Generating expressive timing by combining rhythmic categories and Lindenmayer systems

Carlos Vaquero Patricio [1,2] and **Henkjan Honing** [1,2]

Abstract. This paper introduces a novel approach for modeling expressive timing performance by combining cognitive, symbolic and graphic representations of rhythm spaces with Lindenmayer systems, originally conceived to model the evolution of biological cell structures and plant growth. Logo-turtle abstractions are proposed in order to generate expressive rhythmic performances defined by rule-based displacements through perceptual rhythm categories.

1 INTRODUCTION

In music performance, aspects such as rhythm, pitch, loudness, timbre and memory contribute to our perception of expressiveness [28]. However, the expressive parameters and variables used may vary from one performer to another even within the same piece [2], which is a common cause of disagreement when listeners compare judgments of expressiveness [27]. *Can we then find an intrinsic definition of expressiveness? i.e.* without references to an external score. Are there perceptual constraints on expressiveness? And if so, would it be possible to use them to model performance?

Within the abundant literature on music performance modeling [35] different approaches can be found when defining expressiveness [18]. Davies [7] defines it as the emotional qualities of music perceived by the listener. London [22] identifies the amount of expressiveness the listener expects from the performer. Alternatively, Clarke [6] instead approaches it through the deviations of the performance from notated score durations in the score. And, in contrast, Desain and Honing [8] define expression in terms of a performance and its structural description, *i.e.*, an attempt to define expression intrinsically, independent of a score [11].

For the purpose of the current study we will define expressiveness as *the deviation from the most frequently heard version of a constituent musical element*. This is a reformulation based on the intrinsic definition of expressiveness that was mentioned before.

Previous research [16, 28] shows that even when listeners require no explicit training to perceive expressive timing, memory [14, 33] and expectation [17] play a fundamental role when recognising nuances in music timing [15]. We can therefore hypothesise that the range of expectations and uncertainty in music will be partially determined by our previous exposure to it.

Understanding how our expectations to expressiveness work is a relevant aspect to model the relation between a listener and the music material the listener is exposed to. By studying this process we can find out whether certain domains of expressiveness such as timing can be categorised and following our previous definition, model *how expressive* music could sound to a listener. Using this knowledge we may be able to generate automatic expressive performances to be recognised by the listeners as such.

An example of this creative approach to the expectations of listeners can be found in the way instrumentalists use *ritardandi*. According to Honing [13] performers make use of non-linear models to convey expressiveness using different sorts of *ritardandi*. Non-linearity allows that a player may perform the same music piece differently each time, instead of repeating the same expressive "formula" on each of the performances. These non-linear models can be seen then as a communicative resource to refer to a listener's memory and expectations but also as a way of producing slight deviations adding a certain degree of novelty within the listening and performance experience.

When defining a model of expressive performance we must reckon incorporating within the model the possibility to produce non linear variations within the deviations, defined by the perceptual constraints. This versatility in expressive productions of the model is necessary not only to attend the non linearity in performance but also to respond to our relation to expectancy and uncertainty as listeners. As an approach to model expressive performances within different rhythm patterns and mental representations we propose combining symbolic and graphic representations of rhythm spaces with Lindenmayer systems and logo-turtle abstractions. The model proposed can be used as an exploratory tool of expressive timing for computational creativity music generation.

This paper is divided in the following sections: In §1 an approach to understanding expressiveness as deviations within different perceptual categories is introduced. In §2 a study done by Desain and Honing [10] to collect empirical data on the formation of the rhythmic categories is presented. In §3 a review on Lindenmayer systems and how they can be approached within music applications is introduced. §4 connects the material presented in §2 and §3 proposes a preliminary implementation of the system. In §5 a summary of the previous sections and relevance of this approach is given.

2 RHYTHM CATEGORIES AND EXPRESSIVE TIMING

As explained in §1, factors such as music exposure and music predisposition contribute to our perception of rhythm and consequently in how this affects our relation to musical expressiveness.

In the domain of rhythm, expressive timing is defined by the deviations, or nuances, that a performer may introduce in contrast to a metronomic interpretation of a rhythm. The *ability of listeners to distill a discrete, symbolic rhythmic pattern from a series of continuous*

---

[1] Institute for Logic, Language and Computation, University of Amsterdam, email: c.vaqueropatricio@uva.nl

[2] Amsterdam Brain and Cognition, University of Amsterdam

intervals" [15] requires understanding how the perception of rhythm occurs.

