Aural Identities

Auditive representations of ethnicity in documentary films

MIKI BRUNOU
Aural Identities

Auditive representations of ethnicity in documentary films

Master’s Thesis

Thesis Advisor
PÄIVI TAKALA

Thesis Supervisor
ANTTI IKONEN

Aalto University
School of Arts, Design and Architecture
Media Lab Helsinki
Master’s Degree Programme in Sound in New Media
April 2017
Abstract
In this thesis, I’m discussing the topic of representation in documentary films from the standpoint of audio post production and score music composition. My specific concern is how ethnic and cultural characteristics are expressed and denoted through the narrative conventions of sound design and score music. The case study of this thesis is a feature-length documentary film *Between Rings* (2014), that was filmed in Zambia by a Finnish-Zambian production crew. Although the topics discussed in this thesis are based on a theoretical framework of academic film studies, I’m also examining practical post production techniques and aural denotative practices related to representing social actors and cultural environments in nonfiction films. Here, my purpose is to provide the reader a comprehensive review of frequently employed methods of auditive narration practices found in contemporary film productions, while addressing the characteristic ethical indifference that has been permeating the auditive signification processes in nonfiction films. It is being noted how audio narration has not been receiving the theoretical attention it deserves, contrary to the abundance of critical discussions related to the topic of visual representation.

The aim of the thesis is to present arguments against the persistent notion of sound design and incidental music being essentially non-representative forms of cinematic narration. As long as they are not thought to directly contribute to the depictions of social actors represented in documentary film narratives, their denotative mechanisms will be excluded from the beneficiary influence of critical examination, thus increasing the possibility of misrepresenting individuals and communities depicted in documentary films.

Keywords  documentary film, representation, sound design, score music, audio post production, authenticity, ethics, ethnicity.
**Tekijä**  Miki Brunou  
**Työn nimi**  Aural Identities: Auditive representations of ethnicity in documentary films  
**Laitos**  Median laitos  
**Koulutusohjelma**  Sound in New Media  
**Vuosi**  2017  
**Sivumäärä**  108  
**Kieli**  Englanti

### Tiivistelmä

Yhtenä työtäni ohjaavana pyrkimyksenä on tarjota lukijalle havainnollinen läpikäyttöä äänilaisista kerrontakeinoista, sekä kartoittaa dokumentaaristen vastakouluksien taikaa avata uutta eettistä ongelmankenttää. Tarkastelen lisäksi sotakuvaaään jälkikotouttoprosessien käytännöllisyys joka ongelmaraati kuinka se kulkee työelämään ja autentisoiden kykyyn. Painotan aineistoni analyysissä kuvakerronnan ja äänikerronnan eettista yhteismitattomuutta, sekä että on olemassa asemia, jotta on olemassa sotakuvanderontoon arvoon ja hierarkiaan, mikä heijastaa osaltaan elokuvanterkijöiden sisällöllisiä luontoja ja elokuvan toimintaan rakentumiseen jälkikotouttoväkeessä.

Opinnäytteen kokoelma aikoinaan on kyseenalaistaa ja purkaa käsitteistäääänisuunnittelusta ja elokuvamusiikista ei-representatiivisia kerrontakeinoja. Niin kuin sotakuvaaään ei voita vaikuttavan kuvakerronnan tavoitteen dokumenttaistien eettisiä ulottuvuuksiin sekä kykyyn kykyä autentisesta kodellisuutta tehtävään, että myös musiikin suhteita elokuvanteostoa riippumattomaan kodellisuuteen eivätkä tule ymmärtää. Joissakin tapauksissa tätä saattaa johtaa väärenymmäryksiin sekä dokumenttielokuvassa kuvattujen kulttuureiden ja tosiasiallisten henkilöiden kaltoinkohteluun, kuten pyrin mainitsemissä esimerkkien avulla osoittamaan.

**Avainsanat**  dokumenttielokuva, representaatio, äänisuunnitteluluku, elokuvamusiikki, autenttisuus, etiikka, etnisyyts.
# TABLE OF CONTENTS

1. INTRODUCTION ................................................................. 10

2. CONSTRUCTIONS OF REALITY ............................................ 13
   2.1 Discourses of sobriety ................................................. 14
   2.2 Voices and signifiers ................................................. 22
   2.3 The phantom of authenticity ....................................... 31
   2.4 Undefined ethics ..................................................... 40

3. BETWEEN RINGS ............................................................... 45
   3.1 From rags-to-riches .................................................. 47
   3.2 Finding the right tone .............................................. 50
      3.2.1 Synthesized signifiers .......................................... 53
      3.2.2 Refining the score ............................................. 57
      3.2.3 Appropriations of style ........................................ 60
      3.2.4 Authenticity in virtuality ..................................... 62
   3.3 The broken indices ................................................... 65
      3.3.1 Atmospheric tracks ............................................... 66
      3.3.2 Sound effects and foley recordings ......................... 72

4. CRITICAL APPROACHES .................................................. 78
   4.1 Rendered soundscapes .............................................. 80
   4.2 Authentic appropriations .......................................... 86
   4.3 Exotization and objectification by music ........................ 93

5. CONCLUSIONS ................................................................. 100

REFERENCES ......................................................................... 104
1. INTRODUCTION

In this thesis, I’m discussing the themes of ethnicity and auditively constructed identity in documentary films, and how they are expressed through score music and sound design. Whereas representation as a visual issue has been thoroughly discussed in academic studies dealing with photography (e.g. Sontag 1979), film (e.g. Nicholls 1991), and audiovisual arts in general, the auditive signification practices in contemporary film narratives have not been receiving enough critical attention. Instead of trying to cover the topic from a strictly theoretical perspective, I’m approaching the problematic issue of denoted ethnicity by examining empirical material gathered from the design part of this thesis, examined more closely in chapter three. These examinations will be addressed through a theoretical framework listed at the end of this paper. As I’m addressing the topic from the vantage point permitted by a thorough analysis of a hands-on case study, I’m offering the reader insights related to audio narration techniques in films, but also a critical review of dramaturgically constructed character identities and representations of ethnicity in documentary films.

The design part of this thesis involved composing the original score music and conducting the audio post production process of a feature-length documentary film Between Rings (2014), depicting the story of a Zambian boxer Esther Phiri. The film was produced by a Finnish production company Helmi Films, and directed in collaboration by Jessie Chisi, herself a native Zambian, and Salla Sorri, a Finnish filmmaker. As the supervising sound editor and composer of the film, I encountered numerous instances where the guidelines of truthful representation and authentic qualities of audio narration were called into question. This unusual arrangement included a plethora of artistic challenges, entailing a considerable responsibility in terms
of sheer work and technical considerations, but also in doing justice to the individuals who had agreed to lend their lives for the film.

I will start the strictly theoretical part of this thesis in chapter two by introducing the essential concepts related to this study. I will go through conceptual definitions of documentary films, referring to the writings of filmmakers and theoreticians such as Jouko Aaltonen and Bill Nichols. I'm also addressing the problematic concept of authenticity, as it is constantly appearing in discussions related to documentary films. Here, the emphasis will be on reviewing the discursive practices and denotative methods found in contemporary nonfiction films.

Besides approaching the topic of auditive representation by referring to the academic studies listed at the end of this paper, I'm analyzing the workflow and the artistic outcome of *Between Rings* by engaging into a self-reflexive process, examining the various stages and a myriad of compositional and editorial decisions that took part during the post production of the film. I will be describing these processes in detail in chapter three, starting by summarizing the storyline of the film and outlining the personal history of the protagonist Esther Phiri. From there, I will proceed to explain the stylistic starting points that governed the composition process of the score music, addressing the distinct challenges and preconditions of this particular production on the prospect of signifying ethnic and cultural features through recorded sound. After examining the preconditions and artistic goals of *Between Rings*, I will be discussing the initial demoing phase of the score music, proceeding to analyze and explain why various aesthetic and narrative decisions affecting the representational outcome of the film were made in the first place.

In chapter four, I'm investigating the ethical implications of auditive narrative practices, referring to the examples given in the earlier chapters. I'm beginning my analysis by examining the narrative rendering of nonverbal sounds in films—referring extensively to the writings of Michel Chion. Here, my point is to demonstrate how the perceived realism of audio narration in nonfiction films is *not* dependent on actual profilmic qualities of various aural
environments or physical actions per se, but rather on their artistic rendering in relation to other elements in the narrative. After reviewing especially interesting examples of audio post production practices employed in *Between Rings*, I’m proceeding to discuss the aesthetic decisions made in the composition process of the score music from the perspective of cultural appropriation. I’m specifying the musical references to various African cultures found in the score of *Between Rings*, as well as discussing the general position of music as a self-contained unit within a cinematic narrative. My aim is to demonstrate how seemingly innocuous decisions related to the instrumentation and aesthetic direction of the score music can lead to misrepresentation—especially in the ethically sensitive context of documentary films.

Although score music can be rightfully examined as an integral part of the overall audio narration of film soundtracks¹, I’m reviewing the methods of musical composition separately from practices of nonmusical sound design. This is mostly a practical decision, as the theoretical frameworks for these distinct creative practices rely on different terminology, being dependent on specific discourses of their own. Here, the aim is to avoid unnecessary conceptual confusion, not to reinforce conservative and somewhat redundant dichotomies between music and noise.

¹ In this thesis, the concept of *soundtrack* is referring to the totality of mediated sounds (including, e.g., nonverbal noises, score music, and dialogue) accompanying the visual narrations of a given film.
2. CONSTRUCTIONS OF REALITY

*Art is not a mirror held up to reality
but a hammer with which to shape it.*
—Bertolt Brecht

The purpose of this chapter is to provide the reader a conceptual understanding of the core issues discussed in this thesis. Besides discussing the denotation process of nonfiction narratives from the perspective of recorded and post-processed sound, I will be addressing the key conceptions discussed in this thesis, such as *documentary film, representation, reality,* and *authenticity*. My aim is to demonstrate how difficult it is to find univocal definitions for these concepts, although they are often discussed as if they would be ubiquitously understood regardless of language, individual discursive practices, and specific cultural context. The notions of filmic representation and authenticity in documentary films will be treated thoroughly in relation to the case study of *Between Rings* (2014) in chapter three, whereas in this chapter they will be covered from the broader perspective of analytical film studies.

Instead of discussing sound design in documentary films from a technical point of view, I'm mostly addressing the preconditions of representation through audio: how can sound signify and denote people, cultural environments, actions and other phenomena belonging to a historical reality in the context of documentary films. I'm mostly interested in the genre-specific requirements of authenticity demanded from nonverbal audio narration in documentary films, especially when compared to the expectations set for image-based narration, and their relation to *profilmic* events. The

---

2 The concept of *cultural environment* is referring to the immediate socio-cultural surroundings of a particular individual or a community. It is inherently linked to questions of identity and social cohesion.
interplay of meticulously constructed nonfiction narratives, and their subjects belonging to a historical reality is justly subjected to ethical problema:ization exactly due to the assumed transparency of the documentary form: how a documentary film supposes a direct indexical relationship between a representation (a film) and the represented (a historical occurrence or a social actor). When factual people permit themselves to be portrayed within the medium of documentary films, a compassionate and ethically aware approach is demanded from the filmmakers in order to avoid unintentionally offending or misrepresenting their subjects.

Although the particular topic of this thesis is to discuss how ethnicity as a characteristic is denoted through sound in documentary films, I find it necessary to approach the representative practices and preconditions found in nonfiction films also from a broader theoretical perspective. This is to provide a reasonably comprehensive conceptual backdrop for the topic of auditive representation and audio signification processes within the overall narrative structures of documentary films, and to point out how isolated analyses of individual components of film narratives lead to strictly formalistic treatments that have little significance outside the niche of specialist discussions.

I will not be bringing up individual examples found in nonfiction films or addressing the particular case study of the thesis in this chapter. The critical treatment of *Between Rings* and the analysis of its narrative methods will be thoroughly given in chapter four.

### 2.1 Discourses of sobriety

Documentary films and other non-fictional forms of cinema are often defined through their argumentative tendencies. Whereas the majority of fiction films are aptly defined through their entertainment functions—expressed through high production values and emotionally engaging storylines, scripted with surgical precision—documentary films have a tendency of addressing various
social issues in an investigative and informative way; often at the expense of audiovisual flashiness. Instead of featuring schooled actors and actresses, documentary films are employing social actors belonging to a historical reality. John Grierson, a Scottish filmmaker widely considered as one of the originators of the term ‘documentary film’, saw the medium primarily through its great instrumental potentiality in influencing pressing issues such as poverty, unemployment, and the advancement of democracy in the harsh social climate of the 1930’s (Aaltonen 2006, 36). Grierson (as cited in Nichols 2001, 24) famously defined documentary films as “creative treatment of actuality”, likening them to other visual arts despite their outspokenly practical purpose in addressing sociopolitical issues and educating the public to ensure their capability in participating in the still emerging civil society. Jouko Aaltonen (2006, 36) writes how one of Grierson’s groundbreaking achievements was in gathering the stylistically varied early experiments of nonfiction films under the umbrella of his conceptual definition. Aaltonen (ibid.) points out how Grierson is arguably accountable for the way how contemporary audiences assume the direct indexical relationship of documentary films with the historical reality they’re depicting. Although Grierson’s famous definition of documentary films has been subjected to criticism for its simplistic synthesis of such mutually contradictory terms as “creative” (expressive) and “actuality” (factual), he’s still regarded in relatively high esteem for formulating the ethos that is largely determining the methods available for a documentary filmmaker in the eyes of the general public (Aaltonen 2006, 36–37).

Bill Nichols (1991) defined documentary films as “discourses of sobriety” in his renowned study *Representing Reality*. He expanded the formulations of Grierson by emphasizing the characteristic directness and investigative approach of many documentary films: either detachedly observing their subjects or engaging with them in a dialogous relationship. Documentary films aiming for informatively representing factual occurrences are “scheering”, as Nichols (1991, 3–4) asserts, “because they regard their relation to the historical reality as direct, immediate, transparent. Through them,
power exerts itself.” Nichols (1991, x) further enforces the view of the documentary form as a vehicle for transmitting information, by stating how “a good documentary stimulates discussion about its subject, not itself”, inadvertently likening the form and function of documentary films into newsreel inserts or other journalistic devices. Nichols (1991) appointed the different approaches to documentary filmmaking as modes, each of them standing for a generalized disposition for constructing nonfiction narratives, the taxonomy he later expanded and revised in his *Introduction to Documentary* (2001). The modes proposed by Nichols are poetic, expository, observational, participatory, reflexive and performative. What all of these distinct approaches have in common is the centrality of an argument in the narrative, and the expectancy of indexicality affirming the documentary value of the film. The modes discussed by Nichols comprise specific conventions that a documentary filmmaker may adopt, whether intentionally or not. While they provide “specific expectations viewers anticipate having fulfilled”, the filmmaking modes proposed by Nichols significantly contribute to the way how documentary films are experienced (Nichols 2001, 99). For example, if we’re watching an observational documentary film where the explicit interventions of the filmmakers are seemingly nonexistent, we instinctively expect the events depicted in the narrative to unravel as they would have been unravelling even without the presence and participation of the filmmakers. On the other hand, a film representing the poetic mode might suggest a more interpretative and expressive connection with profilmic phenomena, as the stylistic features of the film are perceptibly foregrounded. Like Nichols (2001) is proposing, adopting certain modes of representation has a profound effect on the way how audiences experience the authenticity effect of the filmic narrative.

Nichols’ arguments presented in *Representing Reality* could be criticized for their reductiveness in front of a polymorphic art form, comprising a diverse array of narrative methods and aesthetical approaches. Although Nichols is admitting the individual stylistic variations of documentary films by categorizing them into specific filmmaking modes, contemporary
documentary films rarely represent a certain mode without featuring discernible traces of neighboring modes. Nichols is stressing the informative aspect as a definitive feature of all documentaries, although it's fairly obvious that many artistic documentary films are not particularly suited for imparting factual information—especially when compared to broadcast journalism and other journalistic devices. To present documentary films as vehicles for 'soberly' addressing issues of the societal sphere, means, by necessity, to exclude and ignore the essential characteristics of many documentary films that have a markedly different relation to the aesthetics of nonfiction filmmaking. Pooja Rangan (2014, 2–3) criticizes the assumed ‘sobriety’ of documentary films and its implications to relations between a filmmaker and her subject:

The discourse of sobriety appears - - not merely as an ideological ruse, but as the symptom and mandate of a humanitarian ethical paradigm—one in which the urgent, immediate task of saving human lives legitimates and even actively defers all other considerations, including the aesthetics and politics of representation. As a medial idiom in which aesthetic concerns are subordinate to the task of touching the real, documentary realism is uniquely aligned with the ethics of humanitarian intervention.

By questioning and criticizing the ideological bravado of neo-Griersonian filmic discourse, Rangan is likening the process of outspoken anti-aesthetic and rigorously subject-driven filmmaking to a colonialist pursuit that reinforces hierarchies and patronizes its subjects as much as it manages to actually aid them. In her reading, the boldly stated commitment ‘to give a voice to the voiceless’ by casting them as the subjects of a documentary gaze is enforced by the modest production values of the audiovisual narration. Like Nichols (1991, 179) asserts; “Engagement is the aim more than pleasure.” Here, the underlying assumption is that in order to avoid distracting the audience from experiencing the authentic and direct presence of the

---

3 Here understood as the tradition of documentary filmmaking that “provides solutions to social problems through overarching narratives that exploit the ‘transparency’ of the camera to promote bourgeois values, beliefs, and interests”, like Lucia Ricciardelli (2014) writes.
represented other (as expressed through the narrative tools and voice of the filmmaker acting as a transmitting vehicle), the narrative mechanisms have to be immediate and untampered by discernible artistic treatment.

Although Nichols defined the filmmaking approach advocated by Grierson as expository, characterized by a conviction that “the world is productive of facts and that those facts can be communicated to others in a transparent way”, like Lucia Ricciardelli (2014) writes, there are many similarities linking the assumedly objective approach to other documentary filmmaking modes, such as the observational mode. Ricciardelli (ibid.) points out how the expository mode is sharing a lot of fundamental assumptions with other “realist approaches” to documentary filmmaking that tend to see the camera as a “window on the world”; as if the indexical nature of the cinematographic image would be sufficient in itself to guarantee objectivity and affirm the filmmaker’s rendering of external reality. Features such as the preference for wide and long shots, shallow focus, hand-held cinematography, sparse and straightforward editing, categorical favoring of sync sound (i.e., the production sound recorded on filming location), and the avoidance of non-diegetic\(^4\) sound design are all seen as testimonial proofs of the sobriety of the documentary discourse at work, characterizing both expository and observational modes of documentary filmmaking. Rangan (2014, 3) continues:

The aesthetic imperfection, poverty, and disposability of the discourse of sobriety indicates that its value is conceived in terms of immediate use rather than exchange or pleasure; it calls for a return to primitive, essential, and basic concerns. Such aesthetic austerity performs the filmmaker’s solidarity and commitment to the grim realities of the other’s (perceived) impoverishment, and demands an analogous sacrifice of the viewer’s spectatorial pleasure.

