

# Well-Structured Vocalizations

*An Attempt To Imitate Birdsong*

Ernests Vilsons

Master's Thesis

Institute of Sonology

Den Haag, 2020

## ABSTRACT

Birds. Thousands of different species, their songs and calls varying in kind and complexity. Both within the individual acts of vocalization and in the way these vocalizations succeed one another, patterns can be observed. Bird vocalizations are produced within and are influenced by their immediate environment – flora, fauna, light, wind, etc., they are a means of communication. But from the recognition of bird vocalizations as fascinating sonic structures, to composition of sound and its organization that would derive from them, a series of intermediary steps are to be taken. My research is concerned with these ‘intermediary steps’ as much as with bird vocalizations.

The thesis consists of two parts that take different, eventually confluent approaches towards the object under investigation – birdsong. Chapter I chronicles my attempt to imitate the song of the common nightingale (*Luscinia megarhynchos*). Displaced from its origin within an environment, the song is situated in an analytical framework within which it is dissected, investigated and reconstructed. Chapter II chronicles my attempt to discover and/or create the place where bird vocalizations are heard. The notion of *place* is explored through observation, reflection and engagement with thoughts of others.

## **ACKNOWLEDGEMENTS**

I am grateful for having had the chance to learn from my teachers: Raviv Ganchrow, Johan van Kreijl, Peter Pabon, Gabriel Paiuk, and Kees Tazelaar.

I am grateful for having had the chance to discuss my research and those of others with my fellow students both within the Research Seminars and in more informal situations.

I owe immense gratitude to Richard Barrett, for his encouragement and patience, for letting me find that path which I had to take.

To my family for supporting me throughout the years, to my friends for their unconditional love, to Lucie for keeping me sound.

## TABLE OF CONTENTS

Introduction	1
The Common Nightingale   A Preliminary Chapter	4
Chapter I – Analysis Interminable	9
I : 1 – The Score	10
I : 2 – Before Classification	14
I : 3 – Classification of the Elements, Their Morphology	17
I : 4 – Formalization of the Structure of Vocalization	24
I : 5 – Phonological Considerations	29
I : 6 – Two Syntheses	36
Chapter II – A Quiet Place	39
II : 1 – Theoretical Considerations	40
II : 2 – Location and Place / Field Diary	44
II : 3 – The Registered	49
II : 4 – A Strangely Indexical Mode of Listening	52
Conclusion	56
Bibliography	58
Appendix	60

## INTRODUCTION

*I leave my room, enter the street, turn right or left, then again turn right or left, then again turn... My movement is an ever-changing sensation; it is tactile, visual, thermic, olfactory, sonorous. The destination is stillness, the cessation of movement. In stillness I open myself to the influence of the field; its flow from without to within is traced, observed by the mind. Mind is the centre of action; the flow from without is incorporated into the flow within; the sounds from without enter a complex field within inhabited by thoughts that are themselves traces of what once came from without. Concepts, percepts, beliefs, convictions, dispositions, images exist side by side or rather are connected, stand in relation to each other constituting that which is the mind, which is the self, which is the experience, which is the foundation for action, for that which proceeds from within to without. This is the productive, the communicable; composition belongs to this domain, as does research. The complexity of the field, the movements within it; if one wished to enter it, one wouldn't know how. But there is no need to enter; the self is already in it, it is the movement itself, one just follows and flows with it. The recognition of birdsong as an area of focus, as a centre from which my research unfolds, around which it revolves, now in narrower, now in much wider concentric circles, is the result of the inclination to listen to birdsong. And this inclination is to be attributed not only to the beauty of the song, but also to the relations that are established between it and other elements that constitute the inner field of my mind.<sup>1</sup>*

The attempt to imitate the nightingale's song is an engagement with the ever-perplexing connection between compositional work and the factors that condition it: the strange combination of being a daydreaming, almost sentimental wanderer and meticulously systematic composer. The research proceeds through isolating a crucial element – birdsong, from the realm of my wanderings and subjecting it to the realm of the systematic in order to reconstitute it; the assumption being that by doing so I would bridge the seemingly insurmountable gap between the two realms. Such an endeavor, however, has to be mindful about the terms in which it is outlined, the presuppositions it harbors, and the limits of its authority in passing judgment about its own unfolding and that which is discovered by means of it. To begin, I discuss three elements the interconnection of which constitute the core of the research – attempt, imitation and birdsong.

Even though, birdsong has accompanied me in my walks for more than a decade, the interest in it was sparked by studies of Olivier Messiaen's music. Once I heard his birdsong transcriptions and the way they shaped sound structures and compositional form, I began to pay closer attention to the almost ubiquitous presence of birdsong in those places that I most willingly visit at times of quiet reflection and repose, and realized that it might have been birdsong itself that brought me there. The attained awareness of the importance of birdsong and its capability to shape a way of listening led me to the decision to pursue the relation between composition and birdsong in form of this research. Furthermore, since first having heard

---

<sup>1</sup> From my notebooks, November 2018.

Messiaen's works involving birdsong transcriptions, it was not the imitative aspect that was of prime importance, but rather their almost algorithmic character; each individual bird – a rule-governed machine of sonic production. This aspect rendered imitation of birdsong as compatible with my compositional interests, while the choice to implement such imitation within the domain of computer sound synthesis offered a possible connection to the non-standard synthesis techniques and algorithmic composition I had experimented with in the past. The choice of the nightingale's song as the object of imitation is partially accidental – I had recorded an extended vocalization of one specimen a few years ago. Yet more important for choosing it as the song to be imitated was the internal complexity of its structure – the nightingale's song provided a challenge that I willingly accepted. Nevertheless, the relation between birdsong and my compositional practice was from the outset of more importance than the prospect of imitative synthesis; it was the process leading up to imitation within which this relation was to be found and examined.

By thinking of the endeavor as an attempt at imitation, the emphasis is laid not on the end result, but on how this end result might be achieved, of what it takes to reach the result. However, the process is shaped by the stipulated end result; if it would not strive to attain it, it would be a different process altogether. The stipulated end result has a normative status as to the way the process unfolds, and as to what would count as a successful completion of it. The goal to be reached, the end result is imitation. But what is an imitation? While there is no single answer, the thesis is my response, and the two threads central to it can be outlined by means of the famous anecdote about the contest between the Ancient Greek artists Zeuxis and Parrhasius:

[Parrhasius], it is said, entered into a pictorial contest with Zeuxis, who represented some grapes, painted so naturally that the birds flew towards the spot where the picture was exhibited. Parrhasius, on the other hand, exhibited a curtain, drawn with such singular truthfulness, that Zeuxis, elated with the judgment which had been passed upon his work by the birds, haughtily demanded that the curtain should be drawn aside to let the picture be seen. Upon finding his mistake, with a great degree of ingenuous candour he admitted that he had been surpassed, for that whereas he himself had only deceived the birds, Parrhasius had deceived him, an artist. (Pliny the Elder 1855)

The two interrelated threads are identified as, first, the production of imitation, and second, the blurring of the boundaries between nature and art.<sup>2</sup> Deception of the birds is a testament to the Zeuxis' mastery of technique, while his misjudgment of Parrhasius' curtain highlights the profoundly circumstantial complexion of our experience. However, in line with the considerations laid out above, it is not so much the deceit itself that interests me, but rather the conditions that have to be satisfied for a deception to take place: how is an imitation produced? how can production of imitation become a technique that alters the ways in which the

---

<sup>2</sup> It is not irrelevant that the story of Zeuxis and Parrhasius is found in an encyclopedia that bears the title *Natural History*, and that art and its production are considered to be relevant to such history.

imitation and its object are perceived and thought?<sup>3</sup> In other words, it is the technique that is of primary importance, not its effects; and to gain traction on what the technique involves one has to attempt attaining that which the technique is supposed to bring about – imitation.

Imitation designates both the process and its eventual effect; and it is the success of imitation-as-effect that allows a recognition of it as an end result of a process of imitation, thus subjecting the process to the normative ideal of its successful attainment. The normative ideal, as initially envisioned, is conditioned by the possibility of deception; because of this, the imitation-as-effect strives to attain the status of being ‘as-close-as-possible’ to its object. Such high standard necessitates precision, exactitude, rigor, and, most importantly, an understanding of the phenomenon to be imitated. Because of these considerations, analysis of the nightingale’s song takes up a central role in the endeavor, while through its capability to produce knowledge, the analysis exerts influence on how the objective of the attempt is conceived. This is the narrative of my research: how through pursual of the intended goals it transformed and reoriented itself, and through doing so allowed me to understand what it really was that I was after, not only in context of this project, but in my compositional practice more generally.

A complementary approach towards the object of imitation surveys the environment within which its sonorous presence is encountered. This environment is not limited solely to empirically observable entities and processes, but includes the observer and its history, for – as Zeuxis’ misjudgment teaches us – we invest the observed with preconceptions. While our gaze obscures, it also allows the observed to be seen in a different light, one that might bring forth a productive engagement reaching further than the limited purview of the gaze itself.

The thesis begins with an introductory discussion of the nightingale and its song from the viewpoint of ornithology. Relevant terms are introduced, the general form of the bird’s vocal display is outlined, thus preparing ground for its more detailed investigation. These preliminaries are followed by two chapters comprised of a number of sections each. Chapter I provides an account of my attempt to imitate the nightingale’s song, while Chapter II is a narrative describing my investigations of the kind of environment in which the nightingale’s song is heard. In Conclusion I reflect on the research, how it changed its course, where it has led me, what has been omitted, and what the future may bring.

---

<sup>3</sup> Chapter I of the thesis seeks an answer to the first question, while Chapter II reflects on the latter.

## THE COMMON NIGHTINGALE | A PRELIMINARY CHAPTER

*On the 6th of June, 2016 I was walking the Meijendel dunes just north of The Hague for the second consecutive night. The night before I had discovered a hypnotizing amphibian chorus by the shore of one of the countless shallow lakes and recording of this chorus was the reason for my return. And now I was standing some ten meters from this very lake, listening, waiting... Twenty four minutes have passed; I pick up my recorder from the wet grass and walk further. There's little wind, no clouds cover the sky, there's enough moonlight and no tall trees. I hear birdsong and stop to listen; a few nightingales in close distance surrounding me. One is in a dense thicket on my left next to a wooden gate ('geen toegang' warns a sign). Slowly, quietly I approach it; it keeps singing. I place the recorder on the wooden gate, and cover it with my hat to mask it. I retreat and stand still for a while, then walk further to explore the surrounding area, looking for another recording. All I find is quietude, no other birds or frogs nearby. Twenty three minutes have passed when I return to pick up my recorder, the nightingale is still singing, the night has just begun, it will sing for a while.<sup>4</sup>*

The common nightingale (*Luscinia megarhynchos*) is a passerine of the Old World flycatcher family (*Muscicapidae*), genus *Luscinia* native to southern and central Europe (including the southern regions of Britain) and central Asia. In winter the nightingale migrates to sub-Saharan Africa. It resides in 'dense, low thicket growth or woodlands with young trees and bare ground underneath' (Song 2008) – a habitat affording protection from its predators, suitable for nest-building and hunting on insects – the main staple of its diet (*ibid.*). The nightingale has plain, brown-colored plumage, its average length is 16.5 cm, wingspan – 22.5 cm, its average mass – 21 grams. It returns from its wintering range around the beginning of April; the mating season – and with it the nightly singing marathons – starting soon thereafter. During this time male nightingales compete for the attention of the female birds, and it is through their singing that the relevant information pertaining to their physical condition is communicated. Singing is a feat of endurance and the ability to persist and outdo one's rivals is directly correlated to increased success in finding a mate (*ibid.*). When a couple is formed, the female bird builds a nest in which it lays four to five eggs that hatch after a fortnight. Both parents take turns in protecting the nest and providing sustenance to the chicks until they can survive on their own. Another two weeks go by until the young birds are fully fledged and it is around the same time that auditory song acquisition begins (Hultsch & Todt 2001, 316). Several weeks pass until the young birds enter the early phase of vocal production (sub-song). Gradually, slowly the nightingale transitions into the next phase of its vocal development – the plastic song. Then, almost a year after hatching the nightingale finally reaches adult vocal competence – full song, just in time for its first mating season (*ibid.*, 318-20).

---

<sup>4</sup> An excerpt of the recording is included in the Appendix under the title 'nightingale'.

In what follows I lay out a description of the nightingale's song, with a specific focus on its structure. The description is schematic and selective, it is by no means exhaustive and ignores many an aspect relevant for understanding the origins, mechanics and functioning of bird vocal activity. What is to be described is the nightingale's song as a sonorous phenomenon independent from the conditions of its production. Nevertheless, some contextual information is provided as long as it has a bearing on the understanding of the structure of the sonorous phenomenon. Before turning to a more detailed description of the nightingales song,<sup>5</sup> a few paragraphs to contextualize birdsong as an instance of an animal signaling system.

Bird vocalization is an animal signaling system. There are only two instances of animal signaling systems that are both vocal and learned – human language and birdsong (Hultsch & Todt 2001, 310).<sup>6</sup> Animal signaling systems can be considered as tools for problem solving through communication and their very existence/persistence point toward their necessity for adaption and survival. According to Hultsch and Todt, two implications follow: (1) diversity of signal systems found across animals is vast due to the diversity of biological niches to which animals have adopted; (2) the sizes of species-typic signal repertoires are rather small because only a limited number of problems (to be solved in order to adapt and survive) can be solved through communication (ibid., 309). The complexity and vastness of many songbird vocal repertoires is paradoxical; they seemingly exceed the spartan requirements of survival and adaption. Due to its paradoxical complexity, birdsong has repeatedly been subjected to comparisons with human language, a comparison I will investigate in a later chapter.

There are two kinds of bird vocal behavior – calls and songs. More than a half of the extant bird species fall under the order *Passerine*, of those the great majority form the suborder *Oscine* (Latin for songbird). All bird species engage in some sort of vocal communication, but only songbirds can be said to engage in singing – a vocal activity that differs from that of calling not only in its greater complexity, but also, and more importantly, in its function. Calls are often shared across a number of species, which is beneficial to, for example, evade common predators. Songs, on the other hand, are species specific (with the exception of the not uncommon phenomenon of mimicry), and is predominantly observed in males. The two main functions of birdsong are those of mate attraction and territorial dominance. The complexity of birdsong differs from one species to another. Some songbirds have rather simple songs consisting of only a few elements and song-types, some have developed highly complex and varied song repertoires. The common nightingale has one of the most complex signaling systems to be found among non-human animals.

---

<sup>5</sup> Full song, that is. I omit discussion of the fascinating process of song acquisition.

<sup>6</sup> However, there is an ongoing research to determine to what extent whale signaling behavior which is vocal is also learned. See Mercado 2018.

Famously the nightingale displays most vocal activity during the night hours. An uninterrupted vocal performance can last for a couple of hours and requires a ‘tremendous amount of energy.’ (Song 2008) The basic form of such performance is relatively simple and consists of a regular interchange of vocal activity (songs) and intervals of rest.

song   rest   song   rest   song   ...

A typical duration of a single song falls within a range from around a second to as much as ten seconds. These extremes are less common than the average duration of around three, four seconds, and are instances of a truncated structure, and an extended whistle song respectively. The average duration of a rest is more or less the same as that of a song. An adult male nightingale has a repertoire of approximately 200 hundred songs, or more precisely, song-types. This repertoire is stable (an acquired song-type does not undergo transformation with time), yet not fixed – new song-types are added to the repertoire as the bird matures (ibid.). Each song-type complies with, what could be called, the song template – the ‘ur-structure’ defined as a set of rules capable of formalizing the regularities observed in the singing behavior of the bird. These three terms – song, song-type and song template, are intimately connected, definitely not interchangeable, and exemplify an explicit methodology. A song is an observable sonorous event – this particular instance of a vocal display. Each song is an instantiation of a learned song-type which in turn is comparable with a design, a pattern giving shape to the sonorous. Ideally, if we accept the metaphor, every instantiation of a song-type would be identical to every other instantiation of the same song-type. However, this is not the case – there is an amount of unessential<sup>7</sup> variation that occurs from one instantiation to the next. If the category of the song-type is deducible from empirically observable recurrence of sonorous patterns, then that of a song template emerges through comparison of the identified song-types as a further abstraction, encapsulating all that remains invariant across the 200 hundred or so types, namely the principles and rules in terms of which each and every song-type can be expressed.

A distinction between two different kinds of song-types can be made: whistle songs and non-whistle songs. The whistles are ‘pure-tone elements with little or no frequency modulation’ that appear at the very beginning of a song (Kunc, et al. 2005, 1078). Non-whistle songs are marked by absence of whistles; they are far more common and allow much greater structural variability (ibid.). It is precisely because these pure whistles are sort of an anomaly, that invites a distinction. In Figures 1.1 and 1.2 spectrograms of whistle and non-whistle songs are given.<sup>8</sup>

---

<sup>7</sup> The accidental interruption, a pattern repeated now three times, now four. These variations are unessential for they do not reconfigure the structural scheme of the song-type in question.