Rhythmic perceptual categories then can be understood as mental clumps by which listeners can mentally relate expressive timing to a rhythmic pattern after having effectively detected it [15]; *e.g.*, the rhythmic pattern that would be symbolically transcribed while doing a music dictation. Fig. 1 shows the process of categorising a possible sequence of expressive timing events into a symbolic representation (perception) and a possible production, or interpretation of the symbolic material while performing it (production). Different interpretations, or performance renditions, of the symbolic representation (musical score) are possible depending on the performer aesthetics, experience and motor skills with their instruments [3].
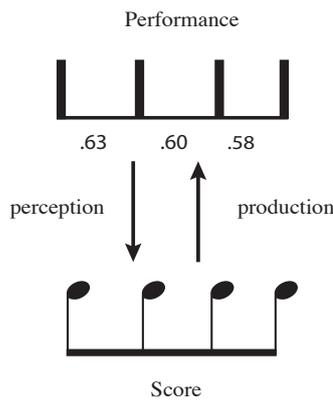

Figure 1. Difference between the perception as a symbolic representation and the production of it within a performance. Adapted from Honing (2013) [15].

Categorization has been studied extensively using behavioral and perceptual experiments [5, 10]. These aimed to answer how a continuous domain such as time is perceived and categorized, as well as represented symbolically in music notation. The two main hypotheses can be resumed in the studies done by Clarke [5] and Desain and Honing [10].

Clarke [5] did two experiments to prove the hypothesis that listeners judge deviations as an element out of the categorical domain. From these experiments it was concluded that rhythm is not perceived on a continuous scale but as rhythmic categories that function as a reference relative to which the deviations in timing can be appreciated.

Desain and Honing [10] did an empirical study using a large set of temporal patterns as stimuli to musically trained participants. By giving an identification task, rhythmic categories were collected through a perceptual experiment in which rhythm on a continuous scale (see example in the top panel op Fig. 1) had to be notated in music notation (see example in bottom panel of Fig. 1). Thus, the participants would have to notate what they heard and guess what would be written in the score a drummer playing that sequence would have in front of him. Repeating this process with every possible combination of four onset rhythms of a one second duration, the authors were able to sample and obtain the perceived rhythmic categories from the whole rhythm space.

Fig. 2 shows two sample rhythms and their location in a chronotopological map or rhythm chart. Each of the sides of the triangle represent an inter-onset interval in a rhythm of four onsets. Fig. 3
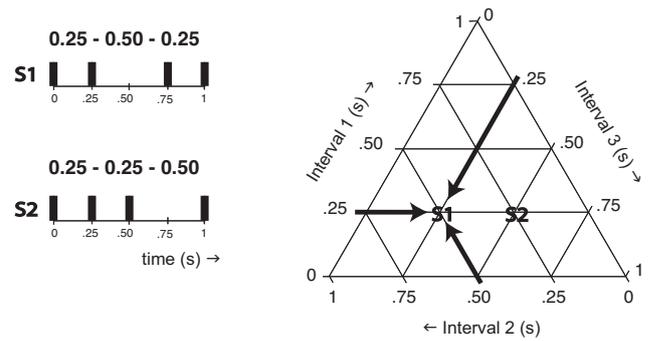

Figure 2. Two sample rhythms (S1 and S2; left panel), and their location in a chronotopological map or rhythm chart (right panel). Adapted from Honing (2013) [15] "

represents a chronotopological map obtained after collecting all the answers belonging to all possible variations of four stimuli within one second (60 beats per minute). Inside this triangle different rhythm categories are demarcated and tagged with a different letter. The black dots represent the modal points which are the points of greatest agreement among the participants when symbolically representing the sequence being heard; which is also the point in which entropy, $H = 0$. When scaled to the unit, the boundaries of each of the categories would represent the values in which $H = 1$.

