consequence of SLH, then, a prosodic tree has a limited depth. As such, prosody is hierarchical but is not recursive.

This view was partially challenged in Ladd (1986, 1992, 1996) by the introduction of the so-called Compound Prosodic Domain allowing for restricted recursion in cases like (6) and (7) and giving a structural account for perceived differences in boundary strength in these structures:

(6) *[A and B [but C]]*
(7) *[A [but B and C]]*

Here, a prosodic domain of a given type X has immediate constituents of the same type X, an evidence for a mechanism to generate an indeterminate depth of prosodic structure (even if only restricted to coordination).

Another way to allow for certain, restricted cases of prosodic recursion was suggested by Selkirk (1995) and Peperkamp (1997), namely, that the constraint against recursion can be maintained by allowing some other constraints to force its violation and produce recursive structure. This position was also adopted by Truckenbrodt (1999) in his analysis of data from Kimatuumbi.

Wagner (2005), following Ladd's proposal, has recently shown that translating recursive syntactic phrase boundaries into recursive boundary length differences (size of a pause) in the production of prosodic segmentation is perceptually valid and indicative of recursive structure. The observation, however, that the boundary length preceding and following an embedded syntactic phrase is perceived and calculated locally, i.e., relative to one another rather than across a complete syntactic hierarchy (cf. Clifton, Carlson, and Frazier 2002) suggests that there can be additional (prosodic) factors that influence the perception and evaluation of embedding. Such a view is further supported by Watson and Gibson (2004) showing that pauses do not always occur at intonational phrase boundaries and, in addition, syntactic structure does not determine the size of a pause, although it may affect the probability of a boundary.

In view of such observations and results we may rightly ask how and to what extent prosody with its limited symbolic inventory (lexicon) of just pitch, intensity and duration (pause) can contribute to the generation of prosodic structure. Since, as we have seen, there is no necessary direct relation between syntactic and pausal phrasing, we might want to turn our attention to prosodic phrasing brought about by variation of pitch and intensity, and, as a result, tonal phrasing and find out how tonal and pausal phrasing participate in their joint venture of representing prosodic (and syntactic) phrases. Finally, as it was shown by Ladd (1996) that prosodic structure is not as flat as predicted by Selkirk's Strict Layer Hypothesis, we may want to test the generative power of prosodic recursion even beyond coordination. By presenting next the results of our experiments, we will suggest that tonal phrasing is mandatory for prosodic phrase formation and that its full generative capacity is based on recursion.

3 Properties of tonal phrase formation vs. recursion

3.1 Experiments

In computational terms recursion is usually defined as a procedure that calls itself; structurally, as a result of recursion a constituent will contain another constituent of the same kind with the potential to generate unbouded sequences. Recursion is always hierarchical, although the existence of hierarchy does not by itself define a structure as recursive (cf. the understanding of the prosodic tree according to the SLH). Recursion always involves repetition of some operation(s), but it is clearly distinct from iteration, the latter being non-hierarchical. Recursion has several kinds (for a systematic typology of recursion see Parker 2006). Among them the most interesting for us will be instances of embedded (nested) recursion as contrasted to tail recursion. In the case of embedded recursion chunks of a phrase/sentence are embedded within a larger material, whereas cases of tail recursion are characterized by their phrase/ sentence-initial or sentence-final position. The latter bear some similarity to iteration but the two are not identical. The distinction between embedded and tail recursion is important for us beyond their descriptive difference. From a computational/procedural point of view, embedded recursion involves access to memory (storage and retrieval): according to Liu and Stoller (1999) (cited in Parker 2006: 175), "[r]ecursion refers to computations where the execution of a function or procedure calls itself and proceeds in a stack fashion." Iteration, on the other hand, has a different relation to memory: it is performed in looping, "i.e., involving keeping track of the times the same operation has been performed." We hypothesize that if the performing of prosodic recursion and iteration has any cognitive relevance, then there will be a meaningful difference in the way (embedded) recursion and iteration access memory, and this difference will be manifested in some prosodic features of the given utterances.

In order to test this hypothesis, instances of utterances of sentences with embedded recursion as well as of sentences with iteration were recorded, analyzed and compared with special emphasis on their tonal and durational (pausal) realization. Below, we will present the results of experiments with recursion (3.2), discuss tonal continuity as the prosodic principle to denote recursion (3.3), show that this principle applies to discontinuous phrases in general (3.4) and, presenting the results of experiments with iteration, discuss how the procedural difference between recursion and iteration is manifested in prosody (3.5).

3.2 The prosodic structure of utterances with central embedded (nested) recursion[3]

The fundamental units of the prosodic structure of an utterance have a strong tonal character. According to Beckman and Pierrehumbert (1986), the edge of the intonation phrase is marked by a phrase tone and a boundary tone, whereas the intermediate phrase by a phrase tone only. Edges of phrase tone contours are, however, often associated also with pauses in a meaningful way: according to Cooper and Paccia-Cooper (1980), the duration of a pause following a tonal break may even reflect the given hierarchical syntactic boundary. It is also widely held that intonation phrase boundaries themselves can be determined on the basis of possible or obligatory pauses (cf. Downing 1970, Selkirk 1984, Taglicht 1998). Next, we will examine the role of phrasal tones and pauses in prosodic segmentation by asking which of the two (if any) has priority over the other in denoting prosodic phrase boundaries. First, we will look at the role of pauses and see how pauses match the structure of a Hungarian sentence (4):

(8) *A macska, amit a kutya, ami megveszett, megharapott, elszaladt.*
the cat that the dog that was rabid bit ran away
'The cat that the dog that was rabid, bit, ran away.'

The above sentence is an example of embedded recursion. Interestingly, as for pausal phrasing, we found a significant difference among the four speakers (the four numbers represent the pauses on the four syntactic phrase boundaries):

Speaker 1: 0.065, 0, 0, 0.158
Speaker 2: 0.050, 0.187, 0.075, 0.075
Speaker 3: 0.164, 0.072, 0.046, 0.217
Speaker 4: 0.844, 0, 0, 0.125

As the data show, no unified way was found to represent all embedded syntactic boundaries by pauses. What they all agreed in was to represent by pause the major embedding only.[4] Three out of the four speakers (Speaker 1, Speaker 2 and Speaker 3) agreed in indicating the boundaries of the major embedding (*A macska ... elszaladt* 'The cat … ran away') so that the pause before embedding (i.e., following the phrase *A macska*) was shorter than the one before de-embedding (i.e., before *elszaladt*), whereas an opposite temporal relation for the same embedding was measured for Speaker 4. For Speaker 2 and Speaker 3, those representing by pause the second level embedding as well, the latter was represented by a pause shorter than the first one. Speaker 1 and Speaker 4 did not appear to represent the inner embeddings by pause at all.

---

[3] For all analyses, the program Praat v. 4.5.16 was used on an Apple PowerBook G4 computer.
[4] This observation coincides with Truckenbrodt (2007), who, following Downing (1970) points out that root clauses (not embedded inside another clause), and only those, are bounded by obligatory intonation phrase brakes.

Hunyadi László: Cognitive grouping and recursion in prosody
Argumentum, 4 (2008), 231-253
Kossuth Egyetemi Kiadó (Debrecen)

Whereas pausal phrasing showed a quite varied picture, tonal phrasing proved to be highly systematic. It was found consistent with all speakers that a. each of the embeddings was performed at a lower pitch level (speaking range) than the phrase it was embedded into and b. the corresponding de-embedding was performed by an upstep, in fact, by a return to the very same pitch level (speaking range) before the given embedding. Lowering the speaking range of an embedding was consistent with the recursive application of embedding: the deeper the embedding in the syntactic hierarchy, the deeper speaking range for the given embedded phrase was chosen.

