Underspecification restrictions in polyrhythmic processing

Stergios Chatzikyriakidis

1 Introduction

There is considerable amount of work focusing on the question of whether there is cognitive overlap (if at all) in the way language and music are processed or/and produced. An answer to this question comes from Patel [2008], who provided a thorough investigation of the experimental evidence available[1] and proposed the *shared syntactic integration resource hypothesis* (SSIRH), according to which both language and music share a common neural resource. Patel argues that even though the two systems have domain-specific representations, they also share common neural resources. Actually, the argument is that these common resources activate the domain-specific representations in syntactic processing [Patel, 2008:297]. Such an assumption is different to the traditional generative syntax view that takes language to be an encapsulated capacity, with individual modules (syntax, semantics) being isolated to each other as well. In this volume, Kempson and Orwin argue for a stronger version of the Patel hypothesis, according to which syntax is entirely domain-general.[2] I would not want, at least for the present paper, to take such a strong stance, but rather start from a weaker position, i.e. the original Patel hypothesis and see whether indeed the framework of DS has anything to offer as regards spelling out in more detail what this shared domain-general (or part of it) might look like.

Assuming that the Patel hypothesis is on the right track, then the next step would be to decide what the domain-specific and domain-general characteristics of both language and music are. One of the natural candidates for being a domain-general process would be a some kind of general mechanism for inducing structure incrementally, a mechanism that can be used in unfolding any type of information incrementally and regardless of the domain-specific mechanisms that might operate simultaneously with it. The framework of Dy-

[1] See Patel [2008] for references as well as detailed review.

[2] See Kempson and Orwin [this volume] for more information on how such a stronger thesis is envisaged and the consequences it will have for linguistic theory.

namic Syntax (DS) [Cann *et al.*, 2005, Kempson *et al.*, 2001] is an incremental, dynamic model with explicit structure building mechanisms for inducing structure, that takes parsing/processing seriously into consideration.[3] The model of DS has successfully managed to deal with a number of very problematic NL phenomena using the concept of underspecification plus update as a basic structure building mechanism. This, along with the incrementality of the model have provided us with satisfactory accounts of problematic NL aspects ranging from purely morphosyntactic phenomena (e.g. the Person Case Constraint (PCC), see Chatzikyriakidis and Kempson [2011]) to phenomena lying within the syntax/pragmatics interface (e.g. dialogue modelling, see Purver *et al.* [2010]). Phenomena like free word-order in head final languages like Japanese and Korean [Cann *et al.*, 2005, Kiaer, 2007, respectively] and elliptical syntactic structures [Kempson *et al.*, 2011] have also been tackled. Structure building mechanisms in DS comprise an explicit machinery that can be seen as being separated from the range of domain-specific NL features (e.g. formula values, semantic types).[4] The DS view of NL syntax can be then seen as the interaction between these general structure building mechanisms and the domain-specific NL vocabulary.[5] Now, if Patel is right and SSIRH is a valid hypothesis, then one would expect that the two systems, i.e. language and music, would share a common set of resources. These common resources according to Patel help in the activation and integration of the domain-specific syntactic representations (different for each system) in each case. The term 'syntactic integration' adapted from the psycholinguistic work of Gibson [Gibson, 1998], refers, in the case of language, to the linking of a given word with its predecessors and the cognitive load it carries in each case. This linking is then what is facilitated via means of common resources. In this context, if a framework like DS has any psychological reality, it should, at least to some extent, reflect this common resource pool assumption in the processing of language, music and potentially other cognitive domains. In this paper, the assumption that DS can indeed function as a framework that can reflect this domain-general vs domain-specific aspect is tested. In order to do this, I look at polyrhythmic processing/perception trying to see whether the notion of underspecification plus update, and in particular a specific restriction on underspecification that follows from the more general notion of underspecifi-

[3] In DS, the parser is roughly the grammar.

[4] Even though, it is not defined explicitly as being so. In principle however, one can distinguish between structure building mechanisms that manipulate domain-specific information (like lambda abstracts or typing information) and this information itself.

[5] This is quite an oversimplification since obviously the context has to be taken into consideration, which it indeed is in DS. In any case, the interaction between general structure building mechanisms and the NL domain-specific features comprises at least a basic part of how the DS model accounts for NL syntax.

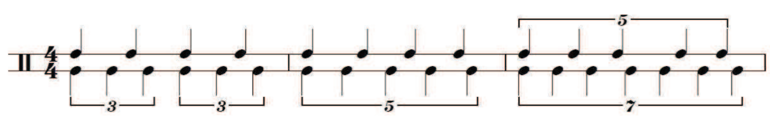
cation, is operative in parsing/perceiving polyrhythms. In this sense, I look at the experimental evidence on polyrhythmic processing and try to see whether evidence for the relevance of the restriction on underspecification, and thus underspecification plus update as a general structure building mechanism, can be found.

The paper is structured as follows: in section 2, I provide an introduction to some basic ideas in the literature on polyrhythmic processing/production and introduce briefly the DS framework. Then, in section 3, the nature of the 'one unfixed node at a time' restriction and its significance in analyzing the PCC are discussed. Lastly, in section 4, the experimental data on polyrhythmic processing are examined and a way of explaining them via the use of the 'one unfixed node at a time' restriction is proposed.

2 Polyrhythms, Polyrhythmic processing and underspecification

According to London [2003] 'a polyrhythm refers to any two or more separate rhythmic streams in the musical texture whose periodicities are noninteger multiples'.[6] Some examples of polyrhythms following this definition are shown below:[7][8]

(1) Examples of polyrhythms



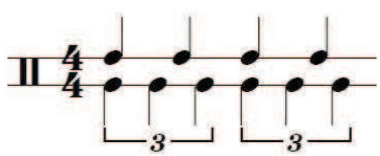
[6]Under the above definition only rhythms with a non-integer ratio are considered polyrhythms. Even though there might be objections to such a definition, we will keep this definition for the purposes of this paper.

[7]Note that the examples are not meant to indicate any pitch differences between the two rhythms. The issue of frequency separation is really important, but it is something that will not be dealt with in this paper. Thus, the examples are notated in the key for percussion, which in general does not indicate pitch.