As it can be observed in Fig. 3, the most frequently identified pattern (marked as modal in Fig. 3) is not aligned with the metronomical interpretation of the same rhythmic pattern. This suggests that deviations within a category do not confirm Clarke's definition of timing being deviations from integer-related durations as notated in a score. Instead, it suggests that the most commonly perceived rendition of a rhythm (modal) is actually not integer-related, but contains a timing pattern (a slight speeding up and slowing down), a rhythmic pattern that seems a more appropriate reference to use than the metronomical version. The latter, in fact, might well be perceived as expressive. [15].
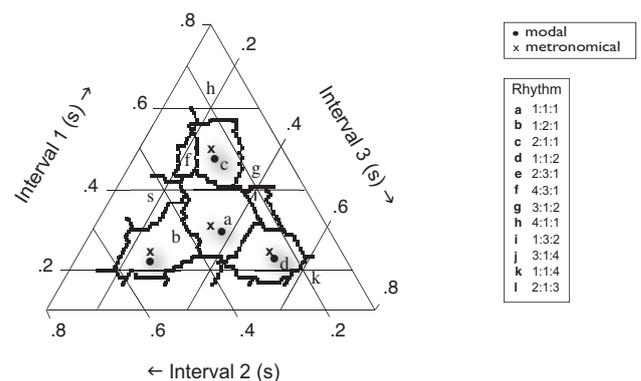

Figure 3. Rhythmic categories, demarcated by black lines in a chronotopological map. Each point in the map is a rhythm of four onsets, i.e. three inter-onset intervals with a total duration of one second; Perceived (modal) and integer related (metronomical) centroids are marked by dots and crosses, respectively; Letters refer to rhythmic categories annotated in the legend. Adapted from Honing (2013) [15]

The results obtained after these experiments [10] explain why traditional software tools in which expressive timing is treated as a result of, e.g., a rounding-off algorithm is often limited in expression and easily differentiated from non-machine generated rhythm [35]. In this study [10] it was also observed that several factors influence the perception of a rhythmic pattern: such as tempo (on this study, specifically, 40, 60 or 90 beats per minute), meter (duple, triple) and dynamic accent. These factors affect therefore the graphical representation of the rhythmic categories, varying the shape and size of each category (e.g. the 40 BPM and duple rhythm category will be different than the 40 BPM and triple). However, at the moment we will focus solely on the temporal aspects of rhythm.

3 LINDENMAYER SYSTEMS

Finding a relation between formal grammars and music syntax has been researched since the publication of the General Theory of Tonal Music [20], a theory inspired by Chomsky's formalization of language [4]. One of the main advantages of Chomsky's formalization is that its approach to the grammar is semantically agnostic. In it a generative grammar G is defined by the 4-tuple:

$$G = (N, S, \omega, p) \tag{1}$$

- N being a finite of nonterminal symbols (or variables) that can be replaced.
- S being a set of terminal symbols (constants) that is disjoint from N.
- ω being the initial axiom, is a string of symbols from N that defines the initial state of the system.
- p being a set of production rules that define how variables can be replaced by variables and/or constants having the axiom as the initial state and applying the productions in iterations.

In 1968, Lindenmayer proposed a similar mathematical formalism for modeling cell development and plant growth, in which a structure, represented by symbols within a defined alphabet, develops over time via string-rewriting [21]. This approach has been applied in many different fields such as computer graphics, architecture, artificial life models, data compression and music. The essential difference between Chomsky grammars and Lindenmayer systems (L-systems) is that in each L-system derivation (i.e. the application of the production rules to re-write the string) all symbols are replaced simultaneously rather than sequentially, which is what happens in normal Chomsky grammars.

In L-systems, structure development is done in a declarative manner according to a set of rules (pre-defined or inferred), each of them taking care of a separate step of the process. There are three types of rules an L-system may use: An essential difference between Chomsky's grammar and Lindenmayer systems (L-system) is that L-systems allow a parallel production of the grammar (instead of sequential) [31], consequently a word might have all letters replaced at once.

L-systems permit therefore the development of a structure of any kind being represented by a string of symbols within an alphabet. This development is done in a declarative manner according to a set of rules (pre-defined or inferred), each of them taking care of a separate step of the process.