<sup>8</sup> Although the function of whistle songs is not entirely clear, a hypothesis suggesting their importance in mate attraction has been tested and a recent study states that ‘whistle songs might play an important role in nightingale mating, as they evoke high arousal in females, and different whistle song features may signal different aspects of male quality to females.’ (Bartsch et al., 2016)

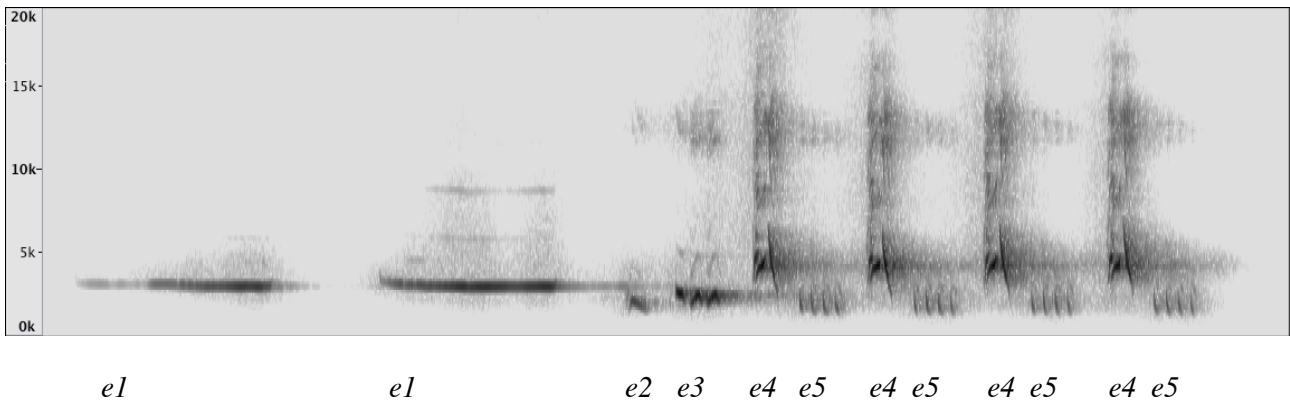


Figure 1.1 – a whistle song

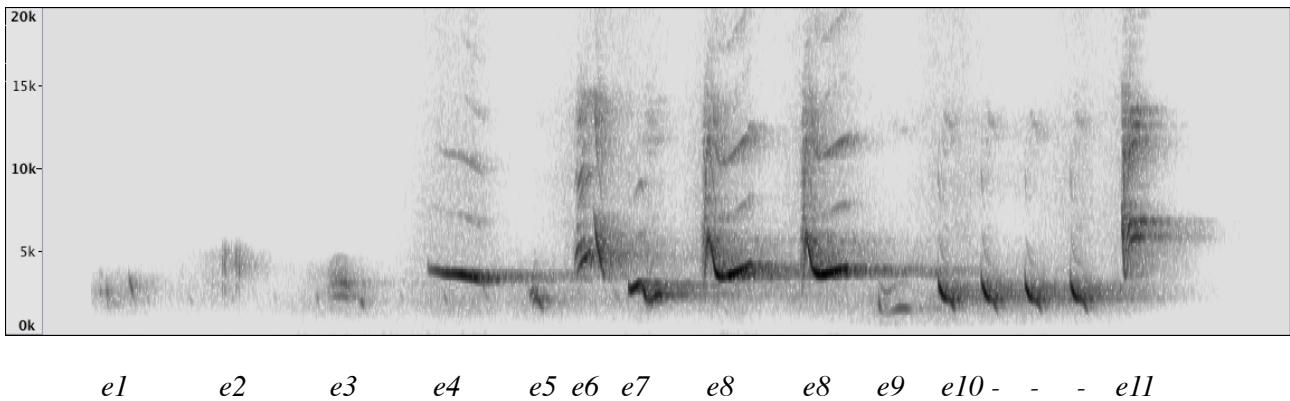


Figure 1.2 – a non-whistle song

There are four generic kinds of phrase – *alpha*, *beta*, *gamma* and *omega* – characterized by the structural role they occupy within the song template.<sup>9</sup> One kind of phrase is not interchangeable with another, they follow a strict sequential order, the only departure from a four phrase sequence model being effectuated through omission.<sup>10</sup>

*alpha* → *beta* → *gamma* → *omega*

The *alpha* phrase is characteristically low in volume, the constituent elements are separated by longer pauses than within any of the other phrase kinds (Figure 1.2, *e1–e3*). The *beta* phrase contains element complexes and motifs of relatively higher volume (Figure 1.2, *e4–e7*); the element distribution is more condensed. In neither *alpha*, nor *beta* phrases repeated elements or patterns of elements occur. Due to the vast variety of different *alpha* and *beta* phrase types, it is rather difficult to find a common criterion which would allow precise delineation between them. The *gamma* phrase is constituted by repeated elements and patterns of elements resulting in a rhythmical structure (Figure 1.2, *e8–e10*), while the *omega* phrase usually consists of a single unrepeatable element (Figure 1.2, *e11*), or a group of two or three. Much clearer distinction can be made between the *beta* and *gamma* phrases, if the latter is defined as beginning with the first instance of

<sup>9</sup> I adopt the terminology employed by Hultsch & Todt 1998.

<sup>10</sup> See Figure 1.1 in which there is no *omega* phrase after the repeated pattern (*e4–e5*); in the non-whistle song in Figure 1.2 the last element repetition within *gamma* phrase (*e10*) is followed by a single element which is its *omega* phrase (*e11*).

element repetition or patterning. Respectively, a more coarse-grained partition divides the song template in two unequal parts – *alpha+beta* and *gamma+omega*, a partition guided by the more perceptible difference in principles apparent in structural organization.

Each phrase is a collection of elements – the smallest structural units of the song. Elements differ from one another as to their morphological complexion and placement within a song-type.<sup>11</sup> An intermediary structural level – the motif, between those of the element and the phrase can be posited due to consistent recurrence of certain element combinations. The existence of observable motifs point toward the non-arbitrary nature of element-to-element transitions. It is precisely this characteristic of the nightingale’s song that is epitomized by the following rule put forward by Hultsch and Todt – ‘some element-type combinations are produced in a predictable manner, whereas others reflect certain degrees of permutational freedom.’ (Hultsch & Todt 2001, 311) As Hurford rightly notes, such a rule can be easily formalized using First-order Markov transition tables (Hurford 2011, 58), and with this formalization another characteristic of the song is made explicit, specifically, that of a one-to-many principle, or as Hultsch and Todt call it ‘diffluent flow’ schema at work in element-to-element transitions (Hultsch & Todt 2001, 311). This, in turn, implies that ‘element-types occurring at the beginning of songs are more frequent than others at later positions.’ (ibid.) To illustrate this point Hurford invokes an analogy of ‘different journeys radiating outward from the same point of origin, with different routes often diverging but never reconverging.’ (Hurford 2011, 59) A phrase, then, is constituted by single elements and their groups, for not all the elements that constitute a phrase can be assigned to a motif of more than a single element. For example, a phrase might consist of two motifs and a single element between them. What legitimates parsing the phrase in this manner is the modular nature of intra-phrase structures. Two distinct phrases might share one or more elements or motifs. What is revealed here, is the compatibility of thinking the intra-song hierarchy in two distinct, yet complementary ways – (1) that of a diffluent flow schema at the level of element-to-element transitions, which does not invoke the intermediary level of the motif; and (2) that of a modular organization of motifs as fixing the possible element-to-element transitions.

What I have attempted to lay out above is a very general schema of the structural principles in terms of which ornithology describes the nightingale’s song. Familiarity with the terms and notions introduced will facilitate the forthcoming exposition of the analysis of the nightingale’s song that I have been engaged with during my research. In the account given, I hope, one can already recognize the reasons why a study of nightingale’s song might have relevance to a practice of sound composition. In the following chapters, it is my aim to show how I ended up thinking of the song *as if* it were composed, thus separating it from its natural habitat in a forest or a field, and placing it within a habitat of a composer’s mind.

---

<sup>11</sup> There are approximately 1000 phonetically distinct elements featured in a nightingale’s song (Hurford 2011, 57-8).

# CHAPTER I

## ANALYSIS INTERMINABLE

‘We know that the first step towards attaining intellectual mastery of our environment is to discover generalizations, rules and laws which bring order into chaos. In doing this we simplify the world of phenomena; but we cannot avoid falsifying it, especially if we are dealing with processes of development and change. What we are concerned with is discerning a *qualitative* alteration, and as a rule in doing so we neglect, at any rate to begin with, a *quantitative* factor. In the real world, transitions and intermediate stages are far more common than sharply differentiated opposite states. In studying developments and changes we direct our attention solely to the outcome; we readily overlook the fact that such processes are usually more or less incomplete - that is to say, that they are in fact only partial alterations.’ (Freud 1937, 228)<sup>12</sup>

---

<sup>12</sup> The quote is taken from Sigmund Freud’s essay ‘Analysis Terminable and Interminable’.

## I: 1 – The Score

*Imitation is a very complicated thing. Imagine speech as intending to produce sounds of this or that frequency, duration, timbre, et cetera. Or intending these precise movements of your vocal folds. You would not be able to speak, for we intend words, or, still further, to communicate meanings. Even if we do give heightened attention to these acoustic characteristics of one's speech when attempting an imitation, we have a recourse to words (for example, 'he lengthens the "o" in "so"'). This is not an option available when imitating animal vocal behavior.<sup>13</sup>*

The first attempt at imitation, begun in the first weeks of the research and grinding to a halt a couple of months later, was guided by a not entirely explicit, often inconsistent methodology. The movement towards the objective, the imitation, was undermined by an at first unacknowledged, and later all too apparent, dissociation between the putative clarity of what would count as a success, and the intuitive procedure thought to be warranted precisely because of the illusory grasp of the desired outcome. The failure of this attempt was illuminating, for it led me to interrogate both the seemingly clear objective of the research, and its inseparability from the methods employed in its attempted attainment. If I were to embark on another attempt, I had to identify all those aspects of the method that had so far relied on the treacherous certainty of intuition and reconstitute them as systematically procedural. A period of (self-)critical assessment and rumination came to a close; insights gained were solidified, turned into prescriptions and brought together in a form of a score. The framework of a score introduced into the research a previously nonexistent separation between conception and realization, while simultaneously prohibiting their interrelation to be considered as given and providing guidelines for its constitution. Additionally, the intermediate nature of the score underlines the indispensability of an interpretation and its dependence on a subject capable of navigating the intersections of theory and practice to eventually effectuate the prescribed realization of the concept. Perhaps, it is somewhat paradoxical to be the subject that produces a score only to become 'subjected' to its strictures at the very next moment. One can question if the coincidence of the 'composer' and the 'interpreter' does not collapse their separation and with it the integrity of the score as the objective intermediary – introduced precisely to avoid such identification. However, to do this would be to value the score's professed objectivity over its status as the site of mediation, for the drafting of the score was first and foremost a process of clarification that resulted in an articulation of the then-current state of affairs. I present the score in its original form, and provide a brief discussion of it, not as a set of prescriptions to be satisfied in order to arrive at the intended imitation, but as a site of concurrence inhabited by intentions, thoughts, proposed methods and unfounded certainties.

---

<sup>13</sup> From my notebooks, 9th of January, 2020.

how to imitate birdsong ( a score )

//

a quiet place: a dune, a park, a field, a forest, . . .

birds

find one whose song can be distinctly (sufficiently, not entirely)  
isolated in listening and in recording

record the song (a quarter of an hour, a third, half an hour)

//

like language – speech, the song exhibits its own structure

guided by this structure, partition the song into a sequence of elements

( what is an element, is for you to determine: a phrase, a gesture, a  
phoneme – all are legitimate; the only measure – compliance with the  
subsequent steps of classification, formalisation and synthesis )

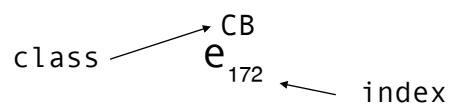
//

( one element differs from another;  
parameter is a gauge for describing difference )

classify each element according to a set of parameters

an element is now denoted by:

- a) an index – position within the temporal structure of the song
- b) a class – position within a parametric space



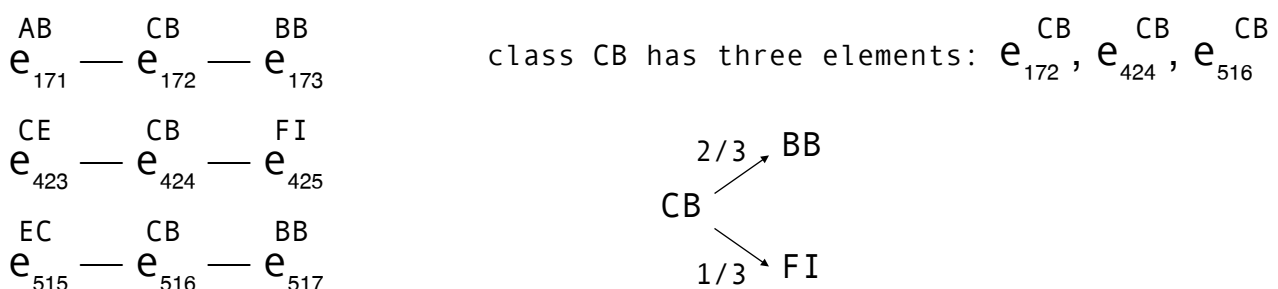
//

a passage through the song – sequential navigation of its structure and the elements of its division – writes a history of a parametric space

a history told in positions (classes) and transitions (between these classes):

- a class – position within the parametric space – has as many instances as there are elements in it
- the number of instances determine the number of possible transitions (to other classes, positions) and their likelihoods

[ an illustration ]



[ an element of the class CB is twice succeeded by an element of the class BB, once by an element of the class FI; accordingly the probabilities for transition from class CB are – 2/3 for BB, 1/3 for FI ]

a history turned into a probability-driven sequential diagram, a map for future navigations

//

design a synthesis mechanism

( classes – types of sound – expressed by numerical difference within the parameter space guide the synthesis )

design an algorithm for navigating the sequential diagram

( a markov chain of 1st, 2nd, 3rd, . . . order )

combination of the two – imitation of the song

//

judge, optimise, refine

All subsequent attempts at imitation issued from an engagement with the score, even if none of them can be called faithful renditions of it. That which the score does not prescribe, indeed, that which it is impossible to prescribe, comes into focus through its absence and pulls one back to that which is external to the score, to the context within which the score is situated. In turn, this context puts into perspective not only the score, as a means to guide oneself towards one's aim, but also the aim itself.

The most relevant feature of the score as presented is its concise delineation of the sequence of actions to be taken, which allows to break down the route towards the attempted imitation in smaller steps, each of which introduce a set of practical and theoretical problems to be addressed:

- (1) birdsong, its isolation through means of recording
- (2) partitioning the recorded vocalization
- (3) classification of the elements derived through partition
- (4) formalization of the vocalization's structure
- (5) construction of a synthesis mechanism based on the formalization
- (6) evaluation of the results, readjustments

While acknowledging the impossibility of complete separation between the steps, in the course of following sections I will address them one by one. The precise prescriptions outlined in the score as to *how* each step should be carried out were not always adhered to in my attempt at imitation; through the attempt itself some of the prescriptions were ruled to be deficient in regard to that which they were to attain, some were dismissed because of an altered perspective. Such readjustments were encouraged by the proposed evaluation and refinement, an invitation which eventually rendered the score itself with all the presuppositions it inscribed inadequate to the task it was intended to accomplish, while the supposed task – imitation of a nightingale's song, turned out to be other than what I thought.

## I : 2 – Before Classification

The score, just as this entire project, begins with a place, a quiet place. The kinds of quietude, are suggested by the enumeration: dune, park, field, forest . . . At first the ‘quiet place’ might seem to be just a *mise-en-scène*, a setting for an action to take place, yet it is more than that. While each of these settings/places share some features, they are significantly different from each other, as in those early Dutch landscape paintings, where ‘a country road’, or ‘a dune’ become genres of their own, for how they ‘compose’ the painting.<sup>14</sup> Likewise here, the place – anterior to one’s encounter with it, has already been ‘composed’, and, in part, determines the ‘composition’ of *this particular* quietude which can only become quietude to someone. This is a reminder of where the structure of birdsong is first encountered, and that this structure is conditioned by another of a different kind, that of a place and its determinants – topology, flora, fauna, to name a few.<sup>15</sup>

The movement toward the internal structure of birdsong begins with a recording – a means of isolation and displacement. A recording imposes limits through an operation of disinterested fixation; it registers a ‘history’ of air-pressure variations, fixed in regard to its position in space – the surface of a microphone’s membrane, and limited in its duration. The mechanics of this operation are far from superfluous and cannot be disregarded; not only because of the practical considerations pertaining to a successful isolation of a birdsong – a sonorous process embedded within (and in interaction with) other such processes, but also for how such an operation conditions the subsequent attempt at imitation and its methodology.

The method put forward in the score involves a three-stage process of analysis, formalization and synthesis, and the recording conducted by the interpreter has a central role in it. The ‘internal structure’ of the nightingale’s song is not given, it has to be uncovered and it is to be done through an analysis of the recording, which, in turn, is presumed to be adequate means for the task bestowed upon it. As noted in the earlier account of the nightingale’s vocal behavior, a typical song repertoire consists of approximately 200 distinct song-types. In a detailed study a group of researchers assert that an hour of nocturnal singing is sufficient to observe each and every song-type of an individual’s repertoire (Weiss et al. 2014). The recording, capturing the vocalizations of the bird which became the focus of my research, is 23 minutes long, and features 206 songs of 161 distinct song-types – a considerable share of the expected repertoire size. The importance of the durational aspect of the recording intended to isolate birdsong becomes manifest; if the object of my attempts at imitation would have been a bird with a less complex, less varied repertoire of songs, then the 23 minutes might have been more than enough to capture the repertoire in its entirety. Yet, the abridged repertoire recorded, even if precluding a complete account of an individual’s vocal display, *might be* sufficient to discover the guiding principles of its organization, or, what is the same, its internal structure.

---

<sup>14</sup> See Stechow, Wolfgang (1968). *Dutch Landscape Painting of the Seventeenth Century*. New York: Phaidon Publishers Inc.

<sup>15</sup> I expand on the notions of place, quietude, listening and their interconnections in Chapter II.

Prior to analysis, the appeal to the nightingale's song having a discernible (thus formalizable) internal structure is, of course, an assumption. However, it is an assumption which to a great extent informs the analytical procedure itself; it is a hypothesis to be tested, and, as such, is supported by the ambiguity concerning the 'object' of the analysis – a circumstance that requires an explanation. The recording is not only the mediator, but also an object in its own right when viewed as a register of air-pressure variations. Thought under this material aspect it presents a complete, finite structure. If nothing else, the structure to be revealed is the fixed structure of the recording as inscribing the sonorous appearance of a nightingale's song.<sup>16</sup> To be clear, such analysis would not warrant an immediate identification of its results with the supposed internal structure of the nightingale's song – there might be no analysis capable of such certainty, yet, it might be sufficient to produce an imitation through subsequent formalization and synthesis, which would be evaluated according to a different standard. In light of these considerations, the statement – '*like language – speech, the song exhibits its own structure*' – which epitomizes the stance I have taken toward analysis, is in need of an explication. It is a simile, and is to be examined as such – birdsong is not a language, but is somewhat *like* a language. I do not intend to reconcile their differences, but rather see what parallels might be drawn between the two phenomena and to what extent they are legitimate and insightful.