This tonal representation of the recursive syntactic structure is shown in the next figures as follows: Figure 1 represents the pitch contour of the overall utterance of (8) by Speaker 1. Figure 2 represents (8a) where the pitch contour of the innermost embedding of the same recording (8) is removed, and Figure 3 represents (8b) with the pitch contour of both inner embeddings of the same recording (8) removed. It is noteworthy that, as suggested above, the removal of an embedding and the subsequent concatenation of the remaining halves of the utterance results in perceptually continuous tone at the point of concatenation, proving that whereas embedding is realized by the lowering of the speaking range of the overall embedded phrase, as a result of de-embedding the tone returns to the final pitch before embedding. Importantly, the return of the tone to the final pitch before embedding suggests that downdrift, the general tendency of the lowering of the pitch contour across an utterance does not apply. It is this effect that, after the removal of the intervening material, results in the perception of the concatenation of the two remaining halves as a single prosodic phrase[5] (in the figures to follow, an arrow indicates the place of concatenation after the removal of an embedded phrase):

(8) *A macska, amit a kutya, ami megveszett, megharapott, elszaladt.*
the cat that the dog that was rabid bit ran away
'The cat that the dog that was rabid, bit, ran away.'

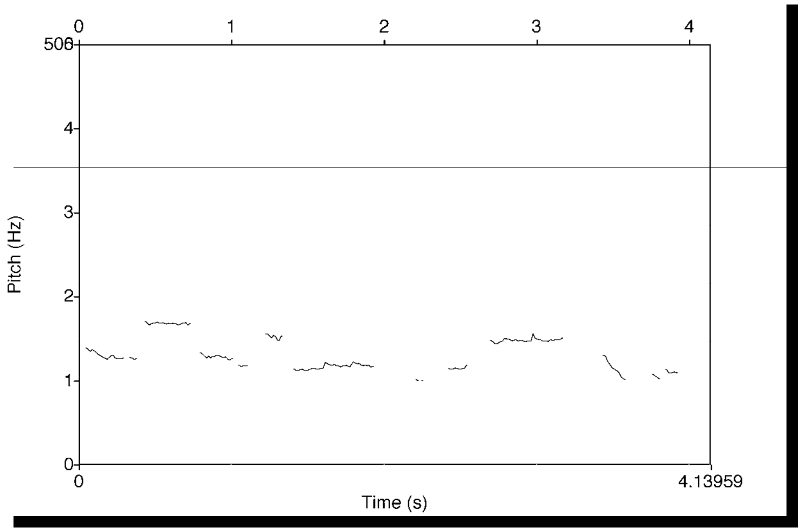

Figure 1

[5] The original recordings and the results of their manipulations can be found at https://ling.arts.unideb.hu/recursion.

Hunyadi László: Cognitive grouping and recursion in prosody
Argumentum, 4 (2008), 231-253
Kossuth Egyetemi Kiadó (Debrecen)

(8a) *A macska, amit a kutya […] megharapott, elszaladt.*
the cat that the dog […] bit ran away
'The cat that the dog […] bit, ran away.'

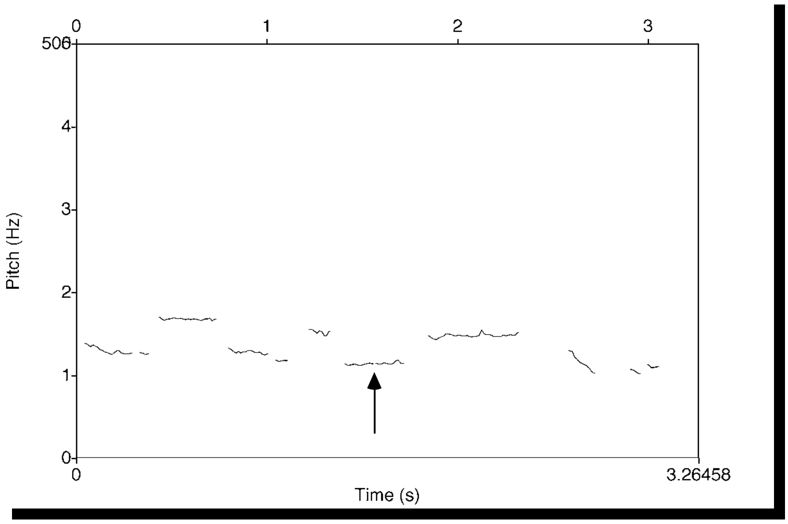

Figure 2

(8b) *A macska […] elszaladt.*
the cat […] ran away
'The cat […] ran away.'

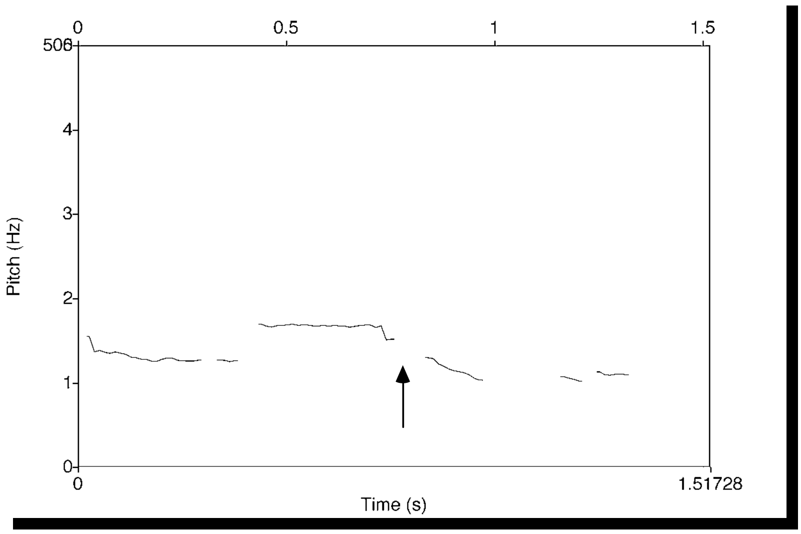

Figure 3

The English equivalent of the above utterance shows the effect of a similar prosodic mechanism; cf. (9) and (9a)-(9b) and Figure 4, Figure 5 and Figure 6, respectively:

