Soundpainting language recognition

EPFL DH Master thesis

By  
Arthur Parmentier

# Introduction

To be written last.

# Table des matières

[I. Introduction 2](#_Toc40090407)

[III. A brief history of Soundpainting 3](#_Toc40090408)

[A. Back in Woodstock 1974: emergence in emergency 3](#_Toc40090409)

[B. A new form of composition 3](#_Toc40090410)

[1. Sign language: from cultural concepts to performance 3](#_Toc40090411)

[2. A multidisciplinary language 3](#_Toc40090412)

[3. Fertility in Europe, worldwide spread in modern societies 4](#_Toc40090413)

[IV. Historical and theoretical context 4](#_Toc40090414)

[A. Gestures and signs for communication: a long history 4](#_Toc40090415)

[1. A very long time ago… 5](#_Toc40090416)

[2. Middle ages: neumes 5](#_Toc40090417)

[3. Rousseau, Wittgenstein.. theory of signs 5](#_Toc40090418)

[4. Creation of modern sign languages for deafs (?) 5](#_Toc40090419)

[5. Attempts or cases in arts before SP 5](#_Toc40090420)

[6. Conduction & other contemporary forms of artistic sign languages 5](#_Toc40090421)

[B. Real-time composition/improvisation/generative music in the XXth century 5](#_Toc40090422)

[1. Cage & co 5](#_Toc40090423)

[2. Algorithmic music 5](#_Toc40090424)

[C. Linguistics 5](#_Toc40090425)

[1. Regular languages 5](#_Toc40090426)

[2. … link with generative music? 5](#_Toc40090427)

[V. (abstract) Mechanisms and concepts of Soundpainting (a model of SP) 5](#_Toc40090428)

[A. Mechanism 1: transformations (projections?) of concepts onto the physical space of the body 6](#_Toc40090429)

[1. Input space 6](#_Toc40090430)

[2. Transformations and output spaces 6](#_Toc40090431)

[B. Structural concepts (the grammar of the language, that results in its language) 7](#_Toc40090432)

[1. Identifiers 7](#_Toc40090433)

[2. Content 7](#_Toc40090434)

[3. Modifiers (content parameters) 7](#_Toc40090435)

[4. Mode 7](#_Toc40090436)

[VI. Soundpainting recognition with Max/MSP 8](#_Toc40090437)

[A. A new configuration: motivations, goals, workflow & challenges 8](#_Toc40090438)

[B. The big picture: general description of the system 8](#_Toc40090439)

[C. Part 1: Posenet & gloves feature building 8](#_Toc40090440)

[D. Part 2: Training & data management 8](#_Toc40090441)

[E. Part 3: Classification 8](#_Toc40090442)

[F. Part 4: Grammar parsing 8](#_Toc40090443)

[G. Part 5: Orchestra simulation 8](#_Toc40090444)

[H. Part 6: Learning mechanism (?) 8](#_Toc40090445)

[VII. Potential & future of the tool 8](#_Toc40090446)

[A. Topic A: what could be improved and how 8](#_Toc40090447)

[B. Topic B: … 8](#_Toc40090448)

[C. The future 8](#_Toc40090449)

[VIII. Conclusion 8](#_Toc40090450)

# A brief history of Soundpainting

## Back in Woodstock 1974: emergence in emergency

Emergence of the first SP signs with the orchestra. Emergent in emergency (*Emergent* means beginning to arise and *emergency* means arising unexpectedly)!

<http://www.soundpainting.com/history/>

## A new form of composition

### Sign language: from cultural concepts to performance

The point here is: SP could grow fast because it is originally built on existing concepts (or latent concepts) in the fields of music, artistic performance, human communication (identifiers, content, parameters…).

Develop cultural aspects examples.

Q: Is it only a mere translation of old concepts into a more efficient sign/gestural language? A: no. It also has emergent properties (think of shapeline).

### A multidisciplinary language

SP not limited to music performance. Signs started to be created by W himself from 90+ for multidisciplinary performance.