Rangan’s critique is supplemented by Aaltonen (2006, 39), when he notes

\(^{4}\) Music that does not occur as part of the action, and cannot be heard by the film’s characters; i.e. background or score music.
how Nichols’ idea of “sobriety” is mostly applicable in the context of conventional nonfiction narratives, as it is bearing clear traces of the expository and observational approaches of addressing the viewer with only minimal stylistic distractions, such as non-diegetic sound design or elaborate cinematography. Aaltonen (2006, 39) writes—referencing Nichols’ trail of thought—how “contemporary documentary film is no longer presenting a uniform world and knowledge”, but instead “incompleteness and uncertainty, recollections, mental images and impressions”\(^5\), and other constructions of clearly subjective nature. This is due to the epistemological re-evaluation of the documentary form that followed the semiotic turn in cultural studies (Aaltonen 2006, 37). The distinctively postmodern idea that everything is reducible to a sign in a language-like structure gained popularity also in film studies during the 1980’s, which lead to a widespread questioning of the uniformity of such fundamental concepts as reality, beauty or truth: all of them related to both epistemology and the aesthetics. Due to this paradigmatic shift, it is none more meaningful to talk about realism in the context of documentary films than it is in the case of fictional narratives. It is no longer plausible to discuss ‘real’ as something fixed and permanent, free from personal values, interpretations, perceptions and cognitive processes of an individual. (Helke 2006, 17.) Instead of a fixed perception of the real that we all share, there are a myriad of realities, each defined by the particularity of the cognitive subject and her perceptions of the world. The definitions get even more elusive and vague when the concept of reality is discussed in the context of artistic representations such as documentary films, constructed by individuals who often have an instinctive drive for unbiased and outspokenly objective stance in depicting occurrences taking place in the historical world. Filmmaker and theorist Susanna Helke (ibid.) is explaining her way of drawing a clarifying line between a filmic representation and its factual source in the world that predated its creation:

As I’m discussing methodological approaches or the

---

\(^5\) Translated by the author.
relationship between a cinematic style and the depictions shown in a film, I’m understanding reality (whether it’s social, historical, physical, immediate reality) as something that is happening in relation to the act of filming despite the act of filmmaking itself.\(^6\)

Here, Helke (2006, 17) is using the concept of *profilmic* to define a historical reality that is autonomous from indexical representations constructed from it. Although Nichols is not employing the concept of profilmic as such, he’s sharing the idea of a ‘directly’ representable profilmic reality by suggesting that “what we see [in a documentary film] would have occurred in essentially the same manner if the camera and tape recorder had not been there” (Nichols 1991, 27). However, this definition of profilmic reality does not mean that the existence of such an autonomous and fixed reality could be ‘transparently’ representable. This is arguably a fundamental precondition of representation, also in regard to sound recording. As anyone who has ever recorded a sound will know, a recording of the original sound is not the same exact sound that was being recorded. When being converted from atmospheric pressures (generated by sound waves travelling through different propagating mediums, such as air) into variations of electric current by the diaphragm of a microphone, the resulting recording is a mediated approximation of the original sound, not a copy of it. What was a profilmic sound event, becomes an audible approximation of the same event when recorded by a microphone.

Helke has been avidly questioning the informative aspect of creative documentary films, stressing the narrative and aesthetical similarities of many documentary films with their fictive counterparts, while simultaneously distancing the conventions of documentary films from journalistic documentaries produced for television. Helke (2006, 18) points out how profilmic events and characters can be considered as expressive material for documentary filmmakers—rather than them essentially constituting the films with their position as ‘subjects’. Nichols reinstates Helke’s view when

\(^6\) Translated by the author.
discussing the poetic mode of documentary filmmaking. He (2001, 102) asserts how poetic documentaries often employ actual people “on a par with other objects as raw material that filmmakers select and arrange into associations and patterns of their choosing”, distancing the act of representation from such ‘sobering’ qualities as directness, immediacy, and transparency, as quoted earlier in this chapter.

Besides examining the expressive facet of poetic documentary films, Nichols (1991, 10) is also pointing out the ideological aspect of all representations by writing how documentary films “are part and parcel of the discursive formations, the language games, and rhetorical stratagems by and through which pleasure and power ideologies and utopias, subjects and subjectivities receive tangible representation”—emphasizing the discursive power of all filmic representations. By admitting the inescapable subjectivity of all documentary films, Nichols is contradicting the very sobriety of non-fiction narratives. Due to this underlying contradiction, documentary films are more easily definable as subjective arguments or artistic studies presented in an audiovisual form, rather than them being direct reflections of an external profilmic reality. However, nonfiction films aligned with the expository and observational modes of documentary filmmaking are often—whether intentionally or not—trying to prove the exact opposite by emphasizing their ‘direct’ and strictly observational approach to representing profilmic occurrences.

The cardinal problem with assumptions of directness and ‘transparency’ lies in the methodic concealment of the filmmaker’s agenda. If the artistic influence of the filmmaker is methodically concealed, the arguments presented in the narrative (whether explicitly or implicitly) are easily perceived as indexical revelations rather than artistic expressions of an individual. This is increasingly problematic due to the inevitable inequalities in power relations between the subject (filmmaker) and the object of representation (a social actor). This disparity of power is further amplified by the privileged position of a filmmaker when entering an impoverished nation or an indigenous community, often backed by prominent Western
institutions—e.g., film boards, production companies, universities, and other establishments of institutional power—granting the filmmaker an upper-status in relation to social actors who might have next to no power as they are targeted by the representative aspirations of cultural outsiders.

2.2 Voices and signifiers

Like suggested in the previous chapter, documentary films are often defined through their argumentative tendencies, whereas sound design and score music are generally thought to address the auditor through imparting affective information and eliciting emotive responses from the audience (see e.g., Chion 1994; Henley 2010; Iversen 2010). How does sound design (i.e., the manipulation and processing of nonmusical\(^7\) audio content within the soundtrack of a film) and score music construe arguments in documentary films? How can audio narration—i.e., the nonverbal noises heard on the soundtrack of a film—be utilized to express the ideological voice of a film and direct the way how audiences interpret scenes and characters presented to them?

Michel Chion (1994) is pointing out the secondary position of auditive narration in films by discussing the *added value* of sound in relation to images. He (1994, 21) asserts how sounds and images—when juxtaposed with each other in an edited sequence—have a deeply reciprocal relationship: “Sound shows us the image differently than what the image shows alone, and the image likewise makes us hear sound differently than if the sound were ringing out in the dark.” This is what Chion refers to as *audiovisual contract*, enabling the narrative power of audiovisual continuity (i.e., the illusory unity of a filmic scene) of sounds and images in an edited sequence. Although sounds and images compliment each other in cinematic narratives, nonverbal

\(^7\) Here, the term *nonmusical* is referring to all other audio content in the soundtrack besides diegetic (source) and non-diegetic (score) music.
sounds are perceived (as designated by cultural and habitual conventions) as secondary to the images in the overall cinematic experience, like the term ‘added value’ strongly suggests. Chion points out how the visual frame is always the focal point of audience’s perceptions—drawing sounds emanating from the speakers towards its magnetizing center. Despite what imagery is presented in its confines, the visual frame is the center of attention in a cinematic experience. Because of this, the accompanying sounds will consequently be perceived in relation to it.

In spite of its position as a secondary means of addressing the audience, dramaturgically treated audio is capable of communicating more than purely emotive or mood-related information in film narratives—no matter if we’re discussing nonfiction documentaries or Hollywood blockbusters. Unfortunately, the narrow conception of sound design as a vehicle for expressing moods and other affective facets of cinematic narration has lead to the overemployment of human voice as a testimonial agent dominating the soundtracks of numerous documentary films (easily encompassing all of Nichols’ modes besides performative and poetic modes). This is endemic in the prevalent filmmaking paradigm, where nonverbal sounds are not considered as narrative devices per se, but rather like affective supplements that mainly serve to add a layer of post production sheen on top of the dramaturgically completed film, as it is handed over to the sound designer. This negligence has lead to a situation where the incorporation and design of narrative audio compositions is not subjected to the same critical assessment as visual storytelling devices—a phenomenon painfully evident in the post production processes of many documentary films (e.g., Aaltonen 2010). The narrative capabilities of sound design are as also underrepresented in the critical academic studies discussing the problem of filmic representation.

Before discussing this anomaly more thoroughly in chapter 2.3, it’s important to examine how the denotation process of cinema works. How and why do we attach certain meanings to sounds and images when they are presented to us in the edited continuum of a film?

Since Sergei Eisenstein developed his revolutionary theory of film
montage\textsuperscript{8} in the 1920's, cinema as an art form has been building its autonomous tradition apart from verbal—or \textit{vococentric}\textsuperscript{9}—forms of storytelling, such as theatrical drama. The montage theory of Eisenstein famously suggested that besides creating a temporal continuum by splicing two different images one after the other, the juxtaposition of the images creates a meaning that is not decipherable from the individual images themselves, but is created by contrasting the images with each other. This principle arguably applies to sound editing as well. When presented in a cinematic montage, images and sounds are not perceived to be appearing next to one another in the edited timeline, but rather on top of the other: every part contributing to the narration in relation to every other image and sound. (Eisenstein 1998, 95–96.) Eisenstein dubbed this mechanism of imparting meaning through contrasting images together as ‘dialectic’, reflecting the Marxist theories that were influencing his thoughts as a filmmaker working in the then recently established Soviet Union. This early realization of the possibilities of images and sounds, brought together by the dramaturgical principles of editing, set a foundation for montage theories that have been pivotal for the historical evolution of cinematic narration. The underlying assumption that images juxtaposed with each other should suffice for an effective cinematic narrative resulted in a common aversion among critics and filmmakers towards films that are wholly dependent on scripted dialogue as their main dramaturgical agent. Films like these were seen as historically backwards in their affiliation to the development of cinema.

However, documentary films based on an expository approach are often constructing their argumentative storylines on spoken commentaries and voice over narrators. Nichols (1991, 50) asserts how the interview in the context of documentary films is an “overdetermined structure”, which “testifies to a power relation in which institutional hierarchy and regulation pertain to speech itself”, thus being an exemplary expression of discourses of

\textsuperscript{8} Partly in collaboration with other Russian filmmakers, such as Lev Kuleshov and Dziga Vertov.
\textsuperscript{9} A term coined by Chion (1994), explaining our genetic and habitual sensitivity to human voice in relation to other aural perceptions, such as noises or musical sounds.
sobriety. Nichols (1991, 50–51) references the writings of Michel Foucault that deal with the discourses found in societal institutions such as medical science, where the interviewed testimonies of patients are translated into the discourse demanded by the disciplinary practices of institutional rule. The verbal narratives of ordinary patients discussing their symptoms become rewritten based on the discursive practices of medical science, a procedure that Nichols (1991, 50) contrasts with the interview practices found in documentary films. Similar mechanics of rewriting are found in many other institutional proceedings, like in anthropology, where the interview is “the testimony of native informants”, who plainly describe their native practices and distinctive customs to an interviewing anthropologist, who will then proceed to “rewrite their accounts into the discourse of anthropological investigation” (ibid.). How this process of rewriting functions in documentary films is directly related to the concept of voice, which is often understood to be the unifying argumentative center of a film or an “intangible moirelike pattern formed by the unique interaction of all film’s codes”, like Nichols (1985, 260–261) poetically describes.

Voice in documentary films shouldn’t be, however, confused with the straightforward verbal accounts given by social actors through interviews, although interviews and other fairly direct statements given by the film’s subjects contribute to the voice of a given film by providing a verbal vehicle for the filmmaker’s argument to take place. Often the filmmaker herself is arguing through the voice of the interviewed social actors by carefully selecting pieces from the interviews and pairing them with other audiovisual content. Marit Kathryn Corneil (2010, 109) emphasizes the problematic definitions of documentarian ‘voice’, by stating how “the connotations of empowerment that are associated with ‘the voice in documentary’ are suggestive of a certain ‘agency’ that we attach to the concept”, pointing towards the filmmaker as an authoritative agent reassembling and orchestrating the recorded voices of the social actors to present a specific argument. Although the actual voice of a social actor heard in the soundtrack of a documentary film is not directly the filmmakers voice or some other centralizing discourse, Corneil (2010, 113)
emphasizes (referring to Chion), how the presence of a discernible human voice sets up “a hierarchy of perception”, acting as an inevitable center in the audience’s aural perceptions in the soundtrack as a whole. This is partly due to vococentrism, our innate cognitive sensitivity to human voice, and partly due to the conventional dominance of speech in cinematic audio narration. Chion (1999, 5) asserts that “if a human voice is part of [the perceived sonic space], the ear is inevitably carried toward it, picking it out, and structuring the perception of the whole around it.” This has lead to a filmmaking paradigm, where voice as a vehicle for verbal expression is given a central position in the soundtracks of films.

The centrality of verbal accounts in documentary films has been subjected to a lot of criticism since the advent of Griersonian filmmaking. Edgar Morin (2005, 187–188) writes how the “fluidity” that differentiates the cinema from verbal language brings it closer to music, since both of them can convey meanings and affects without words, and since their effects remain “ineffable”, likening the mechanisms of cinematic narration to its counterparts found in music theory. David MacDougall (as cited in Crawford 2010, 40) affirms the redundancy of expository narration by asserting how the absence of a voice over narrator makes us more conscious of the nonverbal content of the film:

> When filmmakers renounced narration, it was not to withhold something but to add it. The absence of the narrator was like a phantom image on the retina, making us suddenly aware of what narration had displaced.

What MacDougall is referring to—besides the visual content being made more apparent by the lack of a voice over narration—is the array of nonverbal

---

10 La voix, translated to a voice by Claudia Gorbman in Chion’s The Voice in Cinema (1999). Here, the definitions of ‘voice’ in well-established film studies confusingly diverge. Whereas Nichols is defining ‘voice’ as the governing argumentative discourse of the film, Chion (1999) is using the term as a synonym for a human voice, no matter whether it’s a part of the scripted dialogue, spoken interview, or an off-screen narration. In this thesis, I’m referring to Nichol’s conception of the voice as the main argumentative discourse of the film, although I’m also referring to Chion’s writings in The Voice in Cinema.
auditive signifiers taking place in the soundtrack of a given film. These consist of atmospheric tracks recorded on actual locations (although not necessarily on the corresponding filming locations) conveying consistent environmental information, foley recordings enhancing the impression of physicality of visual movement and actions, sound effects intensifying the dramaturgical impact of onscreen events, both non-diegetic and diegetic music, and other audio material that has a distinctively nonverbal form. These nonverbal components of audio narration could be best described as indirect in their approach to addressing the auditor. As they are imperceptibly decoded by the audience as the narrative unfolds, auditive signifiers implicitly steer the auditor towards a certain interpretation, imparting aural information crucial for fully understanding the events presented. These signifiers are often produced by elaborate recording and signal processing techniques to have a maximal impact with only a minimal amount of attention required from the audience, making them ideal for conveying sensory information without taxing the attention of the audience too much. As an example, atmospheric tracks containing multilayered information concerning geographical location, time of day, weather conditions, off-screen actions, and the spatial or acoustic dimensions of the environment, are employed to impart sensory information that can be pivotal in fully understanding the coded content of the scene. The environmental information inferred from atmospheric tracks gives an impression of an aural landscape extending far beyond the boundaries of the visual frame, enhancing the totality of the experience by non-indexical means. As another example, foley recordings and synthetically produced sound effects are often intensifying visually depicted actions, as seen on the screen. They supplement the accompanying images by directing the focus of the viewer into minute details in the filmic images, contracting the point of attention in the visual frame by aurally highlighting particular components of the images with sound. By directing the attention of the audience, they're undoubtedly participating in the narrative process as active agents of cinematic narration. Aaltonen (2010, 157) affirms the narrative and dramaturgical possibilities of sound design, surpassing the function of providing mere mood and emotional
In addition to being an emotional element, sound is also an interpretative and narrating element. In documentaries, the picture is often realistic, but with sound it is easy to construct more meanings.

Like Aaltonen suggests, sound is perfectly capable of signifying complex abstractions presented in the storyline, such as intentions and causalities not directly decipherable from the accompanying images, as well as intangible concepts vital to the understanding of the visual or verbal narrative. In spite of the deeply-rooted misconception of nonverbal audio narration being capable of transmitting mere affective information, there are numerous other functions for recorded and synthesized sounds (i.e., other than recordings of speech) being frequently employed in films. If we briefly exclude score music from the inspection, nonverbal auditive signifiers contribute to the overall cinematic narration by:

- Indirectly defining the spatial qualities of the location presented to the auditor.
- Contributing to the sense of temporal animation and perceived speed of any given visual action.
- Conveying information about the environment in and outside of the scope of the frame.
- Directing the attention of the audience to particular details in the image that might otherwise go by unnoticed.
- Creating emotional anticipation related to gradually unfolding actions.
- Sensitizing the audience to various physical actions.
- Helping to establish an impression of the filmic reality continuing outside the strict confines of the framed image.
While all of the narrative elements provided by the signification processes of sound design are able to contribute considerably to the auditor’s sense of place and physical setting, they can also direct and alter the audience’s interpretation of the scene as a whole, even attaining a metaphorical (or metonymic) status within the narrated scene. All of the aforementioned functions of nonverbal audio narration are effectively contributing to the ‘voice’ of the film, although—as the quintessential cliché of sound design goes—most of this is happening without the auditor even registering their input to the overall experience.

Peter Crawford (2010, 40) writes about the “inferred knowledge” decipherable from aural information, emphasizing the possibilities of nonverbal narration in documentary films:

No matter how poetic the commentary on a soundtrack may be, it evades one of the points regarding sound that this chapter [- -] is addressing, namely how sound, that is not necessarily verbal, sensually contributes to (inferred) knowledge through contextualization.

The notion of inferred knowledge is crucial in understanding how most of the nonverbal sound content functions within a film narrative. While vision is conventionally associated with notions of rationality and empirical knowledge, nonverbal sounds address us inconspicuously through our hearing, encompassing and efficiently addressing the areas of imagination and sensual intuition. When pouring water into a glass, we don’t have to look at the glass to know when it’s about to spill over the rim. Our hearing efficiently imparts this knowledge without the help of other senses. The research on psychoacoustics has proven that our hearing has an impressive ability to identify highly specific characteristics of various actions and mechanisms that cause different sounds. Studies (as cited in Leman 2008, 166) have shown that auditors have been detecting as specific information as the “shape of struck plates; vessels with different levels of fluid; determination of gender from footsteps; the ascending or descending of staircases; and the trajectory of approaching sound sources”. While hearing is certainly indispensable in
conveying affective stimuli, it is also capable of transmitting more specific information that has a factual nature. It can be indispensable for producing rational conclusions and interpretations regarding other sensory cognitions, like suggested by contemporary psychoacoustic studies (Leman 2008, 166). These aural signification mechanisms are in constant use in film narratives, although we don’t necessarily recognize their workings.

Admitting that these polarizing divisions between visual and aural cognitions are based on simplified views on how our brains process external signals—as transmitted to them by the senses—they manage to illustrate a generalized idea of the differences in our responses to visual and aural stimuli. In audiovisual film narratives, the representative power of nonverbal sounds derives from the tightly-knit cooperation of visual and auditive agents being in a “simultaneous vertical relationship” with each other (Chion 1994, 40). Chion (ibid.) strongly emphasizes the cooperative relationship of sounds and filmic images by asserting how the specific connection a snippet of sound has with an accompanying image is “much more direct and salient than any relations the audio element could have with other sounds”. Both sounds and images are understood in relation to each other, as they appear in the dramaturgically edited narrative. Chion (1994, 137) also makes the important remark of different senses not being isolated from each other in terms of the type of sensory information they are transmitting. He (ibid.) writes how “the senses are channels, highways more than territories or domains”, validating the reciprocal idea of added value.