Birdsong, as was noted in a previous chapter, is an animal signaling system; its being both vocal and learned is what invites the comparison to language, a comparison which has been made by both linguists and ornithologists.<sup>17</sup> In a paper with the intriguing title *From birdsong to speech: a plea for comparative approaches*, Dietmar Todt, once the main expert of nightingale's vocal behavior, lays out some of the more important similarities. First of all, he notes the similarity between the basic form of birdsong and human speech as constituted by alternating moments of vocal activity and rest – an adaptation to modes of vocal interaction facilitating communication: one has to perform, yet also listen (Todt 2004, 202). Further, the phonological organization of both signaling systems provide optimal units of communication in songs and sentences, as the 'intermediate level[s] of a structural hierarchy in which the highest level is given by an episode of singing or speaking.' (ibid.) This intermediate level can be broken down into lower-level structural compounds; in human language these correspond to phrases, words, syllables, morphemes, in birdsong there is more variation across different species, but for the nightingale, as detailed in a previous chapter, the lower level compounds include phrases and motifs. These lower level compounds can, in turn, be decomposed into elements – the smallest structural units of song and speech (in speech these are identified as phones). In birdsong studies, it is the element level which becomes the ground of an analysis 'in which basic units are compared and classified according to parametric features such as measures of sound

---

<sup>16</sup> If one thinks of a recording as a collection of data describing displacement over time, one gives a certain autonomy to the registration independent of that which it registers. This implies the possibility to think of a nightingale's vocalization as generating a wide variety of patterns of displacement over time.

<sup>17</sup> See Bolhuis, J. J. & Everaert, M. 2013. The introduction to the volume is co-authored by none other than Noam Chomsky.

frequency and duration.’ (ibid.)<sup>18</sup> So far, the simile proposed coincides with the methodology at work in an ornithologist’s study, a methodology which, even without the aspiration to a comparison with human speech, through its own design renders birdsong as a formal language and gives it a place within the formal language hierarchy.<sup>19</sup>

---

<sup>18</sup> In the score I parenthetically remark that one can decide on what the basic elements, to be worked upon in analysis, are. However, in all my analytical attempts the element unit has coincided with the smallest discernible unit in a song, just like suggested by Todt.

<sup>19</sup> ‘Within the theory of the Formal Language Hierarchy, a ‘language’ is taken to be nothing more than a set of strings of elements, a ‘stringset’. [...] A formal grammar is a set of precise statements (usually called ‘rules’) which specifies the whole set of grammatical sentences in a language, and nothing but those sentences. The usual formulation is that a grammar ‘generates’ **all and only** the well-formed expressions in the language. The elements constitute the (‘terminal’) vocabulary of the language, and the grammar defines, or generates, all and only the well-formed strings of these elements. The elements are the smallest observed parts of the signals. [...] *Formal Language Theory doesn’t deal with the meanings of the vocabulary elements in languages, nor with the meanings of the strings of these elements which belong in the language.*’ (Hurford 2011, 26-7, [bold in the original, italics are mine])

### I : 3 – Classification of the Elements, Their Morphology

Todt describes classification as a procedure of systematic comparison of elements with respect to their parametric determinants – a procedure not unfamiliar to music analysis and, more importantly, certain approaches to composition. While an explicit comparison of birdsong to music has been absent from my account, it is not because such comparison is of no interest, but rather because it has been quietly implicit all along. It is the notion of a sonorous structure that bridges the two domains, and, irrespective of the framework within which it is analyzed – be it formal language theory, spectromorphological analysis, or information theory inflected serial approach, as long as it is capable of consistent analysis and formalization of the sonorous structure it will not fail to establish the conditions of production of sonorous objects which can be integrated within a compositional framework. It is the moment of classification that brings to fore the (inter)relation between compositional and analytical frameworks and is in need of a detailed investigation.

*The establishing of parameters. One element differs from another; this difference is to be qualified, the qualification becoming ever more complex with an increasing number of different elements to be classified. Two elements are different from one another; this mere difference is enough for classification, if no attention is given to any other elements and differences between those. Because one has gone through the process of partition, one is aware of the immense number of different elements, and thus already has an idea of what these differentiating parameters could be. Likewise, one is not indifferent to tradition, there is a set of well-established parameters linked to sound: pitch, timbre, duration, loudness, (morphology – a parameter describing change within the previously named parameters). The magnitude of a parameter is only known after all the elements have been classified, for any new element brings a possibility of extending the quantitative limits of any one of the parameters.<sup>20</sup>*

If the classificatory procedure flows along the lines just drawn, what begins to emerge is something akin to a parametric space. The nightingale's song – its sonorous structure, is arrested; its temporal flow turned into sequential organization, the unity of its elements dissected and quantified according to the chosen parameters. However, the actual process of classification – conjoining analysis and formalization, is not as smooth and linear as this schema might suggest. No element is capable of inserting itself directly within a parametric space appropriate for it and all others. The parametric space is not a given, it has to be constructed. Determining the relevant parameters according to which a classification is to be done, is a critical step on way to the intended imitation. The choice of parameters influences the success of the classification and, subsequently, the final result. Any given parameter – as, for example, pitch, timbre, amplitude, duration – has to be examined to determine its adequacy as to the task at hand; an appropriate definition of each selected parameter has to be established.

---

<sup>20</sup> From an early commentary on the score, March, 2019.

*Events in time. There is a duration. Duration is the minimum to hold on to. Within this duration, something audible happens. [...] Events are of different durations and they occupy different areas of the sonorous. Some events are more like others, some are unlike others. There are commonalities and differences across any number of events. [...] If there is a parameter I'm certain of, it is duration. All the other possible determinations are to be questioned.*<sup>21</sup>

Duration is a measure extrinsic to that which it measures – this, at once, is its power and limitation; it measures any event, no matter how simple or complex it may be, but in its generality is unable to account for particularity, for the intrinsic character of the event.<sup>22</sup> However, an accumulation of sequential measured durations establish a pattern that is identical to the pattern which it measures, and so captures the intrinsic ‘rhythmical’ character of the inter-event structure constituted by the elements and the rests between them. The delineation between the extrinsic and intrinsic is relative to a given situation, in present case it is drawn according to the unit of classification – that of an element as the most basic uniform sonorous event pertaining to a nightingale’s song. Yet, the secluded element, its temporal extent measured by an extrinsic duration, is constituted by its internal motions which are temporal and thus can be measured. These internal motions are to be measured not according to their constituting the unity of an element – this is done by the extrinsic duration, but in reference to the specific change they register. In other words, they refer to those parameters of the element which are other than duration.

In Figure 2.1 spectrograms of three different elements (*a*, *b*, and *c*) are given. Element *a* could be characterized as preeminently noise-like. It has no clearly definable pitch, while retaining a glimpse of its presence in the unequal distribution of noise respective to the different regions of its frequency range. Element *b* is much simpler, it has a clear, monophonic down-up pitch curve. Element *c* is constituted by two simultaneous, clearly pitched and harmonically independent (that is, not being part of the same harmonic spectrum) tones whose interaction generates a number of sidebands. All three elements are produced by the same bird, the same vocal apparatus operating in different ways; moreover, they are not only three different elements, they are elements of three different *kinds*. Even though, these different kinds of sound emanate from the same source, and are thus necessarily integrated within a unitary mechanism of sound production, it is difficult, if not impossible, to integrate them within a unitary parametric space modeled on the usual determinants of pitch, loudness and timbre. A legitimate question arises: why not model the parametric space as a formalized description of the bird’s vocal apparatus?

---

<sup>21</sup> From my notebooks, 23rd of January, 2020.

<sup>22</sup> Here I use the term ‘duration’ in a colloquial sense as defining a temporal measure. It is not to be confused with the concept of ‘duration’ devised by Henri Bergson and further developed by Gilles Deleuze which I introduce in my discussions of place and its nature.

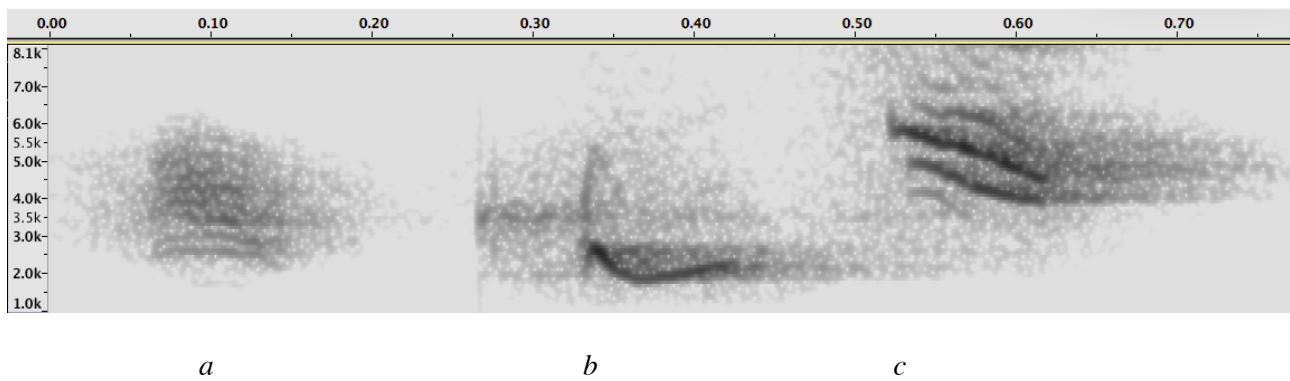


Figure 2.1 – three different elements

A complex interaction of organs and forces within the bird's vocal apparatus – the syrinx, is responsible for the nature of the sounds produced, these sounds are different 'states' of the same apparatus. The scientific investigation of the syrinx, in comparison to the study of the human larynx, is in its nascent stages, and is complicated by the vast number of species-specific variation, the small size of the organ and its placement within the organism (at the very bottom of the trachea, where it forks into the lungs).<sup>23</sup> This introduces the first complication in attempting to map the parametric space of the classification to that of the bird's vocal apparatus – lack of relevant information regarding its operation. However, the analytic environment – that of sound spectrography, within which my attempts at classification are carried out – provides a good source of relevant information as to the sonic effects of the insufficiently understood source and cause of the song – the syrinx. I was able to distinguishing between the minuscule elements present in nightingale's vocalization which the ear might not be able to pick out as separate elements, and, furthermore, to observe not only *that* elements are different, but also *how* they differ from one another (pitch and amplitude curves, formants, etc.). Nevertheless, a recording of the vocalization and its representation as spectrogram – no matter how much information is derived from investigating the spectral appearance of the song and its elements, is not sufficient for designing a physical model of the apparatus responsible for its production. Furthermore, the recording captures a field, one whose most prominent inhabitant might be the nightingale under study, yet a field nevertheless. This field is then the first mediation, through which the signal of the vocalization travels, being, even if slightly, transformed and distorted by it. These slight transformations-distortions might be filtered out when we listen, but a recorder does not listen, it registers. It is not impossible to locate these transformations and distortions and, to a certain extent, eliminate them from the analysis and the subsequent steps, but *the less the clean, undistorted signal of the vocal apparatus is understood, the less consistent are the decisions regarding the clean and the distorted*. To be clear, this last consideration does as much as

<sup>23</sup> 'The basic mechanism of sound generation in birds shares strong similarity with that in the human larynx. In both cases, tissue masses – labia in the songbird syrinx and vocal folds in the human larynx – are set into vibration by a passing airstream. Muscle activity sets the oscillating masses into pre-phonatory position, and the viscoelastic properties of the vibrating masses determine acoustic output. [...] In comparison to much more elaborate information on vibratory behavior and dynamics for a wide range of sounds in the human larynx, *we know very few details about labial dynamics in the songbird syrinx*. Especially in light of the remarkable range of frequencies present in the songs of different species, *more research is needed to elucidate dynamic mechanisms for different sounds*.' (Riede & Goller 2010, [my emphasis]).

reinforce the already-improbable option of an operative classification of the nightingale's vocal behavior based upon a physical model of its vocal apparatus.

The dismissal of the possibility of designing a physical model of the bird's syrinx, and then proceeding to classify the elements and their kinds as different states of this model, does not provide any positive solution to the problem of classification across significantly different element kinds. However, as mentioned in the above paragraph, sound spectrography does allow differentiation of minute details which, in turn, grounds the possibility of accounting for the specific differences between the elements. The incommensurability of the different element kinds does not bar their integration within a parametric space, but rather reveals some fundamental principles of its to-be-established design, most importantly, its nonuniform complexion. This principle, then, motivates the first partition of the parametric space along the lines of its respective element kinds, or, more precisely, their vertical (timbral) structure: monophonic, double-voiced and noise-like (Figure 2.1, elements *b*, *c*, and *a* respectively).

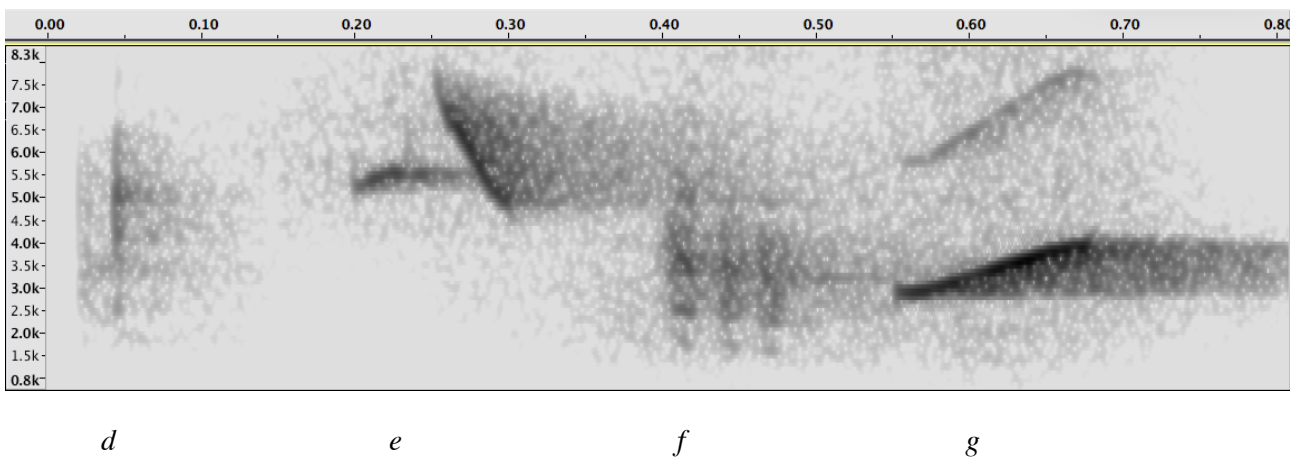


Figure 2.2 – four different elements

Another partition can be made in respect to the horizontal constitution of elements. In Figure 2.2 four elements (*d*, *e*, *f* and *g*) of different horizontal constitutions are shown. Element *d* is what I have labeled an instantaneous element – an element whose duration is sufficiently short as to sound more like a click, than a tone, whether it be monophonic, double-voiced or noise-like. Element *e* consists of two successive monophonic tones separated by a minuscule cesura. Element *f* consists of three noise-like bursts in close succession – a tremolo. Element *g* is a monophonic upwards glide. The most frequent is the uniformly extended horizontal constitution exhibited by element *g*. Repetition is not uncommon and can take as its object any of the three varieties of timbral kinds, while by far the most common is the repetition of instantaneous elements. The two-tone horizontal constitution, exhibited by element *e* consists of two distinct tones – they can be described as sub-elemental. In some instances, like the one in Figure 2.2, both tones are of the same timbral kind, however this is not always the case, there are elements whose sub-elemental

constituents are of different timbral kinds. There is an important difference between elements *d* and *g* on one hand, and elements *e* and *f* on the other. The first pair are of a *simple horizontal structure*, while the elements of the second pair exhibit a *compound horizontal structure*, one in which there is a differentiation between the element as a unit and its constitution as a compound of distinguishable sub-elements. Yet, the latter pair takes the two simple horizontal types – the instantaneous (element *d*) and the extended (element *g*) – as the constitutive sub-element determinations pertaining to the horizontal structure. While acknowledging this significant difference, all four kinds of horizontal constitution displayed in Figure 2.2 correspond to the element level, all are unitary elements due to their structural role within the vocalization.

	simple	compound	
		two-note	tremolo
instantaneous = <i>i</i>	<i>i</i>	$i_1 + i_2$	$l : i : l$
extended = <i>e</i>	<i>e</i>	$e_1 + e_2$	$l : e : l$
		$i + e$	
		$e + i$	

Figure 2.3 – horizontal constitution of elements

What becomes apparent is that a wide variety of element sub-groups can be established through different combinations of vertical and horizontal complexions, and this before any attention has been given to pitch, a parameter to which I now turn.

None of the elements shown in Figures 2.1 and 2.2 has a stable pitch – a characteristic common to almost all elements of the nightingale’s song. One finds pitch glides of various durations, directions, curvatures, covering diverse frequency ranges. Most of the elements can be located within the frequency range extending from 1 to 8 kHz, with an occasional outlier in the upper registers (up to 14 kHz or so), the most prominently featured frequency band laying between 2 and 4 kHz. There are a number of pitch-related determinations to be taken into account in classification: placement within the frequency range, the shape of the pitch curve and its direction, the minima and maxima of the pitch curve, the central frequency, defined as that segment along the frequency axis with most acoustic energy.