### Fertility in Europe, worldwide spread in modern societies

Walter gave his first SP workshop in Europe in the late 90’s and found a very fertile ground for Soundpainting in France, which now has probably the largest SP community over the world.

Today, Soundpainting is used for on every continent and several artists created their own signs for the needs of their own group or configuration.

Important point: W wants to make sure that the language stays normalized and universal. Think tanks, glossary/dictionary of signs, community relations…

Modern societies: to what extent can the SP concepts of performance, content, parameters, identifiers… apply to non-modern groups?

Transition: culture to concepts

Write more about the evolution, in order to introduce the concepts later. It should already give an overall broad picture of what SP looks like at this point

# Historical and theoretical context

## Signs for communication: a long history

Chronological view (?)

### A very long time ago…

### Middle ages: neumes

### Rousseau, Wittgenstein.. theory of signs (linguistics!!)

### Creation of modern sign languages for deafs (?)

### Attempts or cases in arts before SP

### Conduction & other contemporary forms of artistic sign languages

## Real-time composition/improvisation/generative music in the XXth century

How to structure this part? By example? Topic?

### Cage & co

### Algorithmic music

## Linguistics

### Regular languages

### … link with generative music?

Check relevancy: Chomsky, hierarchical models, markov models, study of music within linguistic models.

# (abstract) Mechanisms and concepts of Soundpainting (a theoretical model of SP)

Now that we have seen a bit of the development of SP and the theoretical background it lies in, I would like to propose a model of Soundpainting that would explicit its construction mechanisms as well as the implicit operations that makes it an efficient language for art performance.

## Preliminary remarks

### Context and scope of my personal observations

The reader must be aware that my experience with Soundpainting and the observations that I consider in this section:

* Are rather limited in time (3 years span)
* Are very limited in terms of cultural diversity (most of my participating experience was in Lausanne, Switzerland or nearby (France) with performers I was familiar with, who mostly came from European music education institutions; at the exception of one experience in Rio de Janeiro, Brasil where I could both participate and observe SP practices)
* Are rather limited in terms of configuration (mostly groups of musicians, either in weekly sessions with 3-10 performers or workshops with guest soundpainter).

The reader is invited to compare my observation with his and criticize the models and interpretation I give in this section.

### The categorical and prototypical perception of concepts

The human (innate) categorical perception scheme has been studied in music REF and many other fields and plays a very important role in the construction of basic artistic concepts such as note, pitch, scale, line, hit… by constructing discrete categories out of a continuous set of elements.

We also know from research in psychology[[1]](#footnote-1) that a single concept can be modeled as a category of elements around a prototype, considered as the central point of the category. Moreover, people tend to define the concept itself by the characteristic traits of the prototype, whereas in general, it extends beyond such a definition.  
The prototypical scheme rejects the discrete notion of ‘limit’ or ‘border’, replaced by the continuous notions of graded membership (similarity to the prototype) and the fuzzy edges of concepts.[[2]](#footnote-2) But on top of the prototypical scheme, the categorical scheme introduces a rupture by either accepting or rejection an element inside the category based on its similarity with the prototype[[3]](#footnote-3).

*Let’s take the examples of birds… (does it needs an illustration?)*

## Sign “overloading”

### Motivation

We have seen from the evolution of SP that a single sign can be use to signify a content[[4]](#footnote-4) not only for different instruments of the same discipline (1) but also across discipline (2). This what I call the “overloading” of a sign.[[5]](#footnote-5)

We will try to identify whether the operability of a sign in (1) and (2) involves mechanisms of different nature.

### Differentiation of signified across disciplines

We very commonly observe signs made of one signifier and several signified in oral languages and our everyday life. The signified is understood by an operation of “disambiguation” that depends on the context of the communication. In this section, we will see that a sign in SP can indeed have several signified in the case of multi-disciplinary signs.[[6]](#footnote-6) Moreover, we will discuss the model of a bijection between the disciplines and signified concepts and discuss the differences between discipline and technical apparatus by analyzing the production process of a LT by a percussionist.