Despite the likes of MacDougall and Crawford stressing the importance of nonverbal signification processes in constructing the argumentative content of documentary films, the edited interview is still regarded as the most typical way of addressing the auditor and guiding her towards a comprehensive realization of the ideological content in the film. Nichols (1991, 20) is closely linking this convention to the alleged ‘sobriety’ of nonfiction films, when he points out how “the centrality of argument gives the soundtrack particular importance in documentary”, assuming that the soundtrack of a given documentary film comprises mostly of interviewed speech, voice over
narration and other accounts of clearly verbal nature. Although atmospheric
tracks, foley, and other more complicated sound effects are not always
employed in the context of documentary films (often because of monetary
reasons), their functionality and affective power follows the same basic
principles as in fiction post production, where they are perceived as an
indispensable part of an engaging cinematic narrative.

2.3 The phantom of authenticity

Authenticity as a concept closely connected to the documentary film tradition
is a much debated subject. What has been meant by 'authenticity' in the
context of cinema has been historically changing, and the definitions for the
confusingly abstract term are many. Traditionally, the requisition of
authenticity has been concerning mostly documentary filmmakers. However,
it has also been a concern in fictional filmmaking movements like the postwar
Italian neorealism or the Dogme 95 movement in Denmark, where the
employment of unschooled actors, direct sound\textsuperscript{11}, and unstable hand-held
cinematography were aiming at a heightened authenticity effect. In these
movements, the usual preference for technical fidelity was exchanged for a
stronger impression of cinematic indexicality. The underlying assumption was
that by deliberately sacrificing high production values for greater mobility and
more intimate circumstances in the production process of the film, the
authenticity effect of the film would be more pronounced. (Helke 2006.) This
affirmed the persistent misconception that films marked by a truthful and
'direct' treatment of profilmic events are haphazard in their technical
execution. Chion (1994, 108) confirms the same assumption:

\begin{quote}
Of two war reports that come back from a very real war, the one in
which the image is shaky and rough, with uneven focus and other
\end{quote}

\textsuperscript{11} By direct sound, I'm referring to sounds recorded on filming location, in contrast to sounds applied to
the soundtrack in the post production process of a film production.
‘mistakes’, will seem more true than the one with impeccable framing, perfect visibility, and imperceptible grain. In much the same way for sound, the impression of realism is often tied to a feeling of discomfort, of an uneven signal, of interference and microphone noise, etc.

The anti-aesthetic ethos of the documentary film tradition established in the early 20th century is still evident in the way how outspoken aesthetical aspirations and fastidious processing of sounds and images are often seen as decreasing the probative force and authenticity effect of documentary films (Helke 2006, 211). When discussing the authenticity of a filmic soundtrack, we’re encountering the same problematic issues of representation discussed earlier in this thesis. In post production, all audiovisual material is subjected to consecutive rounds of editing, which involves de- and reattaching snippets of images and sounds, breaking and reconfiguring the chronology of the narrative, adding non-diegetic music and special effects, which all amounts to individual artistic vision taking shape. What could have been a plain audiovisual representation of actions and occurrences taking place in front of the camera is now something vastly more complex. What this means in the framework of sound design is that snippets of sound recorded on location are cut-and-pasted, recontextualized, and digitally processed in the post production phase to suit the dramaturgical demands of the narrative and the chosen aesthetics of the film—a methodical approach encompassing both fiction and nonfiction modes of filmmaking.

Oftentimes, when the production tracks recorded on location are subjected to trimming and polishing to meet the standards demanded by broadcasters and film distributors, the procession of the material can be quite subtle. On a basic level, this means equalizing, multiband-compressing, and using algorithm-based noise reduction plugins to filter out distracting frequencies and adjusting the signal-to-noise ratio12 to make the soundtrack clearer and more intelligible. Often, especially in the case of documentary


12 “Signal-to-noise ratio (abbreviated SNR or S/N) is a measure used in science and engineering that compares the level of a desired signal to the level of background noise.” (https://en.wikipedia.org/wiki/Signal-to-noise_ratio)
films, the process of sound design is more focused on carefully editing and equalizing the production tracks to cut out distracting frequencies and extraneous noises, rather than recording or synthesizing completely new audio material to be added to the overall mix. Here, the aim is to understand what is truly essential to convey the story being told: what kind of aural information does the audience need in order to understand the narrative and thematic dimensions of the film. By focusing on smoothly improving the production tracks recorded on actual locations, the aim of a sound design process is to retain the perceivable indexicality of the captured sounds.

Besides the corrective treatment of direct sounds recorded on location, the post production practices found in contemporary film sound design include deliberate off-sync between sound and image, pitch shifting, filtering, and employing artificial reverberation in order to change the spatial dimensions of the sound image. These are only some of the narrative methods that significantly alter the outcome of a filmic scene; the total amount of digital processing possibilities being almost limitless. Aaltonen (2010) picks up an example from his ethnographic film Kusum (2000) that affirms the expressive freedom given to sound narration in many contemporary documentary films. Aaltonen (2010, 156–157) writes how he used (in collaboration with the sound designer Martti Turunen) various sound effects “which were manipulated, slowed down or speeded up, certain frequencies were cut off, etc.”, concluding how the sound design process of Kusum was “like composing music with sound effects”. Aaltonen (2010, 156) describes the resulting auditive representation of the protagonist’s inner world as expressive; interpreting the emotional states of the protagonist in way that is not as empathetically stated in the visual narration. He (2010, 157) points out, how he processed and treated the sound in Kusum “in a way I would never dare to treat the picture in a film of this genre”:

In Kusum, we stepped out from the tradition of realistic sound representation, which is very typical for ethnographic films. What is interesting here is that although this has been noticed, it has not disturbed the audience or threatened the general authenticity
of the film. The case has even been the opposite. A stronger emotional reaction has led the viewer into a greater sensitivity about the authenticity of the film.

It is important to note how all of the aforementioned techniques can be found in both fiction and nonfiction modes of filmmaking, where they carry out essentially the same narrative functions. However, what is considered an effective cinematic narration might be in conflict with the expectations and demands of authenticity in documentary films. The fundamental difference is thus not of practical nature, but related to the genre-specific conventions of fiction and nonfiction films.

Nichols (2001) revised his view of documentary films and their relationship to historical world in his Introduction to Documentary by asserting how “every film is a documentary”. He emphasized how all films—fictitious or documentarian—can be interpreted as audiovisual documents of their time, culture, and individuals involved in producing them. Every film is some way or the other perceivable as a historical document of its time of conception, no matter if it is produced as fiction or nonfiction. Nichols is still admitting the need to distinguish thoroughly scripted and professionally acted films from films of non-fictitious nature, but he decides to replace the conventional terms with more descriptive ones, calling fiction films “documentaries of wish-fulfilment” and nonfiction films “documentaries of social representation” (ibid.). Documentaries of social representation lend us “the ability to see timely issues in need of attention”, and by doing that, they represent “social issues and current events, recurring problems and possible solutions” (Nichols 2001, 2). It is being emphasized how both fiction and nonfiction approaches to filmmaking are based on similar narrative principles, although the way these principles are utilized in the structure of the film may be drastically different. Nichols (2001, 35) notes how documentaries of social representation engage with the world by building on an assumption that the film’s “sounds and images have their origin in the historical world we share”. Nichols (2001, 35) writes how this assumption is relying on the apparent potency of the photographic image and sound
recording to represent “what we take to be distinctive qualities of what they have recorded”. Here, the assumption of likeness is enforced by technological advancements in cinematic narration, such as lenses, optics, and diaphragms of exceedingly sensitive condenser microphones, and their promise of producing evidentiary representations bearing “the trace of what produced them” (Nichols 2001, 35). It is arguably one of the dominating characteristics of cinema how the improvement of technical tools and production methods has affected the process of depicting recognizable events and characters with ever sharpening fidelity, thus gradually and historically strengthening the assumed indexicality of audiovisual representations.

Bearing this in mind, it seems confusingly contradictory how technical finesse in documentary films is often seen as a sign proving the exact opposite. Researchers like Helke (2006) have written a great deal of how features such as hand-held cinematography or roughly recorded diegetic sound in documentary film narratives manage to assure the audience that what they’re witnessing has an evidentiary value—that the lack of technical quality is a proof of authenticity. Often this assumption is capitalized in documentary film narratives by carefully selecting various qualities of sounds and images to serve the storyline in the dramaturgical context they’re applied in. Paul Henley (2010, 132) stresses the innate connection between argumentative tendencies of documentary films and the assumed transparency of their narrative strategies:

[whether a given documentary film] purports to be a document or a documentary, ‘observational’ or ‘reflexive’, authored or ‘subject-generated’, the filmmaker typically seeks to convince the spectator of the validity of his or her understanding of the subject’s world by representing evidences of that world in a naturalistic manner. Where these various approaches differ is in the degree to which the manipulation of the evidence in the process of representation is considered legitimate and, relatedly, in the degree to which filmmakers flag those manipulations in the filmic text itself.

Historically, the indexical demands for audio narration have not been as rigorously defined as with visual narration. Chion (1994, 22–23) asserts how
most recorded nonverbal sounds are generally not considered to be representative like photographic images, although they bear a certain verisimilitude to their perceived origin. Accordingly, poorly recorded audio is not as likely to generate a similar effect of directness and immediacy as shaky cinematography—contrary to what Chion was arguing earlier in this chapter—as it is not given the same narrative significance as dramaturgically treated images. As recorded nonverbal sounds are not assumed to hold a similar evidentiary value as photographic images, their poor technical quality might be testifying only on the unprofessional level of the sound recordist, rather than on the ‘directness’ of the content. As opposed to filmic images, the referentiality of audio recordings to profilmic phenomena is easily blurred if their technical fidelity becomes compromised. When recorded sounds are presented in the soundtrack of a film, their ambiguity is transformed into decipherable meanings in relation to the accompanying images:

Depending on the dramatic and visual context, a single sound can convey very diverse things. For the spectator, it is not acoustical realism so much as synchrony above all, and secondarily the factor of verisimilitude (verisimilitude arising not from truth but from convention), that will lead him or her to connect a sound with an event or detail. [---] The same noise will be joyful in one context, intolerable in another. (Chion 1994, 22–23)

Aaltonen (2010, 153) asks the rhetorical question of whether sounds can be indexical within the context of documentary films. Although sounds are often utilized to convey affective information, they’re also capable of conveying factual knowledge, like asserted earlier in this thesis. Aaltonen (2010, 152) points out how—as a documentary filmmaker—he’s utilizing nonverbal sounds to present arguments, not any differently than he would employ images to construct the ideological voice of a film. Contrary to the somewhat common misconception, recorded sounds do have documentary when they’re employed in a documentary film narrative. Recorded sounds can undoubtedly bear an

\[\text{[15] In this thesis, photographic images are being understood as synonymous to individual frames in filmic sequences.}\]
indexical relation to the historical reality, although they are often treated as this relation would be considerably less significant than the relation of images with their indexical origins. Aaltonen picks an illuminating example from his own film *Taiga Nomads* (1992), co-directed by Heimo Lappalainen. He (2010, 155) writes how the film can be easily seen as an observational ethnographic film with its “realistic, even naturalistic” approach to depicting characters living in the heartlands of Siberia. Although the film has been receiving recognition especially for its perceived authenticity, a substantial part of its audio content was added to the soundtrack from various external sources in the post production phase. Aaltonen (ibid.) mentions how, to give an example from the post production process of *Taiga Nomads*, the sound of a ‘Siberian’ brook was actually recorded in Finnish Lapland. He also mentions how a pervasive sound of wind—contributing significantly to the sense of natural environment in the film—was picked up from a commercial sample library. Aaltonen (ibid.) emphasizes how the aim of the sound design was to “clarify and and cut off extra elements”. The recordings captured in the actual shooting location—i.e., the original indexical sounds—were depleted of the feel of remoteness and ‘natural’ ambience found in the remote wilderness of Siberian outskirts, that the filmmakers wanted to depict in the story. The field recordings were, in other words, not able to provide the correct dramaturgical signifiers that the directors acutely needed in order to construct the story they were determined to tell. Aaltonen writes, how the field recordings captured in Siberia were hindered by unwanted noises (such as people talking behind the camera) that were constantly reminding the auditor of the presence of civilization in the middle of the untamed taiga. He (ibid.) points out how these dramaturgically redundant sounds would have drawn the attention of the audience “to the wrong details”, or “away from the story or theme”, as scripted by the directors before the shooting took place. It is interesting how the decision to rely on sample library content and recordings done in locations thousands of kilometers away from the actual filming location wasn’t even discussed during the post production of *Taiga Nomads*—at least not from the point of view of authenticity and indexicality, although the filmmakers were
concerted that they were making an ethnographic documentary film. (Aaltonen, 2010, 155.) On the contrary, Aaltonen (ibid.) claims that “the sound effects increase the illusion of reality” in *Taiga Nomads*, stating how “in documentary filmmaking, this kind of hidden support to the dramaturgy of the film is normal work practice.”

What differentiates a recorded sound from a filmic image when considering the questions of authentic representation in documentary films? The double standard of indexicality in documentary films poses an interesting question directly related to the topic of this thesis. It is curious how documentary films that go to great lengths to convince their audience of their truthfulness and transparency can be so indifferent towards the indexical and authentic properties of their soundtracks. As this phenomenon is most likely connected to the common assumption of sound mainly complementing the images with its added value, it is necessary to consider how sounds are made to connect with images that are not referring to the same indexical origin.

Chion (1994) writes about the *synchronism* between sounds and images, largely enabling the effectiveness of cinematic sound design. By *synchresis*, Chion refers to the artificial synchronization of sounds and images that aims to create a “spontaneous and irresistible weld [- -] between a particular auditory phenomenon and visual phenomenon when they occur at the same time” (Chion 1994, 63). By meticulously synchronizing unrelated visual and auditory events together in a temporal filmic sequence, these unrelated events suddenly appear to be referring to the same indexical origin. The effect of *synchresis* to our perceptual abilities is so powerful, that we could easily select any particular sound having a sharp accentuated attack to accompany the image of a hammer hitting a nail, and the end result would be as convincing and realistic as with the actual indexical recording of the event. Because synchronizing unrelated auditive and visual media together in post production is so straightforward and effective, we are led to downplay or even completely ignore the possible consequences of this practice for the authenticity of

---

14 A term coined by Chion (1994), combining the words *synchronism* and *synthesis*. 
documentary film narratives. However, as Chion (ibid.) points out, certain combinations of recorded sounds and images come together more seamlessly than others. The envelope\(^{15}\) of a sound is a crucial factor in determining which sounds synchronize with a given visual accompaniment most convincingly.

Aaltonen (2010, 160) writes, how one of the findings of his doctoral dissertation *Todellisuuden vangit vapauden valtakunnassa* (2006) was that for documentary filmmakers (here understood solely as directors), “the sound is more open for expression than the image and it is not restricted by the demands of authenticity or indexicality, which affect the image.” When considering the reasons behind these “different rules” for sound and image in documentary films, Aaltonen (2010, 161) points out the domination of eyesight in human perception and the history of filmmaking, that began with silent films portraying mute scenes accompanied by live musicians performing in the screenings next to the projected image. Aaltonen (ibid.) remarks how in the years predating the advent of sound films (or ‘talkies’ as they were referred to in the 1920’s) the audiences had already been used to hearing sound effects and voice over narration in conjunction with the images presented in the screenings. Instead of being audio recordings reproduced by loudspeakers, they emanated from sources external to the image-frame, generated by live narrators and mechanical sound generators. Therefore, they were not subjected to the same assumptions of indexicality as the filmic images they were accompanying—being performative creations instead of mediated representations. Douglas Kahn (1990, 73) writes about the rise of mediated sound, made possible by the emergence of both radio and sound films: “For the first time, a diapason of worldly sound encompassing all of its visual, literary, environmental, gestural, and affective context could be displaced, presented, and represented.” This paradigmatic change irreversibly changed the way we hear sounds in general, and how we connect individual sound events to other sensory information presented to us.

To briefly summarize the complicity of the subject, it would be safe to say

\(^{15}\) The *attack, decay, sustain* and *release* of a recorded sound wave as its level of amplitude is fluctuating in time.
that all filmic representations are mediated by practices based on technological affordances and cinematic conventions, where the interplay of the signifier and the signified is constantly in motion. When we discuss filmic narratives from this perspective, the authentic qualities of a documentary film seem suddenly something that have to be actively constructed in order for them to exist.

2.4 Undefined ethics

Despite the similarities between non-fictitious and fictitious approaches to filmmaking, there are many practical and ethical reasons for making a distinction between nonfiction and fiction narratives. Firstly, they are made with different assumptions concerning their purpose. While not every documentary film is involving itself with social issues—i.e., motivated by a humanitarian ethos of ‘giving voice to the voiceless’—documentary films are commonly defined through their informative functions: they “prompt different sorts of expectations from audiences” (Nichols 2001, xi), when compared to films of ‘wish-fulfilment’ (i.e., fiction films with thoroughly scripted narratives performed by actors). Documentary films also involve a significantly more complicated relationship between the filmmaker and her subject, being wholly dependent on the participation of actual people inhabiting the protifilmic reality.

In spite of the many differences, many of the so called ‘creative documentaries’ are openly tapping into the conventions of fiction by relying on scripted scenes and staged enactments, occasionally hiring trained actors to perform in dramatizations of historical events. Even highly stylized animation sequences have become increasingly common within the documentary form. At the same time, there’s an established tradition in fiction filmmaking to shoot on actual locations, with ordinary bypassers acting as unsuspecting extras. Fiction films are also utilizing untrained actors for the
sake of heightened authenticity, even hiring people to perform themselves in scripted scenes depicting things that could happen to them in real life\textsuperscript{16}. The filmmaking practices of both traditions are confusingly overlapping, but the necessity to maintain a distinction between the two is arising from the unavoidable fact that in nonfiction films, we’re always interfering and altering the lives of actual people, no matter how stylized or dissociated may the connection be. The ethically complicated and sensitivity demanding situations that come with representing actual people employed as social actors shouldn’t be taken as a seal of authenticity, but rather as an invitation to a dialogic relationship with the immediate reality as it appears to the imagination and curiosity of a filmmaker.