(9) *The cat that the dog that was rabid, bit, ran away.*

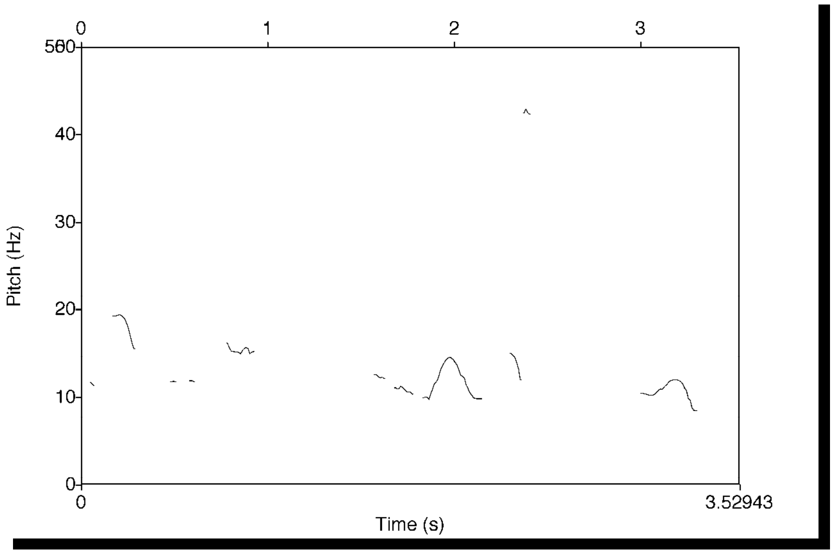

Figure 4

(9a) *The cat that the dog […] bit, ran away.*

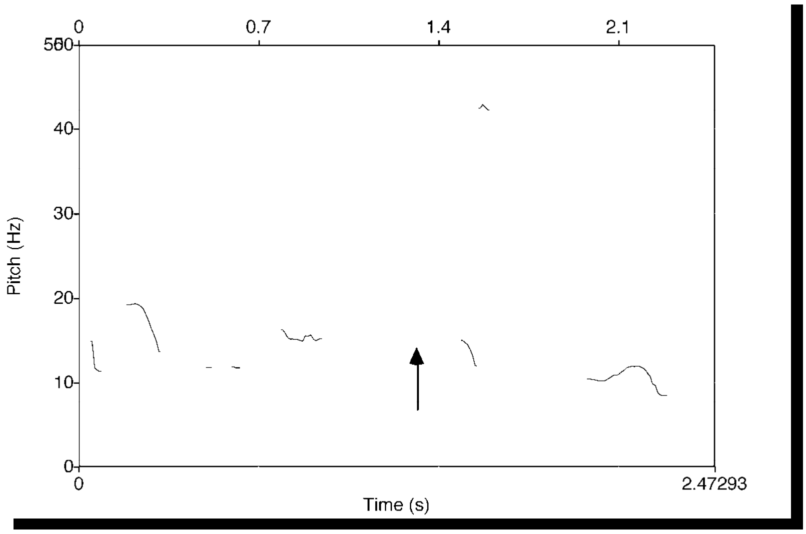

Figure 5

Hunyadi László: Cognitive grouping and recursion in prosody
Argumentum, 4 (2008), 231-253
Kossuth Egyetemi Kiadó (Debrecen)

(9b) *The cat […] ran away.*

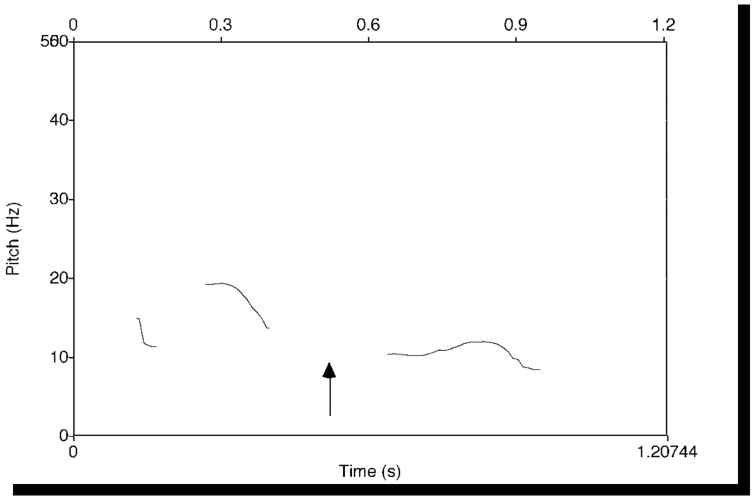

Figure 6

The mechanism that produces the continuation of the tone of discontinuous chunks of an utterance will be called tonal continuity.

3.3 Tonal continuity and the bookmark effect

Tonal continuity is a prosodic principle of long-distance dependency characterizing two discontinuous prosodic segments *A* and *B* as a single prosodic phrase *[A…B]*. It applies when a prosodic phrase *C* of the same type intervenes in *[A…B]*, such as *[ACB]*. The discontinuous segments *A* and *B* of *[A…B]* have tonal properties identical to those of a sequence of *A* and *B* within a continuous prosodic phrase, *[AB]*.

Tonal continuity has three important effects: *a.* downdrift, the general tendency of the lowering of the pitch contour across an utterance does not apply to the intervening phrase *C* within *[A…B]*; *b.* phrase *C* is tonally realized in a different (usually lower) tonal space (the term used in the sense of Ladd 1992), i.e., at a lower overall speaking range), and *c.* de-embedding takes the form of an upstep of *B* to the initial tonal space of *A*.

Since tonal continuity can repeatedly apply to another prosodic phrase of the same type embedded within the embedded *C* itself with yet another similar effect (the relative lowering of the tonal space at the point of embedding and an upstep at the point of de-embedding to the initial tonal space of the given embedding, thus creating a hierarchy of tonal spaces), its repeated application is recursive. As a consequence, the application of tonal continuity to prosodic phrases of the same type shows that prosodic phrases can be generated recursively. Consequently, the mechanism of generating prosodic structure is suggested to include recursion.

We assume that there is a cognitively significant computational process underlying tonal continuity. Tonal continuity, i.e. the principle behind long-distance tonal dependency obviously overrides the principle of downdrift, the default tonal movement across an utterance so that instead of a continuous downdrift the tonal parameters of two discontinuous segments will match: in the above sequence *[ABC] C* is not downdrifted, instead, it matches A tonally

Hunyadi László: Cognitive grouping and recursion in prosody
Argumentum, 4 (2008), 231-253
Kossuth Egyetemi Kiadó (Debrecen)

as if they formed a single continuous phrase. We assume that this process heavily relies on memory: it involves storing the prosodic state before embedding (embedding *C* in our case) and restoring it after embedding. It serves as a sort of a memory aid which, therefore, we will call the *bookmark effect*.

Although tonal continuity is an important diagnostic characteristic of central embedding, it is not restricted to it: it is also the prosodic realization of any kind of insertion, as shown in the next section.

3.4 Tonal continuity in non-recursive discontinuous phrases

The prosodic realization of insertion demonstrates that the principle of tonal continuity is not restricted to prosodic recursion: it equally applies to cases of insertion as well; cf. (10) and (11), examples from Hungarian and English (the corresponding F0-contours are shown in Figure 7 and Figure 8 for (10) and (10a), and Figure 9 and Figure 10 for (11) and (11a), respectively:

(10) *Meg tudnád mondani, hogy – az én órám megállt – hány óra van?*
Vpref you-could tell that the my watch stopped how many hour is
'Could you tell me – my watch has stopped – what time it is?'

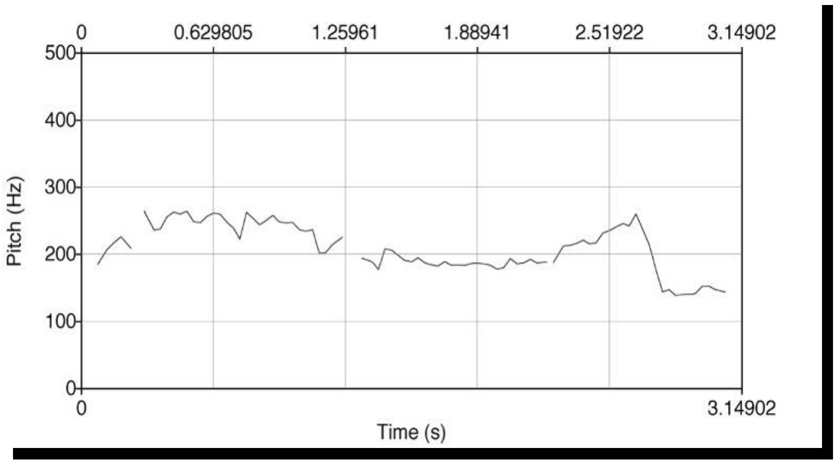

Figure 7

Hunyadi László: Cognitive grouping and recursion in prosody
Argumentum, 4 (2008), 231-253
Kossuth Egyetemi Kiadó (Debrecen)

(10a) *Meg tudnád mondani, hogy […] hány óra van?*
Vpref you-could tell that […] how many hour is
'Could you tell me […] what time it is?'