#### Existence of several signified concepts

To demonstrate the existence of several concepts (the signified) under a multi-disciplinary sign in SP, let’s take the common example of the LT, that we will use all over this section to demonstrate some of the mechanisms of SP.

For a musician, the LT is a concept preexisting to SP with a specific prototype, whose characteristic traits are “constant volume”, “constant pitch” (among others). But is the concept of the LT for a musician the same as the one for a dancer or a visual artist? To answer this question, let’s first remark that soundpainters often explain how to perform a LT differently for each discipline:

* “A fluid movement, without accent” for dancers
* “A freeze on the first syllable of a word” for actors
* “A note with constant volume and constant pitch over time” for most musicians
* …

They also often illustrate those descriptions with a prototypical example.

By looking at the description themselves, we can see that they involve different concepts: a “movement”, “roll” or a “syllable”, which cannot be considered equivalent. Moreover, we know from the history of SP that the concept of a LT was first borrowed from music and “extended” to other disciplines, i.e. that the multiplicity of signified of the sign “LT” is a voluntary construction[[7]](#footnote-7).

Although we illustrated the overloading mechanism with the sign “LT”, we can observe the same mechanism for other signs, for instance multi-disciplinary signs.

#### Bijection, surjection or injection between signified and disciplines?

Although we have shown the existence of several signified and their relationship with the concept of discipline, one can wonder whether their relationship is a bijection (to a unique discipline corresponds a unique signified), surjection (each discipline has one or more signified) or injection (each discipline has either a unique signified, either no signified).

First, we can remark that there is a bijection in the case of the so called “multi-disciplinary signs” whose properties is exactly that each discipline has its own unique interpretation (signified) of the sign.

Then, we can also observe that some signs are very specific to a discipline (sometimes by construction), such that other discipline cannot interpret them. It is enough to conclude that in general, there is no surjection between signified and the set of disciplines.

Finally, on the question of injectivity, my hypothesis is that unlike oral languages, SP can be modeled as a regular language, hence a context-free language, such that the context (signs prior to the last executed sign) does not influence the meaning of the considered sign. We will see in a future section how this approach is successful in modeling the very basics of SP, while it is unclear whether all SP modes, rules and signs could indeed be represented by a regular language. In my hypothesis, there is indeed an injection between signified and disciplines, but I leave to experts in the domain the prospection of counter examples, such as context-sensitive examples that would demonstrate that a single sign can indeed have several signified for one discipline, that can be differentiated by the context instead.

#### The difference between discipline and technical apparatus

In the “Motivation” part, I differentiated the notion of “discipline” and “instrument”, that we could rename “technical apparatus”. Indeed, one could think interpret the significance of signs across instrument and disciplines as the same mechanism. I will try to show in this this section that the significance of a sign across disciplines is a construction of different nature than the significance of a sign across instruments in the same discipline.

We will see further that this relation can be supported by cultural mechanisms, for instance a relation of analogy[[8]](#footnote-8) and/or innate mechanisms such as the categorical human perception scheme.

DIFFERENCE D APPAREIL TECHNIQUE MAIS MEME PROROTYPE VS DIFF DE DISCIPLINE DONC DE PROTOTYPE (MEME SI L APPAREIL TECHNIQUE EST LE MEME).

### Speculative theory of differentiation of signified across instruments (technical apparatus) of the same discipline

Wrap-up: what is a long tone?

By taking the simple example of the LT, we will see how the categorical and prototypical perception scheme plays an important role in the mechanisms of SP.

What is a long tone?   
It would be tempting to define the long tone for each discipline by giving a set of characteristics that all long tone must have, for instance:

* In music, a constant pitch and volume over time
* In dance, a movement without accent…

We know however from the research on our perception scheme that the LT is not (and perhaps, cannot) be defined with such characteristics or traits. It shows us that the LT is not a musical element defined by a finite number of properties but is rather a category constructed by the human perception scheme[[9]](#footnote-11) around a “prototypical” element which exhibits the features that are usually said to define the concepts of a LT.  
In this framework, a content is perceived as “more” or “less” a LT, rather than either a LT or not.