While providing an effective means to draw the auditor inside the realm of a filmic representation, the suggestive and often inconspicuous nature of audio narration also introduces some ethical considerations exactly due to its implicit nature. Nevertheless, documentary film sound design has been an unproblematized area in discussions concerning the ethics of nonfiction narratives, arguably due to its predominant status as an auxiliary part of cinematic storytelling. Because it is seen as subsidiary to images and visual narration, researchers interested in questions of ethics and authenticity in documentary films—such as Nichols or Helke—have not been giving sound narration enough thought. Although the argumentative possibilities of nonverbal sounds are partially included in his concept of voice, Nichols is not discussing nonverbal audio as a narrative agent per se. Nichols later slightly revised his earlier definition of voice in documentary films. He (2001, 46) writes how the voice of a documentary film is actively addressing the audience through every possible way, as allowed by the limitations of its alleged sobriety. By this, Nichols means the editorial selection and creative arrangement of sounds and images, guided by the governing ideological or aesthetical “organizing logic” of the film. Nichols (ibid.) is listing the following directorial decisions defining the individual voice of a nonfiction film, which also the defines the particular mode of the film:

\textsuperscript{16} e.g., Michael Winterbottom’s \textit{In This World} (1998).
(1) When to cut, or edit, and what to juxtapose and how to frame or compose a shot.
(2) Whether to record synchronous sound at the time of shooting, and whether to add additional sound, such as voice-over translations, dubbed dialogue, music, sound effects, or commentary, at a later point.
(3) Whether to adhere to an accurate chronology or rearrange events to support a point.
(4) Whether to use archival or other people’s footage and photographs or only those images shot by the filmmaker on the spot.
(5) Which mode of representation to rely on to organize the film (expository, poetic, observational, participatory, reflexive, or performative).

Here, Nichols is enumerating the fundamental decisions encountered in any nonfiction post production process. By leaning on his list, the concept of an encompassing voice can be summarized as the bulk of editorial decisions and principles building towards a uniform argument.

While all film narratives—nonfiction and fiction alike—are basec on a syntactical play of signifiers bearing an indexical relation to the historical world (i.e., to the signified), the language of cinema is far more ambiguous than the language used in politics, law, medicine, journalism or other governing discourses of sobriety found in a contemporary society. When a documentary film decides to follow the doctrines and 'sobering' principles of such discourses, it is—by necessity—distancing itself from artistic expressions and aesthetic creativity. Likewise, when profilmic occurrences are given an artistic treatment, documentary films are no longer presenting their factual content through the clearest way possible, but instead they start relying on signifiers operating outside the limitations of verbal language. While doing this, they're often trying to grasp meanings that are beyond the explanatory power of conventional language systems.

Aaltonen's earlier example from Taiga Nomads sets a compelling question concerning the double standards of sound and image narration in documentary films. If the sound of a running brook recorded in Finnish
Lapland can be comfortably applied in the context of images depicting Siberian taiga, would it be possible to reverse the procedure without immediate ethical issues arising? What would happen to the authentic value of an ethnographic film, if it would be employing images that bear no actual relation to the narrative setting of the film? Could the resulting audiovisual document be called truthful or authentic? Although clear transgressions in image-based narration (i.e., the fabrication of indexical images and deliberately misleading the viewer by employing images contrary to their implied context) are rare in documentary films, sound editing can easily facilitate a visual misrepresentation by placing unconnected images—detached from their original context—in the same temporal continuum. Chion (1994, 17–18) points out how “the addition of realistic, diegetic sound imposes on the sequence a sense of real time”. This idea of applied temporality—by placing images on a temporal continuity of an auditive backdrop—is crucial for successfully editing disconnected images together in a convincing and cohesive narrative sequence. Nichols (1991, 174) is also recognizing the integrative potential of continuous sound in adhering unrelated images to each other:

In fiction the sound track [sic] often assists in the creation of continuity. Lines of dialogue, music, and sound effects can all carry across a cut, helping to minimize any jarring effect since our attention is given over, in part, to the continuing sound.

The arguments presented in favor of a free associative use of audio material in documentary films wouldn’t be considered ethically sound (no pun intended) if applied to visual signifiers—i.e., filmic images—being utilized without any references to their indexical filming context (e.g., Aaltonen 2010). However, it would be reasonable to claim that either way, the resulting narrative would be essentially the same when regarding its truthfulness, or the perceivable authenticity of the film. This is quite obviously assuming that the unrestrained compounding of sounds and images would be done with a sufficient editorial precision—without completely neglecting the documentary demands of
indexical authenticity.

Even a vague understanding of the ethics of audio narration in
documentary films is yet to be reached. This is arguably resulting from the
subordinate position of sound in the narrative strategies of films. Although
practical ethics are always negotiated in the process of documentary
filmmaking (also in relation to indexical images being employed in the
narrative), if the narrative possibilities of sound would be more pronouncedly
recognized, the possibilities of accidentally misrepresenting or offending
social actors and minority cultures would be less likely to occur. The
persistent underestimation of sound has resulted in an emergence of ethical
double standards, where sound narration is considered too ambiguous to be
concerned with ethical concerns. This has lead to a practical and theoretical
paradigm where a significant portion of filmmaking practices is left
undiscussed. This holds true in the post production processes of documentary
films, and in the critical studies written about nonfiction films.
3. BETWEEN RINGS

In this chapter, I will go through the post production process of *Between Rings* from the perspective of music production and sound design. Unlike in the previous chapter, I will be mostly examining practical examples and focusing on the narrative decisions that directly contributed to the auditive representations of ethnicity presented in the completed film. I will not be addressing the general technical preconditions of audio narration, nor will I try to go through the complete processes of score music composition and audio editing of *Between Rings*, as that would draw too much attention away from the actual topic of the thesis. Instead, I will be discussing the particular instances where the authenticity and ethnic dimensions of the auditive representations were called into question. I will also examine some of the contemporary sampling techniques—e.g., the use of virtual instruments—that partly enabled the representative practices adopted during the post production process of the film. Although the focus of this chapter is to enumerate and discuss the central aesthetic decisions executed in the post process of the film, I will briefly address the concept of ethnicity in order to avoid semantic misunderstandings.

Like the highly problematic concept ‘race’, ethnicity is not a biologically definable attribute, but a social construction (Jones, 2000). Like proven in numerous studies (see, e.g., Isajiw 1992; Samuels 2014; Moubarac 2013; Templeton 2013; Jones 2000), biological racial categories do not exist among groups of people. Because genetic differences between seemingly different racial groups (i.e., ethnic categories based on adaptive traits like skin color\(^{17}\))

\(^{17}\) “Adaptive traits, such as skin color, have frequently been used to define races in humans, but such adaptive traits reflect the underlying environmental factor to which they are adaptive and not overall genetic differentiation, and different adaptive traits define discordant groups. There are no objective criteria for choosing one adaptive trait over another to define race. As a consequence, adaptive traits do not define races in humans.” (Templeton 2013, 262)
are small in comparison to the genetic variation found within those groups, the definition of race or ethnicity cannot be based on biological factors (Jones 2000). Therefore, the biological significance of ‘race’ as a means to differentiate various ethnicities from each other is not based on contemporary scientific understanding (Moubarac 2013).

While race as a biological category differentiating ethnicities has been debunked, it has been suggested that ‘race’ as a social construction could be useful in explaining “social interaction”, and the way people discern and relate to racially perceived adaptive traits (ibid.). Nevertheless, I’m not using race as a descriptive term in this thesis. Instead, I’m referring to ethnicity—mainly as a cultural category having connotations related to physical adaptive traits (e.g., skin color or other physical features) and geographical locale—due to it emphasizing similarities over differences. Horowitz (as cited in Diamond 1987), is stressing the arbitrariness of ethnicity as a descriptive term. He (ibid.) writes, how ethnicity is “based on a myth of collective ancestry”—bringing about “a sense of group affinity”. While I’m admitting the relative arbitrariness of ethnicity as an “ascriptive phenomenon”, I’m mostly using ethnicity as an expressive term that has proven to be useful in referring to the idiomatic characteristics of a particular cultural group. “Ethnic identity”, like Moubarac (2013, 106) asserts, “refers to the expression of ethnicity that is particular to each human group; it is constructed through a process of communalization born out of social relations and based on the belief in a common origin and history.”

If ethnicity is difficult to define in a clear and satisfactory way, it is also difficult to represent as a cultural outsider. However, I will not be addressing the ethical dimensions of representation in this chapter. The ethics of documentary filmmaking—as involving the problematic power relations between the filmmaker and social actors—and the issues of misrepresentation and cultural appropriation are addressed thoroughly in chapter four.
3.1 From rags-to-riches

*Between Rings* was mainly shot in Zambia, Southern Africa, throughout the years 2010–2013. The protagonist, Esther Phiri, is a seven-time world welterweight champion boxer in her late twenties. The film depicts how Esther—after giving up on her marriage plans for a career in boxing—is torn between her family obligations and personal ambitions; trying to maintain her right for privacy while being a national hero. The film’s central narrative is constructed around the apparent difficulty of being a professional female boxer in a subculture defined by masculinity and patriarchal behavior. Besides discussing the prejudice found in global sports culture, *Between Rings* is implicitly addressing the chauvinism in contemporary Zambian society, and how it’s affecting the life of the films’ protagonist on a day-to-day basis. Esther herself describes her affiliation with the film in the official press materials of *Between Rings*:

This is the first time I’m really opening up for people to see me as Esther Phiri, the woman, not the boxer. It’s scary to be this open, but I think it’s important to share my story, so that people can see that even success comes with a price. We work hard, but we also have personal struggles to overcome. I hope my story shows young girls that they can do everything they want to do, and inspire them, no matter their circumstances.

The film begins by following Esther as she’s preparing for a match against her Colombian adversary Lely Luz Florez. After we witness Esther triumph over Florez—her winning the WIBA and WIBO Junior Welterweight titles—the film digs deeper into Esther’s background as an ordinary Zambian girl growing up in a family of eight children, barely supported by their single mother. As a child, Esther ended up living with her grandmother in the poor urban township of Mtendere. She started helping her by working at an early age,

18 [http://www.africavenir.org/event-details/cul/event/detail/2015/12/09/german_premiere_between_rings_the_esther_phiri_story_with_director_jcasic_chisi_in_attend/view-list%7Cpage_id-1.html](http://www.africavenir.org/event-details/cul/event/detail/2015/12/09/german_premiere_between_rings_the_esther_phiri_story_with_director_jcasic_chisi_in_attend/view-list%7Cpage_id-1.html)
selling groceries and second hand clothes in the market, before dropping out of school and becoming a single mother at the age of sixteen. A turn for the better in Esther’s life came about when the international NGO Africa Directions started a HIV-awareness project in the area, mainly targeted for local youth. The project combined health education and sports, and one of the activities established was a training program focused on boxing. Sports journalist Dee Williams writes in her Women Boxing Archive Network’s online bio, how “Esther was the only girl in a physical training program that focused on boxing”. Her gender didn’t fortunately stop a former Zambian amateur champion Anthony Mwamba from discovering her unusual abilities. Since then, Esther has been fighting the prejudiced attitudes of people around her, gaining more and more respect in the male-dominated culture of boxing.

By having a dedicated trainer and premises for training at Mwamba’s Independence Boxing Gym, Esther was soon on her way to become a national celebrity and an unconventional role model for Zambian youth. Home Liwanga, a journalist at Zambia Daily Mail, writes in his article for afrol.com:

Esther Phiri showed that women can compete in spheres most people think least likely when she traded explosive punches with Bulgarian Monica Petrova in an international boxing tournament in Lusaka. Attracting over 8,000 impressed spectators, Zambia’s first ever international female boxing bout challenged gender stereotypes as the two exhibited professional boxing skills rivaling that of male counterparts. The staging of the prestigious event on 18 March was not only good for the two female boxer or Zambia, but for gender activists as well. Boxing is known to be the men’s sport in Zambia, and indeed most of the world. [ - - ] With each blow, one could just imagine how many people watching could never imagine a woman in a boxing ring. 

Besides training and boxing, we’re witnessing her being constantly asked for favors and financial support from her less well-off countrymen, as well as being harassed by abusive men. As the film progresses, we learn that the story is not so much about Esther’s fight for another championship, but rather

19 http://www.xban.org/blog/ephiri.htm
20 http://www.afrol.com/articles/24897
about her struggle for self-reliance and respect as an African woman in a world where she’s expected to serve and submit to the authority of others. Although Esther is not proclaiming herself as a feminist, she is clearly struggling not only for herself, but also for the emancipation of less powerful women around her. Her own history is reflected in the fate of her niece Eunice, who gets impregnated during the film. The father of Eunice’s unborn child cruelly dumps her, and Eunice becomes suddenly dependent on the goodwill of her aunt.

After rising to an international fame, Esther painfully learns how burdensome it can get to be looked up as a celebrity even among her own relatives, who—like we witness during the film—clearly take her support for granted. In the end, we see Esther in the air-conditioned confinement or her husband Sandie Chinombwe’s mansion, alone in the seclusion of wealth and high social status. The ambiguous ending of the film leaves a lot for the audience to interpret, although it is implied that the dream Esther fought so hard to realize might not be as blissful as she was hoping it to be.

**Figure 1.** Esther Phiri in an official promotion picture of *Between Rings*  
**Figure 2.** Mikko Mäkeläinen, Marita Hällfors, and Jessie Chisi on location in Lusaka
3.2 Finding the right tone

Initially, the direction of the score music was decided in cooperation with the directors Salla Sorri and Jessie Chisi. Tahir Aliyev, the executive producer of *Between Rings*, didn’t participate in the process at this point. The first discussions concerning the direction of the score music were held in Zambia’s capital Lusaka while concluding the final leg of the three-year shooting process. In the meetings held in Lusaka, it was agreed that the score music should have some elements of Zambian folk music, although it would be rooted in contemporary European film music aesthetics. To achieve this, I promised to produce a demo theme for the film’s rough cut opening sequence that had been heavily relying on placeholder music since the first versions of the offline edit.

The only direct reference recording I got from the directors was a track called *Catch a Spear as it Flies* by New Zealander composer and musician Peter Wright. The track relies heavily on processed electric guitar, driven through digital delay units and embellished with concrete sounds and subtly treated field recordings. The track was edited to accompany the opening sequence of the film, that was almost fully formed in the first rough cut I received from the picture editor Jukka Nykänen. The sequence comprises close-up images of Esther training at the darkly-lit gym. We see her boxing glove-clad hands furiously working a punching bag, accompanied by direct recordings of heavy breathing and sharp cries of determined aggression. These sounds and images are repeatedly juxtaposed and intercut by poetic shots of a decorative wedding dress and long shots depicting the life in the shantytowns of Lusaka. Esther remains anonymous until the very end of the sequence, when we finally see her face in a close-up, sitting in a car while a heavy rain is washing the windows of the vehicle. At this point of the opening montage, the dreamy music is joined by a voice over prayer in Nyanja—one of the main native languages spoken in Zambia—recited by a clearly agitated female voice that we later learn belongs to Esther herself. The tone of the
prayer carries a hint of desperation. We can understand a lot of the content even without subtitling. The underlying message is emphasized by bleak images of poverty, shot through the window of a car passing through an archetypal African slum and juxtaposed with the dramatic tone of the prayer in the offscreen sound.

When I commenced working on the score in November 2013, I first tried to replicate the timbral qualities of Wright’s composition with my own initial setup. At that time, it comprised a variety of software synthesizers, an electric guitar, a cello bow, and a string of analog effect pedals. After some unsatisfactory experiments, I understood that I would have to find a new approach to the composition process, while still trying to retain some of the elements in the reference material which seemed to contribute something vital to the images they were accompanying. The reference track conveyed a subtle sense of melancholia that was well-suited for the desolate imagery of the opening montage. I felt that the minimalist approach of *Catch a Spear as it Flies* was suitable as incidental music due to the sparsity of its arrangement and a pronounced emphasis on mood. However, I was convinced that it wasn’t in line with the character and the profession of the protagonist—an extremely determined boxer on her way to the top. I strongly felt that the opening sequence demanded bolder gestures and a driving musical energy that would express Esther’s character and her commanding physical presence as a boxer. I followed this idea by trading the dispirited droning and abstract textures of a processed electric guitar for a simple three-note dulcimer arpeggio, that I then proceeded to loop in a digital audio workstation. I decided to use a virtual hammered dulcimer due to its prevalence in different folk music traditions all over the world—from the Appalachian Mountains in the U.S to India and Iraq. Because of its widespread use, I was hoping that the dulcimer would bring a feel of panethnicity to the score, implicitly reminding the auditor of a common ancestral heritage beyond national and ethnic boundaries. As an instrument, dulcimer has a clearly articulated sound that cuts easily through even a busy mix without taking too much space in the middle frequency bands. In many respects, it was a good choice to be employed in a film score
that was aiming to fuse disparate ethnic allusions into a coherent-sounding whole.

After improvising with a grand piano, I came up with a simple chord progression in D minor, played in a sharply articulated portamento. I supplemented the piano track by adding a contrapuntal two-chord progression played with my Telecaster, implementing a sense of austerity and steady advancement to the basic composition. The key element that made the track stand out was a short snippet of looped guitar, recovered from my initial attempts to imitate Peter Wright's playing. The salvaged sound was a half-a-second-long grainy sample, which was looped in a busy sounding 4th note cycle, providing a pulsating rhythmic bed for the alternating grand piano and guitar chords. I processed the sample by adding some digital phase distortion and programmed the ratio between dry and distorted signal to fluctuate in accordance with the arrangement of the composition, automating the distorted signal to slowly rise and fall with the development of the arrangement. Although the looped sound was monotonic and simple in terms of timbre, it gave the theme a rhythmic character that wasn't clearly localized in any specific cultural context. While the piano and guitar tracks were providing a slightly wistful tone to the composition, the pulsating backdrop of the modulating guitar loop gave the theme determination and dramatic pace.

While there certainly was a vaguely 'ethnic' mood established in the composition (with the hammered dulcimer arpeggio providing a perpetual sounding rhythm for the more contemporary leading elements, i.e., sampled grand piano and clean guitar tones), the track still didn't have any references to Zambian or South-African music. In spite of this, I felt that I had found a musical vocabulary to express the slightly tragic tone of the film without losing the dynamism and staccato-like punchiness that was being demanded by Esther's commanding physical character.
3.2.1 Synthesized signifiers

Although I intended to use digitally processed electric guitar in the score music of *Between Rings*, my main instrument in the composition and production processes proved to be Native Instruments’ *Kontakt*-software sampler. It allowed me to employ extensive sample libraries and use the multitimbral \(^{21}\) features inside my digital audio workstation Logic Pro, which was the DAW\(^{22}\) I was mainly working with at the time.

![Figure 3. Native Instruments' Kontakt -sampler (basic editing window)](image1)

![Figure 4. Native Instruments' Kontakt -sampler (wave editor window)](image2)

Just as the decision to work in my own studio was dictated by the budget allocated to the score music production, it was also a matter of monetary limitations to decide to work mostly with virtual sample based instruments. In deciding the instrumentation of the score, I was trying to come up with a clearly defined set of instruments that would characterize the overall aesthetics of the soundtrack besides allowing me to directly allude to the folk music traditions of sub-Saharan Africa. As well as enabling me to work with instruments that are laborious to record with the detail demanded by a professional film score, working with virtual sample based instruments provided the opportunity to employ instruments that are rare and difficult to

\(^{21}\) "An electronic musical instrument may be multitimbral, which means it can produce two or more timbres (also called sounds or patches) at the same time. Instruments which may be multitimbral include synthesizers, samplers, and music workstations." (https://en.wikipedia.org/wiki/Timbrality)

\(^{22}\) A common abbreviation for *digital audio workstation*. 
play without an extensive training. After completing a demo theme for film’s opening sequence, I started to explore the possibilities of implementing authentic Zambian folk flavors into the score. This proved to be a troublesome task.