Due to the aforementioned instability of pitch and the relatively minute durations of the elements, a balance between what can be accounted for through spectrogrammatic analysis and what can be perceived by the human ear needs to be found. There is no necessity for extreme precision, yet not enough precision will render the classification unfaithful to the material classified and correspondingly will prohibit any aspiration to close imitation. My solution to this problem was to devise a grid of optimal grain that would partition the frequency range of the nightingale’s song and place the different elements within this grid. The partitioned frequency range is displayed in Figure 2.4.

band	I	II	III	IV	V	VI	VII	VIII	IX	X
frequency range										
from:	1	1.4	2	2.5	3.2	4	5	6.35	8	11
to:	1.4	2	2.5	3.2	4	5	6.35	8	11	15

Figure 2.4 – frequency range partition (all values in kHz)

The partitioning of the spectrum is not linear, there is more precision around the mean frequency of the song – approximately 3 kHz, while on the outer edges the grid is less precise. The decisions as to the number of bands and their bandwidths was guided by the insights gained during the countless hours of prior unsuccessful classification attempts. The grid allows to effectively trace the pitch curves of the elements, by designating the band within which the pitch glide starts, ends and, if necessary, the minima and/or maxima of the pitch curve which does not coincide with either the start or end points. Furthermore, the central frequency of each element can be described as located within one of the bands, and is of utmost importance when classifying elements without a clear pitch.

Within the curvatures of the pitch glides one can observe an intricate variety. However, this intricacy is more to be seen than heard due to the minute durations of even the more extended of elements. I decided to describe the curves in as simple a way as possible without sacrificing the classificatory significance that their differences provide.

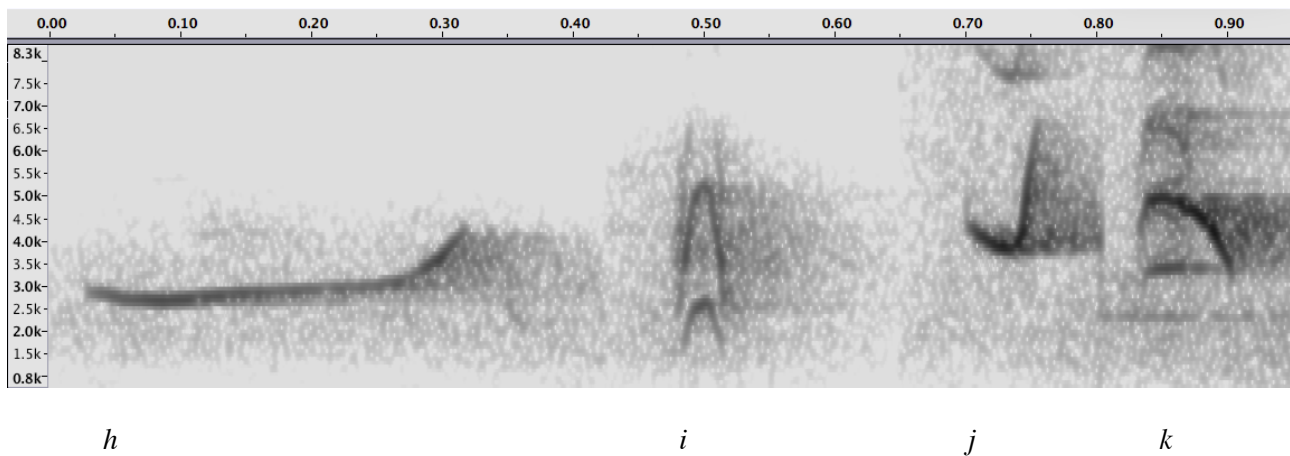


Figure 2.5 – four elements with different pitch curves

In Figure 2.5 the four basic pitch curves can be seen. Element *h* is an upwards glide, element *i* – an up-down glide, element *j* – down-up glide, and element *k* a downwards glide. This basic linear characterization of the directionality of pitch glides ignores the aforementioned variety of curvature types that can be witnessed in Figure 2.5, but is partially corrected for by taking account of the central frequency, the placement of which along the horizontal axis will skew the otherwise linear glide.

element	duration	vertical	horizontal	range	centre	direction
<i>h</i>	<i>29 ms</i>	<i>monophonic</i>	<i>extended</i>	<i>IV - VI</i>	<i>IV</i>	<i>upward</i>

Figure 2.6 – parametric description of an element

The parametric description of the element *h* (Figure 2.5) is given in Figure 2.6. The parametrization on display is sufficient to distinguish each and every of the element-types that constitute the nightingale’s song; further appeal to other parameters, such as amplitude,<sup>24</sup> is not necessary. However, the classification thus produced is incapable of accounting for the structural roles the various element-types occupy within a song, and because of this does not allow expedient formalization as compact abstraction capturing the structure of the song template. For this to become possible one has to look at the positions the established element-types take up within the song structure, find correlations and patterns, and advance to generalize across element-types through forging type-groups and rules as to their function and interchangeability.

---

<sup>24</sup> I have not produced a systematic classification of sound-pressure level differences across the elements, for amplitude as a parameter, in the present case of the nightingale’s vocalization analysis, is more relevant at the intra-element level as describing tendencies along the whole duration of a song.

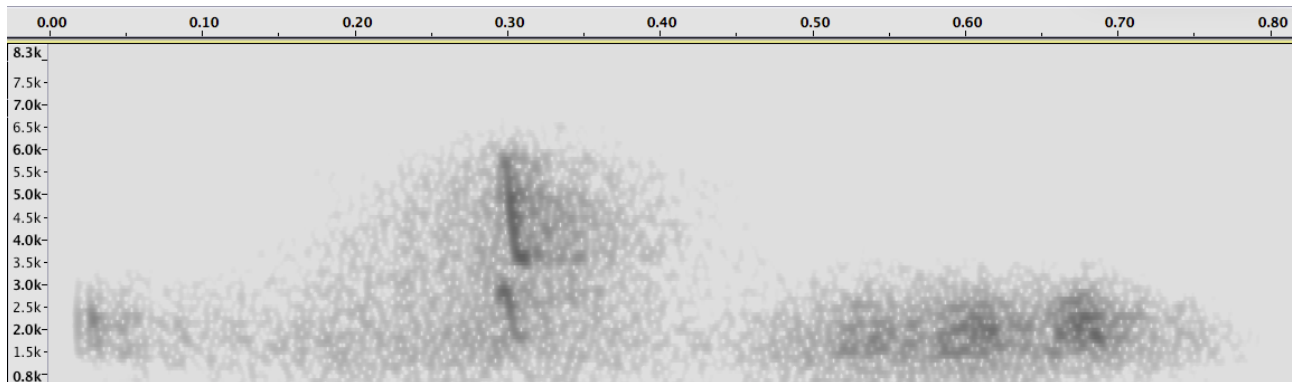
## I : 4 – Formalization of the Structure of Vocalization

In the chapter that described the nightingale's vocal activity from the ornithological point of view, I already laid out the basic structural principles pertaining to it – the basic form of the vocal performance, the distinction between a song, song-type and song template, the phrase structure and the diffuent flow of element to element transitions describable in terms of first-order Markov chain model. In the present section I wish to address the questions regarding the song structure in greater depth and outline the approach I took towards its analysis, which, while in certain respects diverging from the established methods employed in ornithology, offers a method of formalization that inadvertently highlights the difference between structure as discovered and as projected.

Each element-type, classified according to its morphology (as shown in the previous section), is present in one or more song-types, within which the given element occupies a certain position in regard to the song's sequential organization. Each element is part of one of the four phrase types – *alpha*, *beta*, *gamma* or *omega*. Hultsch and Todt in their general account of complex avian vocal behavior, exemplified by the nightingale's song, state that: 'particular types of elements occur at a particular song position only.' (Hultsch & Todt 2001, 311) The statement is ambiguous as to the range of its applicability; it is unclear whether they mean to say that a given element-type (in whichever song-type it is featured) will occupy the same position in respect to all the other element-types it appears together with as being either prior to it (and never posterior) or posterior (and never prior). Or, perhaps, they are making the less ambitious generalization – particular element-types are to be found in particular phrase structures only. The former option is ruled out for it does not align with my observations, the latter, however, is true and implies that an element-type that appears, for example, in *alpha* phrase in one song-type will never appear within any phrase structure other than *alpha* in all the rest of its instantiations. While this observation might not seem particularly astounding, it does establish an important link between the morphology of a given element and the structural position it occupies within the song template.

This relation between an element's morphology and its structural role provides indispensable means for the song's formalization. Any number of similar element morphologies can be said to form an element-type kind only if they occupy the same structural role, thus highlighting the relevant similarities between them and designating their differences as less important, and by doing so introducing a hierarchy within the interlaced parameters. Because of this, it is possible to move from a formalized song-type catalogue, to the abstraction which is the goal of the analytic procedure as such – the song template. This is done through discovering such element-type kinds, consequently prioritizing the similarities of their parametric states – establishing fields of parametric convergence within a limited range. A good example of such grouping is one which includes the most element-types and is the most accessible – that of the initial elements of the song. Invariably these elements are of low volume and are more percussive than tone-like. In my analysis, I

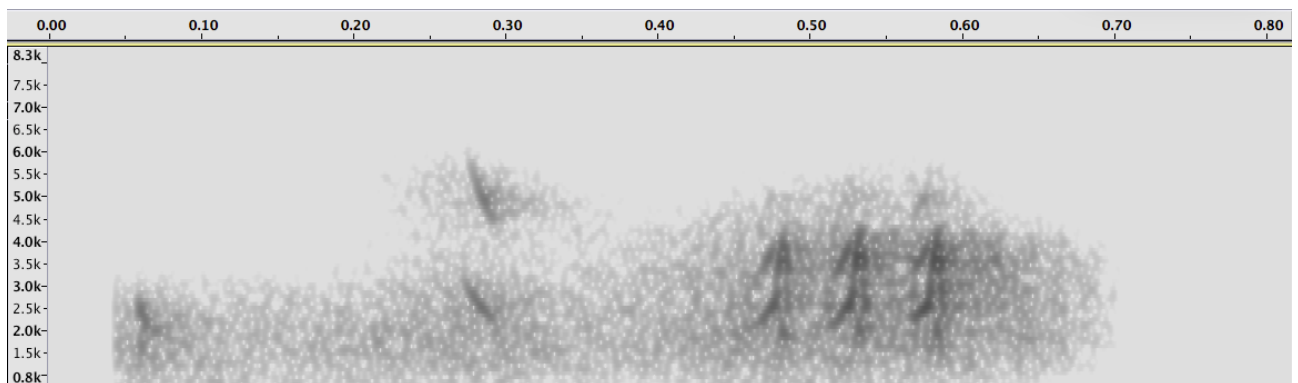
discovered 16 morphologically distinguishable element-types which could be grouped together to form a class of element-types marked by this shared structural position. By proceeding to account for their differences through parametric description I discovered that the variety is achieved through combinations of a limited set of parametric state combinations – an observation which led me to formalize the initial element position as a set of 16 different parametric state combinations. What is achieved through this abstraction is a generalized formula for producing a certain kind of sound-element consistent with the nightingale’s vocal behavior, while not having to formalize each element-type in its own right. The efficacy of this abstracting procedure becomes even more relevant when considering sequential groupings of elements – motifs.



*a1*

*b1*

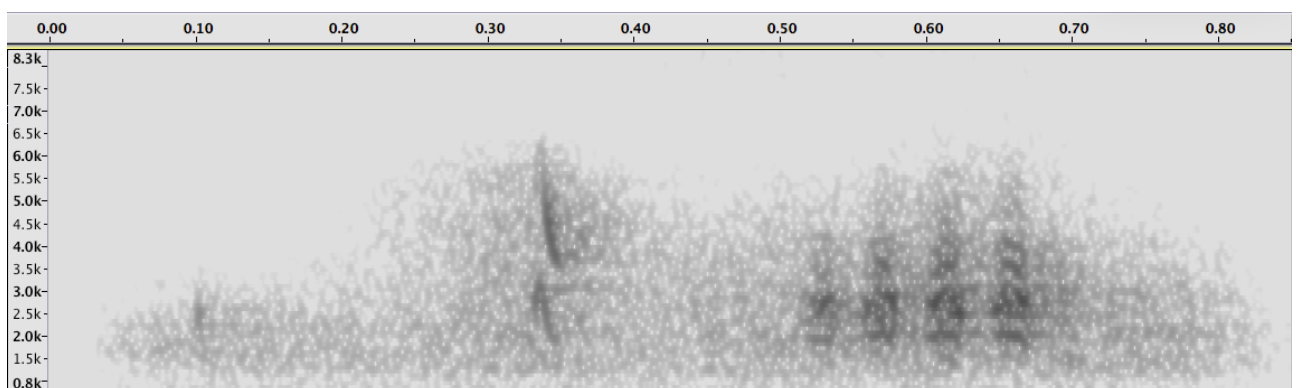
*c1*



*a2*

*b2*

*c2*



*a3*

*b1*

*c3*

Figure 2.7 – three similar patterns

In Figure 2.7 three similar patterns are displayed. The letter designates belonging to the same element-type kind, while the following numeral differentiates between element-types of the same kind. The organization of the patterns is schematized in Figure 2.8. Their general similarity is obvious even if the element-types that constitute the three patterns differ (only the second element – *b1*, of patterns one and three is the same).

index	I		II		III
pattern1	a1	→	b1	→	c1
pattern2	a2	→	b2	→	c2
pattern3	a3	→	b1	→	c3

Figure 2.8 – three patterns schematized

These patterns of a broadly homologous structure can be classified as structurally equivalent, as variants of the same motif *X*, described as a sequence of elements belonging to the element-types *a*, *b*, and *c*, and thus introducing a structural level – the motif – above that of the element, and below that of the phrase. However, the elements that appear within the motif *X*, might, and indeed do, appear in other element groupings of different kinds. This means that, even if at the level of the motif their differences are of little significance in contrast to the identity that the motif *X* retains over its variant instantiations, when correlated to the other motifs these very same elements are constitutive of, the fact of their difference has to be preserved. I proceed with a hopefully clarifying explanation.

The element-type kind designated by the letter *b* features seven different element-types. The formalization of these types follows a process of parametric description, by the end of which the element-type kind is designated by the number of different element-types within it (seven), and the parametric determinations through which different element-types can be produced. Now, because of the process of formalization, the production of these element-types is generative – it may be that none of the generated element-type morphologies is precisely like one of those element-types that provided the parametric determinations for their generation. Yet every possible element-type generated through this formalization will fulfill the requirement of having a morphological structure consistent with the established element-type kind. *This is a compositional idea* – while it allows generation of countless element-types consistent with formalization derived through observation, it clearly is not a principle operative in the structure of the nightingale's vocalization, even if perceptually satisfying the requirements it imposes. This specific kind *b* consists of seven different element-types, two of which are to be found within the variants of motif *X*, schematized in Figure 2.8. Whatever its generated morphology might be, the element-type *b1* has to be identical in patterns 1 and 3, while being different, whatever the specific difference might be, than the element-type *b2* found in pattern2.

These same rules apply when climbing up the structural ladder, for on the intermediary level of the motif, all three patterns are viewed as belonging to the same motif-type  $X$  (as  $X1$ ,  $X2$ , and  $X3$ ), and are interchangeable, for they do not have a fixed content, they are designs of structural relations between morphologically homologous element-types. In other words, both element-types and their kinds and motif-types and their kinds are described by the relations they exhibit to other element-types and motif-types (and their kinds). Thus the nightingale's song template is designed as modular on all structural levels. It is an abstraction formed on the basis of observing a particular (the nightingale recorded) to create a rule-governed schema for generating structures that are not identical, but of the same kind and complexity as those found in the object of analysis.

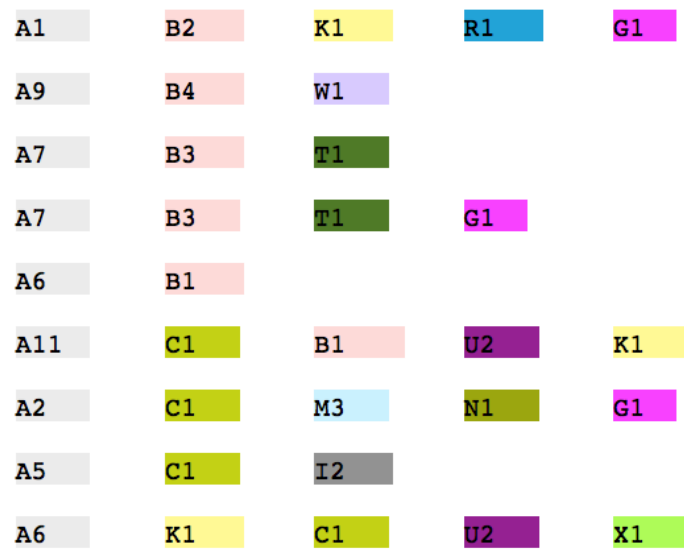


Figure 2.9 – the modular organization of *alpha-beta* phrase motifs within nine song-types

Each of the motifs designated by a capital letter and a numeral in Figure 2.9 is constituted by a pattern of one or more elements. Within each motif the diffuent flow schema is applicable, meaning that the internal structure of the motif is describable in terms of a first-order Markov chain process. The motif as a structural unit segments the chain transition process operative within any given song-type. Because of this, the chain process is transferable onto the level of the motif. As can be witnessed in Figure 2.9, a particular motif of a given motif kind can be followed by a number of different motif kinds (for example, motif kind  $B$ , can be followed by motifs of kinds  $K$ ,  $W$ ,  $T$ ,  $U$ , or a transition to a *gamma phrase* – the empty space after motif  $B1$  in line five). The rules of interchangeability pertaining to element-types are applicable here as well (for example,  $B3$  is interchangeable with  $B1$ ), while being extended even further. This is achieved through surveying all the motifs ( $A$ ,  $B$ ,  $C$ , ...) that are found within the song-types analyzed, then proceeding to determine for each motif which are the motifs that (1) can precede it, (2) can follow it, (3) are incompatible with it – all the motifs that appear in the same song-type that features the motif in question. This allows to determine with which other motif kinds, the motif kind in question is interchangeable. For example, the

motif kind *I*, according to the rules laid out above, is deemed to be interchangeable with motif kind *U*. Such cross-kind interchange is not in operation on the lower level of element-type kinds, and its introduction at the level of motif results in generation of phrase structures that are not to be found in the song-types under analysis. However, it is precisely the more-faithful-to-the-observed formalization on the element level, that allows the cross-kind motif interchange, while retaining the overall form of the phrase structure.