### Sign overloading

I will try to show that that in general, the grouping relations linking those concepts together under a single sign are made possible by cultural schemes (analogies) and innate perception schemes of human mind. I will use the example of the sign LT to illustrate how these schemes allow for translating the sign LT into the concepts that are relevant for each discipline, instrument (or technical apparatus).

#### Innate perception schemes

##### How to perform a LT on percussions?

I often see questions arising about how to perform a sign for a specific instrument or discipline, for instance a percussionist asking how he should perform a long tone with drums.  
While the experienced performer will probably use different possible techniques intuitively (a fast roll, using brushes, playing on cymbals that have a long acoustic response), we can derive from this simple observation that the concept of the long tone doesn’t necessarily have a trivial interpretation and its realization may not be accessible for instruments like percussions.  
Let’s break down the conceptual operations that allow the percussionist to respond to a LT when its prototype is out the set of possibilities offered by his instrument.  
To achieve this, the percussionist would typically translate the concept of LT to another concept, the roll, easily accessible to the performer. The roll will allow him to approach the prototype of the LT by increasing the speed of the roll as much as he can.  
Another way to phrase this conceptually is to say that in the space of musical concepts, the prototype of the LT is the asymptotical, limit point of the concept of roll when its speed goes to infinity.

##### An implicit translation by our perception scheme

The important point illustrated here is that SP involves operations of translation from concepts whose prototype[[10]](#footnote-12) does not exist in a discipline to a concept of the discipline that relates to the prototype. In previous example, we saw that the percussionist could translate the sign LT to the concept of roll that is relevant to its instrument. But what is the implicit scheme allowing for this translation?

I interpreted the fast roll as an “approach” to the prototype of a LT in the space of musical concepts, suggesting that there is a metric that allow us to measure the distance between the roll at a certain speed, volume… and the prototype of a LT. Even though this topic is out of my field of expertise and of the scope of this text, we can remark that the acoustic response of a fast roll “approaches” the performer’s expectation of the acoustic response of a LT, such that the human perception scheme will associate the fast roll with the prototype of the LT.

##### Learning the translation

While the operations of translation are in general not obvious (hence the questions on the subject), I observed that most Soundpainters explicit what operations are valid to beginners in SP.

#### Cultural analogies

## Personal observations during SP practice

My theory is that our categorical and prototypical perception scheme plays an important role on both how performers and soundpainters can interpret the artistic material that is being produced by the group, and how they will respond to the requests, i.e. what content the performer will produce.

I will try to support this theory by analyzing examples from my own experience as a performer, a soundpainter and discussions I had with other performers on the topic of the LT which appears as an easy example to observe some of the underlying schemes of perception and associations of musical concepts.

All 5 experienced soundpainters I have worked with teach the concept of a long tone by giving both characteristic traits and prototypical examples of a LT, showing “how it’s done” in each discipline. Their approach is usually to start defining orally the most characteristically trait of the concept, for instance “a fluid movement without accent” for dancers and then give illustrations using their body, with different speeds for each example. For musician, they would for instance mention that a LT is a sound with constant frequency as first definition of the concept and give examples by singing the prototype of a LT at different frequencies.