In musical L-systems we can differentiate among three different steps or types of rules [23] :

- Production rules: Each symbol is replaced by one or more symbols according to the production rules, which determine the structural development of the model. The production rules are the key of the development of the string and the richness and variety of the output depends on them. Choosing therefore a set of rules or another will define the type and output of the L-system being used.
- Decomposition rules: Decomposition rules allow 'unwrapping' a certain symbol, that is meant to represent a compound structural module, into a set of other symbols or substructures that make this module. Decomposition rules are always context-free and effectively Chomsky productions [23].
- Interpretation rules: After each derivation, interpretation rules must be applied to be able to parse and translate the string output to the desired field and parameter being studied. This parsing and translation is done, as in context-free Chomsky productions, recursively after each derivation. The expressive generative model will focus on these interpretation rules; their mapping is what allows for versatility and richness, while retaining the system's simplicity.

As an example we can study a simple implementation of a Fibonacci sequence using context free L-systems and having, as interpretation rules, the generation of different rhythmic sequences:

- Axiom:
  ω : A
- Production rules:
  $p1 : A \rightarrow B$,
  $p2 : B \rightarrow AB$
- Derivations: We will obtain the following results for derivation steps n:
  $n = 0$: B
  $n = 1$: AB
  $n = 2$: BAB
  $n = 3$: ABBAB
  $n = 4$: BABABBAB
  $n = 5$: ABBABBABABBAB
- Interpretation rules:
  A : quarternote, ♩
  B : halfnote, ♩
- Final result:
  $n = 0$:
  $n = 1$:
  $n = 2$:
  $n = 3$:
  $n = 4$:
  $n = 5$:

L-systems are categorized according to the production rules they use. These can be classified according to the appliance of the production rules, but each of the grammars can be combined with others. According to Manousakis [23], L-system grammars can be: *context-free* (0L systems), *context-sensitive* (1L systems), *deterministic* (DL systems), *non-deterministic* (stochastic) *NDL,bracketed*, *propagative* (PL systems), *non-propagative*, *with tables* (TL system), *parametric* or *with extensions* (EL system).

Originally conceived as a formal theory of development, L-systems were extended by Lindenmayer and Prusinkiewiz [31] to describe more complex plants and branching structures; they also worked on implementing graphical representations of fractals and living organisms.

Prusinkiewicz's approach was based on a graphical interpretation of L-systems by using the logo-style turtle. The turtle movement in a

two dimensions map interpretation consists on a triplet (x, y, α) that includes the Cartesian coordinates (x, y) and the angle (α) that directs its facing. Once the step size (d) and the angle (α) are given the turtle is directed by following rules such as:

- F : Move forward and draw a line. The line should be drawn between (x, y) and (x', y'). (x', y') is defined then by: $x' = x + dcos\alpha$ and $y' = y + dsin\alpha$
- f : Move forward without drawing a line
- $+$: Turn left by angle δ. The turtle should point then according to $(x, y, \alpha + \delta)$
- $-$: Turn right by angle δ. The turtle should point then according to $(x, y, \alpha - \delta)$

4 USING L-SYSTEMS TO GENERATE EXPRESSIVENESS

In 1986, Prusinkiewicz [30] proposed a musical application of L-systems. Since then, several musical approaches have been proposed with purposes such as *e.g.* composing music [34], generating real time evolving audio synthesis and music structures in different time levels of a composition [23, 24, 25] or parsing music structure from scores [26]. However, to our knowledge, L-systems have not being approached yet in combination with perceptual constraints.

A main advantage of incorporating L-systems into a perceptual model of expressiveness is that since its semantic relation to the modeled structure is symbolic, *"there is no topological similarity or contiguity between the sign and the signifier, but only a conventional arbitrary link"* [23]. Due to the versatility in the production or mapping levels within different expressive categories, in any structural or generative level, a parallel development of its symbols (production) will contribute to the generation of expressiveness in music (*e.g.* combining loudness or timbre with expressive timing). This property is essential and the main motivation to use the proposed formalism instead of other algorithmic approaches. By using L-systems we can attend several perceptual categories simultaneously and define or infer rules according to the structure obtained from the musical content.