[8]The first four bars exemplify a 3:2, 5:4, 7:5 and a 15:13 polyrhythm respectively (read as '3 against 2', '5 against 4' and so forth). The next three bars show a case of a dotted quarter note against the quarter note pulse, which is basically a 4:3 rhythm but in a 4/4 rather than in a 3/4 context. In this case, the cycle of the polyrhythm needs more than a bar to resolve. The last case (separated by an empty bar with the previous case) involves a pulse of 5 eight notes against quarter notes in a 3/4 context. This needs 5 bars to resolve.

Work on polyrhythm processing and production has shown that there are two ways polyrhythms can be perceived: a) segregation, where the two rhythms are kept separately and b) integration, where the two rhythms blend together to form one complex rhythm out of the two individual rhythms comprising the polyrhythm (see e.g. Jones [1976], Bregman [1990] and Fidali *et al.* [2011] for a summary).[9] An example of how the same polyrhythm can be perceived either as integrated or segregated is shown below. In the case of (2), the polyrhythm is heard as involving two distinct rhythms, i.e. the two individual rhythms comprising the polyrhythm do not merge into a single rhythm, while in the case of (3) the individual rhythms merge into a single complex rhythm:

(2) Segregation



(3) Integration

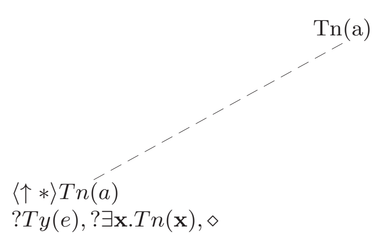
The two ways of parsing polyrhythms do not come in the form of a free option but are rather dependent on a number of factors as a number of studies suggest (see e.g. Jones [1976], Bregman [1990], Beauvillain [1983], Moelants and Noorden [2005] and Fidali *et al.* [2011] for a summary). These are a) presentation rate b) frequency separation, i.e. the pitch difference of the individual rhythms and c) attentional behaviour, i.e. whether the subjects are

[9]This seems to hold for polyrhtyhmic processing but it is not clear in the case of production. As Fidali *et al.* [2011] note, much of the experimental evidence seems to suggest that the only way to produce polyrhythms accurately is via integration. However, this is not conclusive and producing polyrhythms via segregation could be also possible.

instructed to attend to one of the two strategies. In fact, the case of segregation is a little trickier. This is because, one might expect that perception under segregation would involve the perception of the two individual rhythms as going on at the same time, thus being kept separate. However, as London [2003] notes this does not seem to be the case. What really happens, is that the subject focuses on one rhythmic stream entraining to its meter, while s/he treats the other rhythm as noise (see London [2003] for references).

In the domain of syntactic processing there is very strong evidence for incrementality in processing as well as in production [Crocker *et al.*, 2000, Ferreira, 1996, a.o.]. However, the standard assumption is that this should not be reflected in the grammar formalisms, since parsing/production are taken to be independent architectures operating on the same underlying structures provided by the grammar. The grammar formalism of Dynamic Syntax [Kempson *et al.*, 2001, Cann *et al.*, 2005, a.o.], as already mentioned, takes a different stance and assumes that parsing and production are at the core of the grammar itself. In DS, syntax is seen as the progressive accumulation of transparent semantic representations. One of the core mechanisms underlying the incremental nature of both parsing and production in DS is the notion of underspecification plus update, in effect the projection of underspecified structures that get progressively updated/enriched. Specifically, DS uses a concept of structural underspecification expressed via the use of unfixed nodes, i.e. nodes that have not yet found their position in the tree structure. For example, the node denoted with the dashed line below is an unfixed node, meaning that its treenode address is yet to be specified:

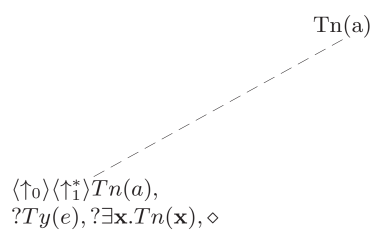
(4) The projection of an unfixed node



Every node in DS carries Formula and Type information, the former referring to its semantic content (e.g. $Fo(George)$) while the latter to its semantic type (e.g. $Ty(e)$, $Ty(e \rightarrow t)$). Furthermore, DS makes use of the concept of requirements, denoted by the question mark (?). The idea is simple and refers to statements that are not yet true but are required to hold at some point later on. Thus, $?Fo(x)$ says that at some point $Fo(x)$ must hold. In the tree above there are two requirements, one requiring a semantic type of type e ($?Ty(e)$)

and a statement requiring a proper treenode address to be found ($?\exists\mathbf{x}.Tn(\mathbf{x})$). As already said, unfixed nodes are not specified as to their treenode addresses, i.e. they involve an address which is not yet specified. In the above example, the unfixed node is not yet specified but however has a requirement that no matter what the actual address, it should be dominated by the top node with the treenode address $Tn(a)$. In simpler terms, the unfixed node here can be interpreted as "I'm somewhere below the $Tn(a)$ node". This form of structural underspecification has been used extensively in the DS literature in order to deal with various syntactic phenomena, notably Hanging Topic Left Dislocation, Clitic Doubling, relative clauses and multiple scrambling among others [Cann *et al.*, 2005, Cann *et al.*, 2004, Gregoromichelaki, 2013, inter alia]. Unfixed nodes can be of different sorts, i.e. there is more than one type of unfixed nodes possible. In particular, one might have unfixed nodes that have reduced update possibilities compared to the one we have seen in (4). One such case are locally unfixed nodes, where the unfixed node must get fixed inside the local propositional domain:

(5) The projection of a locally unfixed node



The difference between the two cases lies in the statements $\langle\uparrow *\rangle Tn(a)$ and $\langle\uparrow_0\rangle\langle\uparrow_1^*\rangle Tn(a)$ in (4) and (5) respectively. Informally the first says that "I'm somewhere below the $Tn(a)$ node" while the latter "I'm somewhere below the $Tn(a)$ node but within the local propositional domain of $Tn(a)$".[10] Now, given LOFT[11], the tree language underpinning the DS system, every node is uniquely identified via a treenode address. This means that each node has a unique treenode address. The important corollary of this fact is that no more than one unfixed node of the same type is possible at any given time at any tree structure. This is because unfixed nodes, even though lacking a fully specified treenode address, do however involve an underspecified address that will potentially get updated into a proper treenode address later on. In case two unfixed nodes of the same type are present, their treenode addresses will

[10] See Cann *et al.* [2005] for the notion of propositional domain in DS.