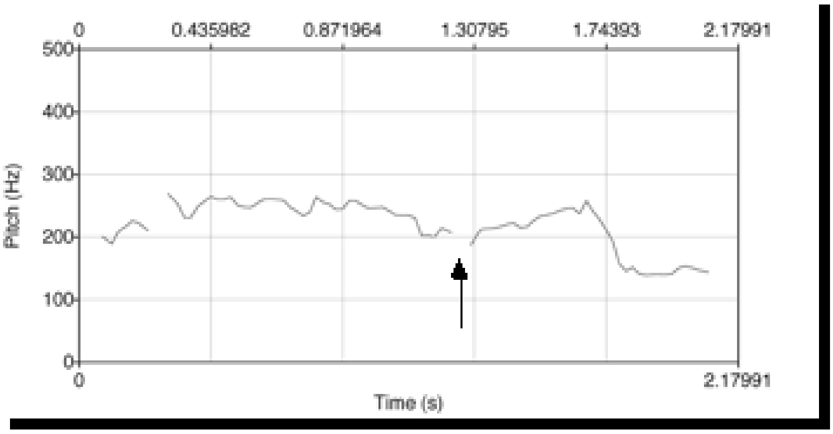

Figure 8

(11) *Why don't you – I said – go and find it?*

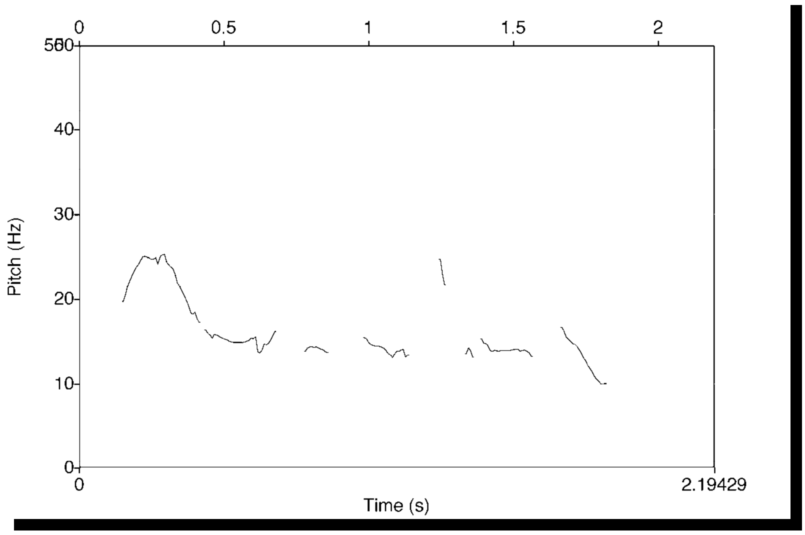

Figure 9

Hunyadi László: Cognitive grouping and recursion in prosody
Argumentum, 4 (2008), 231-253
Kossuth Egyetemi Kiadó (Debrecen)

(11a) *Why don't you […] go and find it?*

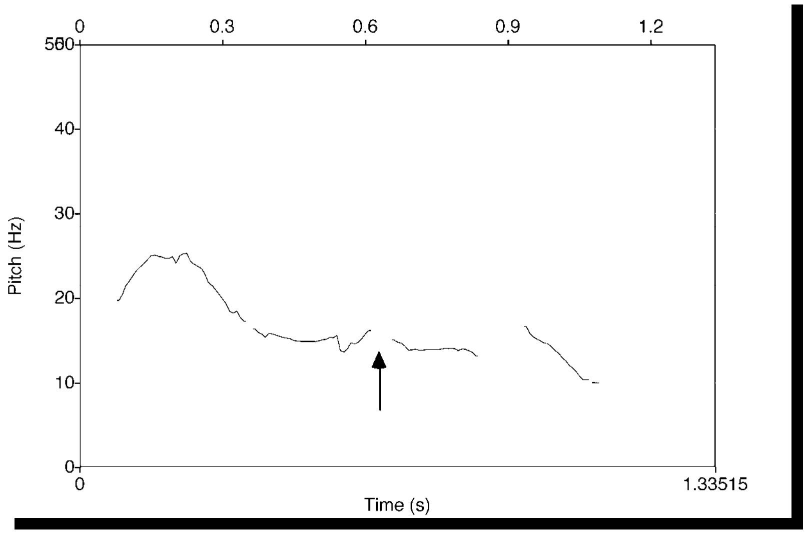

Figure 10

Both (10a) and (11a) clearly demonstrate the bookmark effect: a continuous tone at the point of concatenation, similarly to what was previously found in the case of recursive structures. Since (10a) and (11a) are instances of insertion, we can conclude that tonal continuity is a prosodic principle applied to discontinuous prosodic phrases in general. Its function is to denote that two long-distance discontinuous prosodic chunks are structurally related forming a single phrase. Although insertion cannot be applied recursively, embedding can. As such, tonal continuity, applied to recursive embedding, results in the recursive lowering of the tonal space and the recursive application of upstep to the next higher tonal space.

3.5 Recursion vs. iteration

As mentioned earlier, recursion always involves repetition of some operation(s), but it is clearly distinct from iteration, the latter being non-hierarchical. Since tonal continuity was found to be related to discontinuous structures, iteration being sequential we do not expect the bookmark effect and, consequently, tonal continuity to apply to it. Recordings of iterative structures in Hungarian support such an expectation; cf. (12) and (13) (compare their respective F0-contours in Figures 11 and 12):

Hunyadi László: Cognitive grouping and recursion in prosody
Argumentum, 4 (2008), 231-253
Kossuth Egyetemi Kiadó (Debrecen)

(12) *János egy nagyszerű, jókedvű, jó humorú ember.*
John an excellent cheerful good-humored man
'John is an excellent, cheerful, good-humored man.'

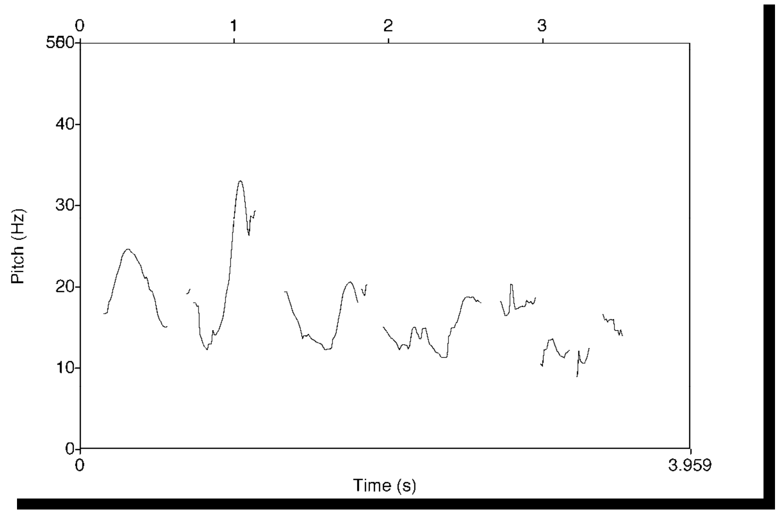

Figure 11

(13) *Egy, kettő, három, négy, öt , hat, hét, nyolc.*
one two three four five six seven eight
'One, two, three, four, five, six, seven, eight.'