To be sure, by formalizing these rules of interchangeability I have clearly moved away from the role of an analyst to that of a composer. No such rules of interchangeability are operative within the nightingale's song as described by the ornithological consensus, even if the results of their application generate structures that are not incompatible with such description. The rules of interchangeability amount to a speculative incursion – the scientific picture as incomplete being supplemented by a makeshift procedure serving the purposes of the composer. The nature of the formalization while in the beginning determined by that which is to be formalized, becomes somewhat autonomous and as a result transcends its conditions of instantiation and proceeds to be guided by the ends which it can bring about, its potential as a compositional strategy. Yet, neither the fine-grained precision of the analysis is lost, nor the structural complexity. The abstraction is not so much a generalizing simplification as it is a formalization of the factual differences between element-types and their kinds, motif-types and their kinds, phrase-types and their kinds, et cetera, in order to generate novel sound structures that for the most part are homologous to those found within the observed song-types.

## I : 5 – Phonological Considerations

If I were to follow the steps laid out in the score, this would be the place to discuss the synthesis mechanism design based on the formalization of the nightingale's song structure and morphology. However, there is no one definitive synthesis mechanism to speak of. Reasons for this being the case are to be found in the preceding steps of classification and formalization, and how through them the stipulated end result – a synthesized imitation of the nightingale's song, lost the significance it had when the project was conceived and brought about a shift in focus, now centered on the *structure-as-designed*, a brief description of which is given in the previous section, while the implications it has on the project as a whole are to be discussed presently.

An important turning point in my attempts at imitation was the discovery of Roman Jakobson's theoretical account of the then-novel branch of linguistics – phonology, as presented in his *Six Lectures on Sound and Meaning*.<sup>25</sup> Jakobson poses a question concerning the relationship between sound and meaning, as the signifier and the signified of a linguistic sign:

The sign has two sides: the sound, or the material side on the one hand, and meaning, or the intelligible side on the other. Every word, and more generally every verbal sign, is a combination of sound and meaning, or to put it another way, a combination of signifier and signified [...] A sequence of sounds can function as the vehicle for the meaning, but how exactly do the sounds perform this function? What exactly is the relation between sound and meaning within a word, or within language generally? (Jakobson 1978, 3)

Having stated the problem, Jakobson immediately offers a suggestion as how to go about answering these questions: one has to identify the 'ultimate phonic elements' of a language – the smallest units within a language that bear a signifying value (ibid.). Because spoken language depends on the material support supplied by the sonorous, and as such can be acoustically analyzed and decomposed to a point where it no longer can bear any signifying value, one has to find the last station before the spoken word bottoms out in the meaningless and to identify the correspondence between the sonorous support and the signifying value it carries. Once this has been achieved, language can be discussed and analyzed on the side of the signified – meaning, without having to invoke the signifier – sound.

According to Jakobson, linguistic sounds 'considered as external, physical phenomena have two aspects, the motor and the acoustic.' (ibid., 5) However, the two aspects are not of an equal importance – when we speak we aim to produce the acoustic phenomenon, and only the acoustic phenomenon is directly accessible to the

---

<sup>25</sup> Delivered in 1942 at the Free School of Advanced Studies in New York, and subsequently published in the original French in 1976, and two years later in English (Jakobson 1978, ix).

listener, and so has an ‘intersubjective, social significance.’ (ibid., 5-6) On the other hand, the motor phenomenon, even if a necessary material prerequisite, is subordinate to the acoustic in the order of understanding and explanation (ibid., 6). However, neither can ever be sufficient to account for the relation between sound and meaning, neither is sufficient for distinguishing the ultimate phonic elements, for ‘speech sounds cannot be understood, delimited, classified and explained except in the light of the tasks which they perform in language. *Motor, acoustic and auditory description of phonic matter must be subordinated to a structural analysis of it.*’ (ibid., 109, [my emphasis]) The primary task of language is communication and as such it is intimately connected to the category of meaning as one of the fundamentals of linguistic communication.<sup>26</sup> Consequently, recourse to meaning is indispensable for a structural analysis that would explain the functioning of sound within language. In other words, it is through differences in meaning that we recognize differences in sound which in turn allow identification and classification of the ultimate phonic elements that subsequently ground the explanatory account of the relation between the signified and the signifier.

A pure classification of the phonic elements without any appeal to the role the classified take up within the structures of vocalization is akin to phonetics – ‘discipline which studies sounds solely in their motor and acoustic aspects.’ (ibid., 20) Phonology, on the other hand ‘studies sounds in their linguistic aspect’, that is, as functioning within a sign system that is language (ibid.). The basic element of language thought phonologically is the phoneme: ‘sounds which have differentiating value, those sounds which are able to distinguish words, have been given a specific name in linguistics. They are called phonemes.’ (ibid., 28) Although here Jakobson refers to phonemes as sounds, the two are not to be identified with one another. The phoneme is a structural entity which depends on its sonorous instantiation (sound is its substratum) but belongs to a regime other than it. It is the *functional aspect of the sound* and understanding how a given phoneme fulfills its differentiating function necessitates a systematic analysis of it (ibid., 37);<sup>27</sup> an analysis that would situate the given phoneme in relation to all the other phonemes within the closed phonetic constitution of the language in question. For the analysis to succeed however, one has to be able to identify pertinent differences in the sonic signifier, and this is far from an easy task:

Motor and acoustic phonetics have proved equally incapable of offering any guidance in this chaos, of identifying the pertinent characteristics, the constitutive and inalienable features of each sound. Acoustics can provide us, in impressive detail, with the micrographic image of each

---

<sup>26</sup> The function of communication and its linkage to meaning introduces a certain teleology: ‘We have said that we speak in order to be heard; we must add that we seek to be heard in order to be understood.’ (ibid., 19)

<sup>27</sup> A good example for illustrating the difference between categories of speech sounds and phonemes given by Jakobson (ibid., 31): ‘The two liquids *r* and *l* have such clearly distinct functions in our languages (cf. *ray-lay*, *fur-full*) that it seems strange to us that in some other languages they are simply two combinatory variants of a single phoneme. Thus in Korean this phoneme is represented by *l* at the beginning and by *r* at the end of a syllable. [...] It is natural that a Korean who is trying to learn English will at first pronounce *round* with an initial *l*, *sell* with an *r* at the end, and will reverse the order of the two liquids in *rule* which will then be confused with *lure*.’

sound, but it cannot interpret this image; it is not in a position to make use of its own results. [...] When, as is always the case, two sounds show both similarities and dissimilarities, acoustics, having no intrinsic criteria for distinguishing what is significant from what is not, has no way of knowing whether it is the similarity or the dissimilarity which is crucial in any given case. It cannot tell whether it is a case of two variants of one sound or of two different sounds. [...] This is not a purely technical difficulty. It is once again the vexing problem of identity within variety; without a solution to this disturbing problem there can be no system, no classification. (ibid., 18-9)

For phonology, the solution to the ‘vexing problem’ outlined by Jakobson is once again provided by the fact of meaning. The difference between two speech sounds is not to be sought in isolation from the specific functions they perform as elements of a temporal chain of signifiers, but in relation to them. Phonemes constitute words and words have different meanings: ‘What corresponds to the difference between two phonemes is solely the *fact* of a difference in meaning, whereas the *content* of these different meanings varies from one word to another.’ (ibid., p. 62-3, [italics in the original]) The task of the phonologist is to identify all the phonemes operative within a language through surveying the apparent differences in meaning that distinguish the elements (words) of a given vocabulary.<sup>28</sup> By doing so not only the *fact* of phonological difference but also its *content* is determined.

	o	a	e	u	ə	i	ɪ	ʊ	f	ʃ	k	z	ʒ	g	m	f	p	v	b	n	s	θ	t	z	ʒ	d	h	ʃ
1. Vocalic/Non-vocalic	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2. Consonantal/Non-consonantal	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
3. Compact/Diffuse	+	+	+	-	-	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4. Grave/Acute	+	+	-	+	+	-	-	-	-	-	-	-	-	-	+	+	+	+	+	-	-	-	-	-	-	-	-	-
5. Flat/Plain	+	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6. Nasal/Oral	-	-	-	-	-	-	-	+	-	-	-	-	-	-	+	-	-	-	-	+	-	-	-	-	-	-	-	-
7. Tense/Lax	-	-	-	-	-	-	-	-	+	+	+	-	-	-	-	+	+	-	-	-	+	+	+	-	-	-	+	-
8. Continuant/Interrupted	-	-	-	-	-	-	-	-	+	-	-	+	-	-	-	+	-	+	-	-	+	+	-	+	+	-	-	-
9. Strident/Mellow	-	-	-	-	-	-	-	-	+	-	-	+	-	-	-	-	-	-	-	-	+	-	+	+	-	-	-	-

Figure 2.10 – the phonemic structure of English Received Pronunciation<sup>29</sup>

The *chef d'oeuvre* of Jakobson's theory is the dissection of the phoneme into *distinctive features* (Jakobson 1978, 109-10). In Figure 2.10 the phonemic structure of English Received Pronunciation is displayed: the

<sup>28</sup> ‘What we recognise in spoken language is not sound differences in themselves but the different uses to which they are put by the language, i.e., differences which, though without meaning in themselves, are used in discriminating one from another entities of a higher level (morphemes, words).’ (Jakobson 1978, 74)

<sup>29</sup> The table is taken from the Appendix of *Preliminaries to Speech Analysis: The Distinctive Features and their Correlates* (Jakobson, Fant, Halle 1963, 43).

vertical column on the left lists the distinctive features (two opposites for each row), while the horizontal row above the table catalogues the phonemes operative within English; beginning with vowels and proceeding with consonants. Each phoneme is classified according to the distinctive features its articulation exhibits: ‘+’ corresponds to the left-side element of the binary pair (for example, ‘vocalic’ in the first row), the ‘-’ to the right-side element (‘non-vocalic’), an empty lane corresponding to a given pair of distinctive features indicates the feature’s absence from the articulation of the phoneme. I will not discuss the specific classification displayed in Figure 2.10, but rather the general idea behind the dissociation of the phoneme.

The proposed analysis affords a precise determination of *how* one phoneme differs from the others, through comparison of the binary opposition pairs of the relevant distinctive features. For example, the sole difference between the first two vowels (/o/ and /a/) in Figure 2.10 is their opposition in regard to the feature labelled ‘Flat/Plain’ – /o/ is a flat, or rounded vowel, while /a/ is plain, or unrounded.<sup>30</sup> By parsing the differences between the phonemes of any given language according to this methodology one arrives at a truly structural analysis of the signifier accounting for how sound is shaped in order to convey differences of meaning. It turns out that not the phonemes, but the distinctive features from which phonemes are composed are the ‘ultimate phonic elements endowed with a sense-discriminating function’ and as such they are ‘clearly and uniquely “oppositive, relative and negative entities”.’ (ibid., 96)<sup>31</sup> The phonological system is rendered as a set of structural determinations that locate the phonemic elements of the language within the matrix of distinctive features, which in themselves have no positive value, for they only exist in relation to one another.

The above account is condensed and simplified, yet it introduces the two main ideas which influenced my approach to identification and classification of element-types present in the nightingale’s song, and the song’s structural formalization. The two ideas are those of dissection of the phoneme into its distinctive features and the importance of meaning in uncovering the structural principles in operation at the level of the signifier.

Parametric analysis of sound is inherently similar to the operation of dissecting the phoneme into its distinctive features. That the dissection invokes a characterization in terms of the phonatory act is inessential, for it might be done in terms of the acoustic characteristics (ibid., 81); both registers allow a consistent

---

<sup>30</sup> ‘Flattening manifests itself by a downward shift of a set of formants or even of all the formants in the spectrum. [...] Flattening is chiefly generated by a reduction of the lip orifice (rounding) with a concomitant increase in the length of the lip constriction.’ (Jakobson, Fant, Halle 1963, 31)

<sup>31</sup> The definition of the distinctive features as ‘oppositive, relative, negative’ is a reference to Ferdinand de Saussure who attributed these determinations to the phoneme (Jakobson 1978, 41). ‘Saussure understood the purely differential and negative character of phonemes perfectly well, but instead of drawing out the implications of this for the analysis of the phoneme he overhastily generalised this characterisation and sought to apply it to all linguistic entities. He went so far as to assert that there are in language only differences with no positive terms.’ (ibid., 64) Jakobson disagrees with this generalization reserving the negativity solely to the register of distinctive features.

analysis but neither is to be identified with the purely structural character of the distinctive feature.<sup>32</sup> Unlike Jakobson's presentation of the distinctive features, my formalization of the element-types invokes the acoustic register as that within which distinctions are drawn. A further divergence from Jakobson's model is that the set of parameters which were chosen to account for differences between the element-types are not entirely oppositive, relative and negative. This status could be attributed to the distinction between the three varieties of vertical constitution: monophonic, double-voiced, noise-like; and the nesting of determinations pertaining to the horizontal: instantaneous/extended, simple/compound. These parameters are defined in discrete terms and their diverse configurations, while not necessarily being oppositive, are mutually exclusive – an element-type, by definition, cannot be both monophonic and double-voiced, or instantaneous and extended.

	<u>element-types</u>			
	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>
monophonic / double-voiced / noise-like	0	[1, 1]	2	2
simple / compound	0	1	0	1
two-note / tremolo		0		1
instantaneous / extended	1	[0, 1]	1	0

Figure 2.11 – a discrete schema of element-type vertical and horizontal constitutions<sup>33</sup>

When the horizontal and vertical constitution of element-types is presented as in Figure 2.11, what is emphasized is the pure fact of difference between the types and its rendition in terms that can be easily separated from the object of the analysis. The parametric configurations – represented by the columns of discrete states under the element-types designated by letters – can be mapped onto a system entirely different from the one which they describe. It is the object which to a great extent influences the formal character of its analysis, while the formalization once established is separable from its object and can be put to a use different from one it serves currently.

Similar considerations led me to partition the frequency range of the vocalization into discrete regions. Even though the discretized frequency axis does not exhibit the simplicity inherent to the binary nature of distinctive features, it presents the parametric states pertaining to frequency as relative to one another over and above their absolute position. It is a further abstraction and as such facilitates the replacement of the

<sup>32</sup> '[Each] differential element exhibits one clear and easily identifiable acoustic feature, and that in analysing phonation precisely in the light of this acoustic effect we are always in a position to separate out from the multitude of phonatory movements a single basic factor which produces the acoustic effect in question.' (Jakobson 1978, 81)

<sup>33</sup> The numerals 0, 1, 2 signify the respective state of each parameter by referring to the index position of that state in the left-hand column. The square brackets are used for element-types that have a two-note horizontal constitution and the applicable parameters are determined for each of the tones comprising the element.

element-type viewed as sound object with its formal description. While the parametric analysis of the element-types does not convert to an explanatory account that would show the necessity of the particular ways they differ from one another, it occasions their integration within a unified parametric space within which the element-types are states and in their terms the unfolding structures of vocalization can be described. The element-types of my analysis are not phonemes, but through their classification and formalization they acquire something akin to the differentiating value of the phoneme, for they do indeed mark a difference and render distinguishable the sound structures they are part of. Additionally, and here another parallel to phonology is to be found, conceiving of the analytical category of element-type as doubly articulated, determined by both its morphology and structural role, is what allows systematic classification.

However, if it is the fact of meaning that drives and supports the dissection of the signifier and by doing so allows its structural analysis to take place, what plays the role of meaning in my endeavor? For even if it is possible to discuss birdsong as having meaning, in the most basic sense of it satisfying the function of communication, this meaning is not to be found within its phonological structure, for it does not have one.<sup>34</sup> It is not meaning as such that I need to find an equivalent to, but rather the influence the conceptual category of meaning exerts when thought in relation to the sonorous signifier which is its support. This influence is what makes possible a structural account of the functioning of sound within language.

The structure of the nightingale's vocalization as observed is unarticulated – the surface structure is accessible but its organizing principles are not known. So long as the element-types are described *as different* without providing an account of *how they differ*, not much insight into the organizing principles is gained. However, if the element-types are to be classified and formalized in relation to both the morphologies they exhibit and the structural roles they play within the contexts of their appearance, it could be objected that there are no grounds on which I can ever speak of a structural role without simultaneously being engaged in a speculative interpretation of the unarticulated structure of the vocalizations; an interpretation that projects rather than discovers the organizing principles. However, there is no measure extrinsic to the surface structure that could affirm or deny the validity of an interpretation as long as it produces results seemingly consistent with that which it interprets. It is in this light that the principles of interchangeability – and the modular complexion of the song's structure they give rise to – are to be seen. The position which in phonology is occupied by the category of meaning is replaced by thinking the structure of the vocalization *as composed*. By doing so, one retains the functionality of the extrinsic measure to distinguish the important from the trivial, the differential from the mere variation, while not conflating discovery with interpretation.

That structure which I sought to discover was not to be found in the object of analysis – the unarticulated song structure as captured in the recording, but rather in the form of the analysis, determined by both its

---

<sup>34</sup> I avoid any discussion as to what 'meaning' in case of animal communication might be, while acknowledging that the pursual of this question might bring about compositionally interesting approaches to imitation.

object and the analyst. This realization came to me late in the process, yet once accepted allowed completion of the analysis on terms which not only disabused me from thinking that a successful analysis would be one which discovered the ‘actual’ organizing principles of the nightingale’s song, but also did not compromise the relevance of compositional perspective. And it was precisely when the compositional perspective was acknowledged and became an integral part of the analytical endeavor that my focus turned away from the intended imitation-to-be-synthesized and towards the analysis itself conceived as a self-revelatory process of coming to terms with my own compositional predilections.

## I: 6 – Two Syntheses

Before arriving at the now-final form of the song's classification and formalization, I did experiment with synthesis, albeit with uneven success. Due to the morphological complexity of a number of element-types the means employed in synthesis were not uniform and the results unconvincing. In part it was because of the insufficient understanding of the nightingale's vocal apparatus which barred precise replication of the timbral characteristics of the more complex signals, in part because of my unwillingness to sacrifice the generic and generative nature of the formalization for a non-uniform, element-type specific approach to synthesis. Furthermore, the realization of the importance compositional considerations played in the process of analysis, extended the range of the ends that such analysis could serve. One such end is the work with whistle tones – a special kind of element-type found in the nightingale's song – I have been engaged with.