[11] LOFT (Logic of Finite Trees) is a modal language to talk about trees underpinning the DS framework. See Blackburn and Meyer-Viol [1994] for more information.

be exactly the same, thus they will collapse into being the same node, i.e. the two nodes will be identified as being the same node. This restriction predicts that no two underspecified structures of the same type should be possible. Indeed, a lot of work in DS has concentrated on this particular issue and has shown that this restriction can give us an account of a number of problematic NL phenomena. One of these, the PCC, will be discussed in the next section.

3 The restriction on underspecification in Natural Language

Underspecification, both structural and semantic, is a core notion within the DS framework. Structural underspecification in DS, as already said, is expressed via means of unfixed nodes, i.e. nodes that have not yet found their position in the tree structure. Dislocation phenomena (long or short distance scrambling) are prototypical cases which may be accounted for by assuming unfixed nodes. Let us see the way unfixed nodes work by examining the two cases of scrambling (short and long-distance respectively) exemplified by the Modern Greek (MG) examples shown below:

(6)	*Ton*	*Giorgo*	*ksero*
	the.ACC	George.ACC	know.1SG

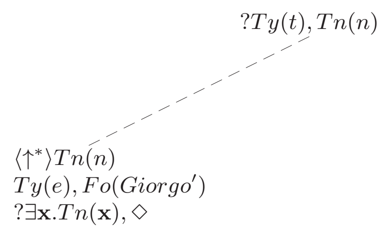
‘It is George that I know.’

(7)	*Ton*	*Giorgo*	*mu*	*ipan*	*oti*	*kseris*
	the.ACC	George.ACC	me.CL-DAT	said.3PL	that	know.2SG

‘They told me that it is George that you know.’

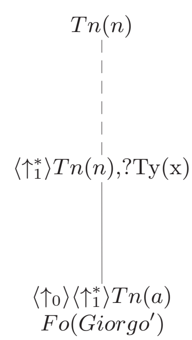
In (6) the NP is interpreted locally, while in (7) non-locally (it is the direct object of the embedded clause), appearing somehow dislocated from the place where it is supposed to be interpreted (the embedded clause). A way of accounting for both cases is to assume that the NP *ton Giorgo* is parsed on an unfixed node, given the possibility to appear in both the positions in (6) and (7). This node will be further specified when more information about the clause has been provided. The idea is that *ton Giorgo* will appear on a node with an underspecified node address as shown below:

(8) Parsing *ton Giorgo* in *ton Giorgo ton ksero* or in *ton Giorgo mu ipes oti kseris*



The underspecified treenode address $\langle\uparrow^*\rangle Tn(n)$ (which reads as 'somewhere above me there is a treenode bearing the label $Tn(n)$) is compatible with fixing of the node both in the local or in a non-local domain. This is because both these nodes satisfy the requirement that, informally put, somewhere above there is a node bearing the label $Tn(n)$.[12] In cases of short scrambling, i.e. scrambling within the local domain like (7), one might use the other unfixed node variant, i.e. locally unfixed nodes. In this case, examples like (7) can not be captured given that fixing of the node must be done inside the local propositional domain and not outside of it. Note also that locally unfixed nodes involve two nodes, one that is fixed (the $\langle\uparrow_0\rangle$ part of the treenode description statement) and one unfixed (the $\langle\uparrow_1^*\rangle$ part):

(9) Parsing *ton Giorgo* in *ton Giorgo ton ksero* using a locally unfixed node

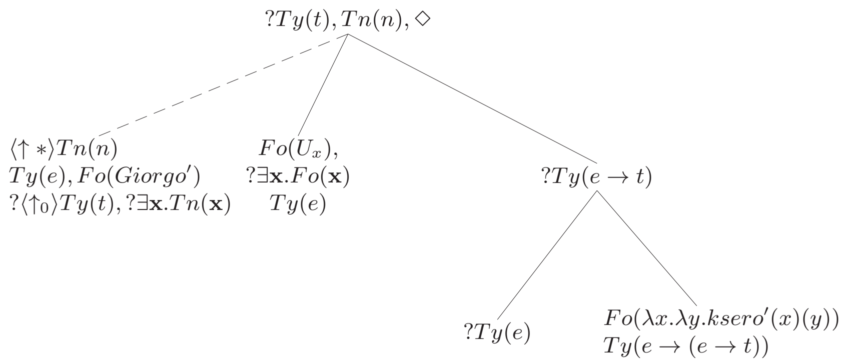


Fixing of the nodes in both cases (unfixed and locally unfixed) must be done at some point during the parsing process. For example, in (6), the NP is parsed first as an unfixed node as shown in (8). Then, the verb comes into

[12] For more details on how fixing of the unfixed node is done see Kempson *et al.* [2001], Cann *et al.* [2005] as well as Chatzikyriakidis [2010] for the specific MG examples.

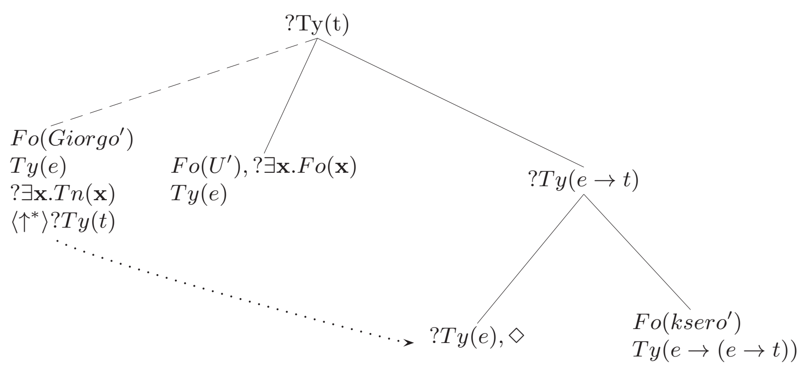
parse which unfolds the predicate argument structure. It introduces a formula and type value for the verb, builds and decorates the subject node with a type and formula metavariable[13] and further builds the direct object node adding a requirement for a type e expression to be found an this node:

(10) Parsing *ksero* 'know' in *ton Giorgo ksero* 'It is George I know'



At this point, the unfixed node can merge with the direct object node. By doing this, the unfixed node receives a proper treenode address while the direct object node a type and a formula value:[14]