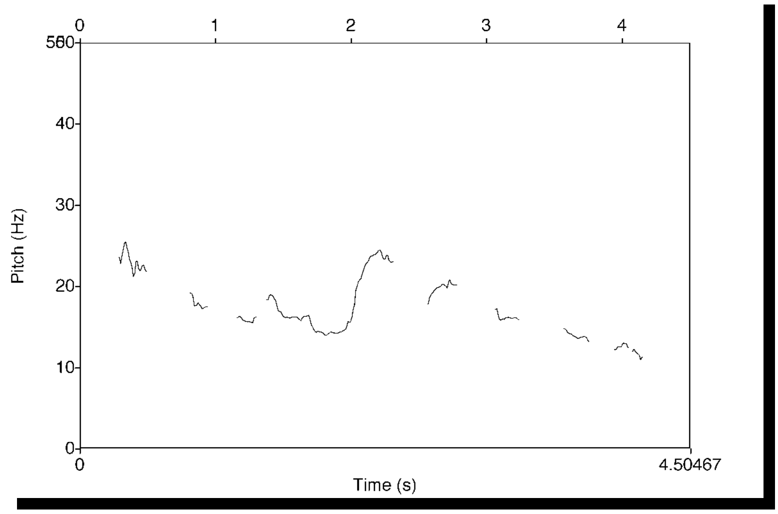

Figure 12

As we saw in 3.3, prosodic recursion involving central embedding follows the principle of tonal continuity, characterized by the lack of downdrift at the juncture of A and B inside the phrase [A…B], the presence of a change (usually lowering) of the tonal space for embedding, and the presence of upstep at the point of de-embedding to the original tonal space at B. Iteration, as shown in both (12) and (13), does not have any of these properties: first, there is a continuous downdrift across the whole utterance, second (as a consequence of the overall downdrift), there is no change of the tonal space, and third, the slight upstep observed at the beginning of each phonological word is not an upstep to the very same pitch level of a previous, discontinuous chunk of prosody but an upstep to the downdrifted initial pitch of each copy of the iterated prosodic segments (looking at it from a different angle, one can observe here the effect of downstep, the stepwise lowering of pitch at specific pitch accents; cf. Pierrehumbert 1980).[6] We can thus conclude that prosodic structures with central embedded recursion and those with iteration are derived by different prosodic mechanisms: such recursive prosodic structures (and non-recursive discontinuous structures in general, as shown in 3.4) are derived by tonal continuity, whereas iterative structures follow downdrift.

Tail recursion is similar to central embedded (nested) recursion in that it also involves embedding and generates a hierarchical structure. But in an important aspect it is different: its assumed computational mechanism does not need to store information on a stack in the way central embedding does, since there is no need to return to the pre-embedding state. As for its prosodic realization, then, we can expect that it does not display the bookmark effect, i.e. it does not involve a change (lowering) of (and return to) the initial tonal space by a pair of a specific downstep and an upstep. Instead, we can expect the sequence of prosodic material to follow downdrift across the utterance. This is exactly what we find in (14), a case of tail recursion in Hungarian (see its corresponding F0-contour in Figure 13):

(14) *János találkozott Anna férje barátjának a lányával a moziban.*
John met Anna husband (poss) friend (poss) the daughter (poss) the cinema (in)
'John met Anna's husband's friend's daughter in the movie.'

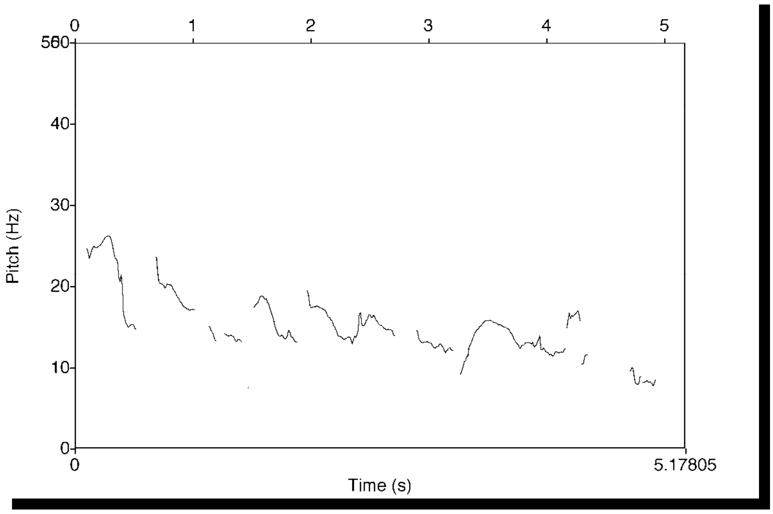

Figure 13

The observed downdrift at points of embedding takes exactly the same prosodic form as the one observed in iteration (cf. (12) and (13)) – thus giving a prosodic support for the assumption that iteration and tail recursion are not structurally distinguishable.

---

[6] Carefully observing the pitch contour of (13) we can note an upstep with continuation rise after négy 'four'. This upstep, however, not associated with a previous downstep is not the effect of tonal continuity. Instead, it is the effect of inherent grouping, a property of grouping in general, to be discussed in Section 4 below.

Hunyadi László: Cognitive grouping and recursion in prosody
Argumentum, 4 (2008), 231-253
Kossuth Egyetemi Kiadó (Debrecen)

3.6 Recursion and pausal phrasing

Although it seems to be sufficient to differentiate nested recursive and iterative prosodic structures on the basis of their tonal properties alone, there is still one more aspect of prosody, duration (pausal phrasing) that is worth considering. The difference in their use of pausal phrasing is equally telling.

We find that in nested prosodic recursion, pausal phrasing is structural (it follows hierarchical relations), whereas in iteration it is rhythmical. Accordingly, in recursive structures at least the major (first level) embedding is generally preceded and followed by a pause (the same for cases of single insertion) with some tendency for the pause before the embedding to be shorter than after de-embedding (in three out of our four cases; cf. Section 3.2). We found pausal segmentation at structural positions only, even though pauses were less articulate (if at all) at lower levels of the hierarchy. In contrast, an utterance containing iteration was found to be characterized by a eurhythmic organization of the sequence of iterated constituents denoted by the rhythmical distribution of pitch accents rather than the use of pauses (more on eurhythmics cf. Hayes 1984, see also Liberman and Prince 1977).

A computational model of recursion, as mentioned earlier, involves access to memory: one has to store information on a stack and it has to be retrieved later on. In recursive speech prosody there are two kinds of information to be taken care of in memory (and to be shown in the bookmark effect): structural (to keep track of the hierarchical position of the given constituent) and phonological (to keep track of the tonal position of discontinuous segments following the principle of tonal continuity).[7] The storing and retrieving of both kinds of information need access to memory and, consequently, require processing time. One might want to suggest that pauses in recursive structures are then reflections of such a mental operation.

However, the fact that pausal phrasing in speech utterances only occurs at some but not all structural positions indicates that access to memory is too short to be directly reflected by a pause. As such, the presence of a pause cannot then be the consequence of such a memory access, it is rather a formal means to express some structure. However, since pauses do not necessarily occur at all hierarchically significant positions whereas tonal phrasing is present in all utterances, we may conclude that pausal phrasing is, to some extent, secondary to tonal phrasing.

If the above mentioned obvious prosodic difference between nested recursion and iteration (pausal vs. rhythmic phrasing) cannot be justified on grounds of differences in memory requirements, we have to suggest that pausal and rhythmic phrasing are actually two different formal means for the prosodic representation of two different types of structure, together with their differences in tonal phrasing as well. The question remains whether these prosodic means of expressing structure are purely linguistic or have some more general, cognitive basis. The aim of the following two sections is to offer a possible answer to it.