When visiting Zambia during the final leg of the production process, I was eager to hear local music and try traditional Zambian instruments. Unfortunately, most of the music I heard during my seven-day stay was heavily autotuned plasticky pop that had little traces of originality in terms of ethnic or regional qualities. I felt that the music that was being played in Lusaka at the time, in July 2013, could have been produced anywhere in the world—from Finland to South-Korea. When I brought this up with Jessie, the Zambian director of the film, she said that Zambian folk music had been in decline for as long as she could remember. She told me that there was hardly any living tradition of folk music in contemporary Zambian culture beyond appropriating elements from other African traditions and blending them into a radio-friendly fusion that was now filling the local FM-stations. After being exposed to Zambian radio pop, I was convinced that it didn’t have a lot of inspiring things for me in terms of instrumentation or tonal vocabulary.

After returning to Finland and doing some research on the subject, I found out that there actually was a struggling folk music scene in Zambia, but the musicians upholding the tradition and playing the instruments associated with Zambian folk—like the budima-drums, kalimbas or the silimba xylophones—are mostly based on the rural outskirts of the country. When trying to find reference recordings online, most of the performances I managed to discover were choir works heavily based on the gospel tradition of the 20th century, with call-and-response -style choral singing often backed by simple drum rhythms. I acknowledged that this was probably the kind of folk music that actually had some relevance in the protagonist’s life, considering that Esther is a practicing Christian herself. Incorporating choral gospel singing into the compositions would have been impossible without actually recording an authentic South African choir. I briefly pondered the possibilities before deciding that it would be beyond the scope of this production. I also
considered using sampled choir singing snatched from published commercial recordings, but realized how tedious it would be to start bargaining for the rights to use those samples in the soundtrack of *Between Rings*. Although I managed to find some gospel-style choir samples that were distributed with a Creative Commons license, I didn’t really consider it to be a viable option to rely on stiff sampled material in the score.

Because of the aforementioned budget limitations, I started to feel that it would be best to rely on sampled virtual instruments in alluding to Southern African folk music tradition. I was surprised to learn how limited the supply was regarding high quality African sample libraries. Also, most of the few alternatives I managed to find were recordings of West African instruments, such as *kora*, *djembe* or *balafon*, which are not found in any Zambian folk tradition. After discovering Native Instruments’ *West Africa*-sample library, I knew that I wanted to use this particular set of virtual instruments. I especially liked the sound of sampled kora and ngoni. They had the warmth and liveliness that was acutely needed to achieve a convincing synthesis between the more polished tonal qualities of Western instruments, such as grand piano and clean electric guitar. Although NI’s *West Africa* instruments were carefully recorded and sampled in acoustically refined recording studios using high end preamps and condenser microphones, they still boasted a certain grittiness which, in my personal experience, tapped nicely into the easy-going attitude and *joie de vivre* found in local Lusakan communities, such as the township of Mtendere where Esther is hailing from. The library also had a good selection of percussive instruments that are found in many West African cultures, e.g., *dunun*, *djembe*, *krin*, and *calabash*. Most of the percussive instruments were applicable in either single or percussion ensemble formats, making them highly suitable for arranging complex and authentic sounding percussion compositions. With the sampled percussion ensemble kits, the composer is able to MIDI-control pre-arranged drum ensembles; covering a host of traditional rhythms encountered in West Africa. Some of these distinct polyrhythmic patterns—such as *Baga Giné*, *Kounbe* or *Soboninkun*—are encountered in many West African cultures, while some of
them are only found in particular villages or among certain tribes. While all of
the featured percussion ensembles are ready to be used as such, every note in
the pattern can also be modified. The plugin also provides the user a group of
parameters—e.g., variation, swing, feel, intensity and reverb—to enhance the
feel and timing of each programmed pattern.

![Figure 5. Native Instruments’ West Africa -percussion ensemble settings.](image)

![Figure 6. Native Instruments’ West Africa -percussion ensemble mixer view.](image)

While I had found a great sounding sample library to be employed in the score
music, using Native Instruments’ West Africa -library was a problematic
decision due to the fact that the instruments featured within the compilation
had very little to do with Zambian culture. This was possibly the most acute
problem encountered in the score music production—as it was directly related
to questions of truthful representation and cultural appropriation. How could
I justify using a collection of instruments belonging to a culture that is roughly
8000 kilometers apart from Zambia? It was obvious that I was making a crude
generalization by appropriating West African cultural content to represent the
life of a protagonist living on the opposite side of the African continent.
Despite the possibility of misrepresenting the culture and cultural
environment depicted in the film, I decided to utilize the sampled sounds of
West Africa to their full extent. I will be discussing the ethical implications of
this decision thoroughly in chapter 4.2.
3.2.2 Refining the score

After deciding the right instrumentation for the score, the composition process started to progress rapidly. I wrote and recorded most of the cues without knowing their actual position and exact dramaturgical function in the storyline. I ended up producing nearly completed mixes with full instrumentation before sending them to the picture editor Jukka Nykänen, who would then place the individual cues in a specific narrative context that he thought would benefit the film the most. This arrangement was governed by necessity, as the film was being edited at the same time while I was composing the score.

After incorporating West African instruments from Native Instruments’ library into the score music, the discussions I had with the two directors revolved mostly around issues of mood and individual instruments included in the mix: whether an individual element was demanding too much attention, or whether a particular mood of a cue was not exactly suitable for a given scene or a dramaturgical context in the storyline. One cue that generated a lot of discussion—plainly dubbed as the Fighter Theme—went through many consecutive rounds of rearranging and remixing. The first time the theme appears in the film, it is employed to create a tense backdrop for an archive sequence comprising video footage from various television documentaries and talk show inserts. The montage is explaining Esther’s past and her rise to national stardom in a distinctively informative way, facilitated by a voice over female narrator describing Esther’s background as a teenage single mother who had to fight for her survival. It is being made fairly clear that the voice over recording is extracted from the same set of archival material we see in the video footage. Meanwhile, the accompanying Fighter Theme is carried by a throbbing sub bass pulse, rhythmically accompanied by a tinny-sounding electric guitar loop. The cue carries on in an intentionally dragging tempo, until it suddenly comes into prominence through a piano break with an ascending chord progression. The break creates expectations of
the arrangement finally evolving and building up towards a climax, but instead of answering to those expectations, the track recedes back to it’s monotonous pace. At the same time, we’re being shown newspaper clippings and talk show-videos from Esther’s past, while she was still an unknown beginner struggling to gain recognition as a serious boxer in a culture dominated by dismissive males. Every time Esther is shown to have some kind of a breakthrough in her career, there’s always a setback in some form or the other, dragging her back into hardships. The radiant piano break is repeated twice in the arrangement, and after each time, the theme falls back into the repetitive loop led by the throbbing bass. There’s not a single element in the composition reminding the listener of Esther’s ethnic heritage or her cultural background as a native Zambian. The instrumentation and mood of the cue is gently pointing towards Esther’s estrangement in a reality indifferent to her aspirations.

Later, when the theme is reintroduced in the film—right before the dramaturgical climax of the storyline—the tempo of the cue is considerably faster. The rhythmic background is now supplemented by a driving djembe-pattern, embellished with snappy percussive fills played with a Guinean krín. With the new percussive tracks added, the theme suddenly mutates into a groovier version of it’s former self, although the repetitiveness of the bassline remains unchallenged. This time we hear the theme in the dramaturgical context of seeing Esther preparing for an important match against her old rival Monalisa Sibanda. The two contestants confront each other in a press conference, both of them boasting on their superior abilities as formidable boxers. We see Esther at the top of her game, calm and self confident. The faster tempo and the agile djembe rhythms emphasize her physical fitness and hard-earned position as a world-class boxer, having clearly overcome the difficulties that shadowed her rise to stardom. The renewed Fighter Theme is also possessing clear ethnic influences, suggesting that Esther is now standing on her own two feet. In spite of its more confident disposition, the new version of the theme is not entirely triumphant: while the faster tempo and the familiar major-key piano breaks are suggesting a victorious outcome, the
mood of the theme is interrupted twice by a transformation of the pia30
breaks from E major to a wistful chord progression in E minor, played with a
reverb-laden electric guitar. The reverberated sound of the guitar is gently
dissociating the listener from the adrenaline-fuelled events seen on the screen.
This sudden darkening of the mood implies the presence of an underlying
tragedy, questioning the self-assured disposition of Esther while she’s
preparing herself for the fight of her life.

The ambiguous shifts in key from major to minor raised some doubts
with the directors, both slightly confused by the divergent associations
aroused by the theme and its impact on the internal dramaturgy of the scene.
It seemed that the dramaturgical execution of the composition was too
complex; as if the music was driving the overall narration of the scene too
forcefully. I tried to justify the relative complexity of the theme by appealing to
the length of the theme. The editor had been asking for a three-minute theme
to support and strengthen the lengthy preparations leading to the match that
proves to be a crucial watershed in Esther’s boxing career. I felt that the only
way to keep the theme engaging for the listener would be to build
transformations and narrative development within the composition. I also felt
that the instrumentation should signify the vigorous physics of Esther’s
appearance, and the exceptional formidable of her presence, which the
throb of the bassline supplemented by the pounding djembes arguably
achieves.

After making some final mixing adjustments to balance the individual
elements of the cue, the directors eventually gave their approval for it. The
Fighter Theme is possibly the most seamless synthesis between European and
West African elements I ended up composing for the film. It has an aura of
thoughtful ambivalence provided by the piano melody, which is poignantly
contrasted by the muscular timbral characteristics of the percussion tracks—
suggesting of a fateful resolution looming on the horizon. The conflicting
elements were implying that Esther’s fate wasn’t necessarily dependent on the
outcome of this particular match, as the hardest fight would be fought in the
ring of her own troubled mind.
3.2.3 Appropriations of style

Regarding the more ‘African’ sounding score music, I ended up composing a cue that directly appropriated the distinctive aesthetics of highlife, a musical style developed in Ghana in the early 20th century. While highlife itself was based on appropriations from various highly different music cultures—incorporating and synthesizing elements as diverse as Caribbean rhythms and the harmonic structures of Christian hymns—it’s arguably one of the genres that quintessentially captures the upbeat ‘Africanism’ in many outsider discussions of contemporary African music (Matczynski 2011). The cue itself plays only a minor role in the film, as we only hear it for the duration of nine seconds in conjunction with a subjective point of view shot depicting Esther’s car ride to dramaturgically climaxing boxing match. The brief musical cue is accompanied by a direct recording of an overenthusiastic preacher giving a sermon on the radio, providing a convincingly authentic accompaniment to the music. It’s worth examining how the cheerful cue is presented as diegetic music in the scene, although it was not recorded on location in Zambia. I applied both low-pass and high-pass filters in the finished piece of music to create an impression of a song being played through a set of low-fidelity speakers inside of a car. In addition to the filtering, I ended up processing the perceived acoustic properties of the cue with a convolution reverb plugin, to further heighten the diegetic impression of the music. I assumed that by introducing the cue as diegetic, it wouldn’t feel imposed or gratuitous despite the exceptionally short duration of the scene. Without the faux-diegetic music, the scene felt like an insignificant transition between more eventful scenes, considering that the short scene is comprised of a dusty road and a row of unremarkable-looking residential houses in a plain subjective shot; as seen through the windows of a moving car.

The cue itself is comprised of a fast 4/4 drum pattern played with a standard Western drum kit, accompanied by a bouncy bassline and a repeating melody played with a sampled kora. The main melody is a short
ostinato arpeggio, clearly referencing the upbeat style of Ghanaian highlife. Only the unorthodox decision to replace the traditionally employed clean electric guitar as the main melodic instrument with a kora sets the cue formally apart from the stereotypical highlife aesthetics (although the sound of the sampled kora is very similar to a sound of an acoustic guitar).

The selection process of the main melodic instrument in this short cue poses an interesting question related to authenticity. The decision to use virtual kora instead of an electric guitar—which is a clear transgression from the norm in contemporary highlife (Matczynski 2011, 13)—was based on the timbral qualities of kora being more ‘African-sounding’ than the sound I was able to produce with my Telecaster guitar. Kora brought a distinctively exotic flavor to the composition, although it is not inherent in the Ghanaian folk tradition where the highlife aesthetics were appropriated from. I made the decision solely based on the timbre of the sampled instrument, as it carried easily distinguishable allusions to sub-Saharan musical traditions. This decision—like nearly all the other instances where I ended up employing the sampled instruments from Native Instruments’ West Africa collection—was met with acceptance and enthusiasm by the directors. I understood from the discussions with Chisi and Sorri, that the inclusion of sampled kora (even though it was introduced in the aesthetic framework of Ghanaian pop music) clearly reinforced the authenticity effect of the score, although kora is not known in Zambian folk music tradition. I will be discussing the ethical implications of this observation in chapter 4.2, as it relates to the practice of cultural appropriation. Before that, it’s worth examining the preconditions of employing virtual instruments in film music, and how they might be connected to questions of authenticity in documentary films.
3.2.4 Authenticity in virtuality

One of the decisive qualities in selecting the virtual instruments used in *Between Rings* was a culturally coded feel of the sounds that the selected instruments were capable of producing, alluding to a specific ethnic context and cultural milieu. The associations generated by the instruments were meant to transport the audience to a musical environment that was neither Zambian nor European, but had clearly articulated traces of both cultural realities. Although the score music of *Between Rings* was never intended to accurately represent Zambian folk music, the perceived authenticity of the music was a significant concern for both of the directors. Besides selecting instruments such Guinean *krin*, Malinese *djembe*, or West African *kora*, it was crucial that the chosen instruments were woven into the musical compositions in a way that wouldn’t feel superficial or out place when applied with contemporary Western elements, i.e., looped rhythm tracks, pulsating synthesizer sequences, and digitally processed electric guitar tones. Oftentimes, the optimal balance was attained through fastidious mixing, gluing the disparate elements together by employing compressors and algorithmic reverberation. However, it proved to be essential that the West African virtual instruments used in the soundtrack retained some of their roughness that was presumably preserved on purpose while recording and compiling the *West Africa* sample library by Native Instruments’ recording engineers. This procedure emphasized the feel of materialness in the virtual instruments, by selectively retaining audible traces of physical resonances and timbres characteristic of the actual acoustic instruments, including various minor deficiencies found in the physically modelled instruments themselves. Sometimes this meant accepting or even cultivating minor imperfections in the recordings, for the sake of emphasizing idiosyncratic details and deviations in the tone colors of individual samples.
Chion (1994) writes about the materializing sound indices when discussing the perceivable material qualities in recorded sounds. He (1994, 114) points out how “the materializing indices are the sound’s details that cause us to “feel” the material conditions of the sound source, and refer to the concrete process of the sound’s production”, transmitting us information concerning the substance and resonant properties of a given sound source. Materializing sound indices have an influence on our perception of various sounds, whether we are discussing the sound of a bow resonating the body of a violin or the sound of a specific type of gravel under the feet of a character walking across the frame. It was deemed crucial for the score music of Between Rings to retain an indexical relation to the tangible wooden bodies and distinct structures of the sampled acoustic instruments, while relying on the use of sample libraries and virtual instruments. Occasionally, certain audible deviations strengthening the timbral referentiality (i.e., perceived timbral characteristics referring to a particular physical instrument) of the sampled sounds had to be boosted by multiband equalization to further emphasize the impression of physical imperfection. This arguably augmented the feel of authenticity in the score music by directly tapping into the Western preconception of African music as something distinctly ‘rough’ and ‘unpolished’ in its perceived timbral characteristics.

Although the slight imperfections accentuated by equalization were contributing to the authentic feel of the music, there were other, perhaps more prominent factors at play. Native Instruments’ West Africa provides the opportunity to choose from two different mapping modes that define the organization of individual samples when played and triggered with a regular keyboard. The manual (Discovery Series – West Africa 2011, 18–19) of the sample library is describing these modes as authentic and chromatic. While set to the authentic mode, the melodic instruments found in West Africa are set to a heptatonic scale, consisting of seven pitches per octave, all mapped to the white keys of the keyboard.
The manual explains how the black keys—not dedicated for playing the tuned octaves—are assigned to trigger special articulations like trills and percussive hits characteristic of the musical tradition of the sampled instrument. The chromatic mode in turn assigns the pitched samples to an ordinary Western 12-tone scale, discarding the special articulations altogether. I understandably decided to employ the authentic mapping mode when utilizing instruments such as kora and ngoni. In a cue dubbed as Eunices’ Preparations, the backbone of the composition is formed by a simple three note melody played with a sampled kora. The melody is accompanied by a plucked atonal percussive sound, picked from an assortment of unusual articulations associated with the sampled kora. The sharply plucked and gently muted sound is played in unison with the backing melody, providing a sharp-attacked textural element to alter the perceived timbre of the tonal notes. When combined with the pitched samples of kora providing the backing melody, the muted percussive sounds produce an impression that is implying a slightly haphazard playing, although the melodies were played coherently in time. The melody would arguably sound much less energetic without the snapping percussive accompaniment, now adding a palpable sense of immediacy to the composition. While the cue is being embellished by a more complex kora melody—played on top of the snapping backing track—the real character of the cue is accomplished by the supplementary percussive plucks that blend with the melody due to them being played with the same physical instrument (albeit in a sampled virtual form) as the main melody.

Besides employing mixing techniques and special sampled articulations of the pitched instruments, the heptatonic scale applied in the compositions guided the compositions to a more ‘African’ direction. While heptatonic scales
are common in many African music traditions, they are also frequently featured in Western and Indian music. Although not being distinctively African as such, the decision to employ heptatonic tuning limited the melodic possibilities of the virtual instruments, thus enforcing a stricter melodic framework when compared to a regular 12-tone chromatic scale.

All of these aforementioned factors contributed to the feel of ethnic authenticity in the score music of *Between Rings*. They also efficiently compensated for the lack of haptic feedback in playing the virtual versions of traditional West African instruments, such kora, ngoni, or dunun. Leman (2008, 161) writes, how controlling a musical instrument is “realized in a closed loop with haptic, sonic, and perhaps visual feedback.” He (ibid.) continues by explaining how “in the mind of the performer, this physical interaction can be enhanced by corporeal imitation processes that translate the sensed energy back into the action-oriented ontology”, thus producing meaningful musical gestures that are capable of expressing and representing various phenomena when decoded by the listener. Without receiving apt haptic or visual feedback characteristic to the acoustic instrument being played—as in the case of virtual instruments employed with a mouse or a plain keyboard—the musical performance will inevitably produce different results than a performance executed with the timbrally corresponding physical instrument. This means that these deficiencies or deviations in the performance of virtual instruments have to be compensated to attain an authentic sounding musical result.

### 3.3 The broken indices

As the aesthetic direction of the score music in *Between Rings* was based on a synthesis between different ethnic references—consisting elements of traditional Zambian music juxtaposed with allusions to contemporary Western film music—it was bound to remain ambiguous in its relation to exact
geographic setting and cultural environment of the storyline. The score was
not intended to be directly referencing any specific region or culture for the
sake of ethnographic authenticity. *Between Rings* is not an ethnographic film
aiming for an anthropological accuracy, nor was there any explicit rules being
imposed on the score music regarding ethnic authenticity. The impression of
‘authenticity’ in the score music was established by employing folk
instruments that were clearly referring to African cultural heritage, albeit not
tapping into Zambian musical traditions per se. The focus of the score music
was to connect with Esther’s personality—or moreover, our interpretation of
Esther’s personality—and her impressive story of rising from rags to riches.
The ethnic allusions were constructed to depict Esther’s background and
character: they were subordinate to the demands of the narrative. This
assumedly relieved the composition process from the constraints of objectivity
and ‘sobriety’ in relation to ethnographic accuracy. Having said that, the score
wouldn’t have been successful without referring to Zambian, or at least
vaguely African musical tradition (as I came to realize during the composing
process), because it needed a certain amount of geographical and cultural
contextualization to reflect Esther’s position as an empowering representative
of a Third World country: defying the unspoken expectations set for her by
triumphantly fighting her way into international stardom.