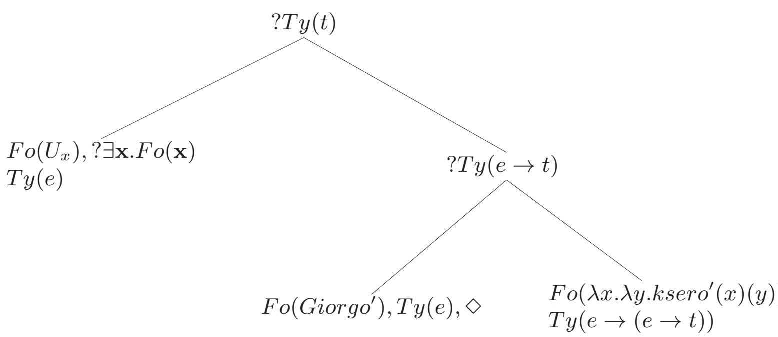
(11) Before and after MERGE of the unfixed node



[13]This takes care of the pro-drop properties of Greek. See Cann *et al.* [2005] for the notion of pro-drop in DS and Chatzikyriakidis [2010] for Greek specifically.

[14]Merging of the two nodes is done via a computational action called MERGE (see [Kempson *et al.*, 2001, Cann *et al.*, 2005] for details).

(12)



Given the way treenodes are defined in LOFT, and specifically the requirement that every treenode is identified by a unique treenode address, it is not possible to have more than one node with the same underspecified address. This is because having two such nodes will result in them collapsing into a single node via treenode identity. This can be seen as a restriction on underspecification, a hard-wired restriction of the system that does not allow more than one unfixed node of a type at a time. In what follows, the applicability of such a restriction in straightforwardly explaining the Person Case Constraint (PCC) is shown.

3.1 The Person Case Constraint

The PCC is a clitic co-occurrence restriction, which states, in its "strong PCC" variant, that a dative clitic cannot co-occur with a 1st/2nd person accusative clitic. The restriction is found across a remarkable number of both related and unrelated languages, from Romance and Greek to Kiowa and Basque (see Rezac [2008] for Basque data and Adger and Harbour [2006] for Kiowa). The Spanish and Standard Modern Greek (SMG) data are illustrative:[15]

(13)	**Le*	*me*	*ha*	*dado*
	it.CL-DAT	me.CL	has	given

'S/he has given me to him.' [Spanish]

(14)	**Mu*	*se*	*exi*	*dosei*
	me.CL-DAT	you.CL-ACC	has	given

'He/She/It has given you to me.' [SMG]

[15] There are other variants of the constraint, e.g. the weak version of the constraint [Bianchi, 2006, Bonet, 2008]) or the version exhibited by Pontic Greek [Chatzikyriakidis and Kempson, 2011].

The PCC has been puzzling syntacticians for quite a while and a number of accounts have been proposed in the literature in different frameworks.[16] It is interesting to note that the restriction is definitely not due to semantic reasons, since it is well-known that by substituting one of the two clitics of the illicit cluster with a strong pronoun grammaticality is restored, a phenomenon noted in the literature as a repair [Bonet, 2008, Rezac, 2008, a.o.]:

(15)

Me	*sistisan*	*se*	*sena*
me.CL-ACC	introduced	to	you.ACC

‘They introduced you to me.’ [Greek]

(16)

Je	*t’*	*ai*	*presenté*	*à*	*lui*
I	you.CL	have	introduced	to	him

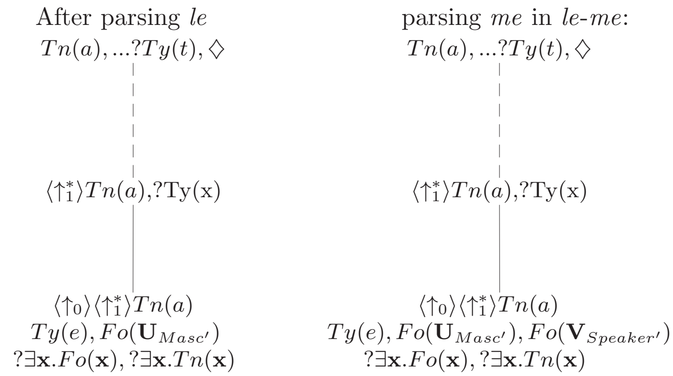
‘I introduced you to him.’ [French]

Kempson and Cann [2008] have argued that this idiosyncratic morphosyntactic restriction can be straightforwardly explained given the ‘one unfixed node at a time’ restriction. The idea is fairly simple. In a PCC language like e.g. Spanish or Italian, dative clitics as well as 1st/2nd clitics can be seen as projecting locally unfixed nodes, i.e. their lexical entries will specify the projection of such nodes. This underspecification will capture the fact that both dative and 1st/2nd person clitics in these languages can be structurally underspecified. For example, datives can be arguments or different types of datives (ethical, possessives, benefactives/malefactives etc.) and 1st/2nd person clitics can be both direct and indirect objects[17] (plus all the different datives). On the other hand, 3rd person accusative clitics are not structurally underspecified and as such project fixed structure. These assumptions in combination with the restriction on underspecification already discussed, will give us a straightforward solution to the PCC. Remember that the illicit combinations involve a dative clitic plus a 1st/2nd person clitic. Notice however that both these cases are analyzed as projecting an unfixed node. If this is the case, then these clitic combinations will be ruled out via the ‘one unfixed node at a time’ restriction. Parsing the illicit cluster *le me* in Spanish will involve the projection of two locally unfixed nodes. However, these two will collapse into being the same node via treenode identity. This node will now have conflicting formula values (one projected from *le* and one from *me*):

[16] It is not my intention to go through these proposals in this paper but the interested reader is directed to Anagnostopoulou, [2003, 2005], Béjar and Rezac [2003], Bianchi [2006], Adger and Harbour [2006], Nevins [2007] among others for a minimalist approach to the problem as well as to Bonet [1995] and Heap [2005] among others for a templatic one. See Chatzikyriakidis [2010] for an overview.

[17] This is due to syncretism, i.e. one form corresponding to both the dative and the accusative function.

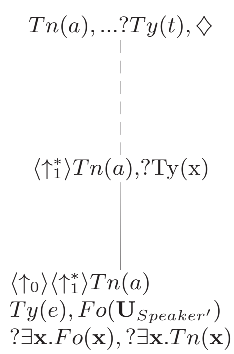
(17) Parsing the illicit cluster *le me* in Spanish



It has to be noted at this point, that if one wants to be precise, the actual reason for the ungrammaticality is not the 'one unfixed node at a time' restriction per se, but the fact that due to this restriction, we end up with one node with conflicting formula values. This might provide useful when we will discuss the relevance of the restriction in polyrhythmic processing.