Surface differences apart, all sentences agree that any syntactic structure is based on the grouping of words and phrases of a given sentence. The prosodic segmentation of a sentence is also based on a similar grouping. Similarly to the fact that the number of underlying types of syntactic phrases is finite (and limited), one may also assume that the infinite prosodic surface variation of utterances can also be generalized as instances of a finite (and limited) number of underlying prosodic phrases. It will be assumed that these abstract underlying

[7] Our observations based on the role of tonal continuity both in instances of recursive and non-recursive long-distance dependences suggest that, contrary to Parker (2006: 185), a single center-embedding also has the memory requirement of keeping track in a stack fashion.

Hunyadi László: Cognitive grouping and recursion in prosody
Argumentum, 4 (2008), 231-253
Kossuth Egyetemi Kiadó (Debrecen)

prosodic phrase types represent the fundamental types of grouping in general (coordination and subordination) and that rules of generating complex prosodic phrases include but are not limited to recursion. In order to identify the underlying, abstract properties of prosodic grouping, we carried out two kinds of experiments: those involving the grouping of abstract prosodic patterns with no linguistic (syntactic or semantic) content and those involving the grouping of abstract visual patterns even with no prosodic content. With these experiment we expected to identify those abstract tonal and pausal (durational) means of grouping which are general enough to underlie grouping in speech prosody.

Furthermore, we assumed that if the tonal and pausal properties of speech prosody will be found to share in common the basic corresponding properties of abstract prosodic/visual grouping, then we can make the conclusion that prosodic phrasing is based on more general principles of cognition rather than language alone. Below, we will examine the grouping of abstract prosodic patterns (in Section 4) and the grouping of abstract visual patterns (in Section 5) (the data and analyses referred to here are taken from a more extensive study on grouping; cf. Hunyadi 2006).

4 Grouping in abstract prosodic patterns

A group of 23 university students were presented with a sequence of linearly arranged letters such as *A*, *B*, *C*, and *D*. They were instructed that in some patterns pairs of parentheses were used to denote various kinds of grouping. They were asked to observe the given patterns and then, following the pattern, read them out pronouncing the names of the letters. Each reading was recorded and its tonal and pausal phrasing analyzed.

A strong consistency was found among subjects in the tonal representation of patterns; cf. and (16):

(15) *A B C D*
- a. rise rise rise fall (14)
- b. fall fall fall fall (6)
- c. rise fall rise fall (2)
- d. fall rise fall fall (1)

The above pattern had a simple internal structure: the coordination of four abstract elements. Whereas the tone of *D* was determined by its pattern-final position to be a fall, *A*, *B* and *C* were also pronounced using the same rise or fall tone in 20 out of the 23 cases representing the equal structural status of the four elements. It was only in 3 cases that there was a tonal variation on *B*, suggesting a structural division between *B* and *C*.

(16) *(AB) (CD)*
- a. fall rise fall fall (15)
- b. rise fall fall fall (5)
- c. high_level high_level low_level low_level (2)
- d. rise fall rise fall (1)

The parentheses in this pattern suggest two groups: *AB* and *CD*, and this grouping is tonally represented by most (18) subjects applying a change in the direction of the tone or level of the pitch at the group boundary. Even in the remaining 5 cases (response type b.) there was a significant lowering of the tonal space following the first fall (i.e., after *B*).[8]

Tonal phrasing in (17) with a recursive pattern resulted in more variation:

(17) *A(B(CD))E*
a. rise rise fall rise fall (12)
b. rise rise fall fall fall (6)
c. rise fall fall fall fall (3)
d. fall rise fall fall fall (2)

In 12 of the cases tonal phrasing was used to represent embedded structure (with rise at a phrase boundary). The rise on *B* and the fall on *C* were initiated at an overall pitch level lower than the previous tonal space, corresponding to what was found earlier for recursive embedding in speech prosody as well. In 6 cases the main recursive embeddings between *A* and *B*, and between *B* and *C* were tonally represented in a similar way, but the de-embedding following *D* was tonally "missed". The last two groups of subjects used tonal phrasing less consistently for the denotation of this abstract recursive structure.

If we look at pausal phrasing manifested in the above patterns, we find a strong consistency among all subjects in representing the given patterns by temporal structure (cf. Figure 14 for (15) and (16) and Figure 15 for (17):

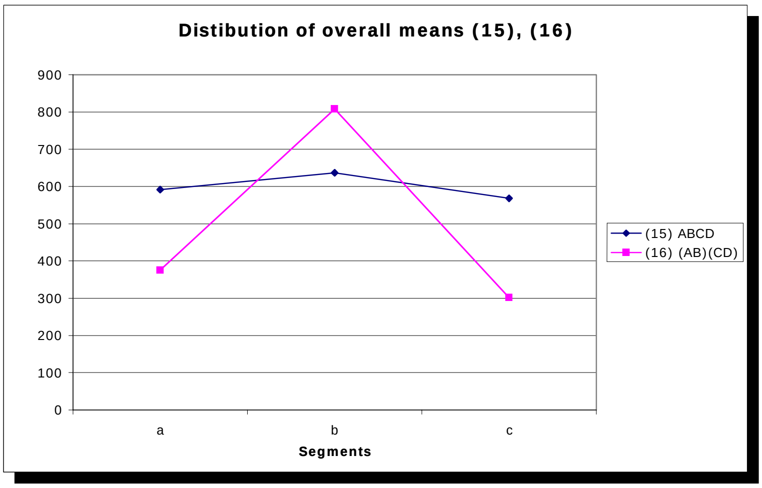

Figure 14

[8] Ladd (1996) notes a similar tonal effect at a major boundary: the initial peaks of a clause at a major boundary were higher than those at a minor boundary. See also H-tone insertion in Truckenbrodt (1999).

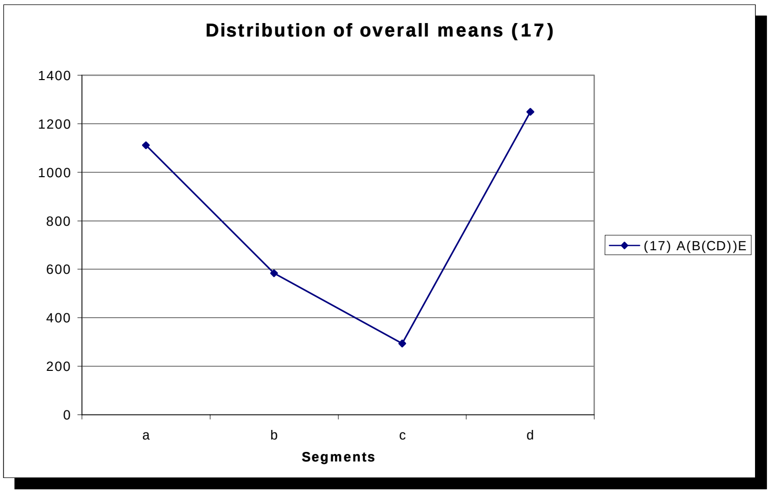

Figure 15

The figures above demonstrate the expected pausal phrasing of (16): the segments (*AB*) and (*CD*) are both short and their values fairly close, whereas the pause between B and C is represented by a significant increase of segment length (in other words: boundary strength). Although we would expect equal durations for all segments in (15), there is still a slight increase of duration found between *B* and *C* (essentially due to the 3 subjects who also performed a tonal segmentation in (16). This is no mistake, however, but it indicates that grouping is, as a default, inherently present in our perception of structure in general.

