

City University of New York (CUNY)

## CUNY Academic Works

---

Dissertations, Theses, and Capstone Projects

CUNY Graduate Center

---

2-2023

### Speakers and Addressees as Creative Interpreters

Svitlana Novikova

*The Graduate Center, City University of New York*

[How does access to this work benefit you? Let us know!](#)

More information about this work at: [https://academicworks.cuny.edu/gc\\_etds/5210](https://academicworks.cuny.edu/gc_etds/5210)

Discover additional works at: <https://academicworks.cuny.edu>

---

This work is made publicly available by the City University of New York (CUNY).

Contact: [AcademicWorks@cuny.edu](mailto:AcademicWorks@cuny.edu)

**SPEAKERS AND ADDRESSEES AS CREATIVE  
INTERPRETERS**

by

SVITLANA NOVIKOVA

A dissertation submitted to the Graduate Faculty in Philosophy in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2023

© 2022

SVITLANA NOVIKOVA

All Rights Reserved

Speakers and Addressees as Creative Interpreters

by

Svitlana Novikova

This manuscript has been read and accepted for the Graduate Faculty in  
Philosophy in satisfaction of the dissertation requirement for the degree of  
Doctor of Philosophy.

---

Date

---

[Gary Ostertag]

Chair of Examining Committee

---

Date

---

[Nickolas Pappas]

Executive Officer

Supervisory Committee:

Stephen Neale (Advisor)

Noël Carroll

Daniel Harris

Gary Ostertag

THE CITY UNIVERSITY OF NEW YORK

# ABSTRACT

Speakers and Addressees as Creative Interpreters

by

Svitlana Novikova

Advisor: Stephen Neale

Philosophers in the past have argued that the way the word “interpret” is used within creative fields unhelpfully conflates the notions of deriving and creating content. Arguing against this, I propose that the blurring of these two notions accurately describes how addressees interpret speakers, performers interpret scores, and audiences interpret art and music. Even speakers can be described as creative interpreters in this sense, as articulating a thought into an utterance requires an interpretive effort. I develop the idea that interpreting straddles the divide between deriving and creating content within the framework of the ostensive-inferential communication proposed by Sperber and Wilson (1986/95) in *Relevance: Communication and Cognition*. I also suggest that Sperber and Wilson’s framework should be modified to recognize both communicative and interpretive intentions. The notion of interpreting that I advocate supports the view of ostensive communication as spanning from regular everyday communication to creative and artistic expression in arts, music, and literature.

## ACKNOWLEDGEMENTS

I am deeply indebted to many people for making this project possible. First, I would like to express my deepest appreciation to my advisor Stephen Neale for inspiring and challenging me, for making me renew my enthusiasm when my goals seemed elusive, and for his patience and humor. I have grown tremendously by pursuing this journey under Stephen's guidance. I would like to thank Daniel Harris for being not only an outstanding academic mentor but also an understanding friend. Dan's sharing his experiences as a graduate student helped me pull through in times of self-doubt. I am incredibly grateful to Gary Ostertag for his thought-provoking commentary on my writing, sharing his experiences as a musician, and introducing me to a whole slew of literature on musical cognition. My sincere thanks to Noël Carroll for his generously provided knowledge and expertise as a member of my defense committee. I also immensely appreciate the guidance of John Greenwood. I have come to rely on his helpful advice and encouragement throughout my time as a Ph.D. student.

This endeavor would not be possible without the support of my family and friends. My husband, Shi Nisman: this project is your accomplishment as much as it is mine. My children and my mom: thank you for providing the fodder for thinking about communication (and miscommunication), but mostly for your love and patience. My many cherished friends: thank you for cheering me on during this grueling but rewarding intellectual marathon.