In the case of licit clusters, only one of the nodes projected by the two clitics will be unfixed. For example in the case of the licit cluster *me lo*, *me* is first parsed, projects an unfixed node and decorates this node with a formula metavariable and a type value. The metavariable needs to be later on substituted from context or from the NL string itself.[18]

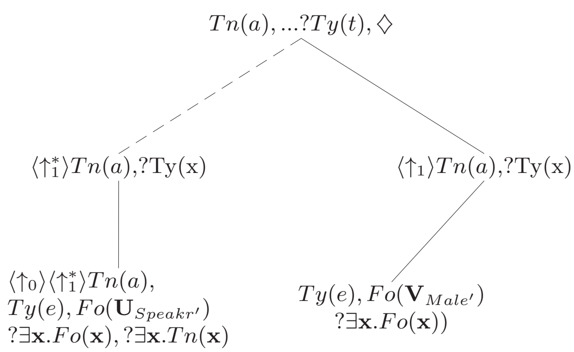
(18) Parsing *me* in *me lo*



[18]This is the DS analysis of pronouns. These are assumed to decorate the node with a metavariable to be later on substituted. For more details see Cann *et al.* [2005].

Then, the accusative clitic (*lo*) comes into play, which projects fixed structure and as such, only one unfixed node exists in the structure (the one projected by *me*):

(19) Parsing *lo* in *me lo*



Thus, such constructions are predicted to be well-formed according to fact. There are a number of further details as regards the PCC, notably the case of Standard Modern Greek (SMG) where syncretism is not present in the morphological forms of 1st/2nd person singular clitics (i.e. they involve distinct morphological forms for dative and accusative), as well as different variants of the PCC like e.g. what has been dubbed as the weak PCC version, but I will not dwell on these in this paper. The interested reader is directed to Chatzikyriakidis [2010] and Chatzikyriakidis and Kempson [2011] for a full-blown complete account of the PCC. It is however interesting to note that the whole idea in the minimalist accounts on the PCC is that the PCC is about person. However, evidence from Pontic Greek (PG) has shown that PCC effects can be found in clusters of 3rd person clitics which under most (if not all) minimalist accounts should not be the case. The relevant data from PG are shown below:

(20) **Edek aton ato/a*
gave.GAVE him.CL it/these.CL
'I gave it to him' [Chatzikyriakidis, 2010]

(21) **Edek ats ato/a*
gave.GAVE them.CL it/these.CL
'I gave it to them' [Chatzikyriakidis, 2010]

PG clitics are syncretized across the board and as such 3rd person clitics will be taken to project unfixed nodes as well according to the DS story. Thus, the data in (20) and (21) are naturally precluded according to fact. There are some further details in the case of PG as well, but for space reasons I will not discuss them, redirecting the interested reader to Chatzikyriakidis and Kempson [2011]

4 Underspecification restrictions in polyrhythm processing

4.1 What counts as an underspecified rhythm?

The first question naturally arising in a context discussing underspecification in polyrhythmic processing, is what counts as an underspecified rhythm, in effect how is an underspecified rhythm defined. With the term 'underspecified rhythm' I refer to a rhythm that has not yet been defined in relation to a metric context or in relation to another rhythm that has received its interpretation in reference to a metric context. This does not imply that there is no meter involved, but that the meter might not have been decided yet by the listener.[19]In these cases, resolving the underspecified rhythm and assigning a definite interpretation to it, gives rise to the sort of metre involved given that the interpretation will have to be defined in relation to this metric context. To illustrate this, let us assume an empty metric context (in the sense that the metric context, i.e. metre, has not yet been defined or identified by the listener) and given that no harmony, melody or any other kind of musical layer is salient in the musical context (however this is defined). In such an environment imagine a sequence of isochronous rhythmic pulses with an inherent accent every three of them. This rhythm can be characterized as underspecified, since it can receive a number of interpretations. Some of them are shown below:[20]

[19]One might very well argue that there are musical traditions where meter is not present, like e.g. Javanese music or Persian chant music. This is certainly true [Clayton, 1996]. These traditions involve rhythmic structure but this structure does not show any kind of periodicity. However, periodicity is a necessary condition for polyrhythms. As such, in cases of polyrhythms, some sort of meter needs to be established. Free rhythm, i.e. rhythm with no metre, is an issue that needs to be taken into consideration when providing an account of music, but it is out of the scope of the current paper.

[20]These are only some of the interpretations that might arise. Of course, depending on the hearer's musical and social background, some of these might be highly improbable while some others not presented here more probable. This is just a list of the parsing possibilities for the given underspecified rhythm.

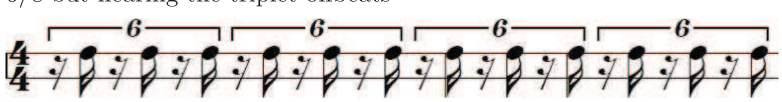
(22) Triplet interpretation in 4/4

(23) 6/8 interpretation

(24) Sixteenth note with three note groupings in 4/4

(25) 6/8 but hearing the offbeats

(26) 6/8 but hearing the triplet offbeats



(27) Sixteenth notes with three note groupings in 3/4

In (22), the rhythm under consideration receives a triplet interpretation given the 4/4 metric context. In (23), an eighth note interpretation in a 6/8 context is given, while (24) receives a sixteenth note interpetation with accents every three beats against a 4/4 context. Examples (25) and (26) receive off-beat interpretations and lastly, in (27), a sixteenth note interpretation with an accent every three sixteenth notes is given but this time in a 3/4 context. Of course, some of the interpretations are more probable than others and usually this sort of underspecification is easily resolved by the listener choosing one of the most salient options.[21] Resolving the underspecification too early can be

[21] For a western listener this would be options (22) and (23) most probably.

successful in most cases, but in other cases it can lead to a re-evaluation of the rhythm's interpretation once the metric context has been established contrary to the listener's initial assumptions. For example, imagine a case where the listener is given a three beat rhythmic unit and assigns it the interpretation in (22). Then, the metric context is established, which suggests the interpretation given in (24). In this case, the listener is forced to re-interpret the rhythm, given that otherwise s/he is just unable to parse the rhythm under the new metric context.[22] It is not my intention to check how and in what way listeners resolve rhythm underspecification but to rather claim that underspecification is always possible in cases of an empty metric context. To recap, rhythm underspecification can arise in cases where no established metric context or any other part of the rhythmic or musical context in general does not provide the listener with the information needed to assign a definite interpretation to a given rhythmic phrase/unit. In these case, a number of interpretations are possible, and as such the rhythm can present itself as underspecified until further resolved.