The whistle tones are almost pure, extended sine-wave-like signals.<sup>35</sup> While a couple of harmonics can be detected in a spectrographic image of a whistle tone, they are rather weak. The frequency glides vary in range, but are invariably narrow. As I mentioned in my discussion of the nightingale's song from the perspective of ornithology, the presence or absence of whistle tones marks a distinction between two separate kinds of song-types. Furthermore, the fact that whistle tones appear in groups and always at the beginning of a song,<sup>36</sup> separates them from the other phrase structures and singles out the whistle tone element-types as fundamentally different from the rest, thus facilitating their isolation in listening and motivating their close imitation as an independent compositional strategy.

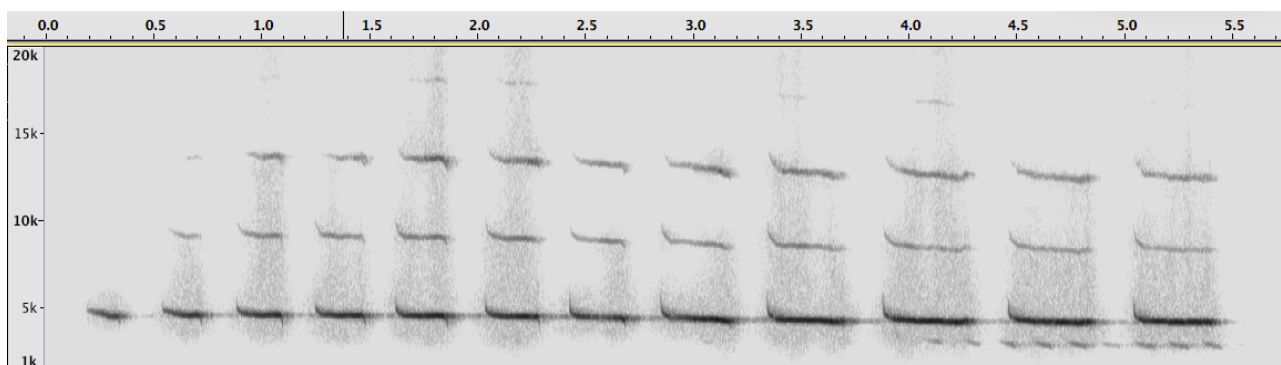


Figure 2.12 – an observed whistle phrase

I carefully analyzed thirteen whistle phrase types found in the repertoire of the particular nightingale captured in my recording. Although the morphology of a whistle tone is simpler than that of an average non-whistle element, the ability to model the subtle changes of frequency and amplitude across its extended

---

<sup>35</sup> No other kind of element-type comes close to the extended duration – up to a half of a second – of a whistle tone.

<sup>36</sup> The duration of a whistle phrase not uncommonly exceeds an entire duration of non-whistle song-types. The longest whistle phrase featured in my recording is eight seconds long.

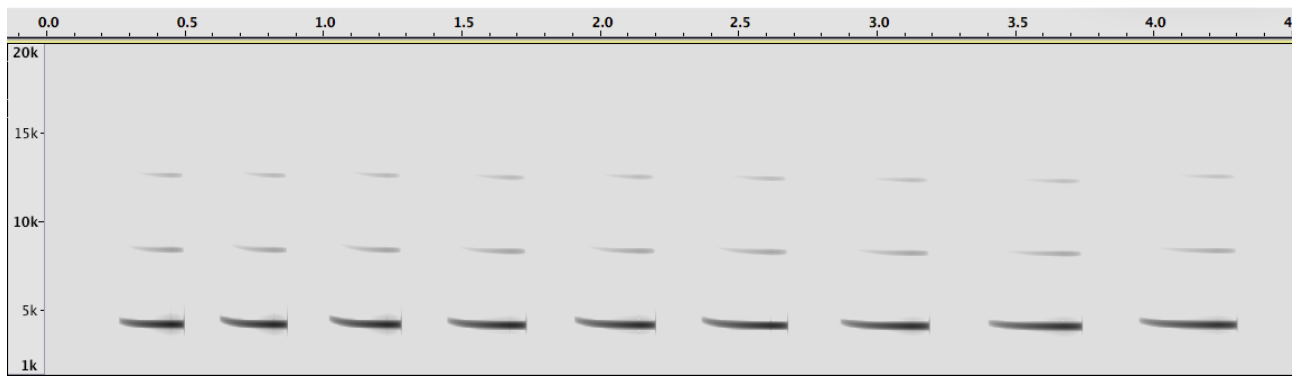


Figure 2.13 – a synthesized whistle phrase

duration are integral in attempting a close imitation.<sup>37</sup> The more general formalization of the element-types described earlier does not admit such degree of precision when it comes to, for example, the subtleties of pitch glide curvatures and their placement along the frequency axis. Further, as can be seen in Figure 2.12, even if the internal structure of a whistle phrase can be reduced to an iterative repetition of the same element-type, the element-type in question undergoes modification from one iteration to the next. The specific character of the modification across the length of the phrase, the whistle element-type modified, and the number of its iterations is what distinguishes one whistle phrase from another. For each of the thirteen I designed an algorithm that would recreate the principles of modification: shortening or lengthening of the element or the gaps between its iterations, widening or narrowing of the range of the frequency glide, the rising or falling of the the mean frequency of the element, et cetera.

The ability to model imitations of the whistle phrases with high degree of precision, allows a compositional approach centered on the ambiguity between the recorded and the synthesized, which I have experimented with by embedding the modeled whistle phrases against a backdrop of field recordings.<sup>38</sup> If the model of a whistle phrase is unaltered, thus approximating the phrase it is a model of, its appearance will most likely be perceived as part of the soundscape it is artificially inserted in,<sup>39</sup> while systematic alteration of the parameters determining the model's output, and thus its perception vis-à-vis the soundscape, can be employed as a compositional method.

Returning to the general formalization of the non-whistle song-types. As noted, the formalization underwent a change in status through its own unfolding; from being considered as means to reach the goal that would be a close imitation of the nightingale's song, it became an end in itself, and by doing so occasioned a scission between the analysis and the synthesis it was supposed to bring about. No longer, did I thought of synthesis as having to be imitative, but rather as based on the structural principles encapsulated in the formalization.

<sup>37</sup> This is mainly because of the whistle's near purity, for if the subtle change is not present in the imitation, the resultant signal is too pure, too mechanic and as such reveals its synthetic nature.

<sup>38</sup> An example of whistle-tone integration within a field recording is provided in the Appendix under the title 'whistles'.

<sup>39</sup> Of course, this in large part depends on the constitution of the soundscape in question.

As a result, the particular constitution of the synthesis mechanism is not determined and can take a variety of forms, as long as the formalized structure exerts an influence on the sound processes produced. In the near future I plan to experiment with applying the structural principles derived from the analysis in writing for acoustic instruments, a possibility opened up by the generic form of the final formalization. For now, however, I have limited my experiments to a kind of non-linear synthesis approach involving both acoustic feedback and feedback loops internal to the synthesis mechanism, as well as band-pass filters.<sup>40</sup> The results of these experiments are to be showcased alongside the aforementioned studies in whistle phrase imitation as the compositional supplement to the theoretical questions addressed in the thesis.

---

<sup>40</sup> An example of this kind of synthesis is provided in the Appendix under the title ‘synthesized nightingale’.

## CHAPTER II

### A QUIET PLACE

‘The amount of quiet I need does not exist in the world, from which it follows that no one ought to need so much quiet. [...] [It] is about seven o’clock in the evening; I am lying in the reclining chair at the margin of a three-walled hut, with two blankets, fur coat and pillows. Outside the hut is a clearing in the forest, about one-third the size of the Ringplatz in Zürau. The meadow is all yellow, white, violet with familiar and unfamiliar flowers. All around is ancient spruce forest; behind the hut the brook murmurs. I have already been lying here for five hours, today slightly disturbed, yesterday and day before yesterday entirely alone, only with the bottle of milk beside me. One must really be grateful for this, and today I am going to keep silent about things for which one need not to be grateful. Anyhow, if every afternoon were like this and the world left me here, I would stay until they had to carry me away on the reclining chair. In the meanwhile, you would come once to visit me, wouldn’t you?’ (Kafka 1982, 73-4)<sup>41</sup>

---

<sup>41</sup> The quote is taken from Franz Kafka’s letter to his sister Ottla and her husband Josef David, written in June 1921.

## II : 1 – Theoretical Considerations

I find myself in ‘a quiet place’, perhaps in a forest or a field, and I hear birdsong. Although this might be sufficient for proceeding to listen, isolate and record a song of a particular bird, not all has been settled, at least not entirely, for a question persists – how did I get here? This location, this ‘place’. The score remains silent on the particularity of an interpretation, and if this particularity is to be understood, accounted for, then one has to examine not only the score, but the ‘examining of the score’ or its interpretation. One has to attempt the acrobatics of standing on one’s own shoulder. Furthermore, the procedure, the action of interpretation, cannot be satisfied by the mere arrival at a particularity; it has to concern itself with how this particularity comes about, because only through this understanding will it be able to judge whether it is or is not successful, whether it is or is not legitimate.<sup>42</sup>

The research has no definite beginning; the score gives one to it retroactively by way of a schematic description, equation – *a quiet place: a dune, a park, a field, a forest, ...* – an accidental past turned into a binding prescription, a reciprocal rendering of the accidental as the prescribed origin. Yet to leave it at that was something I could not do. The English naturalist Gilbert White, a dear companion throughout the research, once wrote: ‘all nature is so full, that that district produces the greatest variety which is the most examined.’ (White 2016, 45) The ‘quiet place’ is that ‘district’ which I had to examine; its nature was to be observed – not only in the narrow sense as an environment formed by its topography, flora and fauna, but also in the more general, overarching sense, one that would be true to its being a quiet place – a determination necessarily involving a subjective component, someone to perceive the place as quiet. In other words, the examination had to concern itself with the nature observed, and its observer as being part of this nature – observation as a natural process embedded within a nature it observes.<sup>43</sup>

A quiet place. Quietude and place – their coincidence within an environment (a dune, a park, et cetera). ‘Quiet’ is a qualification, it marks off a certain type of a place; the precise location of the cut separating the quiet from that which is not so is neither given, nor fixed, nor arbitrary. The subsequent enumeration – ‘a dune, a park, a field, a forest, ...’ – exemplifies but does not determine. A quiet place might be encountered; one may look for it in a forest or a field. Quietude, however, is not a self-sustaining being, it is consequent upon something that it is not, upon a multitude of processes and their relations. Thought is one such process, for within it the concept of quietude arises. Sound is another, for it is sound that is conceived within this thought. To be sure, the constitution of quietude is not exhausted by the relation of these two processes, not

---

<sup>42</sup> In following this line of thought, I accept a fundamental insight by Hegel, who insists on the necessity of understanding not only the end, but also the path traversed en route to this end: ‘[The] subject matter is not exhausted in its aims; rather, it is exhaustively treated when it is worked out. Nor is the result which is reached the actual whole itself; rather the whole is the result together with the way the result comes to be.’ (Hegel 2018, 5)

<sup>43</sup> ‘[The] nature that is thought does not issue from the thought of it; rather the thinking of that nature has the character it has precisely insofar as nature is the ground of which its being thought is the consequent.’ (Grant, I. H. 2013, 39)

least because of their own composite nature; however, they are crucial for comprehending quietude and its relation to place in the specific context of my research. It is not the totality, not thought or sound as such, but particular, even if not easily discernible, manners of operation within thought and sound that compose the quality I have named quietude – a ‘composition’ to be found in the interrelation of sound and thought. Listening is a process that forges such relation; through it one begins the approach toward comprehending quietude and the manners of operation underlying it.

For quietude to be, is for it to be situated within an unfolding experience – first and foremost marked by the act of listening, and as such it is always coincident with a particular location in space and a duration in time. Thus, a quiet place is recognized, established through the coincidence of quietude and a location. Yet it is important not to take coincidence as a ‘mere coincidence’, but as a ‘coinciding’ – of place and quietude being coextensive and in a relation that is not arbitrary, since it is mediated by the listener. It is precisely the never-entirely-extricable complexity of the processes underlying both quietude and place that does not allow coincidence to be mere coincidence. Through listening, one can discern some processes that are equally part of both – thought and sound. These considerations do not amount to a definition, if one was ever attainable, but they are of help in examining the district that is the quiet place.

In her essay *Series and Place* (Grant, M. J. 2011), Morag Josephine Grant traces the two concepts and their interrelation in the practice of composers associated with the Wandelweiser collective. The focus of her analysis is directed towards a project – *3 Jahre—156 musikalische Ereignisse—1 Skulptur* [3 Years - 156 Musical Events - 1 Sculpture] – conceived by the composer Carlo Inderhees and the artist Christoph Nicolaus. Grant describes the project thus:

[Every] Tuesday evening from the beginning of 1997 until the end of 1999, a ten-minute long musical event for one performer was premiered in the Zionskirche in Berlin. The performances took place beside a sculpture made up of ninety-six bore stones, the positions of two of which were swapped weekly (in advance of the performances) according to chance operations. These changes, as well as details of the performances, the number of visitors, the temperature in the church and the degree of sunlight, were documented by the organisers. (Grant, M. J. 2011, 529)

These 156 musical events were not composed by Inderhees alone, but by a number of composers directly or indirectly related to the Wandelweiser group. Further Grant quotes a description of the project by the artists themselves:

A period of time is constituted through regular changes.

A place is constituted through regular changes.

Within a time period of three years, a sequence (Folge) is established.

The elements in the succession each last 168 hours and succeed one another directly. At the beginning of each element in the succession there is an event of ten minutes' duration. The 156 elements in the succession begin on Tuesdays at 19:30.

The period in which the succession is projected begins on 1.1.1997 and ends on the 31.12.1999.

The place in which this succession is established is the Zionskirche in Berlin Mitte.

(ibid., 529-30, [my underlinings])

There is an ambiguity regarding the notion of 'place' (*der Ort* in German) in the above description. 'Place' is said to be constituted through regular changes, and as such to be be-coming and necessarily temporal, while at the same time a 'place' – the Zionskirche – is where the project is to be carried out and thus already is and remains what it is – Zionskirche. This ambiguity is present also in Grant's own discussion of the concept:

[What] makes a place a place are its non-temporary qualities, particularly that we return to it time and again. *Places become places through a serial process: places are not one-off things, but arise from repetition.* (ibid., 536, [my emphasis])

Grant does not so much define a 'place' as propose conditions for a 'place' to become. Places become through a process, and this process is repetition. Places are created, and, importantly, places are constituted by temporary and non-temporary qualities. In other words, a distinction is made between that which *takes place* – the temporary, the changing, and the non-temporary *against which* temporality and change are registered, while at the same time both kinds of quality constitute a single place. However, there is an emphasis on the non-temporal qualities as being those which 'make a place a place', through their permanence that allows a return 'time and again'. The *permanence* is what allows *us* to return, to carry out the serial process of repetition through which place becomes place. Accordingly the measure that decides what is and what is not temporal is the subject who returns, its temporality, or duration.

The concept of *duration* as devised by Henri Bergson, and further developed by Gilles Deleuze, offers a framework through which the constitution of a place and the part we play in it can be thought, especially because of the emphasis on the interconnection between being, temporality and individuation. In Deleuze's account 'everything is duration' (Deleuze 1991, 76), or in other words everything that is participates in *duration*; however *duration* is 'dissipated in all these differences in degree, intensity, relaxation (*detente*), and contraction that affect it' (ibid.); *duration* does not appear as a unified whole, but as a multiplicity of durations all coexisting within this one unified duration in which they partake. We access *duration* through intuition, through our own psychological duration; in it we attribute to what we perceive as other to ourselves a duration independent from ours, yet we integrate it within our own duration (we perceive it). However the

coexistence of these two durations – mine, and that of another – is only possible if there is a third, overarching duration within which both of these, and indeed all, durations are integrated (Deleuze 1991, 76-80).<sup>44</sup>

The permanent, the non-temporary – that which endures – is not absolutely so, but is so relative to a measure.<sup>45</sup> A place becomes a place for us when our temporality-duration is embedded amidst the multiplicity of temporalities-durations that constitute a place; our duration coincides with them, while simultaneously experiencing the multiplicity of the temporalities within itself; it integrates the degrees of temporality and experiences them within its own duration. So it might be said that the place that becomes is one which we synthesize, create within our own duration, while at the same time affirming the existence of the place from which it becomes as independent from this construction within our own duration. Importantly, through this synthesis, what comes to constitute a place is not only that which we perceive as other to ourselves – other beings, but also all the threads of our own psychological duration – thoughts, memories, sensations. The place which is created is inhabited by trees and birds, thoughts and memories, all of them coexisting and coinciding. This does not mean, however, that the embedding of our duration amidst others in no way changes the constitution of a place; it does, our presence changes this constitution, but this change is nevertheless not synonymous with the synthesis we carry out within our own duration. The importance of recognizing the role of temporality, or duration, as crucial for understanding the concept of place, allows one not to mistake something for what it is not, or, in other words, not to project onto a place one's synthesis of it.

Having introduced the theoretical considerations at play in thinking of place and its constitution, I offer a short preliminary description of their practical counterpart – a project of observation, registration and reflection carried out in spring of 2019. The afore-cited sentence by Grant – *Places become places through a serial process: places are not one-off things, but arise from repetition*, supplied a method – repetition, for this endeavor. During the first half of April, I made repeated visits to a particular location that was to become the 'quiet place'. The choice of the location – within the dunes of Meijendel, just north-east from The Hague – was informed by the score, and by it being not far from and similar to the location where the recording of the nightingale's song was made three years earlier. Extended field recordings were produced, which were listened to back at home, making notes on what I heard and thought while continuing to wander through this place, constructing it by this very movement.