4.1 Implementation

A practical implementation based on the above theoretical framework is currently being developed. The purpose of this implementation is to verify that the hypothesis proposed can be empirically validated as a cognitive exploratory framework and a computational model of generative expressive performance (or musical composition). We therefore focus on using the rhythmic categories as a conceptual space through which a 'logo-turtle' will move to generate different sorts of expressive timing within a musical bar consisting on four onsets and according to the prediction rules previously defined by our L-system. Due to the versatility within the different steps of L-systems explained in §3, several approaches can be further developed. In the following subsections a possible implementation is presented within the different phases necessary to attend a possible generative system:

4.1.1 Geometrical approximation

In an implementation scenario, a first issue when using data from perceptual rhythm categories, is how to approach the complex geometrical shapes of each category. While finding a fitting function through each of the samples that forms the geometrical shapes it is

a more precise solution, a simplistic alternative can be the approximation of the complex geometrical shapes to simple ones. Since the shape of the rhythm categories the simplest geometrical forms that we can visually approximate them to are the circumference or the ellipse. Since we aim to cover as much space of each category as possible, an ellipse seems as the best approximation.

Obtaining measurements manually from the graphical representations of the categories [10], we have defined the position in the geometrical space as well as dimensions (axis lengths) and angle inclination of each of the ellipses being used. The result of this hand-aligned approximation to ellipses for all rhythms with a duration of one second (cf. 60 BPM) can be observed in the upper panel of Fig. 4.

4.1.2 Object Mapping

The formalisation of L-system mapping typologies has been first introduced by Manousakis[23]. Following this formalisation, each rhythmic category can be represented by a letter of the L-system dictionary and this abstraction can be used simultaneously with different production rules attending different expressive aspects (in addition to rhythm). From a generative perspective of a compositional system, once we have mapped the different rhythm categories we can define some production rules to alternate (or "jump") from a rhythm category to another generating different rhythmic patterns.

4.1.3 Movement mapping (relative spatial mapping)

Another strategy is to use a direct logo-style mapping, mapping the turtle's trajectory in 2D spaces to a path within a perceptual category. We'll use a simple L-system with a 3-letter alphabet, interpreted as movement and angle commands, and a single production rule. Lets illustrate this with an example:

- Alphabet:
  $V : F, +, -$
- Production rules:
  $p1 : F \rightarrow F+F--F+F$
- Axiom:
  $\omega : F$
- Derivations:
  $n = 0: F+F--F+F$
  $n = 1: F+F--F+F+F+F--F+F--F+F--F+F+F+F--F+F$
- Interpretation rules:
  F : move forward with a distance x
  $+$: turn right with angle θ
  $-$: turn left with angle θ

According to the example presented, in the first derivation, the turtle abstraction will advance one step, turn right, advance another step, turn twice left, advance one step more, turn right and advance another step. In order to warranty that the turtle will respect the size of the category approximation (ellipse in this case) a normalisation of the distance from the centre of the ellipse to the perimeter of it is being applied. Considering that the distance advanced by the turtle on each step might be determined by the degree of expressive deviation we want our system to produce the production possibilities are greatly determined by the amount of derivations and that the expressiveness and musical style coherence will depend on the interpretation rules being used. For instance, it would not be sensible to allow the system to have very big distance values on each of the steps to be advanced when the music style being reproduced would

not allow to have much rubato. This step distance can be easily set by adjusting the distance variable x; thus the same L-system can produce quite different results. In Fig. 4 it is shown an example of a hypothetical trajectory of expressiveness generation (using the turtle) through different points of a rhythmic category.

A combination of the two mapping strategies above can be implemented through modular mapping, in which some symbols of the L-system string select perceptual categories while others create self-similar trajectories within those categories.
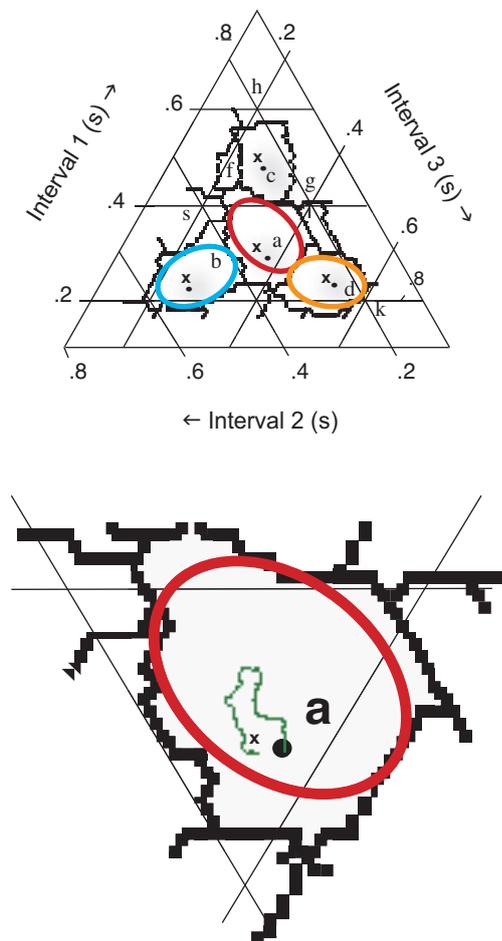