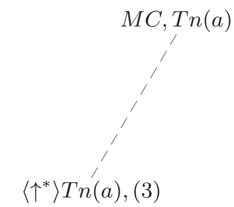
4.2 Accounting for the experimental evidence

As already mentioned, the experimental evidence as regards polyrhythms suggests two ways of polyrhythmic processing and production: a) integration and b) segregation. If in principle polyrhythms can be presented out of the blue in a context where no metric context has been established, then these polyrhythms can potentially be underspecified. If furthermore, the 'no more than one unfixed node at a time' restriction is domain-general, one should expect such restriction to be relevant to polyrhythm processing as well. In this connection, the first thing to be checked is the compatibility of the experimental evidence with the presumed existence of the restriction in polyrhythmic processing. The question thus is how these two ways of polyrhythmic processing fit (if at all) the story of the 'one unfixed node at a time' restriction. Let us first try to visualize the case where a three note rhythmic unit with an inherent accent every three notes comes into play. Given that no established metric context exists yet, this rhythm is potentially underspecified:[23]

[22] An interesting such case concerning the song *murder by numbers* by the British rock band *The Police* is reported in Vazan and Schober [2004]. The song starts with what might seem as a 3/4 beat and gradually shifts into 6/8 (or 2/4), forcing in this way listeners to re-interpret their initial assumptions about the metre. This song is a good case of using metric ambiguity to produce tension in a song, i.e. for aesthetic reasons. The paper studies the behaviour of listeners as regards detection and resolution of the song's metrical ambiguity over five hearings.

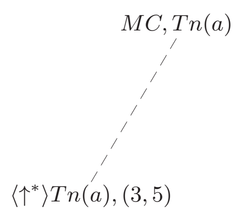
[23] MC stands for metric context. Note that the domain-specific vocabulary needed for the analysis of rhythm and music in general has not yet been decided. Actually, there is no work of this kind in DS. I tried to have as few decorations of this kind as possible here. I just note the individual rhythms with the number corresponding to the number of pulses. This

(28) Parsing the three note unit as an unfixed node



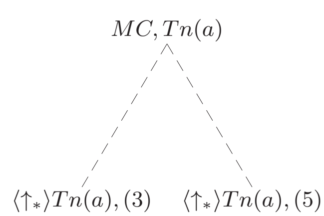
The question is what happens when a second rhythmic unit, let us say a five note rhythmic unit temporally equal to the three note one is present, i.e. in effect a case of a 5:3 polyrhythm. In the case of integration, the two rhythms are heard as one composite rhythm, so in this case what we get is a composite rhythm in the exact same node:

(29) The case of integration



In the case of segregation, and given the restriction on underspecified structures of the same type, one would expect such cases not to be parsable given the restriction.[24] In effect, what one would end up with would be the following structure:

(30) An illicit structure with two unfixed nodes



issue is at the core of providing a DS account of music but this cannot be done here.

[24] Again this is based on the assumption that the two rhythms are underspecified.

The above two nodes will then collapse into one via treenode identity if the restriction is active. However, the experimental evidence preclude a case like the above, since, as we have already seen, they suggest that in the case of segregation the subject focuses on one rhythmic stream entraining to its meter while treating the other rhythm as noise rather than parsing/processing both streams independently [London, 2003]. This is compatible with the idea of the 'one unfixed at a time' restriction, given that in this case only one of the two unfixed nodes is present. Whether the impossibility of parsing the two rhythms independently that the experimental evidence suggests is (partially) due to the 'unfixed node at a time' restriction is something that cannot be decided on the basis of the existing data alone.

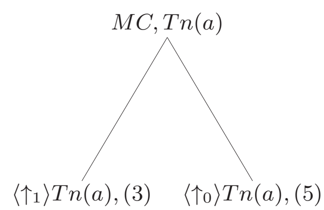
What is rather difficult as regards rhythm processing in comparison to the cases where the restriction seems to be operative in NL, is the way in which one traces violation of the restriction. In NL violation of the restriction leads to ungrammaticality. Even though the notion of ungrammaticality is far from clear cut, it is however a means that can be used (at least in some cases) as a diagnostic when an assumed restriction seems to have been violated in NL. Indeed in cases like the PCC violations, the grammaticality judgements are so clear-cut across all speakers that they cannot really be doubted.[25] On the other hand, there is no notion of grammaticality (and thus of ungrammaticality as well) in the case of music. Thus, the question arising is what counts as a violation of the restriction on underspecification in the case of polyrhythmic processing. The experimental evidence might be compatible with the restriction, however things are not that straightforward as they are in the case of language. The first problem has to do with the fact that DS is a fully-fledged grammatical framework, designed for NL. There is no analogue of a similar system designed for music (or for modules of music as such, i.e. rhythm, harmony and so on) and as such there are a number of basic things that need to be taken into consideration in giving an account of polyrhythms that have not been looked upon yet.[26] Another issue, which is also quite problematic, is the one already alluded to: testing for positive evidence. Again, as we have said, the experimental evidence is very well compatible with the account proposed but there are issues that are far from settled. For example, in the case of segregation, one can assume that the listener always resolves the underspecification by assigning the most salient metric context right from the start. If this is true, one might argue that there are no unfixed nodes at play and thus, the

[25] There are PCC constructions in some languages exhibiting variants of the constraint that do present variability as regards grammaticality judgments but in principle, in languages exhibiting the strong version of the PCC constraint, judgements seem to be uniform across all speakers.

[26] These include questions as to how rhythm is going to be represented in the system, what is to be represented etc.