---

<sup>44</sup> It seems to me that Deleuze's interpretation of Bergson is influenced by his reading of Spinoza. In Deleuze, Bergson's concept of duration is analyzed in terms that resemble Spinoza's metaphysics: a single substance (the overarching duration) as differentiated through modifications (the individual durations) immanent to it. Having been influenced by Spinoza's rationalism myself, I find the primary importance Bergson attributes to *intuition* as a method highly problematic. However, as long as one brackets out Bergson's insistence on the immediacy of intuition, the concept of duration serves the current purpose.

<sup>45</sup> 'If things are said to endure, it is less in themselves or absolutely than in relation to the Whole of the universe in which they participate insofar as their distinctions are artificial.' (Deleuze 1991, 77).

## II : 2 – Location and Place / Field Diary

Meijendel dunes, stretching 6 km along the coast and 3.5 km inland, cover 2000 hectare area (van der Meulen, Bakker & Houston 2008, 268). It is a nature reserve, and for almost 150 years has been a drinking water catchment zone (ibid., 272). Additionally the dunes ‘play a major role in sea defence by protecting the low-lying western part of the country against the sea.’ (ibid., 269) All these factors have influenced the biotope.<sup>46</sup> The location in which I carried out my observations is situated 2.2 km away from the seashore and falls within the most densely forested area within Meijendel. A broad, sandy trench marks its south-east border, on the other side of which a more sandy, dune-like patch is located. The other ‘borders’ coincide with walking paths, which in turn adapt to the natural topography of the site. The location covers an area of approximately 7 hectares and is around 10-12 meters above sea level. A wide variety of different trees, shrubs, grasses and herbs intertwine in irregular patterns forming the flora. Almost all of the area is covered in thorny vines inhibiting free movement off the paths and aiding the accumulation of plant litter which consequently enriches the soil and furthers the growth, while providing protection for a number of bird species that nest close to ground – the nightingale being one of them. The choice of this particular location within the dunes is due to a combination of two factors – practical considerations (distance from the main walking routes and bicycle paths, relative seclusion), and the location’s ambience. The practical considerations led me to the location, while the ambience experienced there ratified it as the location for the endeavor.

date	time	number of entries	
		on location	listening to the recording
03/04	early morning, sunrise	11	19
05/04	evening, sunset	14	19
06/04	evening, an hour before sunset	15	25
08/04	late afternoon	17	13
10/04	midnight	8	11
11/04	evening, two hours before sunset	19	12
14/04	late afternoon	18	12

Figure 3.1 – an overview of my visits

---

<sup>46</sup> For example, the scattered shallow lakes that have become home to a number of water birds and amphibians are a product of the water catchment activities (van der Meulen, Bakker & Houston 2008, 272), while the importance of the dunes as a coastal protection system has prompted planting of marram grass to prevent coastal erosion, although, this is no longer practiced (ibid., 273). Similarly, the more forested areas, like the one which became the location of my observations, are, in part, formed by human intervention (ibid., 269).



Figure 3.2 – on location, six different days in first half of April

Once I had ‘entered’ the place-to-become, I would begin taking notes while looking for a spot to position the recorder – a different one for each visit. From the moment a recording was begun, I would freely wander around the place observing whatever caught my attention and making entries in the field diary. I did not have an objective; there was no program that would determine what was or was not relevant to observe and note. The only two limitations were time and space restrictions – each visit would last for approximately an hour, and I would not wander away from the place by making sure that at any given time I could perceive it aurally. When forty-five, fifty minutes had passed, I would stop the recording and leave. Afterwards I would listen to the recording and take notes again thus producing a second layer of observations, which at a later time I would re-read while making comments on the themes found therein.

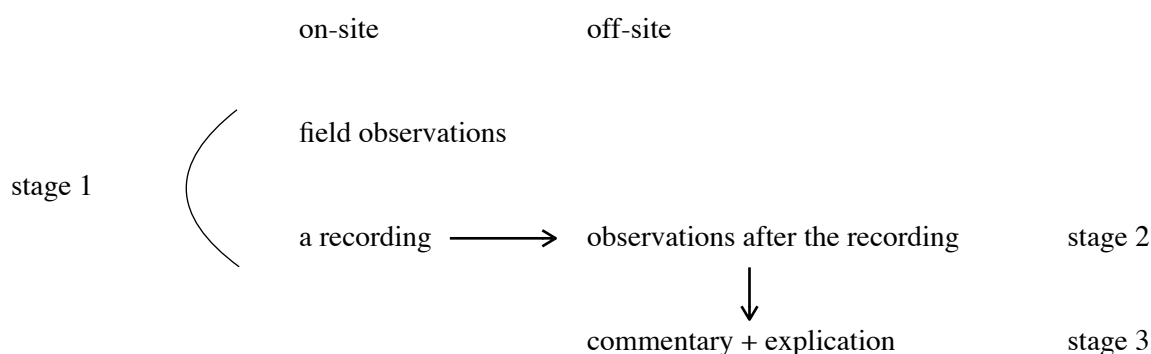


Figure 3.3 – a schema displaying the three-stage process of observation

Seven visits resulted in 306 minutes of field recordings, 102 entries in the field diary, a further 111 remarks made while listening to the recordings, and thousands of words developing notions found within the latter. There was no further plan, no intent to make use of the gathered materials. Yet the process of reflection did not end here; it could not, for the accumulation of thoughts, their connections, juxtapositions, contradictions, continued to continuously power ever more thinking. After having absorbed within itself the observations, thoughts, reflections that came to constitute the ‘quiet place’, it began to burst open, to project its contents outward, taking form of compositional ideas, critical insights, self-revelations. Most importantly, however, the ‘quiet place’ came to cast a different light on the main objective of the research – imitation of the nightingale’s song. I present the observations made on one of my visits and offer a loose reading of them to exemplify the ongoing thought process that these observations give rise to.<sup>47</sup>

<sup>47</sup> A sequence of excerpts from the recordings made on location can be found in the Appendix under the title ‘quiet place’.

field diary: 11th of April, evening

17:58 — approach along a new path, a pigeon just above my head, a crow in the distance

18:00 — church bells from afar

18:01 — a crow lands on a big barren tree calls or rather screams, takes flight

18:02 — trees crackling in the breeze

18:03 — I stop to listen in, again crows and pigeons, further afar the clear, yet subdued sounding of 'the place'

18:11 — a recording is begun

18:13 — a couple of birds (various species) plucking the green sprouts of the new leaves, flutter of wings, subdued, occasional vocalization

18:15 — a clear, precise trill

18:17 — a bird of prey silently flies above my head landing on a high pine tree, I move closer

18:23 — an airplane (a beautiful noise)  
(I wonder how many birds native to this place are killed in accidents each year)

18:28 — a deer feeding (it might be the same one I've encountered here twice before)

18:32 — it is rather small, freezes when I come closer, looks directly in the eye, still, not moving

18:37 — a small, dark rodent in last year's leaves

18:38 — (I wonder how many birds are still to return to nest, to sing and call through the summer)

18:41 — cawing again (I don't have a proper vocabulary for describing all this)

18:45 — navigation through

18:46 — three geese, only one calls

18:54 — I found the tree from which 'the soloist' rolls out its song

18:56 — a woodpecker

*The entries are short, laconic; they seemingly do not reveal much about what they account for. Yet, through abstracting from the concise particular observations, it is possible to derive some constitutive processes and elements of the place under surveillance. The crow's cawing and the pigeon's cooing – two different kinds of vocalization, two sound sources loosely located as distant (the crow), as sounding above (the pigeon) relative to the observer... Church bells from afar – a great distance, a signal from without enters the scene, its regular pattern and fixed, stable sonority in stark contrast to the quasi-chaotic processes of bird vocalization, wind induced shivering of leaves. Juxtaposition, reconfiguration of what is heard and how it is heard through this intrusion... The placement of a place. What allows one to hear a bell from a great distance has only partially to do with the bell itself, it also depends on the environment in which the hearing of a bell is possible – in this case a rather quiet evening hour... The crow again, its movement – flight; interruption – landing; articulation – calls and screams; continuation of movement – taking flight. A description of a pattern of actions, temporally, spatially, sonorously instantiated against the background of the continuous wind-induced crackling of the trees, its intensification and decline, and autonomy – a slow, unpredictable rhythm against which other rhythms are sensed and gauged... Sounding of the place in distance – a perception of it as separate from the larger territory within which it is embedded, a perception in part determined by conceiving of the place as separate from its surroundings... Flutter of wings; a trajectory of a sound source, the possibility of the birds vocalization to be heard in a while from a distance, the others, remain near by... A clear precise trill – an explicitly delineated sound event, effecting convocation: a sense of some sonic process becoming a centre of the audible; against it, in harmony with it, the other soundings are heard. A hierarchy is established within the place itself, through increased intensity of the now-central or its marked difference from the rest... The silent glide of a predator, its landing on a high tree, absence of vocalization, draws my interest and initiates a purposeful movement through the place... The above-ness of the bird of prey is matched by an airplane crossing over far above; its glide is far from silent. Airplanes through their incursion are as much part of the place as its flora and fauna. They appear as separate from it, as distanced, yet they appear as such within the place itself... The presence of a doe, it affects my behavior, especially movement, to a greater extent than most other inhabitants of the place. Its size, silence, gaze while not inducing fear, ordains respect... A rodent, heard, not seen, retreats away from the path I walk; rustling of dead leaves... This evening is marked by the caw of the crow, the same one I encountered upon my arrival; I will pass by it again when taking leave... Navigation through (the place, its sonorous activity, the corresponding trains of thought) is the manner in which I operate... Geese, they intrude violently, never landing, never having landed here. A place, even if its boundaries are mutable, has an identity constituted by what happens within it and its relation to what is outside of it – a distinction constituting the included and excluded. The geese are part of the place while never having been truly within it, they happen to the place... Even without objectives, objective discoveries are being made. Locating a tree that has become the stage of the daily performance of the soloist – a lone blackbird I encounter here time and again – will determine the placement of the recorder on my next visit... A woodpecker signals the end of my stay.*

## II : 3 – The Registered

Each entry found in the field diaries is preceded by an indication of the time at which it was made; thus they present an account which is not only sequentially organized, but also marks the intervals of time that have passed between them – a temporal structure occasioned by observation. This, however, is not the case with the notes made while listening to the field recordings – they are sequentially, but not temporally articulated, even though their production was likewise brought about by the temporal processes captured in the recordings. The entries are short, sometimes cryptic, sometimes almost nonsensical, sometimes of questionable relevance – an apt state of affairs given the lack of a systematic methodology. Yet, by penetrating through their apparently accidental nature, one can find schematic descriptions of those processes that constitute the place, and especially its sonorous appearance. Through abstraction the terse entries are broken down; the register and its components – objects (entities and processes), characterizations of them – are isolated, schematized, aligned with the like and juxtaposed to the contrary – a process not dissimilar to that of analysis guiding the imitation of the nightingale's song.

processes: pecking, walking, flying, vocalization, learning, change, running, moving, pulsing, being still, dodging, receding, approaching, taking turns, appearing, disappearing, returning, increasing, not knowing, remembering, swishing, searching, whistling, ascending, playing, flutter, laughing, descending, covering, decay, investigating, surveying, screaming, calling, falling, resonating, singing, sounding, reappearing, resting, squeaking, rustling, turning silent, stumbling, striking, ringing, blowing, hallucinating, noticing, landing, crackling, plucking, trilling, droning, feeding, staring, not moving, nesting, wondering, crowing, navigating, finding, phasing, hocketting, waving, settling, crossing, not singing

entities (perceived and thought): woodpeckers, tree, geese, volume, register, wood, sun, deer, shrubs, a breeze, blackbird, an unknown bird, something little, trees, birds, jets, boundary, what is not, not a nightingale, a small bird, skeletal tree, branches, a bird, its tail, one place, and another, two unknown birds, their wings, Messiaen, something, monumental, heavily branched trees, a territory, memory, dunes, settlement, mist, thorns, soil, leaves of another summer, children, beautiful, elusive, rather large bird, wind, people, unknown birds, leaves, the sea, grass, a forest, a bell, a bird, an owl, fallen trees, a pigeon, a crow, church bells, big barren tree, various species, new leaves, wings, a bird of prey, a high pine tree, last year's leaves, a path, flora, fauna, an array of yellow-breasted birds

characterizations: distant, surrounding, peripheral, quiet, irregular, alarming, aerial, again, quick, close, rapid, present, low, sudden, tangential, as if, new, quicker, louder, involuntary, more than, active, not far, out of, adjacent, not within it, same as, accidental, earlier, high, just like, different, slowly, swiftly, from-to, before, certain, absent, not present, voiceless, from without, short, as if delegated, cold, colder, like smoke, saw-like, probably, directly, for a moment, rather far away, not even close, from one side or another, loud, unusual,

above, in the distance, from afar, in, further afar, clear, occasional, precise, silently, beautiful, the same, twice, closer, through, sparse, uncommon, subdued, behind, in front, very quiet, with it, more, like sea, unsettled, not threatened, nearby, percussive, sometimes synchronous

I find connections between entities and processes, whether they are material or ideal; I find their characterizations, or how they appear to be. Something appears ‘distant’, something appears ‘clearly’, something appears ‘beautiful’, nothing that does not appear is documented. What appears, appears as and appears through. The mechanics of appearance are driven by hearing and listening, seeing, sensing, and thinking. The place that is described appears as described; it is composed through this description. It coincides with that which is described, yet is never identical to it. A place recognized as a place is already more than, and different from, that which is recognized. The place as created includes not only birds, and their vocalizations, trees and their slow movements in a breeze, but also the thought which coincides with these vocalizations and movements. This thought that does not know if what it thinks is a blackbird or a mistle-thrush, the thought that recognizes a woodpecker’s laugh, because of its realistic rendition by Messiaen in his *Réveil des oiseaux*. And just as I wander through what I believe to be a place, my mind wanders through perceptions, memories, associations; as my ear traces the pattern of a nightingale’s song, my mind unfolds a thought, it wonders, imagines how this song was taught, contemplates nature and its manner of operation, and extends the fluid boundaries of this place by doing so.

As noted above, each observation entered into the field diary is marked by the time of its inscription; the duration of my stay partitioned into segments of uneven extent and articulated by the observations. When through abstraction the relevant characteristics of each observation are derived, the textual record begins to resemble a register of approximate determinations pertaining to the generalized sonic processes encountered. These registers, supplemented by the reflective notes made while listening to the respective field recordings, offer loose outlines for possible compositional forms – a strategy I intend to explore in the near future. However, another more direct application of these registers for compositional purposes can be envisaged: one could treat the text as a script to be recited following the implicit temporal demarcations it contains. This, in its most basic form, would result in extremely sparse spoken-word piece, a strategy I explored in the early stages of the research when during the workshop *listening to the infra-ordinary* (led by Tao G. Vrhovec Sambolec), I created a registration/spoken-word piece entitled *4 hours (in which little is said, but much is meant and understood)* (see Figure 3.4). On the 31st of October, 2018, beginning at 10:30 every five minutes I would note the first process/thing relating to sound I would notice and in the form in which I would notice it: there is an entry ‘a monologue’ – a formal characteristic of the noticed, there is an entry ‘bless you!’ – a direct registration of someone’s speech. The entries were made wherever I happened to be at the moment, whether it was a classroom or a bench in the park. Two days later I proceeded to read out loud each entry at the precise time of the day it was made, again wherever I happened to be at the moment.

10:30 - TWO SIMULTANEOUS CONVERSATIONS, DIALOGUES  
 :35 - HEADPHONES, SOFT MUSIC (IS TURNED OFF)  
 :40 - TYPING, SILENCE (OR DISINTEGRATES IN LAUGHTER)  
 :45 - A MONOLOGUE  
 :50 - A DIALOGUE  
 :55 - A MONOLOGUE  
 11:00 - (SHY) (APPLAUSE) (AN INTENTION, NO EFFECT)  
 :05 - SOFT HANDCLAPS (APPROACHING)  
 :05 - FOOTSTEPS FROM WITHOUT THE RAMP, DOORS OPEN, SPEECH, DOORS SHUT.  
 :10 - ~~ASBEN~~ (DISTANT, APPROACHING, RECEDING)  
 :15 - A MONOLOGUE  
 :20 - SILENCE (TERMINATES IN SPEECH)  
 :25 - RAIN  
 :30 - A DIALOGUE (OR TWO SIMULTANEOUS MONOLOGUES)  
 :35 - LAUGHTER  
 :40 - A SPEAKER'S MOUTH SQUEAKING  
 :45 - SPEECH  
 :50 - BLESS YOU!  
 :55 - A MONOLOGUE, A POLYLOGUE  
 12:00 - A CROISSANT (CRUNCHING)  
 :05 - A PLASTIC BOTTLE, A WOODEN SURFACE  
 :10 - SUBVOICED CONVERSATIONS  
 :15 - A SIGH  
 :20 - SHOES (SLIDE OVER THE FLOOR)  
 :25 - PAPER (RUSTLING)  
 :30 - A DISTANT CONVERSATION  
 :35 - CAR ENGINE (CLOSE BUT QUIET)

(12) :40 - A TREE (DISTANT)  
 :45 - A VERY DISTANT CONVERSATION (MURMUR)  
 :50 - PIPES, RUNNING WATER  
 :55 - AIR CONDITIONING  
 13:00 - AN ENGINE  
 :05 - BIRD CALLS (CROWS?)  
 :10 - A TRUCK, PASSING BY  
 :15 - BIRD CALLS (CROWS)  
 :20 - A CAR, PASSING BY  
 :25 - CARS, PASSING BY  
 :30 - FOOTSTEPS, FOOTSTEPS  
 :35 - CONVERSATION ((IN) DUTCH)  
 :40 - FOOTSTEPS (RECEDING)  
 :45 - CONVERSATION (BEHIND CLOSED DOORS)  
 :50 - A HUM  
 :55 - A QUESTION  
 14:00 - WATER, SWALLOWING  
 :05 - CLOTHING, SWISHING  
 :10 - FOOTSTEPS (FROM WITHOUT)  
 :15 - A VOICE (HESITATING)  
 :20 - A LAUGH  
 :25 - FIST, CHEST, RHYTHM.