Figure 4. Top panel shows the full rhythm map of perceptual categories corresponding to four onsets stimuli played at a tempo of 60 beats per minute. Ellipses represent an approximation to the complex shapes of the categories. Bottom panel shows a 'zoomed-in' version showing category "a" with an elliptical approximation of its perceptual boundaries. The green line marks a possible turtle path on that map after using an L-system

4.2 Evaluation

As already explained in §2, the perceptual categories in which the expressive timing is generated were obtained through empirical experiments. From this perspective we have a ground to understand that the material over which the expressiveness will be generated should be perceptually valid for a human listener. "Yet, since the use of L-systems can vary much depending on the different rules and alphabets being used, to validate the hypothesis presented in this paper,

further experiments with other listeners should be carried for each of the alternative systems being developed.

4.3 Practical and conceptual challenges in the implementation of the model proposed

Some pitfalls from turning a reductionist approach into a microworld have been previously addressed by Honing [12]. Consequently, in this microworld abstraction of music and, in particular, rhythm, the formulation of the rules and the assignment of their production properties will need to attend a perceptual scenario also coherent with the music theory grounds and style specifics that our generative model is dealing with.

Based on the study done by Desain and Honing [10], Bååth et al. [1] implemented a dynamical systems model making use of Large's resonance theory of rhythm perception [19]. This implementation might be a solution to generate the data of other tempo values or inter-onset intervals durations of the rhythm categories in case empirical data is not available.

In the current microworld two issues have to be addressed to arrive at an exploratory model of expressive timing:

- The first issue is whether tempo and perceptual categories can be scaled proportionally by keeping a centroid relation derived from a morphological inference between categories. Having the results of centroids and categories for the BPM values of 40, 60 and 90 we could define an optimisation of the model to infer shapes and sizes of rhythmic categories belonging to other BPMs. However, as suggested by Sadakata et al. [32] the hypothesis is that while score intervals scale proportionally with global tempo, the deviation of the performed onsets with respect to the score is invariant of that.

- The second issue to be addressed is concerned with how to correlate positions of the turtle movement within the rhythm perceptual spaces being explored. We must clarify that this paper is concerned on how to generate expressiveness within a bar, hence no structural form of the music piece and its relation to musical style is being considered at this stage. Solving the possibility of correlating positions between categories is essential when applying this model in a real scenario since music often has different rhythmic patterns to be alternated and combined through several bars. In order to address this issue, the expressive deviations of one rhythmic category should be consistent with the deviations of the category following or preceding it. This can be done by locating these deviations according to the relative position of the 'turtles' within the different categories. The trajectory of the turtle, defined also by the length of the step, should be coherent with the rhythmic category in which it is being developed. Even when expressive timing is often oscillating between interpretations within an average of 50 to 100 ms, there is evidence that timing varies depending on tempo [9]. Having then a bigger or smaller definition of the path of the turtle might mainly make sense to be able to define concentrically its movements around the centroid; to avoid great deviations at the same time that we are aiming to achieve variation.

Scaling the modal centroid to a fitted or approximated area of the category will allow the turtle to jump in a continuous music line from a category to another one ("mirroring" these positions), being coherent with the degree of expressiveness among them; also

when approaching expressiveness complexity in musical passages in which variation is needed. Considering the continuity and progression of time in music being produced by the model we can establish mirror positions of the turtle within different categories that would follow the turtle positions within the ellipse, depending on the predetermined context (music style, performer). In the case of representing scored music, it would be determined by a score follower to choose the appropriate category representing a the rhythmic pattern, and placement of the turtle before jumping between categories (different rhythmical patterns).