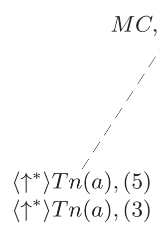
restriction should not be operative anymore. If again this is true, one would expect that processing of more than one rhythm at a time would be possible. In effect, if fixing of the nodes is possible, we would expect the following to be possible as well:[27]

(31) The two individual rhythmic units as receiving a fixed interpretation



However, one must keep in mind that the two rhythms are underspecified when first heard, even in case the listener assigns a metric context early on. In this sense, there is no way we can end up with a structure like the one shown in (31). This is because the 'one unfixed node at a time' restriction will not allow the two nodes to be present by the time fixing of these (via establishing a metric context) is done. When fixing of the nodes is about to take place, the two nodes will already have collapsed into one node as a result of the restriction on underspecification:

(32) Structure after collapse of the two unfixed nodes into one

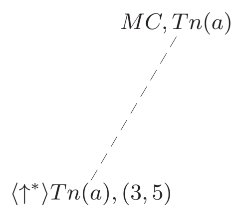


Note that in the case of NL, is that two conflicting formula values that are the result of two individual nodes having collapsed into one are not allowed. It is useful to think whether this is so in the case of polyrhythms as well, or a composite rhythm can be formed out of the two individual rhythms. In

[27]The positions in the tree are arbitrary. Where in the actual tree as well as how the actual trees will look like in the case of music is something that is not yet resolved.

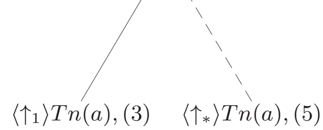
effect, if the latter is assumed, then the prediction is that the two rhythms must be parsed as integrated, given that they appear in the same treenode. In the example above, this would mean that the two rhythmic units comprising the polyrhythm will end up as being one, in effect a similar result to that of integration:

(33) Collapse into one node and integration of the two rhythms



This issue needs to be further discussed, but for the moment and given that the domain-specific DS vocabulary for music is pending, this cannot be done here. It has to be mentioned however, that whatever the decision we make, segregated structures where two underspecified rhythms are heard as being separate, are not possible. This is again in accordance with the experimental evidence that dictate that in the case of segregation there is perception of one of the two rhythms. The other option is for the rhythm to be heard as one composite rhythm, thus integrated. In effect, it must be kept in mind that the 'one unfixed node at a time' restriction per se does not cause the parser to fail. This is due to the fact that collapsing the two nodes into one will create a situation which is highly probable to include contradictory values of various sorts. Thus, it remains a possibility that structures where the restriction applies but however lead to successful parses might exist in other domains rather than language.

However, what the present account would undoubtedly allow, is the case where one of the rhythms is presented on its own (no polyrhythm), the listener assigns a metric context to it, and then the second rhythms is presented, in effect ending up with a structure where only the second rhythm is underspecified:



However, and as far as I know, there are no experimental evidence for such cases, and as such at the present the assumptions made in this paper cannot be tested against such a situation.

4.3 The need for experimental evidence

This last observation in the previous section leads us to a more general issue that has been sketched already in this paper: we need new experimental evidence that will help us determine whether such a constraint is operative or in any way useful in understanding polyrhythmic processing/perception and production. The experimental evidence might be compatible with the assumptions made in this paper, however one must be cautious in concluding that any kind of restriction like the 'one unfixed at a time' restriction is operative in polyrhythm processing. This is because being compatible with the existing data and the actual presence of such a restriction are two separate issues and the first does not necessarily entail the second. This is not strange in itself, given that the experiments presented in the literature on polyrhythmic processing/production were not designed to test any of the assumptions in this paper and as such they are not "looking" for any kind of evidence (whether positive or negative) on the existence of the restriction. Thus, our preliminary conclusion is that if one wants to really test whether the assumptions as presented in this paper can be in any way useful in polyrhythmic processing/production, experiments designed for this purpose should be designed. What such experiments might look like is another issue that cannot be discussed here. However, some things should be obvious as regards the content of these potential future experiments. For example, one test could concentrate on presenting one of the rhythms to the listener first, followed by the presentation of the second one, rather than presenting the two rhythms comprising the polyrhythm simultaneously right from the start. This type of an experiment will allow the listener, if the restriction on underspecification has any relevance whatsoever, to fix the interpretation of the first rhythm prior to the presentation of the second one. Such tests will prove to be useful in deciding whether some form of the restriction is operative. Furthermore, cases of polyrhythms with more than two rhythms involved should be examined. Is it again the case that segregation would involve parsing of only one rhythm or do we get different results in this case? Is integration possible when the number of rhythms involved is more than two?[28] These questions and many others can however be answered once specific such experiments are designed. Thus, to conclude, it needs to be stressed once more that there is a need for experimental evidence to further

[28] Work on polyrhythms comprised of more than two individual rhythms have been studied by Handel [1984]. The conclusions seem to be compatible with segregation working in the same way when more than two rhythms comprise the polyrhythm, i.e. one is rhythm is extracted, the rest are heard as noise.

back the claims made in this paper.

5 Conclusion

In this paper I looked at the issue of polyrhythmic processing/production under the assumption that structural underspecification is a domain general rather than a domain specific restriction. To this end, I looked at one of the most common structure building mechanisms, that of underspecification plus update. Specifically, I looked for evidence that the 'more than one unfixed node at a time' restriction, a corollary of structural underspecification, is relevant for other domains as well by looking at polyrhythmic processing. What was found, is that such a restriction is very well compatible with the existing data on polyrhythmic processing/production and potentially offers a new way of explaining the experimental data. However, this latter fact does not guarantee the existence of such a restriction. For this reason, it would be wise to have more (new) experimental data specifically designed to test whether underspecification in polyrhythmic processing is relevant or not, in order to be conclusive on the issue.

BIBLIOGRAPHY

[Adger and Harbour, 2006] D. Adger and D. Harbour. The person case constraint. *Syntax*, 4, 2006.

[Anagnostopoulou, 2003] E. Anagnostopoulou. *The Syntax of Ditransitives: Evidence from Clitics.* Mouton de Gruyter, 2003.

[Anagnostopoulou, 2005] E. Anagnostopoulou. Strong and weak person restrictions: a feature checking analysis. In L. Heggie and F. Ordonez, editors, *Clitics and affixation*, pages 199–235. Benjamins, 2005.

[Beauvillain, 1983] C. Beauvillain. Auditory perception of dissonant polyrhythms. *Attention, Perception and Psychophysics*, 34(6):585–592, 1983.

[Béjar and Rezac, 2003] S. Béjar and M. Rezac. *Person licensing and the derivation of PCC effects.*, pages 49–62. John Benjamins., 2003.

[Bianchi, 2006] V. Bianchi. On the syntax of personal arguments. *Lingua*, 116:2023–2067, 2006.