Figure 3.4 – 4 hours (in which little is said, but much is meant and understood)

What I find interesting about such 'music', if this label is appropriate, is the richness of the semantic content the sparse speech-events – which are the immediate sonic material of the piece – carry within them. In particular, the way that something said before casts a different light on what is said later – a very musical characteristic, for music, quite often, makes sense of itself through development, or, to put it in other terms, the 'sense' at any given time relies on what came before. Furthermore, the spatiotemporal placement of each speech-event contributes to the piece as much as its semantic content, for this content is juxtaposed with the conditions within which it appears. These conditions can be in alignment with conditions of observation, or they might not, the relationship between the two contributing to the resultant experience.

## II : 4 – A Strangely Indexical Mode of Listening

Something you cannot register as sound recording is the strangely indexical mode of listening that I time and again enter within the place. It might be a bird call that attracts my attention, I might respond by moving closer to the source; this shifts the auditory scene and marks a change within the place itself. Your movement can be perceived as dangerous by a nearby bird which you have not noticed; it might fly away – the wings flutter, your attention is caught again – an iterative process. The walking around and wandering is a way of listening, in part determined by the sudden appearance of a sonic process that was not present a moment ago. Its opposite, sudden termination, is more subtle and effects stillness rather than solicits movement. What happens when a certain sound-pattern disappears is not just a loss of acoustical information; rather, a redistribution of attention and restructuring of relations between the remaining constituents takes place and allows the heretofore inaudible to appear. Yet the liminal, that on the edge of inaudible, once heard becomes a point of attraction; it teaches you to be attentive, to listen more carefully, and this in return changes how that which is not liminal is perceived, you begin to perceive the minute within it, you notice ever more slight differences within the already familiar.

The two principles of appearance and disappearance, even if they effect differing responses, can be reduced to a more fundamental principle – both are kinds of change. And change has the quality of rendering discrete that which is continuous; the act of listening segments that which is heard into number of scenes, one after the other, interpenetrating; duality of continuance and discreteness. One grasps an identity through difference, the sudden realization that *this* is not the same as *that* which was before, so the discreteness of the ‘scene’ is established retroactively – it does not matter if it happens ‘right after’ or ‘some time after’ the scene thus grasped is no longer present – as a thought which for its object has something that no longer has the same existence as before. Yet, and this is crucial, this retroactive identification is made possible through the continuity and interpenetration with which the listener sides and tunes into when listening, as if attempting to become one with the act of hearing. One identifies oneself with the continuity – sounding, hearing – and so with change, which then is grasped retroactively as a memory of a past moment within the continued process of sounding. So, in a way, listening depends not on this pure identification with the sonorous through hearing, but on the inability to do so, on the resurgence of memory that intrudes in the mind that attempts to side with the ear.

There is no decision on where the cut between one scene and another is to be made, the cut becomes apparent only after it has taken place. This partially depends on the nature of that which is cut: the sonorous through being heard enters one’s mind and the changes within the former are changes within the mind that perceives. So a change in the sonorous can instantiate a cut. Yet the moment of the cut also depends on the natural history of the cutter: the sonorous is never the only content of a mind listening, and is never pure, but

is always more than itself when perceived, and the cut does not necessarily have to correspond to a change within this sounding content, but can originate in the change within one's mind as such.

The kind of listening just described illuminates an approach to composition that grounds the research as a whole. This listening, as manifested in the observations found in the field diary, proceeds through identifying – even if in a little too terse a manner – those processes that effectuate a shift within listening itself. The identification allows one to proceed and isolate these processes and deal with them on their own terms – separating them from that scene within which they are encountered, only to reintroduce them as autonomous elements within a scene re-composed: relations observed within listening are replaced by relations of another kind – they are constructed and thus more permanent, allowing the listening to cease while maintaining that ‘scene’ which listening created. Only to an extent, of course, for as noted earlier, many of the processes observed involve the observer as the interpreter of the inter-subjective, while many others depend solely on the subject’s ‘inner life’ – these processes are part of the compositional link, albeit in a different register. For example, when in the ‘quiet place’ at night, no birdsong was to be heard, except for a sporadic call every few minutes. The call itself is inter-subjective and as such permits its isolation, re-constitution and its eventual injection within a compositional schema. However, the appearance of the call as experienced by me on that night – not surprising, or even attention-grabbing, delicate, distant, non-intrusive – enters compositional practice as a question: how to introduce, contextualize appearance of something so unlike that within which it appears, while not creating a juxtaposition, but an (immediate) inclusion.<sup>48</sup>

*‘ of course, the light becoming ever dimmer is an aspect (plays a role) in listening ’*

*Alteration of lighting brings forth a number of significant changes: (1) the influence of lighting conditions on the flora and fauna is an influence on the sounding, its production; (2) a shift in the relations between senses, a change in functionality, a re-delegation of responsibilities. Both could be viewed as the same kind of process, for, in a way, I am part of the fauna, and a change in lighting conditions alters my behavior. However, the second kind of change involves something more than that, hence the emphasis on perception, the aesthetic. When listening to a recording, everything pertaining to the other senses – in operation when on location, is bracketed out, while retaining the influence on the sounding that these other modalities exert.<sup>49</sup>*

When listening to the field recordings, a different kind of listening is in play. Gone is the uninhibited movement through a place, often conditioned by chance encounters, and with its departure the ear is fixed in a certain position within the field, dependent on the placement of the recorder. A more limited domain

---

<sup>48</sup> This question is explored through experiments in nightingale’s whistle-tone imitation described in the previous chapter.

<sup>49</sup> From notes made while listening to a recording of the place (5th of April, 2019). The recording was made just after sunset.

sharpens the focus; happenings within it can be monitored over a more extensive duration, while the operation of framing a recording engenders divorces what it captures from the incessant sounding that is its object. The reason a recording instills a higher degree of composedness on that which it registers is, I presume, the fact of repeated mediation: first mediation – production of the recording, second mediation – the listener listens to a recording. Even though the mechanics of the two are necessarily different – recording is an instance of technological mediation, while listening could be described as phenomenal mediation – both are operations effecting spatial and temporal displacement.

It is through the operation of displacement that a recording lends itself to possible further alterations, and through them one can attempt to propose a certain kind of listening, difficult to capture in a recording as such. The above mentioned principle of change, the foundation of that kind of listening I was engaged with on location – more specifically, change in guise of disappearance – became a phenomenon I set out to formalize, so as to develop it into a compositional principle. I did so through a technique of subtle spectral manipulation of otherwise unaltered field recordings. The basic schema: a number of band-pass filters dissect the vertical dimension of the recording and alters the spectrum through gradual removal of these vertically differentiated streams.

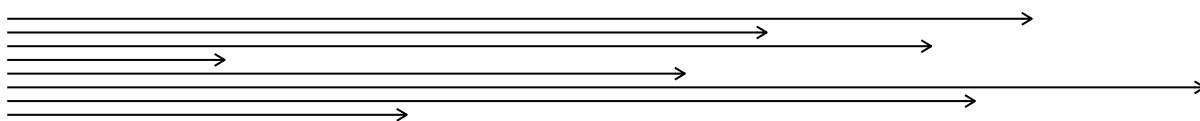


Figure 3.5 – an illustration of ‘subtractive filtering’

The number of band-pass filters, their centre frequencies and bandwidths, are all variable.<sup>50</sup> Different configurations of these variables produce diversely ‘colored’ subtractive processes. A variation of the procedure involves gradually shifting both the frequencies and bandwidths of the separate streams to effect a transfiguration of their relations, either through employing systematic and precise trajectories with predetermined start and end points, or through chance-governed, disaligned movement of the bands. At the beginning of the subtractive process, the original recording is not altered much. While its frequency spectrum has been partitioned into a number of bands, all these bands are present, adjacent to one another, resulting in a faithful synthesis. If the frequency axis is partitioned in, for example, bands of 1/6 of an octave across 9 octaves with the lower frequency limit at 35 Hz – and correspondingly the uppermost band filter having the frequency of 15965 Hz – a set of fifty-four bands is obtained. As shown in Figure 3.5, the subtraction process does not follow a regular pattern; each band is given a time bracket within which it is to drop out, while the actual moment of subtraction is left to a chance procedure. There is no coordination across the bands, and

<sup>50</sup> Although I mostly utilize different equal octave divisions (from 1/3 of an octave to 1/6 or 1/8) and calculate the bandwidths accordingly, while the number of bands used, in part, depends on the chosen octave division.

neither is there coordination across the two stereo channels – a band of a given frequency might drop out on the right channel, while remaining present on the left.

Each field recording produced on location – the quiet place – is more or less forty-five minutes long. Given that the frequency spectrum is divided in fifty-four bands and each of the bands is present in both channels, one hundred and eight subtraction events occur throughout the duration of the field recording – statistically, one subtraction every twenty-five seconds. However, the relative significance of a band dropping out increases as the number of bands present decreases, this results in an uneven distribution of significant change. Additionally, the relative significance of each band is conditioned by the frequency distribution captured in the recording. Experientially, the process unfolds through initial habituation to the field recording in its near-unprocessed state, within which subtle instantaneous shifts begin to be noticed, then, perhaps, a relatively significant frequency band drops out and reconfigures the auditory scene. By this time, the fact of spectral manipulation has become apparent, while the effect of reconfiguration is not lost. The subtractive process dissects the continuous stream of sound into a number of scenes, thus mirroring, albeit in crooked symmetry and with apparent technological mediation, the kind of listening I found myself in at the quiet place.<sup>51</sup>

---

<sup>51</sup> I employed the techniques described above in composing the piece *a corporate company for the propagation of beautiful but unreliable imaginings* for non-standard synthesis mechanism, field recordings and band-pass filters, premiered at the Sonology Discussion Concert on the 11 December 2019.

## CONCLUSION

The more closely we examine actual language, the greater becomes the conflict between it and our requirement. (For the crystalline purity of logic was, of course, not something I had *discovered: it was a requirement.*) The conflict becomes intolerable; the requirement is now in danger of becoming vacuous. – We have got on to slippery ice where there is no friction, and so, in a certain sense, the conditions are ideal; but also, just because of that we are unable to walk. We want to walk: so we need *friction*. Back to the rough ground! (Wittgenstein 2009, 51e [italics in the original, underlining is mine])

Wittgenstein's rejoinder to his past self resonates with the path my research took from its overly ambitious beginnings to its modest conclusion. I never imagined being caught up in an unending process of analysis, classification, formalization. These procedures were supposed to be expedient means *en route* to composition; the nightingale was supposed to be the first specimen of a populous synthetic bird choir. Yet something in form of a 'requirement for a crystalline purity' happened to be my stumbling block – the endeavor to discover the 'true' organizing principles responsible for the structure of the nightingale's song turned out to be, if not vacuous, then certainly somewhat misguided. Nevertheless, the failure of the 'requirement' was not without its merits, for it revealed the unacknowledged preconceptions of mine, and by doing so made me aware of both its inadequacy for the task at hand, as well as the reasons for its allure.

Investigations of the 'quiet place' began at a critical moment. After the initial failures at imitation I drew up the score and following its opening lines proceeded to retroactively establish the inaugural moment of the research by (re)turning to that place within which its origins were to be found. At the time, I thought it would be no more than a side-project, however, it introduced a shift in my conception of the attempt at imitation. The investigations effectuated a distinction between the place *as observed*, and the place *as constructed*; and the transposition of the distinction to the domain of imitation resulted in thinking the song's structure under the two modes. I recognized that the analytical apparatus – employed in the presumed discovery of the song's structure – in fact, constituted a certain projection of structure onto the phenomenon, one that would never be entirely reconciled with the song's structure *as unarticulated*, indifferent to my interpretation, objective. The analysis of the nightingale's song turned into an analysis of my own unacknowledged preconceptions; the 'true', not entirely revealed principles of the bird's vocal display were replaced by principles operating across the intersection between composition and observation. *My formalization of the nightingale's song – a speculative proposition that expresses the way I would have composed it.*

Much has been left unsaid. I have not engaged in any discussions regarding the practice of imitation in art and music, something I initially thought I would do. The names of Messiaen, David Tudor, Peter Ablinger, John Cage and many others, are absent from my account, yet they are not absent from my thoughts, and it is

to their work that I owe gratitude. Neither have I discussed the symbol-nightingale that has inhabited the poetic production of our civilization from times immemorial; again, not for a lack of interest, but rather because I myself fell victim to the nightingale's enchantment. Although the research has reached a tentative conclusion, its results have not yet been fully utilized in compositional production. Now I am in a position to do so, and it is the compositional work that will pass the last judgment as to the success or failure of the attempt at imitation and its method.

## BIBLIOGRAPHY

- Bartsch, C., Hultsch, H., Scharff, C. & Kipper, S. (2015). "What is the whistle all about? A study on whistle songs, related male characteristics, and female song preferences in common nightingales." *Journal of Ornithology*. Vol. 157, No. 1: 49-60.
- Birdsong, Speech and Language: Exploring the Evolution of Mind and Brain (2013). Edited by Bolhuis, Johan J. and Everaert, Martin. Cambridge, MA and London, England: The MIT Press.
- Deleuze, Gilles (1991). *Bergsonism*. Translated by Hugh Tomlinson and Barbara Habberjam. New York: Zone Books.
- Freud, S. (1937). "Analysis Terminable and Interminable." *The Standard Edition of the Complete Psychological Works of Sigmund Freud, Volume XXIII (1937-1939): Moses and Monotheism, An Outline of Psycho-Analysis and Other Works*, 209-254. London: The Hogarth Press and The Institute of Psycho-Analysis.
- Grant, I. H. (2013). "How nature came to be thought: Schelling's paradox and the problem of location." *Journal of the British Society for Phenomenology*, Vol. 44 (1): 24-43.  
Available: <http://eprints.uwe.ac.uk/19351>
- Grant, M. J. (2011). "Series and Place". *Contemporary Music Review* Vol. 30, No. 6: 525-542.  
DOI: 10.1080/07494467.2011.676899
- Hegel, G. W. F. (2018). *The Phenomenology of Spirit*. Translated and edited by Terry Pinkard. Cambridge: Cambridge University Press.
- Hultsch, H. & Todt, D. (2001). "Developmental Trajectories of Complex Signal Systems in Animals." In *Approaches to Bootstrapping: Phonological, Lexical, Syntactic and Neurophysiological Aspects of Early Language Acquisition Vol. 2*. Edited by Weissenborn, J. & Höhle, 309-311. Amsterdam: Benjamins.
- Hurford, James R. (2012). *Language in the Light of Evolution, Volume 2 – The Origins of Grammar*. Oxford: Oxford University Press.
- Jakobson, R. (1978). *Six Lectures on Sound and Meaning*. Translated by John Mephram. Cambridge, MA and London, England: MIT Press.
- Jakobson, R., Fant, C. G. M. & Halle, M. (1963). *Preliminaries to Speech Analysis: The Distinctive Features and Their Correlates*. Cambridge, MA: The MIT Press.
- Kafka, F. (1982). *Letters to Ottla and the Family*. Translated by Richard and Clara Winston. Edited by N. N. Glatzer. New York: Schocken Books.
- Kunc, H. P., Amrhein, V. & Naguib, M. (2005). "Acoustic feature of song categories and their possible implications for communication in the common nightingale (*Luscinia megarhynchos*)." *Behaviour*. Vol. 142, No. 8 (August): 1077-1091.

- Mercado, E. III (2018). "The Sonar Model for Humpback Whale Song Revised." *Front. Psychol.* 9:1156.  
DOI: 10.3389/fpsyg.2018.01156
- Meulen, F. van der, Bakker, T.W.M. & Houston, J.A. "The Costs of Our Coasts: Examples of Dynamic Dune Management from Western Europe." *Ecological Studies*. Vol. 171: 259-277.
- Pliny the Elder (1855). *The Natural History*. Book XXXV, Chapter 36. Translated by Bostock, John and Riley, H. T. London: Taylor and Francis.  
available: <http://data.perseus.org/citations/urn:cts:latinLit:phi0978.phi001.perseus-eng1:35.36>
- Riede T. & Goller, F. (2010). "Peripheral Mechanisms for Vocal Production in Birds – Differences and Similarities to Human Speech and Singing." *Brain and Language*. Vol. 115, 1 (October): 69-80.  
available: <https://doi.org/10.1016/j.bandl.2009.11.003>
- Song, H. 2008. "*Luscinia megarhynchos*" (On-line), Animal Diversity Web.  
Accessed May 22, 2020 at [https://animaldiversity.org/accounts/Luscinia\\_megarhynchos/](https://animaldiversity.org/accounts/Luscinia_megarhynchos/)
- Todt, D. (2004). "From Birdsong to Speech: A Plea for Comparative Approaches." *An. Acad. Bras. Ciênc.* 76 (2): 201-208.  
available: <http://dx.doi.org/10.1590/S0001-37652004000200003>
- Todt, D. & Huttsch, H. (1998). "How songbirds deal with large amounts of serial information: Retrieval rules suggest a hierarchical song memory." *Biological Cybernetics*. Vol. 79: 487–500.
- Weiss M., Hultsch H., Adam I., Scharff C., Kipper S. (2014). "The use of network analysis to study complex animal communication systems: a study on nightingale song". *Proc. R. Soc. B* 281: 20140460.  
available: <https://doi.org/10.1098/rspb.2014.0460>
- White, Gilbert. (2016). *The Natural History of Selborne*. Edited by Anne Secord. Oxford: Oxford University Press.
- Wittgenstein, Ludwig (2009). *Philosophical Investigations*. Translated by G. E. M. Anscombe, P. M. S. Hacker and Joachim Schulte. Edited by P. M. S. Hacker and Joachim Schulte. Chichester, UK: Blackwell Publishing Ltd.

## APPENDIX

The appendix includes four sound files:

1. 'nightingale' – an excerpt from the recording of the nightingale's song I analyzed
2. 'whistles' – a sound file that includes the synthesized whistle-tones in isolation and against a backdrop of a field recording
3. 'synthesized nightingale' – a short excerpt presenting the non-linear synthesis mechanism guided by the nightingale song's structure
4. 'quiet place' – a sequence of one minute excerpts from the field recordings made in the 'quiet place'