# CONTENTS

<b>TABLES</b> .....	viii
<b>CHAPTER 1: INTRODUCTION</b> .....	1
1.1 Two senses of “interpret” .....	1
1.2. How underdeterminacy inverts traditional descriptions.....	3
1.3 Some crucial terms and assumptions .....	8
1.4 The roadmap.....	9
<b>CHAPTER 2: APPROACHES TO COMMUNICATION</b> .....	12
2.1 The preservationist approach .....	12
2.2 The anti-preservationist approach .....	15
2.3 Approaches to communication implied in art and music .....	22
<b>CHAPTER 3: CONNECTING MUSIC AND OSTENSIVE COMMUNICATION</b> .....	26
3.1 Perceived differences .....	28
3.1.1 <i>What</i> is conveyed: propositional versus non-propositional content .....	29
3.1.2 <i>How</i> the content is conveyed: <i>meaning</i> vs. <i>showing</i> .....	30
3.1.3 What it means to understand: inferring vs. experiencing .....	32
3.1.4 Absolute music .....	33
3.2 Communication as a matter of degree.....	35
3.2.1 Degrees of propositionality .....	36
3.2.2 Degrees of <i>meaning</i> and <i>showing</i> .....	38
3.2.3 Understanding as a matter of degree .....	39
3.3 Erasing distinctions between music and regular communication .....	41
3.3.1 Absolute music versus utterances.....	42
3.3.2 Timeless meaning: musical and linguistic.....	48
<b>CHAPTER 4: MUSIC</b> .....	53
4.1 The failure of traditional models to describe music .....	53
4.2 Musical performers: prying creativity from inferring .....	56
4.3 The musical audience .....	63
4.4 The composer .....	68
4.5 Musical performance as a relay of interpretations .....	71
<b>CHAPTER 5: OSTENSIVE COMMUNICATION</b> .....	74
5.1 Creative intermediary: reporting on the speech of others .....	74

5.2 Audience as a creator of content .....	89
5.3 The speaker as an interpreter of thoughts.....	97
<b>CHAPTER 6: CONCLUSION</b> .....	103
6.1 Summary .....	103
6.2 Implications.....	106
<b>BIBLIOGRAPHY</b> .....	110



## TABLES

Table 1: Comparison of a reporter's mental representation of what the speaker means and the reporting utterance, as traditionally understood.....	77
Table 2: Proposed comparison of a reporter's mental representation of what the speaker means and the reporting utterance.....	79

# CHAPTER 1: INTRODUCTION

## 1.1 Two senses of “interpret”

How would a theory of communication describe the role of a musical performer? In interpreting a musical composition, a performer both infers what the composer meant to be played and expresses their own creative intentions toward the audience or potential audience. Musical performers are not typically considered mere middlemen or facilitators between the audience and the composer; instead, people recognize them as artists and their work as a creative pursuit underpinned by technical mastery. However, traditional theories of communication neatly divide the communicating parties into those who create content and those who derive it in some fashion: speakers and addressees, authors and readers, artists and the audience. This is the core of what is often called the producer-consumer view of communication. The role of agents whose actions involve creating and deriving content does not correspond to the traditional descriptions given by this view of communication.

The way the word “interpret” is used within the creative fields is also discordant with traditional descriptions of communication. In the arts and humanities, to interpret is not only to derive meaning but also to create or express meaning. For example, an artist interprets a scene in a painting, a dancer interprets music with her movements, a director interprets Othello, and a chef interprets a blueberry pie. Some have argued that such uses of the word interpret are no more than unhelpful coincidences (Levinson, 1993) and that the act of expressing meaning should not be conflated with the act of deriving it. In this work, I argue against this contention. A notion of

interpreting that straddles the divide between producing and deriving content is needed, not only in the arts, but also in an adequate theory of verbal communication, in both its formatics, which is meant to characterize the cognitive processes and information bases involved on the speaker side, and in its pragmatics, which is meant to characterize those involved on the audience side. Indeed, it is my contention that a proper understanding of the relationship between formatics and pragmatics requires the notion of interpreting I am advocating.

Traditionally, the speaker has been understood exclusively as a content producer with the addressee a mere passive copier of content. Cognitive science has associated their actions with different types of mental processes: the process underlying the speaker's act has been seen as active, deliberate, and intentional, while the process underlying the addressee's act as largely automatic, mandatory, and intention-independent. Instead, I argue that neither manifest intention nor automaticity should serve as marks that distinguish the mental processes underpinning the speaker's and addressee's acts. Utterances can be performed "on autopilot," without deliberation, with the speaker later regretting not paying attention to what they say. On the other hand, the process of comprehending the speaker can sometimes be deliberate, even laborious, requiring creativity and imagination, and even with self-directed intention. Therefore, I suggest that the speaker and the addressee are both interpreters in this sense: both *create* content, often with manifest intention, within constraints so tightly interrelated that there is a perfectly good sense in which we can talk about a *single* set of bidirectional mental processes.

In Section 1.2, I present my proposal and suggest that the requisite notion of interpreting is an outcome of pervasive underdeterminacy inherent in human communication. In Section 1.3, I

explain some crucial terms and assumptions, particularly my use of the terms “content” and “proposition.” Section 1.4 gives a roadmap to the rest of the dissertation.