The pausal phrasing in (17) appears quite systematic: on one side of the embedding, the deeper the embedding the shorter the segment (boundary length), and on its other side, de-embedding is represented by a boundary longer than the corresponding embedding. Accordingly, embedding denoted by segment *b* between *B* and *C* is shorter than the embedding above it denoted by segment a between *A* and *B*, whereas segment *c* between *C* and *D* is the shortest with no embedding inside. Finally, the longest segment *d* represents the boundary between *D* and *E*, a boundary of de-embedding corresponding to the highest embedding between *A* and *B*.

The comparison of tonal and pausal phrasing in speech prosody and abstract prosodic grouping shows important similarities and certain differences as well. They are similar in that, first, they agree in using both tonal and pausal phrasing to represent structural relations, and, second, and most importantly, they agree in applying recursion in structure generation. They differ in the extent they use tonal and pausal phrasing. Whereas in the case of abstract prosodic patterns pausal phrasing was highly consistent with all subjects, and tonal phrasing, however strongly present, had a secondary role, in speech prosody there was a reverse relation between the two kinds of phrasing: the use of tonal phrasing was systematic in all utterances, whereas pausal phrasing was, in some cases, less articulate. We will offer an account for this difference after showing how grouping in abstract visual patterns is realized.

Hunyadi László: Cognitive grouping and recursion in prosody
Argumentum, 4 (2008), 231-253
Kossuth Egyetemi Kiadó (Debrecen)

5 Grouping in abstract visual patterns

A group of 50 university students were presented a set of patterns consisting of a series of dots like "•" arranged linearly. (For proper comparison, these dots were arranged in patterns essentially identical to those presented in the experiment on abstract prosodic grouping.) Subjects were instructed that in some patterns pairs of parentheses were used to denote various kinds of grouping. They were asked to observe the given patterns and then, following these patterns, represent them by mouse clicks. The onset time of each click was recorded on the computer and the resulted durations analyzed.

A strong consistency was found among subjects in the durational representation of all patterns; cf. (18), (19) and (20) (below, Figure 16 shows the corresponding temporal segmentation of (18) and (19) and Figure 17 that of (20), respectively:

(18) ••••
(19) (••)(••)
(20) •(•(••))•

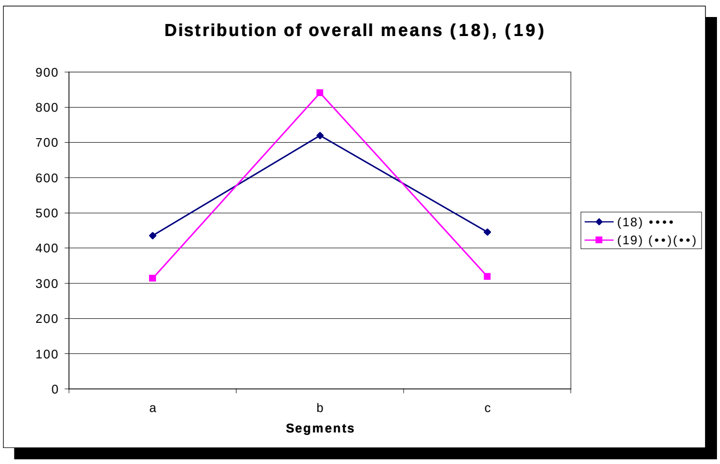

Figure 16

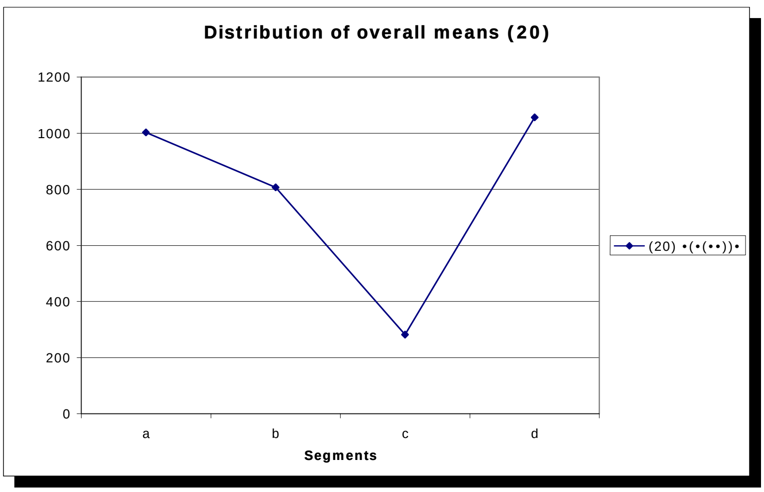

Figure 17

Hunyadi László: Cognitive grouping and recursion in prosody
Argumentum, 4 (2008), 231-253
Kossuth Egyetemi Kiadó (Debrecen)

A comparison of the above data on pausal phrasing in abstract visual grouping with those in abstract prosodic grouping show important similarities and certain differences as well. They agree in the systematic role of pausal phrasing and, most importantly, the use of recursion in the durational (pausal) representation of recursive structure. They differ in two aspects. First, pausal segmentation in visual grouping appeared to be more exact: the two short segments (boundaries) in (19) proved to be virtually identical in abstract visual grouping, more than in abstract prosodic grouping. Although we might find an explanation for it in the more precise nature of mouse clicks, it does not account for the second difference: inherent grouping observed in (18), where no overt grouping was denoted in the visual pattern, was shown more articulate than in the case of abstract prosodic grouping. These two differences may receive the following common explanation: if a modality has a single means of grouping (such as pausal phrasing in visual grouping), this means needs to be applied as closely as possible. However, if a modality has more than one means of grouping (such as tonal and pausal phrasing in abstract prosodic grouping), they can share their responsibility in denoting grouping. As a consequence, these means do not necessarily participate in denoting grouping equally. Our experiments allow us to suggest that a means more specific to the particular modality assumes a more prominent role in grouping. Accordingly, pausal phrasing, universal across the three modalities studied (speech prosody, abstract prosodic grouping and abstract visual grouping), is less specific to prosodic modalities since what makes the latter, prosodic modalities specific is their use of tones in phrasing. However, tonal phrasing was found to have more preference over pausal phrasing in speech prosody than in abstract prosodic grouping. (The specific role of tones in speech was also shown by Lickley (1994) who, examining disfluencies found that, in case pause was absent at the point of disfluency, both phoneme-level and tonal cues could still contribute to identifying disfluencies.) This finding can be related to the increased overall role of tones in speech: whereas in abstract prosodic grouping (where the use of tones is an added specificity as compared to the highly abstract nature of visual grouping) tones are only used to denote abstract grouping relations, tones in speech prosody, while preserving their role in denoting abstract grouping relations, also assume a language-specific role, that of expressing language-internal modalities. This increased specificity of tones in speech prosody is what gives tonal phrasing a stronger preference over less specific pausal phrasing than what we find in the case of abstract prosodic grouping.

6 Summary and conclusions: recursion in prosody and its cognitive status

We have shown that the computational mechanism of recursion can be identified in prosody in two forms: in tonal and pausal phrasing. As for tonal phrasing, the principle behind nested recursive prosodic phrase generation is tonal continuity. Since tonal continuity involves a change (usually the lowering) of the tonal space of the embedded phrase, applied recursively it results in the recursive unidirectional change (continuous relative lowering) of the tonal space.[9] It was shown that the computational difference between nested recursion and iteration has its prosodic correlate. Namely, recursion is characterized by the lack of downdrift,

[9] It is useful to note that such a change of the tonal space for recursive embedding appears to correspond to key change modulation in recursive musical phrase formation observed in Hofstadter (1980). We may also note that a special case of key change in music is fifth tone change, an ancient way of modulation in pentatonic music, among others in Hungarian.