However, the musical soundtrack was only one (albeit possibly the most
prominent) piece in the soundtrack of *Between Rings*, the other elements
being nonmusical noises, either recorded on location in Zambia, or applied in
the mix in the post production process of the film.

### 3.3.1 Atmospheric tracks

The atmospheric tracks (also referred to as ambiances) of *Between Rings* were
mostly recorded on actual filming locations in Zambia. Ideally, atmospheric
tracks are recorded in between shots on actual locations (including both
interior and exterior environments), to capture the natural ambience of the environment. They are usually recorded in stereo to provide a broader and more convincing representation of actual geographic surroundings. Different atmospheric tracks can also be layered on top of each other to produce a denser and more varied aural impression. The main purpose of atmospheric tracks is to construct a consistent aural backdrop for a filmic scene, and to mask the differences in the background noise of individual recordings comprising the entire content of the soundtrack. While doing this, atmospheric tracks are establishing a convincing representation of an actual aural environment expanding beyond the confines of the visual frame.

Due to the limitations of the recording gear available, I ended up recording only monophonic atmospheric recordings while working in Lusaka. These recordings were done in the actual locations where the profilmic events took place. In *Between Rings*, the atmospheric tracks acted as both geographical and cultural signifiers, providing the audience an inferred sense of being in a densely inhabited place—encompassing a mixture of agricultural phenomena and contemporary urban life mashed together in an unmistakably un-European soundscape. Here, the recordings signified a certain cultural environment, effectively contextualizing the protagonist and her immediate cultural surroundings in terms of aural information. The atmospheric tracks recorded in Lusaka comprised a mixture of individual sound sources, featuring a loud array of different motor vehicles, domesticated animals, the chirping of insects distinctive to South-African fauna, kids shouting and playing somewhere in the distance, and a baffling hodgepodge of languages spoken by the various ethnic groups living in Lusaka.

While all of these wildly heterogeneous sounds—each referring to their own particular source—were recorded by directional microphones, they transformed from a diverse collection of autonomous sound events into a cohesive atmospheric whole. When combined and mixed into the background of the production tracks, they provide an array of loosely unified signifiers attaching separate shots and recordings into a consistent geographical and cultural soundscape, masking minor inconsistencies in the background noises.
of direct indexical recordings. While doing this, they’re transporting the audience into the perceived environment where the protagonist is living, inconspicuously persuading the auditor of the authenticity of the scene with their idiosyncratic qualities. While being exposed to the affective influence of atmospheric tracks—each of them containing a plethora of sound events alluding to a specific cultural and geographical environment—the audience is persuaded to believe that because the consistency of the signifiers feels seamless, the representation of events depicted in the scene must be substantially authentic. Like Aaltonen (2010) described in referring to the post production process of *Taiga Nomads*, the unfettered treatment and combination of audio material—recorded possibly far away from the factual location of filming—with factually unrelated but indexically corresponding images, is often reinforcing the feel of authenticity rather than hindering it. “Paradoxically”, Aaltonen (2010, 155) writes, relying on non-indexical atmospheric recordings might increase “the illusion of reality”.

I had successfully recorded several atmospheric recordings while the film was being shot in Lusaka, but in the exterior scenes depicting the courtyard of Esther’s home, the restless and traffic-laden qualities of the recordings didn’t support the feel I wanted to convey to the audience. While Esther’s life (as depicted in the film) is in constant tumult, I wanted to present her home as a safe haven for her extended family consisting of cousins, housemaids, and adopted children, besides her own biological daughter Eunice. While searching for suitable stereophonic ambiances recorded in Zambia, I came across a collection of sounds recorded in Burkina Faso by sound designer and production mixer Rubén Durán Rebato. After going through the collection, I picked several atmospheric recordings that catered these narrative needs adequately. While I had to do slight editing to cut off the parts where the local Mossi-language (spoken widely in Burkina Faso) was being heard through other noises in the recordings, Durán’s field recordings captured a sufficiently placid atmosphere, featuring distant sounds of children playing and bicycles driving by. These recordings fitted effortlessly to the images shot in Esther’s courtyard, providing the necessary auditive signifiers for evoking a feel of
pleasant suburban atmosphere with distinctly un-European flavors, e.g., sharp buzzing insects, passing mopeds, and the chirping of various exotic sounding birds. Although all of these sounds could be associated with European aural environments as such, when mixed together in a single stereophonic recording, they form a foreign-sounding whole—an aural impression of a coherent cultural environment—that transports the auditor to a distinctly non-Western aural space.

I briefly pondered the possibility that the employment of these recordings—with their indexical origins in Burkina Faso—could undermine the authenticity of the scene. But due to the sounds corresponding offhandedly to the accompanying images, I decided to stick with them—valuing their suitable narrative properties more than problematizing clear transgressions of indexical authenticity.

**Figure 9.** Detail from an advertisement in an online store www.asoundeffect.com

**Figure 10.** “Authentic ambiances” being sold at www.asoundeffect.com

There were also other instances where I decided to employ atmospheric recordings from commercial sample libraries and Creative Commons databases. These instances included scenes set in downtown Lusaka (where detailed stereo recordings of cityscapes were needed to broaden the mono ambiences recorded on actual filming locations), and scenes set in around Exodus Boxing Stables, where we’re frequently visiting in the film to see Esther training with her promoter Anthony Mwamba. The non-indexical recordings employed in these scenes are providing a consistent aural ambience, with only a few lone auditive signifiers standing out from the flat environmental noise of chirping birds and a distant drone of traffic. The
decision to employ these recordings—as selected from various ‘inauthentic’ sources, like mentioned above—was not even discussed with the directors Chisi and Sorri, nor did I perceive it as a noteworthy transgression from the usual working practices in documentary film audio post production.

Chion (1994, 109) writes about rendering in film sound, asserting that instead of trying to vainly replicate the actual sounds of a given historical location or audible event, film soundtracks are technologically processed and layered in order to present a rendered impression of an actual environment or represented action:

The film spectator recognizes sounds to be truthful, effective, and ‘fitting’ not so much if they reproduce what would be heard in the same situation in reality, but if they render (convey, express) the feelings associated with the situation. This occurs at a barely conscious level, for film viewers (in which we must include most critics and theoreticians) have little more than a fairly crude and immediate understanding of the cinema’s figurative nature.

Chion problematizes the notion of ‘authentic recordings’ by claiming how the process of recording irreversibly disconnects sounds from their original context of hearing, destabilizing claims of originality and authenticity. When heard in conjunction with edited images and other audio events, recorded sound is not experienced or evaluated in relation to its indexical origin, but to the mediated context where it’s being heard, i.e., a dramaturgically treated filmic scene. Partly because of our cognitive tendencies to disregard the actual aural characteristics of a given sound while describing and interpreting sound events, the effectiveness and credibility of the soundtrack in the flow of the filmic narration is not produced by plain reproduction of actual sounds as they appeared during the event of recording, but by rendering of the auditive qualities as they are implicitly associated with the accompanying images. Chion (1994, 107) points out how the perceived ‘truthfulness’ of a given sound in the soundtrack is not dependent on any external realism, nor is it bound to the original moment of recording:
First of all, sound that rings true for the spectator and sound that is true are two very different things. In order to assess the truth of a sound, we refer much more to codes established by by cinema itself, by television, and narrative-representational arts in general, than to our hypothetical lived experience.

Although Chion is chiefly examining the audio narration techniques of fiction films, all of the methods and practices he's discussing apply to contemporary documentary post production processes as well (Tiller 2010, 55). This perceptual propensity to evaluate the ‘truthfulness’ of various sounds primarily in relation to other elements in the audiovisual sequence enabled me to convincingly combine atmospheric recordings captured in Burkina Faso with filmic images shot in Zambia, despite the geographical and cultural distance between the two countries being close to 7000 kilometers. How did this arrangement affect the authenticity of the film? If we are taking an idealistic stance, applying atmospheric recordings that do not refer to the same profilmic context established in the scene surely hinders their value as authentic ethnographic documents. However, if our attention is shifted from individual indexical recordings to a dramaturgically treated documentary film, the authenticity of the individual fragments (e.g., indexical audio recordings and video images) have to be evaluated in the context of the film as a whole. If we're agreeing with Chion in that the authenticity of a filmic experience is not dependent on the connection of recorded media to its factual place of recording, but rather on the perceived feel of sounds and images in connection with the associative flow of the audience’s experience, the notion of documentary authenticity becomes much more complicated. The assumption made by Chion is referring to the same contradiction discussed by Aaltonen in relation to *Taiga Nomads*. If we're accepting the assertion that an audio recording of a running brook recorded in Finnish Lapland might reinforce the perceived authenticity of images shot in Siberian taiga, were admitting that 'authenticity' in documentary films is an artistic abstraction. Hence, it is not reasonable to continue treating authenticity as an inherent quality affirmed by the indexical bond between a recording and a historical occurrence. Accepting this notion will challenge Nichols’ definition of documentary films, as it is
seriously undermining the ideals of ‘discourses of sobriety’ as a transparent mode of filmmaking, devoid of overt artistic intervention in the dramaturgical organization of audiovisual material.

3.3.2 Sound effects and foley recordings

If the function of atmospheric tracks is to provide a consistent aural background for a filmic scene, foley recordings and sound effects are commonly employed in film narratives to accentuate and enhance both physical and intangible elements appearing in the visual frame. In Between Rings, the most urgent need for synchronized sound effects was in designing the sounds of Esther’s boxing gloves-clad fists, as she’s fighting her adversaries in the boxing ring. Here, the original indexical sounds were not delivering the necessary weight and impact demanded by the intended feel of the boxing scenes. During the sound design process of Between Rings, I ended up selecting appropriate sound effects from a commercial sample library: exactly as I would have done in the context of fiction film post production.

Where the indexical sounds were devoid of the much needed physicality due to them being recorded afar (resulting in the loss of both low and high frequency content), the sample library punches carried an inferred sense of brute force—courtesy of them being captured from a close distance without any extraneous sounds masking their full recorded frequency spectrum. Without their facilitative power, Esther would have seemed far less vigorous in the representation presented in the film. Chion (1994, 60) writes how punch sound effects are “emblematic synch points” in film narratives. Their appearance and employment in film soundtracks is so thoroughly conventionalized, that their perceived realism is not judged according to the actual sound emanating from a fist hitting its target in the profilmic reality. Instead, their realism is dependent on the rendering of the punch in relation to the accompanying image and conventions of cinematic sound design. Chion
(1994, 60–61) writes how “the punch becomes the moment around which the narration’s time constructed”, stressing the physio-temporal effect of sound design to the overall narrative:

In real life a punch does not necessarily make noise, even if it hurts someone. In a cinematic or televisual audio-image, the sound of the impact is well-nigh obligatory. Otherwise no one would believe the punches, even if they had really been inflicted. (Chion 1994, 60)

In Between Rings, the real challenge was to mix the high definition punch recordings to blend in with other audio content in the soundtrack, including the direct recordings captured on the actual filming locations of the boxing scenes. Although I was employing sound effects selected from commercial sample libraries in several occasions throughout the film, the vast majority of the audio content in the finished film was comprised of direct indexical recordings, i.e., the production tracks recorded in Lusaka. Therefore, the rendering of the accented punch effects had to match the documentary qualities (e.g., handheld cinematography, accompanying direct sound, occasional breaches of continuity editing) of the film—not just the dramaturgical demands of the narrative. Despite these limitations related to the conventions of documentary film post production, I ended up layering many different punch recordings on top of each other, even employing the LFE-channel of the 5.1 setup to augment the synchronized punches with subtle amounts of sub-bass frequency content. The balance of the punching sounds heard in the film was refined in a final mixing session in studio Kalevala, where the soundtrack of Between Rings was finalized with the assistance of mixing engineer Sami Sarhamaa. In the final mixing session, the weighty punching effects of the boxing scenes were brought higher in the overall mix. This was partly resulting from the spacious acoustics of the mixing studio. As the control room in Kalevala is modelling a small film

25 “The low-frequency effects (LFE) channel is the name of an audio track specifically intended for deep, low-pitched sounds ranging from 3–120 Hz. This track is normally sent to a speaker that is specially designed for low-pitched sounds called the subwoofer.” (https://en.wikipedia.org/wiki/Low-frequency_effects)
theatre, the initial level of the added punch effects—as they were first mixed in a much smaller room—didn’t feel quite adequate to convey the intended feel in a theatrical listening space. In other words, their rendering was not matching the wide-screen visuals and diffused sound of a film theatre listening context. The end result can be still described as realistic in its relation to the visual narration. What significantly differentiates the weighty sound effects of Esther’s punches (as heard in the finished film) from the original indexical recordings, is the highlighted materiality and physical weight of the punches that would have been impossible to capture during the time of filming.

Chion (1994, 99) points out how the conventions of sound design in film narratives demand a high level of definition from the recordings employed in the soundtrack, especially concerning the high frequency content of the recordings:

High frequencies reveal a new multitude of details and information, contributing to an effect of greater presence and realism. [---] In the ‘natural’ world sounds have many high frequencies that so-called hi-fi recordings do capture and reproduce better than they used to. On the other hand, current practice dictates that a sound recording should have more treble than would be heard in the real situation [---] This proves that it’s definition that counts for sound, and its hyperreal effect, which has little to do with the experience of direct audition. (Chion 1994, 98–99)

This hyperreality is efficiently contributing to the dramatic effect of cinematic narratives, intensifying the sense of drama with emphasized material indices and accentuated physicality. Although sound design is often utilized for the purposes of heightening the impression of reality in cinematic narratives, it can also preoccupy itself with imparting metaphorical meanings that have no indexical origin in the corresponding images. Example of this can be found from the opening sequence of Between Rings. In the dramatic montage, we see darkly-lit images of Esther’s fists furiously pounding a punching bag. The image is accompanied by direct recordings of sharply accentuated punches
and Esther’s frantic utterances; both being captured on location while filming. I decided that because the images were so stylized, the sound design needed to match their dramatic rendering. Consequently, I opted to add samples of growling hyenas (as selected from a commercial sample library) to supplement the indexical sounds—viciously gnarling in perfect *synchresis* with the forceful punches seen in the closeup images. The recordings were slowed down a bit to change their timbral qualities. I also ended up processing them with moderate amounts of tape delay to introduce artificially perceived acoustic properties that would *not* blend with the acoustics of the direct recordings. In this way, the recordings retained their metaphorical disposition despite them being mixed with the foregrounded indexical sounds. Although the extraneous sounds of growling beasts are mixed relatively low in the overall mix, they can be effortlessly heard underneath the indexical sounds, introducing a semantic layer of new meanings to the audiovisual narration.

What did these metaphorical sounds suggest when placed in their specific narrative context in the film? When synchronized in conjunction with the direct punching sounds, they detach the narrative momentarily from indexical realism, adding a sense of metaphorical expansion to the scene. With the added value of the hyena recordings, Esther’s character is infused with bestial connotations that explicitly remind the auditor of the brutal nature of boxing. Although it was certainly not intended while deciding to utilize these recordings, the animalistic connotations can easily remind the auditor of a thoroughly racist conception of ‘savage Africans’ as a derogatory description that has been used to justify the pillaging of natural resources, slavery, and colonialist rule in the history of practically every African nation. Like this example efficiently proves, seemingly innocent decisions in the audio post production processes (including documentary and fiction films alike) can imbue the overall narrative with explicit connotations that might affect the outcome of filmic representation in an irreversible way. This is assuming that the filmmakers are *not* aware of the full extent of auditive signifiers they’re employing, which is unfortunately too often the case. It is worth noting that the exemplified decision was not even discussed with the directors of *Between*
Rings, although it undoubtedly has a definite impact on the representation of the protagonist as a person of color. This is evidentiary of the common indifference towards the connotative processes of sound design in film narratives, as discussed earlier in this thesis.

Despite the connotative mechanisms of sound design being susceptible to misrepresentations, they are also capable of facilitating empathy towards the represented social actors. Chion (1994, 34) asserts how “sound works on us directly, physiologically”, pointing out how recordings of breathing—to provide an illustrative example—can directly affect our own rate of respiration as auditors. A noteworthy example testifying on the empathetic possibilities of sound design was encountered during the post production process of Between Rings. When the film was already being mixed in studio Kalevala, it was suddenly realized that one of the closeup images of Esther felt lifeless without a sufficient aural accompaniment. Sorri, being present in the mixing session as a director, lent her breathing to supplement an image of anxious looking Esther in a scene predating an important showdown against Lely Luz Florez. Although being a seemingly diminutive part of the overall narrative, the long closeup shot—accompanied by the slightly distressed sounding breathing recorded from a close distance—brings Esther empathetically closer to the audience. After we witness her training fiercely, trying to keep up her appearances in the front of her male colleagues during the first part of the film, the moment we see her silently staring away in an intimate closeup feels like a tacit turning point in the dramaturgy of the film. It is as if we’re suddenly being introduced to a completely new aspect of the protagonist. Here, the added value of the breathing foley contributed significantly to the resulting impression.

The empathetically produced identification with representations of social actors and fictional characters alike is one of the most powerful aspects of cinematic narration. When such identification is enhanced—or thoroughly facilitated, as in the case discussed above—by the affective power of mediated sound, it is effectively reminding the audience of the corporeality of the characters seen on the screen. Such identification is perceived as vital for
conventional documentary films, as they are often occupied with explicit humanitarian aspirations, e.g., raising awareness about various social issues related to ethnic and indigenous minorities. Although combining a pronounced humanitarian ethos with expressive artistic practices is far from being unproblematic, it would be foolish not to exercise the empathetic possibilities of sound narration to their full extent.
4. CRITICAL APPROACHES

*The danger of present-day cinema is that it can crush its subjects by its very ability to represent them.*
—Walter Murch

As documentary films are depicting factual people inhabiting the historical reality, the ethical issues that accompany representative artistic processes are amplified when compared to processes based on fictive dramatizations relying on professional actors. While it is perfectly reasonable to examine cinema as a semiotic language—with its specific audiovisual narrative conventions (e.g., the categories of shot sizes and types such as point of view and establishing shots, foley recordings directing attention to details in the frame, atmospheric tracks providing consistency etc.) as its grammar, and the principles of montage as its syntax—it is necessary to note how documentary films are based on bodily interactions and representations of factual people, likening the filmmaking process to a physical performance. Film narratives differ from textual depictions through their corporeality and emphasized indexicality, making them more susceptible to representational abuse, such as misrepresentation and political propaganda. The abusive power of nonfiction films is even greater due to the assumed transparency and ‘sobriety’ of the documentary form, already discussed thoroughly in chapter 2.1. When writing about the questions of gender in documentary films, Nichols (1991, 177) is addressing the bodily dimensions of representations employing social actors:

What subjectivity can be attached to [a physical body] when it is not an imaginary construct to begin with (a fictitious character) but a participant in the historical world (a social actor)? What elements of sexism pervade not only the roles and subjectivities made available to women in the world but the representations of the body as image and Other (male or female)? And what responsibilities accrue to the filmmaker when people, made more widely known by their exposure in the film, resume their lives
after the film, possibly subject to insult or injury as a result of the film? (Nichols 1991, 177)

When we’re examining documentary films that are truly giving a “creative treatment” to actual events and people inhabiting the profilmic reality, we’re eventually admitting that creative documentary films comprise the same narrative practices as their fictive counterparts. Does the acceptance of the structured and plastic nature of all artistic creations mean that social actors represented in films should be treated with the same ethical ease reserved for professional actors working in the confines of a fictional story? It may seem obvious that there is a fundamental ethical difference in employing factual people who will be personally bearing the consequences of lending their lives as material for the filmmakers to use, when compared to the employment of trained actors who are not assumed to have an indexical bond to the personalities they’re embodying while acting in a film production. However, when we’re examining distinctly stylized documentary films that are dealing with social actors and profilmic occurrences as mere raw material for their aesthetically-driven narratives, the lines regarding the ethics of representations became confusingly blurred.