[Blackburn and Meyer-Viol, 1994] Patrick Blackburn and Wilfried Meyer-Viol. Linguistics, logic and finite trees. *Logic Journal of the Interest Group of Pure and Applied Logics*, 2(1):3–29, 1994.

[Bonet, 1995] E. Bonet. Feature structure of Romance clitics. *Natural Language and Linguistic Theory*, 13:607–647, 1995.

[Bonet, 2008] E. Bonet. pages 103–128. Mouton De Gruyter, 2008.

[Bregman, 1990] A.S Bregman. *Auditory scene analysis: the perceptual organization of sound.* Cambridge: MIT Press., 1990.

[Cann *et al.*, 2004] Ronnie Cann, Lutz Marten, Masayuki Otsuka, and David Swinburne. On the left and on the right. In D. Adger, C. De Cat, and G. Tsoulas, editors, *Peripheries*, pages 19–47. Kluwer, Dordrecht, 2004.

[Cann *et al.*, 2005] Ronnie Cann, Tami Kaplan, and Ruth Kempson. Data at the grammar-pragmatics interface: the case of resumptive pronouns in English. *Lingua*, 115(11):1475–1665, 2005. Special Issue: On the Nature of Linguistic Data.

[Chatzikyriakidis and Kempson, 2011] Stergios Chatzikyriakidis and Ruth Kempson. Standard modern and pontic greek person restrictions: A feature-free dynamic account. *Journal of Greek Lingusitics*, pages 127–166, 2011.

[Chatzikyriakidis, 2010] Stergios Chatzikyriakidis. *Clitics in four Dialects of Modern Greek: A Dynamic Account*. PhD thesis, King's College, London, 2010.

[Clayton, 1996] R.L. Clayton, Martin. Free rhythm: ethnomusicology and the study of music without metre. *Bulletin of the School of Oriental and African Studies*, 59(2):322–332, 1996.

[Crocker *et al.*, 2000] Matthew Crocker, Martin Pickering, and Charles Clifton, editors. *Architectures and Mechanisms in Sentence Comprehension*. Cambridge University Press, 2000.

[Ferreira, 1996] Victor Ferreira. Is it better to give than to donate? Syntactic flexibility in language production. *Journal of Memory and Language*, 35:724–755, 1996.

[Fidali *et al.*, 2011] B.C. Fidali, E. Poudrier, and B.H. Repp. Detecting perturbations in polyrhythms: effects of complexity and attentional strategies. *Psychological Reserach*, 77(2):183–195, 2011.

[Gibson, 1998] E. Gibson. Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 78:1–76, 1998.

[Gregoromichelaki, 2013] Eleni Gregoromichelaki. A dynamic perspective on left-right asymmetries: Clitic left dislocation and clitic doubling in greek. In *Rightward Movement in a Comparative Perspective*. John Benjamins, 2013.

[Handel, 1984] S. Handel. Using polyrhythms to study rhythm. *Music Perception*, 1:465–484, 1984.

[Heap, 2005] D. Heap. Constraining optimality: Clitic sequences and feature geometry. In L. Heggie and F. Ordonez, editors, *Perspectives on Clitic and Agreement Affix Combinations*, pages 81– 102. John Benjamins, 2005.

[Jones, 1976] M.R Jones. Time, our lost dimension: toward a new theory of perception, attention, and memory. *Psychological Review,*, 83:323–355, 1976.

[Kempson and Cann, 2008] Ronnie Kempson and Ruth Cann. Production pressures, syntactic change and the emergence of clitic pronouns. In R. Cooper and R. Kempson, editors, *Language in Flux: Dialogue Coordination, Language Variation, Change and Evolution*, pages 179–220. College Publications, London, 2008.

[Kempson and Orwin, forthcoming] R. Kempson and M. Orwin. Language and music as mechanisms for interaction. Communication, Mind and Language. College publications, forthcoming.

[Kempson *et al.*, 2001] Ruth Kempson, Wilfried Meyer-Viol, and Dov Gabbay. *Dynamic Syntax: The Flow of Language Understanding*. Blackwell, 2001.

[Kempson *et al.*, 2011] Ruth Kempson, Eleni Gregoromichelaki, Wifried Meyer-Viol, Matthew Purver, Graham White, and Ronnie Cann. Natural-language syntax as procedures for interpretation: the dynamics of ellipsis construal. In A. Lecomte and S. Tronçon, editors, *Ludics, Dialogue and Interaction*, number 6505 in Lecture Notes in Computer Science, pages 114–133. Springer-Verlag, Berlin/Heidelberg, 2011.

[Kiaer, 2007] J. Kiaer. *Processing and Interfaces in Syntactic Theory: The case of Korean*. Phd, KCL London, 2007.

[London, 2003] J. London. *Hearing in Time: Psychological Aspects of Musical Meter*. OUP USA, 2003.

[Moelants and Noorden, 2005] D. Moelants and L. Noorden. The inxuence of pitch interval on the perception of polyrhythms. *Music Perception*, 22:425–440, 2005.

[Nevins, 2007] A. Nevins. The representation of third person and its consequences for person-case effects. *Natural Language and Linguistic Theory*, 25:273–313, 2007.

[Patel, 2008] A.D. Patel. *Music, Language, and the Brain*. Oxford University Press, 2008.

[Purver *et al.*, 2010] Matthew Purver, Eleni Gregoromichelaki, Wilfried Meyer-Viol, and Ronnie Cann. Splitting the 'I's and crossing the 'You's: Context, speech acts and grammar. In P. Łupkowski and M. Purver, editors, *Aspects of Semantics and Pragmatics of*

Dialogue. SemDial 2010, 14th Workshop on the Semantics and Pragmatics of Dialogue, pages 43–50, Poznań, June 2010. Polish Society for Cognitive Science.

[Rezac, 2008] M Rezac. The syntax of eccentric agreement: The person case constraint and absolutive displacement in basque. *Natural Language and Linguistic Theory*, pages xx–yy, 2008.

[Vazan and Schober, 2004] P. Vazan and M. Schober. Detecting and resolving metrical ambiguity in a rock song upon multiple rehearings. In *Proceedings of ICMPC8*, pages 426–432, 2004.

Stergios Chatzikyriakidis
Royal Holloway University of London, UK.
Open University of Cyprus
stergios.chatzikyriakidis@cs.rhul.ac.uk