Hunyadi László: Cognitive grouping and recursion in prosody
Argumentum, 4 (2008), 231-253
Kossuth Egyetemi Kiadó (Debrecen)

whereas iteration by the presence of downdrift. Upstep is also conditioned by different processes: it is a phonetic requirement for tonal continuity in recursion and a phonetic consequence of downdrift in iteration. It was also shown that the principle of tonal continuity is not restricted to prosodic recursion: it underlies any structure involving discontinuous prosodic phrases, including insertion and appears as the bookmark effect. As for pausal phrasing, it was found to be structure-dependent and, as such, suitable for denoting hierarchical phrase boundaries in recursive embedding. It was also found that iteration, being non-hierarchical, does not make use of pausal phrasing, instead, it uses a rhythmic organization of the sequence of iterated constituents based on the rhythmic (metrical) distribution of pitch accents. It was also found that tail recursion that is computationally similar to iteration has the same prosodic realization as iteration: the principle of tonal continuity does not apply to it.

The question that is of theoretical importance is whether recursion in prosody is the effect of an interface-relation between syntax and prosody or has at least some other sources for its origin. Experiments involving grouping in abstract prosodic patterns as well as grouping in abstract visual patterns demonstrate the recursive use of tonal and pausal phrasing in a fashion similar to those found in speech prosody. Their differences, especially regarding the relative preference of means of grouping across different modalities based on their modality-specificity suggest that grouping in prosody has its cognitive basis in the grouping of less specific, more abstract, nonlinguistic elements. That recursion in prosodic grouping is exactly the same mechanism as the one found in the grouping of less specific, more abstract, non-linguistic prosodic and visual elements suggests that recursion in prosody cannot be the effect of an interface relation between syntax and prosody, instead, it is the manifestation of a more general, more universal computational mechanism found beyond linguistic structure. Suggestions of the similarity of tonal continuity (the recursive handling of tonal space) to a change of key modulation as well as pentatonic fifth tone modulation) in music point in this direction.

In their reply to Pinker and Jackendoff (2005), Fitch, Hauser, and Chomsky (2005: 201) consider what consequences it would have on the theoretical and evolutionary status of the narrow faculty of language (FLN) if phonology were eventually found recursive. They ask the following questions: "is it [a recursive mechanism in phonology] the same as or different from that in phrasal syntax?", "is it a reflex of phrasal syntax perhaps modified by conditions imposed at the interface?" and "is phonological recursion the same as or different from that in musical phrases?" They conclude that "If the answer to all of these questions were "same", we would reject our hypothesis 3, possibly concluding that FLN is an empty subset of FLB, with only the integration of mechanisms being uniquely human."

Even though some of these questions cannot be answered within the confines of this article, based on the fact that recursion is identified in prosody and important properties of prosodic recursion are found in cognitive systems beyond language, the present article may have offered further arguments for a possible answer in this direction.

References

Beckman, M.E. & Pierrehumbert, J. (1986): Intonational structure in English and Japanese. *Phonology Yearbook* 3, 255-310.

Chomsky, N. & Halle, M. (1968): *The Sound Pattern of English.* New York: Harper and Row.

Clifton, Ch.Jr. & Carlson, K. & Frazier, L. (2002): Informative prosodic boundaries. *Language and Speech* 45.2, 87-114.

Cooper, W. & Paccia-Cooper, J. (1980): *Syntax and Speech.* Cambridge, MA: Harvard University Press.

Hauser, M. D., Chomsky, N. & Fitch, T. W. (2002): The faculty of language: What is it, who has it, and how did it evolve? *Science* 298, 1569-1579.

Downing, B.T. (1970): *Syntactic structure and phonological phrasing in English*. PhD. Dissertation, University of Texas.

Fitch, W.T. & Hauser, M.D. & Chomsky, N. (2005): The evolution of the language faculty: Clarifications and implications. *Cognition* 97, 179-210.

Hayes, B. (1984): The phonology of rhythm in English. *Linguistic Inquiry* 15, 3-74.

Hofstadter, D.R. (1980): *Gödel, Escher, Bach: An Eternal Golden Braid.* London: Penguin.

Hunyadi, L. (2002) *Hungarian Sentence Prosody and Universal Grammar.* Frankfurt & New York: Peter Lang.

Hunyadi, L. (2006): Grouping, the cognitive basis for recursion in language. *Argumentum* 2(2006), 67-114.

Jackendoff, R. & Pinker, S. (2005): The nature of the language faculty and its implications for evolution of language (Reply to Fitch, Hauser, and Chomsky). *Cognition* 97, 211-225.

Ladd, D.R. (1986): Intonational Phrasing: The Case for Recursive Prosodic Structure. *Phonology Yearbook* 3, 311-340.

Ladd, D.R. (1992): An introduction to intonational phonology. In: Docherty, G.J. & Ladd, D.R. (eds.): *Papers in Laboratory Phonology. Vol. II: Gesture, segment, prosody*. Cambridge: Cambridge University Press, 321-334.

Ladd, D.R. (1996): *Intonational Phonology.* Cambridge: Cambridge University Press.

Liberman, M. & Prince, A. (1977): On stress and linguistic rhythm. *Linguistic Inquiry* 8, 249-336.

Lickley, R.J. (1994): *Detecting Disfluencies in Spontaneous Speech.* PhD. Dissertation, University of Edinburgh.

Liu, Y.A. & Stoller, S.D. (1999): From recursion to iteration: What are the optimizations? *ACM Sigplan Notices* 34, 73-82.

Nespor, M. & Vogel, I. (1982): Prosodic domains of external sandhi rules. In: van der Hulst, H. & Smith, N. (eds.): *The structure of phonological representations, Vol. I.* Dordrecht: Foris, 225-255.

Parker, A.R. (2006): *Evolution as a Constraint on Theories of Syntax: The Case against Minimalism.* PhD Dissertation, University of Edinburgh.

Peperkamp, S. (1997): *Prosodic words. HIL Dissertations 34.* The Hague: Holland Academic Graphics.

Hunyadi László: Cognitive grouping and recursion in prosody
Argumentum, 4 (2008), 231-253
Kossuth Egyetemi Kiadó (Debrecen)

Pinker, S. & Jackendoff, R. (2005): The faculty of language: What's special about it? *Cognition* 95, 201-236.

Pierrehumbert, J. (1980): *The phonology and phonetics of English intonation.* PhD Dissertation. Massachusetts Institute of Technology, Cambridge.

Price, P.J., Ostendorf, M., Shattuck-Hufnagel, S. & Fong, C. (1991): The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America* 90, 2956-2970.

Selkirk, E. (1984): *Phonology and Syntax: the relation between sound and structure.* Cambridge: MIT Press.

Selkirk, E. (1995): The prosodic structure of function words. In: Beckman, J., Welsh Dickey, L. & Urbanczyk, S. (eds.): *Papers in Optimality Theory.* University of Massachusetts Occasional Papers 18. Amherst: Graduate Linguistic Student Association, 439-469.

Taglicht, J. (1998): Constraints on intonational phrasing in English. *Journal of Linguistics* 34, 181-211.

Truckenbrodt, H. (1999): On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30, 219-255.

Truckenbrodt, H. (2007): The syntax-phonology interface. In: de Lacy, P. (ed.): *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press, 435-456.

Wagner, M. (2005): *Prosody and recursion.* PhD. Dissertation, Massachusetts Institute of Technology, Cambridge.

Watson, D. & Gibson, E. (2004): The relationship between intonational phrasing and syntactic structure in language production. *Language and Cognitive Processes* 19.6, 713-755.