* From the examples given by the soundpainter, unexperienced performers are usually able to internalize the prototype of the concept of a LT for their discipline and produce examples of their own at different frequencies/speeds. I observed that the examples produced by unexperienced performers at first are very often close to the prototype.  
  But as performers become more experienced, I observed that they tend to increase the span of produced examples not only by broadening the distribution to the parameters that have been introduced to (volume, frequency, timbre…) but also by exploring different “dimensions” of the sound, such as vibrato, micro-distortions, extended techniques etc. In other words, they progressively “detach” from the prototypes of the LT concepts[[11]](#footnote-13) and their characteristic features by exploring more features of the content and reaching more extremal points. In other words, if we take the N-dimension space (one dimension per parameter of the LT) of all LTs accessible to a performer, whose boundaries are determined by the technical and imaginative limitations of the performer himself, the distribution of the production of a learning performer should first span a limited volume around the centroid of the space that can be considered the prototype of the concept of a LT. Then, the volume covered by the distribution and its variance would increase with experience and artistic research.  
  This expansion process was discussed especially in my experience with a year-lasting Soundpainting group, in which we had sessions dedicated to explore new generative processes and dimensions for LT and other very prototyped concepts. From the discussion, it was clear that our production of LT was largely prototypical and that extending the range of production required dedicated work and one explanation that I remember was that it takes a lot more cognitive load to produce a LT far from the prototype than a LT close to it. Therefore, under the constraints of immediate play, it was hard to propose something original. I conceived this training as a way to reduce the cognitive cost of the production of less-prototypical LT, therefore bringing more diversity to the responses of the performers.  
  During one week or shorter workshops that mixed both beginners and experienced performers in Soundpainting, I observed that experienced performers were responding to requests with a wider variety than beginners in Soundpainting as one could expect from previous observations, but also that performers with a greater technical level would also respond with a greater variety. My interpretation of this observation is that
  1. the cognitive load of the production of a content depends not only on the experience of the performer with SP, but also with his discipline
  2. the cognitive load is a key metric for understanding how far from a prototype a given performer can respond to a SP request

From my experience in Brazil, I can add to the latter the following remark:

* 1. the desire or willingness of a performer to respond to the request in a certain way is very cultural. In the previous discussion, I have interpreted the expansion of the variety of responses as a consequence of the decrease of its cognitive load for the performer, but it is important to remark that this expansion may not be observed at all if the performer himself is satisfied by a certain type of response, should it by prototypical or not. In fact, this expansion relies on the motivation of/relevancy for the performer to vary its responses for artistic reasons that depend on the context of the performance and the background of the performer, including his cultural background.[[12]](#footnote-14)
* Another consequence of the categorical nature of the concepts beneath SP signs is the inexistence of a clear frontier between the concepts themselves. For instance, one could argue that silence can be considered as an extremely low volume long tone, and purposefully respond to a request of LT with silence. My observation is that during learning phases, Soundpainters prefer that beginners show that they have understood the concepts by responding with prototypical examples instead of “extreme” examples.

*Goal: describe the concepts that SP deals with and propose a categorization of these concepts based on an interpretation of the mechanisms of SP (SP as a set of operations – transformations).*

## Mechanism 1: transformations (projections?) of concepts onto the physical space of the body

Key idea: forming a sign/gestural language means creating a mapping between concepts from different sources (fields?) and the physical space of the human person who signs. How is this done in SP?

OVERALL TRANSFORMATION SCHEME (illustrate the scheme of the transformation process: source concepts to signs that represents them, gestures that evoke them…)

### Input space

We can identify several repertoires (sources) of concepts in the input space of this transformation scheme

#### Concepts from artistic disciplines

Long Tone, minimalism…

#### Concepts from oral languages

Logical elements

#### “High/low” cultural representations of quiet/loud (volume), slow/fast (tempo), pitch…

Even though the concepts of volume, tempo or pitch may be universal, their mapping onto a low/high axis is defined culturally in modern societies.

#### …

### Transformations and output spaces

#### Creation of a sign

… Describe here more about the process of creating a sign (examples, general rules…)

#### Creation of a potentiometer

Volume, tempo…

#### Creation of a “pointer” (as identifier)

PTP, scanning

#### Creation of a gesture (!= sign, == mere evocations of ideas through cultural representation)

Shapeline example is the most relevant one. => using suggestive gestures that are interpreted by performers thanks to their cultural equipment/knowledge.

## Structural concepts (the grammar of the language, that results in its language)

Those concepts are used to describe the structure of the language: its grammar. They are emergent from SP, in the sense that they have been formalized dozens of years after the creation of SP and refer to its particular grammar rather than objects that are also found in other contexts. I would call the latter ones “borrowed” concepts.

One should really think of these concepts as an equivalent of the “noun”, “verb”, “adjective” … that are used to describe oral languages that modern societies are familiar with in their oral languages.

It is important to note that sign languages (among which, SP) do **not** use the same concepts as modern oral languages to describe their grammar.