How could the ethical outlines of filmic representations employing social actors be explicitly defined? Could it be that if the varying level of artistic expression in nonfiction narratives is made clear—or if the creativity in treating actuality is articulated clearly enough—we’re instinctively understanding that what we see and hear in an audiovisual representation presented to us is not ‘reality’ objectively captured by film cameras and multitrack recorders, but an assemblage bearing a relation of artistic semblance to the events depicted. This will not relieve the filmmakers from ethical responsibility when constructing representations of factual people, but it will help clear out the confusion around documentary films that has been plaguing the form since the likes of Grierson confounded their audiences by advocating the assumed transparency of documentary films—mistakenly likening them to journalistic devices.
When it comes to constructing representations through sound design and score music, the inconspicuous nature of conventional sound design can be seen as an agent concealing the aims of the filmmakers, thus making its methods questionable in closer ethical scrutiny. Chion (1994, 33–34) points out, how sound (when compared to purely visual narration) has the innate capability to “saturate and short-circuit our perception”. Because of this sensory and cognitive reality, sound narration in films, “much more than the image”, can be easily harnessed for “affective and semantic manipulation.” (ibid.). If the audiences are not allowed to realize how audiovisual narratives are addressing their senses, they are not given the proper chance to take a critical stance to the presentation. While this is arguably not a pressing ethical issue in the context of fiction films, documentary films—justifying their technical deficiencies and lower production values with promises of sobriety and socially conscious treatment of actual events—are compromising their alleged sobriety when utilizing methods that are based on them not being easily recognized by the audience.

On the other hand, sound can be utilized to empathetically connect the audience with the events presented to them through audiovisual narration. The affective power of sound narration can transmit direct bodily sensations—such as the recorded sound of slightly anxious breathing produced on a foley stage, as discussed in the previous chapter—efficiently bridging the gap between the textual dimension of all filmic narratives and the bodily reality of social actors performing themselves in front of the camera.

4.1 Rendered soundscapes

The majority of nonmusical and nonverbal sounds in film soundtracks—as treated by the sound designer—are employed to enforce and expand the actions depicted in the visual frame (see, e.g., Chion 1994; Sonnenschein 2001; Aaltonen 2010). The design processes of various nonverbal sounds are
mostly concerned at intensifying the materiality and physical impact of visual
cues by calling on (and even exaggerating) the indexical linkage between
sounds and images. The original indexical sounds captured on location may
even be completely replaced by sounds recorded on the foley sound stage.
Noises accompanying visual cues in a synchronous relationship can impart
information about the physical substance producing the sound—e.g. wood,
metal, paper, cloth—as well as inform the auditor of the way the sound is
being produced, whether by friction, impact, periodic movement back and
forth, and so on. Chion (1994, 155) writes how studiously recorded nonverbal
noises are “re-introducing an acute feeling of the materiality of things and
beings”, thus advocating “sensory cinema”. Curiously, in fiction films, this is
perceived as adding a healthy dose of realism to the narration (e.g., Rose
2008, 80), whereas in nonfiction films, the employment of extraneous sound
effects—added to the soundtrack in post production—is sometimes thought to
undermine the authenticity of the film if executed too lavishly.

Annette Davison (2009, 140) suggests that the intensified materiality of
contemporary film sound design has the power to momentarily disrupt the
unfolding of the narrative, by diverting our attention to the timbral qualities of
the sound effects heard in conjunction of the visuals:

[---] moments of sonic spectacle, as generated by high definition
sound that is also spatialized, have the potential to divert the
audience from the onward trajectory of the narrative as their
bodies respond to a ‘set piece’, the object of sonic spectacle. [---]
Thus, moments of sonic spectacle hold within them the potential
for the incoherence and contradiction of jouissance by
interrupting our connection with the narrative, albeit
momentarily.

This affective phenomenon gives materially intensified sound design the
power to disrupt the argumentative motivations of the narrative by
momentarily releasing the auditor from the cognitive pull of the unfolding
storyline, introducing pockets of pure sensory space within the centralizing
force of the narration. The auditor is drawn, to cite Davison (ibid.), “from the
sound’s semantic sense into its ‘grain’”. These sensory spaces could be utilized
to encourage identification that is not based on premeditated argumentation, but rather on the immediate sensuous engagement between the auditor and the 'voiceless' representation (i.e., not directly defined by the author) of a social actor. What could this mean in relation to the usual denotative practices of documentary films? If the centrality of the narrative—constructed around the discursive formations and ideological motives in the filmic text—is destabilized and challenged by the *jouissance* of aural identification, the resulting voiceless space in the disrupted narrative is permitting the auditor to reflect the intuitive knowledge inferred from the sensual perceptions. It is important to note how the inferred knowledge—as it is based on 'discourseless' assemblages of sounds and images—is not retraceable back to a consistent representation. When the ideological *voice* and directorial presence of the auteur is thwarted, the sensuous knowledge permitted by the absence of a centralizing argument is not lending itself (as a mere replacement) to didactic explanations or totalizing representations of social actors and profilmic phenomena.

In the case of *Between Rings*, the type of rendering of various non-indexical sounds with the direct sounds recorded on filming locations was often dictated by the dramaturgical needs of a given scene. Like suggested in chapter 3.3.2, as the decision to replace the faint indexical sounds of Esther's punches with weightier recordings from a commercial sound library was made to heighten the dramatic impact of the scene, it was similarly required by the dramaturgical needs of several scenes to employ atmospheric tracks recorded in less noisy environments—as defined by the aural presence of contemporary urban life—than the actual filming locations allowed. This decision was made to accommodate the soundtrack to match the visual portrayal of the film's setting—not as it appeared to be in the profilmic reality, but how it was represented in the artistic rendering of the cultural environment.

Chion argues against the notion of cinematic realism as a mimetic approach. He (1994, 108) proposes instead that the perceived realism of audiovisual narratives is understood in relation to the cinematic conventions
themselves, in lieu of “our hypothetical lived experience” in the profilmic reality:

The codes of theater, television, and cinema have created very strong conventions, determined by a concern for the rendering more than literal truth. We are all thoroughly familiar with these conventions, and they easily override our own experience and substitute for it, becoming our reference for reality itself. (Chion 1994, 108)

In spite of this conclusion making a great deal of sense in light of an engaging cinematic experience, it is also worth considering how the inferred qualities in various recordings might be interpreted differently according to the cultural backgrounds of the auditors. Chion’s arguments concerning the cinematic realism of audio narration are based on an assumption that we are all perceiving the audio content presented in a soundtrack more or less the same way—as habituated by the conventions of television and cinema. It’s also highly possible that many subtleties in the soundtrack will communicate different things to different people. This is especially true in documentary films, where the lack of scripted narratives and high production values is replaced by understated narrative connections and small but meaningful gestures, brought forth and framed by careful editing of sounds and images.

R. Murray Schafer coined the term ‘soundscape’ as an aural equivalent for the concept of landscape. As soundscapes “consists of events hear not objects seen” (Schafer 1994, 8), they direct our perceptions toward the actual qualities of aural perceptions per se—instead of interpreting them in relation to a visual accompaniment provided by eyesight. Schafer (1994, 9) writes about the “authenticity of the earwitness”, referring to the myriad of particular qualities and their meanings in a given aural environment. Those qualities—as opposed to the logic of contemporary narrative practices of cinematic sound narration—are only decipherable to those that expose themselves thoroughly to a particular soundscape. Schafer emphasizes how a soundscape is not reducible to generalizations based on the common denominators between different acoustic environments, directly questioning the cinematic practices
of free exchange of recorded ambiences from a context to another to enhance
the narrative with their idiosyncratic auditive signifiers. Instead of endorsing
the exchange and associative combination of different aural qualities, as they
are mixed together in a filmic soundtrack, soundscape recordings cherish the
untampered aural characteristics found in an individual soundscape,
encouraging the auditor to attune to the sonic environment as an enveloping
totality. Schafer (1994) distinguishes three distinctive aspects that constitute a
soundscape as keynote sounds, signals and soundmarks. He (1994, 9) writes
how “keynote sounds do not have to be listened to consciously”, as they are
being “overheard” as an aural background behind other more prominent
sounds. Although Schafer is mostly interested with soundscapes as actual
aural environments in their unmediated form, the concept of keynote as an
encompassing ambience reminds us from atmospheric recordings employed
in films. Keynote as a concept refers to the ubiquitous sounds of a landscape
as conditioned by various geographical and climatic qualities, e.g., “water,
wind, forests, plains, birds, insects and animals”, possessing “archetypal
significance” to the people inhabiting the particular soundscape (Schafer
1994, 9–10). The concept of soundscape is not directly applicable to the
practices of recording and processing atmospheric tracks, as it is
unsusceptible to recontextualization exactly due to it possessing archetypal
significance. Whereas in audio post production processes, the employment of
atmospheric recordings outside of their original indexical context is
considered a normal working practice, the habit of casually disregarding the
uniqueness of each soundscape (albeit in its mediated form as an audio
recording) would probably seem preposterous to Schafer. He (1994, 10) writes
how certain archetypal sounds that constitute the heart of a soundscape “have
imprinted themselves so deeply on the people hearing them that life without
them would be sensed as a distinct impoverishment.” Could it be that the

24 Although I’m only discussing keynote sounds in this thesis, signals and soundmarks are referring to
“foreground sounds” that demand a more active listening when compared to keynote sounds. The term
soundmark is the aural equivalent for landmark, and “refers to to a community sound which is unique
or possesses qualities which make it specially regarded or noticed by the people in that community”
(Schafer 1994, 9–10).
most essential inferred qualities of atmospheric recordings are lost when they are employed outside of their original context? This is hard to prove conclusively without the insights of an auditor inhabiting and intimately connecting to the said soundscape. When we’re exposed to the sounds of a particular soundscape for years, the intricate aural qualities of our surrounding environment start to gain meanings that would be indecipherable to outsiders. In spite of this, recordings of actual organic soundscapes are constantly employed in film narratives as though they would be ubiquitously understood, regardless of the specific cultural background of the auditor.

J.C. Carothers (as cited in Cavell 2003) points out how “rural Africans live largely in a world of sound—a world loaded with direct personal significance for the hearer—whereas the western European lives much more in a visual world which is on the whole indifferent to him”. This is very much true for the most of us. We are so preoccupied with images and the flux of visual information that the vibrant soundscapes that surround us affect our daily lives very little. For a representative of a culture obsessed with visual stimulus, the notion of a living aural environment imbued with inferred knowledge might seem impossible to fully grasp. Although it might seem fairly innocuous to employ atmospheric recordings despite not realizing the intricacies of their coded content, for an auditor who is capable of decoding the inferred qualities of the recording, it might seem utterly confusing when those qualities are utilized in a haphazard and negligent way. This is where the questions of target audiences come into play. Considering that the vast majority of documentary films produced throughout the world are aimed at Western audiences, it seems somewhat understandable that the questions of inferred knowledge in non-Western atmospheric recordings are not of top priority in the busy post production schedules of film productions. However, it’s hard to justify this negligence in light of the attention given to the authenticity of documentary images. Considering how easy it would be to exchange a small portion of the time reserved for filming various exotic locations for recording atmospheric recordings native to the location, it seems like plain indifference to customarily rely on recordings lifted from
commercial sound libraries whose inferred meanings even the filmmakers themselves are not fully realizing.

4.2 Authentic appropriations

Erich Hatala Matthes (2016, 343) defines cultural appropriation in his article Cultural Appropriation Without Cultural Essentialism? as follows:

Cultural appropriation in the arts is a diverse and ubiquitous phenomenon. It might plausibly be thought to include occurrences as varied as 1) the representation of cultural practices or experiences by cultural “outsiders” (sometimes called “voice appropriation”); 2) the use of artistic styles distinctive of cultural groups by non-members; and, 3) the procurement or continued possession of cultural objects by non-members or culturally distant institutions.

Matthes describes how cultural appropriation is related to varying occurrences of misrepresentation, theft, and abuse, covering an assortment of ethically problematic instances related to the power structures between different cultures and ethnicities, from fairly benign cases of cultural outsiders expressing their personal style by appropriating designs or paraphernalia from minority cultures, to more harmful instances, such as casting white actors to enact the stories of ethnic minorities due to them being more ‘bankable’, i.e., drawing bigger audiences and generating more revenue.

James O. Young (2010), specifies five different types of cultural appropriation:

A) object appropriation
B) content appropriation
C) style appropriation
D) motif appropriation
E) subject appropriation
While object appropriation is referring to tangible objects (usually works of art or religious artifacts) being appropriated from a culture to another by individuals such as art collectors, archeologists and ordinary tourists, the other forms of appropriation discussed by Young are related to intangible cultural property, such as stories, design patterns, music, visual motifs, and other artistic elements and practices that have been evolving throughout generations within a certain culture.

A common feature for all cases of cultural appropriation is the condition of uneven power relations between minority and majority cultures, enabling the exploitation of minority cultures by appropriating cultural capital that is vitally important to the well-being and functioning of the abused culture. This might lead to a destructive assimilation of a minority culture into a majority culture, or to the weakening of a minority culture due to it’s cultural capital being appropriated. It has also been suggested (e.g., Carroll 2006; Gaut 2007), that works based on appropriated content are bound to have aesthetic flaws due to them lacking the vital cultural experience of the insiders. Young (2010) discusses the matter thoroughly, finding many arguments that reaffirm the assumption. For him, one of the decisive factors in determining the harmfulness of various types of cultural appropriation is in the inventiveness of the appropriative act: whether the appropriation is innovative or non-innovative. This division is useful in evaluating the aesthetic value of content appropriation, and whether the resulting work could be seen as surpassing mere outsider imitations. (Young 2010, 36.) Young (ibid.) writes, how “artists who engage in [innovative content appropriation] appropriate a style or a motif from a culture but use it in a way that would not be found in the culture in which it originated”, whereas non-innovative content appropriation is bound to remain derivative compared to aesthetically similar works produced by cultural insiders.

For critically reviewing the soundtrack of *Between Rings* from the perspective of cultural appropriation, I’m especially interested in the category of content appropriation with its subcategories of style and motif appropriation. Young (2010) defines content appropriation as reusing an idea
that was initially produced and expressed in the work of an artist from another culture, possibly representing it in a context foreign to the cultural source where the idea was appropriated from. Typical examples of content appropriation are stories and musical styles adopted from minority cultures into majority cultures, such as blues and jazz music developed by Afro-Americans within the majority culture of white Americans, both being appropriated by the oppressing establishment and transformed into more marketable aesthetic variations suited for the taste of the majority culture.

Young (2010, 25) mentions one possible outcome of content appropriation by discussing the possibility that once a majority culture has appropriated cultural content such as a musical style from a minority culture, it’s perfectly possible that members of the minority culture, “exposed to the performances by outsiders”, will adapt their original way of playing to reflect the outsider interpretation of their music. While some might say that this is just ordinary cultural evolution taking place, it will also directly contribute to the depletion of the minority culture and its gradual assimilation into the dominating majority.

What were the ethical implications of appropriating West African rhythmic motifs and playing styles for the score music of *Between Rings*? Is it possible that the synthesis of West African and contemporary Western aesthetics in this Finnish-Zambian film production was negatively contributing to either West African or Zambian culture? Answering to this question is further complicated by the fact that the musical motifs appropriated for *Between Rings* were not appropriated from Zambian folk music heritage, but from various West African musical traditions—although they were ultimately represented in the film in a context implying otherwise.

Software developer Native Instruments advertises on their website\(^2\), how they “carefully sampled regional percussion and melodic instruments from Mali and Guinea” for the sample library dubbed as *West Africa*. The online advertisement describes how the “authentic percussion patterns”

provided by the accompanying sample player give the user customizable instant rhythms and complex polyrhythmic textures, allowing for an “easy recreation of traditional ensemble playing with one touch of the play button.” Although NI’s West Africa provided me the opportunity to directly tap into Malinese and Guinean music traditions by employing preprogrammed percussion polyrhythms and meticulously sampled pitched instruments, my aim was to blend the sampled rhythms to pure electronic sounds and the tones provided by Western instruments as seamlessly as possible, creating an aesthetical hybrid between disparate musical references. This required modifying the preprogrammed and ‘authentic’ drum patterns by careful MIDI-editing, and altering the timbre of the samples by using equalization, filtering, and algorithmic reverberation (in addition to adjusting some basic parameters like tempo and swing). The tonal alterations were mostly subtle mixing adjustments made to make the drum and instrument tracks of West Africa to ‘sit in the mix’ without them sticking out from the rest of the tracks, that comprised digitally processed electric guitars, synthesizer basses, grand pianos, rhythmic loops, and other instruments and sounds found in contemporary Western film music scores. The employment of authentic sounding drum samples and preprogrammed polyrhythmic patterns was not to convince the audience of the ethnic authenticity of the music itself. Rather, it was to create an impression that the music was reflecting the cultural environments of the factual personalities depicted in the film. In other words, the commercially produced drum samples were acting as auditive signifiers pointing towards certain cultural contexts vaguely identified as ‘African’, instead of more accurate geographical localization.