This scaling however implies the need to discretise the category being represented. Using entropy (as pointed in §2) as a measure to allow comparing categories and to estimate the amount of complexity in performance before the boundary of a category is reached by our turtle abstraction, seems as an optimal solution. Following the work of Sadakata et al. [32], a more thorough study of the relation of centroids to absolute tempos would be to fit a bayesian model to the data, separating the probability of identifying a performance as a certain rhythm (score), into the prior distributions of scores, and a gaussian (normal) distribution of a performance given a score. The last distribution is expected to be off-center by an amount which is independent of global tempo [32].

In addition, moving through each of the rhythmic categories (*e.g.* using just the first three inter-onset intervals in a 4/4 bar) implies the necessity of defining a model to estimate the duration of a fourth inter-onset interval to be able to move onto a next bar through an score.

In order to determine the duration of a fourth inter-onset interval and applying this model to generate expressiveness with symbolic music scores we can use a weighting scheme such as the one proposed by Pearce et al. [29]. Weighting the distribution of the first three inter-onset intervals within a bar, we can effectively infer the duration of the 4th inter-onset intervals. A method to extract the distribution of weights within the musical structure of the piece could be done by using a parsing algorithm such as *Sequitur*, proposed by Nevill-Manning [26].

5 SUMMARY AND CONCLUSIONS

Despite much research having been done in the field of music expressiveness generation, little attention has been paid to the possibility of using data from perceptual experiments to generate expressiveness. In order to embrace the necessary versatility to produce expressiveness in music, we have presented in this paper a novel approach to modeling expressive timing performance by combining cognitive symbolic and graphic representations of rhythm spaces with Lindenmayer systems.

In §1 an approach to understanding expressiveness as deviations within different perceptual categories has been presented. In §2 the study done by Desain and Honing [10] to collect the data empirically and the formation of the rhythmic categories has been presented. §3 introduced a resume on what Lindenmayer systems are and what the state of the art on musical applications is. In addition, it has been described how by means of a symbolic abstraction we can construct rules, dictionaries or axioms using different L-systems types depending on the requirements of the music that wants to be generated. In order to follow a scientific method in §4 a preliminary implementation of the system has been presented together with a solution for further validation of the system being implemented. Nevertheless, it remains

a challenge to scale from a microworld approach (as was presented in this paper) to a more realistic model of expressive performance and, in addition, all of the proposals made in this paper still await proper evaluation, validation and empirical support. Yet, the initial steps done on this expressive cognitive model seem promising to develop automatic music performance systems as well as to understand the cognitive aspects being involved in expressiveness perception and generation of music.

ACKNOWLEDGEMENTS

REFERENCES

[1] R. Bååth, E. Lagerstedt, and P. G.ärdenfors, 'An Oscillator Model of Categorical Rhythm Perception', mindmodeling.org, 1803–1808, (2010).

[2] E. Cheng and E. Chew, 'Quantitative analysis of phrasing strategies in expressive performance: computational methods and analysis of performances of unaccompanied bach for solo violin', Journal of New Music Research, 37(4), 325–338, (December 2008).

[3] E. Chew, 'About time: Strategies of performance revealed in graphs', Visions of Research in Music Education, (1), (2012).

[4] N. Chomsky, 'Three models for the description of language', Information Theory, IRE Transactions on, 2(3), 113–124, (1956).

[5] E.F. Clarke, 'Categorical rhythm perception: an ecological perspective', Action and perception in rhythm and …, (1987).

[6] E.F. Clarke, 'Rhythm and Timing in Music', in The Psychology of Music, ed., Diana Deutsch, Series in Cognition and Perception, chapter 13, 473–500, Academic Press, (1999).

[7] S. Davies, Musical meaning and expression, Cornell University Press, 1994.

[8] P. Desain and H. Honing, 'The quantization problem: traditional and connectionist approaches', in Understanding Music with AI: Perspectives on Music Cognition, ed., & O. Laske (eds.) M. Balaban, K. Ebcioglu, 448–463, MIT Press, (1992).