### Identifiers

### Content

### Modifiers (content parameters)

### Mode

All these structural concepts (grammar) are not culturally dependent.

DRAFT of ideas

* “Traditional” Soundpainting vs “Alternative” Soundpainting: a question a configuration (soundpainting as a language but not a configuration) means that it is not necessarily linked to the frontal relation between a composer (human) and an orchestra (of human performers). My use of the word “traditional” refers to the transmission and definition of SP by Walter, with his own use & configuration of Soundpainting performances. There is no “modern” Soundpainting but rather an expanding field of research and experimentations in how the language and its concepts can extend to very different contexts and configurations. Develop with Brasilian version of SP & baila baila.
  + The concepts depend on the configuration. For instance, a “note” in music is very different from the concept (or its absence!) of “note” in visual arts.
  + SP early concepts were built for music performances and translated later for other disciplines. (Q: why do we even use this word “discipline” in arts?)  
    -> Concepts inside SP are “stolen” from the configuration/context it deals with.
  + SP acts as a kind of translation from the concepts of different practices to signs
  + It is therefore important to understand that to my point of view, there is not a fixed set of concepts that would be incorporated by the SP language
  + but rather a fluid, dynamic and varying grammar

Modes and defaults: attempt to define particular grammars inside a “universal” one?

# Soundpainting recognition with Max/MSP

## A new configuration: motivations, goals, workflow & challenges

## The big picture: general description of the system

## Part 1: Posenet & gloves feature building

## Part 2: Training & data management

## Part 3: Classification

## Part 4: Grammar parsing

## Part 5: Orchestra simulation

## Part 6: Learning mechanism (?)

# Potential & future of the tool

## Topic A: what could be improved and how

## Topic B: …

## The future

# Conclusion

SIGN = something that stands for something else; signified & signifier  
3 != types of signs: symbols, icons and indices

No synonyms in SP

Are there homonyms? YESS

Meronymes… to check/discuss : sense relations

Sign overloading: what is an analogy?

Quantum-like theory of concepts to model human interaction (interference) in responding to SP signs? https://www.frontiersin.org/articles/10.3389/fpsyg.2016.00418/full

1. Eleanor H. Rosch. Natural Categories. Cognitive Psychology 4, 328-350 (1973), introducing the “prototype theory” [↑](#footnote-ref-1)
2. For a concise explanation of the ideas behind the prototype theory, check https://www.youtube.com/watch?v=mff\_sPnz\_gs [↑](#footnote-ref-2)
3. Ref M.R. class on pitch categorisation. Emphasis on the complexity of the categorisation process (non linear wrt similarity) [↑](#footnote-ref-3)
4. Note that the whole discussion of this section is only relevant to the signs used to signify contents. [↑](#footnote-ref-4)
5. In reference to the concept of overloading in programming languages. [↑](#footnote-ref-5)
6. There may be other cases in which a sign has several signified, even inside one discipline, but we won’t discuss this possibility. [↑](#footnote-ref-6)
7. We will discuss the motivations of this construction later. [↑](#footnote-ref-7)
8. For instance, the analogy between the concept of a LT in music and the concept of a line in visual arts. [↑](#footnote-ref-8)
9. it can also be manipulated by cultural schemes, but that is not the point here. [↑](#footnote-ref-11)
10. or the whole concept itself, but as we already saw, defining the frontier of a concept is not always possible. [↑](#footnote-ref-12)
11. I use plural here because I am assuming from the beginning that there is not a single concept of a LT for all disciplines but rather a set, a group, a of concepts, some relevant for only one discipline, some relevant for several disciplines, under the sign « LT ». In fact, I am following Maurice Bloch by thinking of signs and words as decoupled structures, where the sign LT can be thought as an alias for several underlying concepts (a line in painting, a fluid movement in dance…) : this is exactly what I am calling the overloading of signs. [↑](#footnote-ref-13)
12. In my case, from my experience in Europe, producing a large variety of contents was personnally exciting and part of the interest I and others share for the technique ; therefore it was an important consideration that may not be shared at all by other groups with different interests and motivations to use SP. [↑](#footnote-ref-14)