It’s plausible to argue that due to lacking the cultural experience of a fully participating member in a given culture, a cultural outsider is bound to produce inauthentic works with detectable aesthetic defects when appropriating aesthetical styles and motifs from cultures other than his own. The great saxophonist Charlie Parker (as cited in Young 2010, 35) was quoted on saying, “If you don’t live it, it won’t come out of your horn”, emphasizing the importance of personal experience transcending mere musical training in
playing jazz. Is it really crucial to live the life of an ethnic African-American in order to produce legitimate and aesthetically compelling jazz? Possibly in the case of recorded music, few aficionados might prefer jazz made by African-Americans due to it sounding more authentic, but in incidental music such as film score music, the success or failure of the music is not evaluated with the same criteria. Because incidental music is composed to enhance the dramaturgy of a narrative and contribute to the overall mood of the scene, its qualities will be evaluated in relation to the context it is presented in. This also applies to the perceived authenticity of the music cue. In *Between Rings*, whether audiences considered the appropriated percussive content provided by NI's *West Africa* to heighten the impression of authenticity is a difficult thing to measure. I'm arguing, however, that by structuring a musical hybrid consisting of both Western and West African motifs and elements, I was able to reach a synthesis that was trying to imitate neither West African nor Zambian folk music aesthetics, although it certainly appropriated artistic content from West African cultural traditions. It is noteworthy how Chisi, the Zambian-born director of the film, didn’t initiate any ethical discussions when I presented her the idea to employ West African sampled content to compensate the lack of suitable Zambian folk music elements to draw from. Despite her not problematizing the matter, it's reasonable to suppose that some Malinese or Guinean natives might consider it inappropriate or even offensive how their cultural heritage was being appropriated for *Between Rings*. When giving it further thought, Chisi’s views on appropriating from West African cultural traditions might not be any more relevant than Sorri’s (herself a native Finnish), considering that Chisi’s cultural background had not much to do with Malinese or Guinean cultures. It is fairly irrational to suppose that an adaptive trait—like the color of one’s skin—would give a person better judgment in matters related to cases that are *not* affiliated with one’s own immediate cultural heritage. This assumption is based on crass racialized dichotomies, such as black/white, as it’s suggesting that certain ethnic qualities would inherently give a person a better position for ethically
sound judgments related to misrepresentation of neighboring (or otherwise loosely-affiliated) ethnicities and cultures.

The cultural experience argument is implying that “all representations of insiders by outsiders”, like Young (2010, 59) writes, have to be considered as misrepresentations. Young (ibid.) himself is a little skeptical that this would be the case, emphasizing that which is common to human experience regardless of cultural or ethnic divisions; available to us through imagination and empathy. Young (2010, 61) also reminds that in some cases a work of art can be exceptionally valuable because it benefitting from the perspective of an outsider. Such works provide insights into a specific culture exactly due to them being produced by cultural outsiders. “We can learn something about ourselves from seeing how others see us”, like Young (ibid.) suggests. Arguably this idealistic principle applies to hearing as well.

One could easily claim that documentary films are especially apt examples of works that contribute beneficially to the understanding between different cultures and ethnicities, despite being often produced by cultural outsiders and utilizing appropriated subject material for their own artistic means. Despite the humane intentions of such aspirations, it’s necessary not to forget the problematic implications of participating in representative practices that prompt to ‘give a voice to voiceless’, as if the subjects wouldn’t have a voice of their own. The writings of researchers like Rangan (2014) suspect the validity of an outsider perspective on the basis that it is reaffirming the hegemonic Western influence among indigenous cultures, and in developing nations already devastated by a history of colonization. As these fragile nations and communities are already struggling to get their voices heard over the discursive practices and depictions produced by cultural outsiders, it is called for to criticize the validity and need for yet another outsider representation. Young (2010, 59) is also admitting the problem from an ethical point of view, by noting how “there may be an asymmetry between the representation of a minority culture by a member of a dominant culture and the representation of a dominant culture by a member of a minority culture.” This asymmetry in the power relations between competing cultures
is affirming the need to consider questions of cultural hegemony when defining ethical guidelines for appropriating content from foreign cultures. This is also a valid concern in constructing representations of minority cultures by members of dominant cultures.

When critically reviewing filmic representations where the themes of ethnicity are weaved into the narrative fabric of the film—either through the protagonist or the overall cultural setting of the narrative—it’s necessary to consider whether the filmmakers (including directors, picture editors, sound designers, cinematographers, producers etc.) are openly admitting their position in relation to the culture or ethnicity they’re representing. In the case of *Between Rings*, there was clearly no doubt whether Chisi, sharing the directorial responsibility with her Finnish colleague Sorri, was a cultural insider—her being a native Zambian and a relative of the protagonist Esther Phiri. However, like I’ve suggested in this thesis, there are many features in the film which can be considered disputable in terms of authenticity and ethnographic accuracy. It might be beside the point to discuss ethnographic accuracy in relation to the outcome of *Between Rings*, considering the film is not meant to be an ethnographic film as such. Nonetheless, when a documentary film is depicting the struggle of an orphaned protagonist rising from a poor African township to an international boxing stardom, the dramatized juxtaposition of African and European realities brings us inevitably vis-à-vis with questions of inequality and otherness in the form of a quintessential *rags-to-riches* narrative. When carefully reviewing the production process and the documentarian outcome of *Between Rings*, it no longer seems significant whether or not there was a cultural insider involved in the directorial decisions of the film. The more we inspect the signification processes of documentary films, the more plausible it seems that the demands of artistic integrity and the conventional requirements of dramaturgy are dictating the production (and post production) processes of nonfiction films to a significantly greater degree, than a zealous drive for ethnographic authenticity or representational transparency.
4.3 Exotization and objectification by music

Considering that documentary films are often based on dramatizations of subject matter adopted from foreign cultures, documentary filmmakers are exceedingly liable to issues related to cultural appropriation and misrepresentation. The mere practice of producing an interpretative artistic representation without engaging in a dialogic relationship with the represented can be viewed as inviting exploitation and abuse. The connection between artistic appropriation and documentary film tradition is so tightly knit, that the very film often considered to be the first feature-length documentary, Nanook of the North (1922) by Robert Flaherty, is a prime example of subject appropriation and misrepresentation. The film is following the life of an Inuit hunter Nanook and his family, while they’re struggling for a livelihood in the harsh natural conditions of early 20th century Canadian Arctic. Flaherty shows Nanook as a resilient survivor, fiercely fighting against the odds at surviving in a hostile environment, hunting walrus with handcrafted harpoons and building igloos out of blocks of sheer ice to provide a shelter for his family. Flaherty had to laboriously enact most of the scenes seen in the film to tell his story of primeval struggle against the unforgiving elements of nature. He also decided to change the name of the protagonist from Allakariallak to Nanook, assumedly because he considered Allakariallak too difficult to pronounce for Western audiences.

Flaherty (as cited in Helke 2006, 27) famously said: “Sometimes you have to lie. One often has to distort a thing to catch its true spirit.” It is necessary to ask whether Flaherty had the knowledge and perspective as a cultural outsider to define the “true spirit” of indigenous Kalaallit people depicted in the film. It might lead to pointless speculation to imagine how a cultural insider would have depicted the daily life of Greenlandic Inuits in the early 20th century—assuming that she would have had an access to the same technology and filmmaking experience as Flaherty. Film theorist Richard M. Barthes (as cited in Helke 2006, 35) has criticized Flaherty’s methods for
simplifying and westernizing the cultures portrayed in his films. Helke (2006, 35) writes, how Flaherty constructed the representation of Inuit people on “mythical conceptions of indigenous people as innocent and untainted ‘noble savages’”26, actively setting up events, and projecting his personal fantasies of ‘primeval struggle’ on the actual culture of pre-war Inuits.

Despite Nanook of the North being probably the most discussed documentary film in the history of film, analytical studies related to its direction of score music are hard to come by.

![Figure 11. The official poster of Nanook Of the North (1922).](image1)
![Figure 12. The official poster of Between Rings (2014).](image2)

Being a product of the silent era of filmmaking, Nanook of the North has little to examine in terms of audio narration. During its time of existence, the film has been supplemented with three autonomous musical soundtracks. All three versions—first composed by Rudolf Schramm in 1947, and followed by renditions by Stanley Silverman in 1976 and Timothy Brock in 1997—share a noticeably European approach, devoid of any allusions to Inuit culture of the

---

26 Translated by the author.
Canadian Arctic. The latest and the most widely available score at the time, composed by Brock, features a chamber music orchestra comprising a string section, an oboist, flutist, clarinetist, harpist, and a bassoonist\footnote{http://www.timothybrock.com/joomla/music/original-scores/13-music/original-scores/44-nanook-of-the-north-1922-score-1997}. While Brock’s score is based on a romantic and occasionally melodramatic execution, it has also traces of 20th century modernism, reminding the listener of the harmonic ambiguity of early Schoenberg or Shostakovich. Although meticulously composed, the score chooses to emphasize—whether inadvertently or not—the documentary gaze of Flaherty—the director being an educated and wealthy American among poor indigenous people. With it’s utterly Western realization, the score is affirming the underlying depiction of the protagonist as adorably unaware of the ways of the civilized world. The colonialist attitude of Flaherty is glaringly apparent in the scene where Nanook is inspecting a gramophone, acting bewildered in front of a seemingly incomprehensible mechanical gadget. According to William Rothman (1997, 9–11), the scene is a prime example of filmic misrepresentation, as Allakariallak knew well enough what a gramophone was. The scene can be easily seen as deliberately reinforcing the colonial preconceptions of indigenous people as intellectually and culturally inferior to Western cultures, such as the one represented by Flaherty himself. Nanook is even shown biting a phonograph record handed to him by a white fur merchant, underlining his primal behavior when faced with the spoils of the civilized cultures. The ensuing impression is fortified by the playful chamber music theme composed by Brock, comprising a string arrangement accompanied by a clumsy-sounding bassoon and a whimsical flute supplementing the main melody, as delivered by the strings.

Mark Brownrigg (2007, 307) writes how the most basic procedure in constructing a sense of “geographical specificity” in film music is selecting and writing music for an instrument that’s commonly associated with a certain culture or country. He (ibid.) points out, how the broad majority of films are habitually composed for conventional Western orchestral instruments, such as strings, brass, woodwind, percussion, and grand piano, resulting in a film
music paradigm that is oblivious to the actual geographic and cultural context of the instruments. The prevalence of traditional Western orchestral instrumentation in film music is so strong, that while composers are incorporating distinctly European instruments and writing music that’s explicitly referencing specific periods in Western classical music, the resulting music can still feel geographically and ethnically neutral when presented in a film music context. Brownrigg (2007, 309) picks an example from one of the most celebrated soundtracks of our time to prove his point:

Even though it’s largely written following the conventions of Nineteenth Century European and Twentieth Century American Romanticism, when the march that opens Star Wars (1977) begins the audience is not intended to think ’Ah, here we are in the Western hemi-sphere’; to Western ears this music carries with it little sense of place. But over the neutral backdrop of the symphony orchestra, the introduction of an exotic instrument stands out and begins to articulate a sense of place.

This creates an interesting backdrop for expressions of cultural attributes and ethnic allusions in conventional film music. When the decision to create an synthesis between African and European aesthetics was made in the post production process of Between Rings, it seems questionable whether the aim was to actually bridge the gap between European and African music conventions—as equal cultural traditions brought together by artistic curiosity—or whether the inclusion of grand pianos and slick synthesizer sounds was an attempt to dilute the African influences to a more easily digestible form, so that they wouldn’t destabilize the hegemony and prevalence of Western gaze in the finished film. It is arguably true that the perspective in the film is European, despite Chisi being undoubtedly an insider in Zambian culture. When critically reviewing the score music of Between Rings from the prospect of cultural appropriation, it seems plausible that the inclusion of West African musical references was not reinforcing an insider perspective on the depicted culture, but rather confirming the otherness of Zambian reality by casually appropriating elements from West
African music traditions—for the sake of producing a sufficient effect of ‘documentary authenticity’ without destabilizing the European perspective in the narrative.

It might lead to a cycle of endless speculation to try to delineate other possible approaches for scoring a film such as Between Rings. When considering the authenticity of score music in documentary films, the self-contained quality of practically all film music is setting it apart from other kinds of audio narration in soundtracks. Although it might be clearly signifying and enhancing visually narrated events, efficiently tapping into dramaturgical turning points in the overall narrative (even directly alluding to specific ethnicities and cultural characteristics), the abstract being of music is felt as complete and self-contained in itself. This is arguably true even as it is being mixed with other audio content, e.g., dialogue, atmospheric tracks and digitally designed special effects. This leads to a predicament where music composed for a film will always—by inherent necessity—feel more or less insulated from other narrative devices, such as nonmusical auditive signifiers and sequenced images. Contrary to score music cues, filmic images and nonmusical noises are ingested and absorbed with each other in the overall narrative due to their intertwining indexical connections. Film music, however, is more often than not left to its own devices, referring to the dramaturgy of the narrative itself. As score music is mostly denoting fairly abstract and intangible features in the narrative, such as emotions and moods implicitly suggested by the concrete actions of social and professional actors, it is in a dissimilar position when compared to other auditive signifiers in the soundtrack. Douglas Kahn (1990, 67) writes, how the ideals of modernism lead to a failure to question the “representational operations” of contemporary Western music, despite the technological advancements in acoustic and electronic recording technologies offering previously unimaginable possibilities in auditively representing various worldly phenomena. The resulting “auto-referentiality”, according to Kahn (1990, 67), damaged the system of aural signification—the musical mechanisms enabling aural representation by musical gestures, “whereby the associative characteristics of
sounds, their attendant social and imaginative domains, were reduced, trivialized, or eradicated.” Kahn (1990, 68) points out, how one of the most prominent proponents of modernism in music, Karlheinz Stockhausen, preferred pure electronic sounds over the sounds produced by acoustic instruments exactly due to them being free of extraneous associations to tangible objects and their physical mechanisms in the sound production processes.

Would it be plausible to argue—based on the writings of Kahn—that by refraining from the use of acoustic instruments and by methodically favoring synthesizers in the processes of score music composition, the composers of documentary films would be exempt from accusations of cultural appropriation? Would a more pronounced employment of synthesizers in the score music of Between Rings have presented a less problematic representative framework for the story of a Zambian boxer, as depicted by a crew of filmmakers consisting mostly of white Europeans? The question will probably seem absurd from the prospect of practical filmmaking conventions, but it’s worth pondering nevertheless. Considering that most documentary films are practically dependent on subject appropriation, it’s necessary to outline applicable practices that are not oblivious to ethical problems related to notions of cultural property and direct (non-innovative) style or motif appropriation.

While the score music of Between Rings is comprising a hodgepodge of convoluting musical allusions to diverse cultural traditions, it is not trying to represent any particular musical tradition accurately as such. Although it is clearly referencing African influences—all the way from Senegalese folk to South-African urban dance pop—most of the appropriated elements and influences have been blended into a mixture that bears only a vague similarity to any specific folk tradition found in sub-Saharan Africa. It remains debatable whether the mere appropriation of individual musical motifs—such as rhythm patterns or the combination of certain instruments in a given composition—could be considered harmful to an associated culture or ethnic minority. Like Young (2010) is pointing out, all cultures have more or less
been developing by appropriating aethetcal content and motifs from other cultures, which makes the act of appropriation seem like an essential mechanism for individual cultures to remain in motion. This also applies to the development of documentary films, as they are through and through dependent on the practice of ‘telling other people’s stories’.

*Figure 13.* Nanook inspecting a gramophone in *Nanook of the North.*
5. CONCLUSIONS

Representing social actors and their immediate cultural environments requires constant ethical deliberation. In documentary films, the ethical responsibility of the filmmaker is often linked to questions of authenticity, i.e., whether the filmic representation is regarded as 'authentic' in its depiction of profilmic occurrences. Like asserted earlier, authenticity in dramaturgically treated narratives (fiction and nonfiction films alike) is a construction, inherently relying on the perceived indexicality of the recorded media to profilmic events and people—as recognized and understood through the probative force of verisimilitude, i.e., audiovisual likeness to 'real life' phenomena. However, indexicality in itself is not enough to produce a sufficient authenticity effect: the perceived authenticity of a documentary film being the sum of a myriad of artistic decisions made throughout the consecutive stages of a film production. Although enumerating the comprehensive methods of reinforcing the feel of authenticity in nonfiction narratives through the employment of sound (e.g., score music, foley recordings, atmospheric tracks, digital signal processing) is ultimately beyond the scope of this thesis, the perceived authenticity of the soundtrack is often thought to be the result of methodical exclusion rather than the outcome of an additive creative process. This assumption is implicitly reinstating the ideological approach of aesthetic austerity, as discussed in this thesis. Whereas the balance between informativity and aesthetics is obviously dependent on artistic decisions—such as the employment of non-diegetic music to build mood and punctuate emotive content—proponents of aesthetic asceticism are reflexively rejecting the audio post production practices of fiction filmmaking as ethically suspicious or otherwise redundant. However, in contemporary documentary film productions, these methods and techniques are in common use, like pointed out by Aaltonen (2010). The fact that their contribution to the resulting filmic representation is not openly
discussed by film theorists and filmmakers creates an air of confusion around those very practices. This is an issue that this thesis hopefully manages to address. Like suggested in earlier chapters, instead of checking arbitrary authenticity boxes when comprising nonfiction narratives in post production processes of documentary films, it's arguably more important to determine how the aforementioned narrative devices are employed in the film, rather than whether one should employ them or not.

When we examine the narrative outcome of Between Rings, it seems that the film is trying to assure us that everything is possible—even for a poor woman of color raised in an African ghetto. Because this emotive message is presented to us partly through dramaturgically edited verbal accounts of the protagonist herself, it feels genuinely empowering rather than an aphoristic neoliberal morality. The outcome would be obviously different if the authentic voice of Esther wouldn’t be acting as a mediating agent; if her words would be replaced by the voice of the director or a trained actor narrating a semantically identical message.

Despite the films’ heartfelt empowerment, it’s well worth asking whether the story of Esther would have struck an even deeper note if the same narrative ambiguity defining the final scenes of the film would have been utilized in the film more comprehensively. In the penultimate scene of the film, we witness Esther wandering alone through the marble hallways and polished corridors of her fiancé’s mansion. The metallic clatter of her high-heeled shoes echoing in the empty rooms is accompanied by a meaningless racket of a television show, reverberating somewhere in the background. The inferred feel of these diegetic sounds is evoking an impression of emptiness (facilitated by the unusually reverberant acoustics inherent in the direct indexical recordings), and an unexplained sense of solitude (only lifeless gadgets keeping her company). It is as if the scene is gently implying that Esther might have been happier living the simple life predating her troubled boxing stardom. The suggested message—with its inferred sense of detachment and solitude induced by the nonverbal noises—is affirmed when juxtaposed with the following scene. The scene is comprised of a single wide
shot of a slightly run-down looking Lusakan neighborhood, commanded by an ancient looking tree covering the view with its protective branches. We see a group of local people walking along a footpath, chattering with each other while they move further away from the fixed camera position. After a short while, two boys enter the frame, jointly driving a rickety bicycle. The bicycle clanks and creaks, as it gives an offbeat rhythmic backdrop to the laughter of the boys while they windingly pedal towards the sunny vanishing point of the frame. No poignant score music or introspective monologue is needed to emphasize the immediacy of these final scenes of the film, as they are intuitively felt and recognized by the audience as authentic moments bearing an identifiable familiarity. As everyday glimpses of longing and intrinsic beauty—devoid of imposed sentiments and ideological arguments presented in the form of didactic verbalizations—these seemingly uneventful scenes convey a sense of universal stature despite their specific cultural setting. The authenticity of the moments is felt within a heartbeat—a feat not hindered by the fact that every sound in the final scene was either recorded elsewhere in Zambia, drawn from a commercial sample library, or performed as foley recordings in a Finnish post production studio.
REFERENCES


Ricciardelli, Lucia. 2014. American Documentary Filmmaking in the Digital...


Articles


Eisenstein, Sergei. 1998. The Dramaturgy of Film Form (The Dialectical Approach to Film Form). In The Eisenstein Reader. London: British Film Institute, 93–110.


Online references


*Low-frequency effects*. https://en.wikipedia.org/wiki/Low-frequency_effects


Films


Recorded music