[9] P. Desain and H. Honing, 'Does expressive timing in music performance scale proportionally with tempo?', Psychological Research, 285–292, (1994).

[10] P. Desain and H. Honing, 'The formation of rhythmic categories and metric priming', Perception, 32(3), 341–365, (2003).

[11] H. Honing, 'Expresso, a strong and small editor for expression', Proc. of ICMC, (1992).

[12] H. Honing, 'A microworld approach to the formalization of musical knowledge', Computers and the Humanities, 27(1), 41–47, (January 1993).

[13] H. Honing, 'Computational modeling of music cognition: A case study on model selection', Music Perception: An Interdisciplinary Journal, 365–376, (2006).

[14] H. Honing, Musical cognition: a science of listening, volume 25, Transaction Publishers, 2011.

[15] H. Honing, 'Structure and Interpretation of Rhythm in Music', in The Psychology of Music, ed., Psychology of Music D. Deutsch, D. (ed.), chapter 9, pp. 369–404, London: Academic Press / Elsevier., 3rd edn., (2013).

[16] H. Honing and WB de Haas, 'Swing once more: Relating timing and tempo in expert jazz drumming', Music Perception: An Interdisciplinary Journal, 25(5), 471–476, (2008).

[17] D.B. Huron, Sweet anticipation: Music and the psychology of expectation, volume 443, The MIT Press, 2006.

[18] P.N. Juslin, A Friberg, and R. Bresin, 'Toward a computational model of expression in music performance: The GERM model', Musicae Scientiae, (2001), 63–122, (2002).

[19] E.W. Large, Neurodynamics of music, volume 36 of Springer Handbook of Auditory Research, Springer New York, New York, NY, 2010.

[20] F.A. Lerdahl and R.S. Jackendoff, A generative theory of tonal music, volume 7, MIT Press, 1983.

[21] A. Lindenmayer, 'Mathematical models for cellular interaction in development, Parts I and II.', Journal of Theoretical Biology, 18(3), 280–315, (1968).

[22] J. London, 'Musical Expression and Musical Meaning in Context', in 6th International Conference on Music Perception and Cognition, Keele, UK, August 2000., (2000).

[23] S. Manousakis, Musical L-systems, Master thesis, Koninklijk Conservatorium, Institute of Sonology, The Hague, 2006.

[24] S. Manousakis, 'Non-Standard Sound Synthesis with L-Systems', Leonardo Music Journal, **19**, 85–94, (December 2009).

[25] S. Mason and M. Saffle, 'L-systems, melodies and musical structure', Leonardo Music Journal, **4**(1), 31–38, (1994).

[26] C.G. Nevill-Manning and I.H. Witten, 'Identifying Hierarchical Structure in Sequences: A linear-time algorithm', Journal of Artificial Intelligence Research, **7**(1), 67–82, (1997).

[27] S. Nieminen and E. Istók, 'The development of the aesthetic experience of music: preference, emotions, and beauty', Musicae Scientiae, **16**(3), 372–391, (August 2012).

[28] C. Palmer and C.L. Krumhansl, 'Pitch and temporal contributions to musical phrase perception: Effects of harmony, performance timing, and familiarity', Perception & Psychophysics, **41**(6), 505–518, (1987).

[29] M. Pearce and G. Wiggins, 'Improved methods for statistical modelling of monophonic music', Journal of New Music Research, (2004).

[30] P. Prusinkiewicz, Score generation with L-systems, 1986.

[31] P. Prusinkiewicz and A. Lindemayer, The algorithmic beauty of plants, volume 31 of The Virtual Laboratory, Springer-Verlag, 1990.

[32] M. Sadakata, P. Desain, and H. Honing, 'The Bayesian way to relate rhythm perception and production', Music Perception: An Interdisciplinary Journal, **23(3)**, 269–288, (2006).

[33] B. Snyder, Music and memory: an introduction, MIT Press, 2001.

[34] M. Supper, 'A few remarks on algorithmic composition', Computer Music Journal, **25**(1), 48–53, (March 2001).

[35] G. Widmer and W. Goebl, 'Computational models of expressive music performance: The state of the art', Journal of New Music Research, **33**(3), 203–216, (September 2004).