## **1.2. How underdeterminacy inverts traditional descriptions**

My proposal is not a figurative expression or a metaphor; it is an upshot of the ostensive-inferential model of communication taken to its natural conclusion. The model was formulated by Dan Sperber and Deirdre Wilson (1986/95) in their groundbreaking book, *Relevance: Communication and Cognition*, but its philosophical foundations were laid by Paul Grice beginning with his paper “Meaning” (Grice, 1957) and expounded in his later works (Grice, 1968, 1969, 1989). The model describes the most prevalent way humans communicate – not by simply coding and decoding content, as was commonly understood, but rather by offering and interpreting evidence of communicative intentions.

When communicated by code, content can be duplicated or nearly duplicated. When communicated by offering evidence, there is pervasive underdeterminacy: there are multiple possible states of signal for every thought the producer intends to communicate and multiple possible ways for the addressee to interpret the signal. The speaker’s thought constrains but does not determine the resulting utterance. The utterance constrains but does not determine the addressee’s interpretation.

Every thought thus allows a variety of equally suitable expressions<sup>1</sup>. I can communicate my desire to leave the restaurant to the server by saying, “I’m finished!” or “May I have the check please?” or by making a signing motion with my hand. My understanding of the context and of the background assumptions and abilities of my addressee constrain my utterance further. What I utter or do to convey my thought might be different depending on whether I am at a diner or a fancy restaurant or whether the waiter can hear me. However, even after all such situational constraints are satisfied, there remains an indefinite number of different sentences I could use or actions I could produce to communicate my desire to leave the restaurant: “I am done for tonight,” “I don’t want to keep you here all evening,” “Why don’t we free up the table for the next party.”

Likewise, even a straightforward utterance, such as “Can I have the check?” allows many equally suitable interpretations. Constrained by the linguistic meaning of the words and their view of the context, the addressee may understand that the speaker is finished with the meal, would like the table cleared, or is ready to leave. Additional cues, such as the speaker’s intonation and body language may further constrain the addressee’s understanding that the speaker is impatient with slow service or upset and will not order dessert. Still, equally suitable options will remain even after all the relevant constraints on the addressee’s understanding are fulfilled.

---

<sup>1</sup> This assumes a method for individuating thoughts. However, in this context, the only notion of sameness of thought that is needed is the sameness from the perspective of the thinker or *de jure* sameness. (The notion of *de jure* sameness is borrowed from Schroeter, 2012). For the thinker, the appearance that they are thinking the same thought or thinking about the same subject is always obvious and primitive: “the subject matter is simply served up to your conscious attention as *the same*. There is no need to engage in any reflection or calculation in order to recognize their sameness: it seems self-evident that there’s just one subject matter in question” (Schroeter, 2012). *De facto* identity claims are not obvious, incontrovertible, or primitive in the same way. Saying that the same thought allows for a variety of equally suitable expressions implies the same thought from the perspective of the thinker.

How do the speaker and addressee overcome such underdeterminacy to arrive at a concrete result among the many options? A leap of creativity is required of both the speaker and addressee in addition to fulfilling all the relevant constraints. The process by which each one arrives at the result—my proposed notion of interpreting—crucially *blurs* creating and deriving content.

In forming an utterance, the speaker cannot simply duplicate and send the thought they are trying to convey. A further medium is required, a public language “into” which the thought can be represented. A creative-cum-inferential effort is therefore required to form the best surrogate or proxy for the thought in a different medium within the constraints imposed by this medium, the circumstances of the utterance, and the audience’s needs. We can think of this as interpreting a thought “into” an utterance in much the same way we can think of a musician interpreting a score “into” a performance. Like any interpreting, it can be unreliable. Just as a composer may be unhappy with how the performer interprets their score into a performance, the speaker may be unhappy with how the resulting utterance reflects the thought it is meant to convey. One can sometimes struggle to articulate their thoughts and miscommunicate what is really on their mind. The success of interpreting thoughts into utterances depends on the factors that determine the success of other creative efforts—partially on the skill of the interpreter which includes navigating the constraints of the genre (i.e., how masterful they are in using words or playing a musical instrument), partially on the creative spark, and partially on chance.

For the addressee, underdeterminacy means that they cannot duplicate the original (the speaker’s) content but rather must create content within the constraints, imposed by the linguistic meaning of the words, the context of the utterance, and the beliefs and assumptions assumed to be shared with the speaker. However, the things not shared with the speaker are inevitably

reflected in the final result: the addressee's unique background, perspective, values, biases, and preferences. The restaurant server may take my remark differently depending on whether they are temporarily moonlighting or have years of experience, whether they are closer to the beginning or the end of their shift, their mood, and even on the stereotypes they may consciously or unconsciously assign to me as a person. Utterances that involve figurative language or weak implicatures require the addressee's outright imagination and creative choices to be understood, while heavily theoretic speech and texts require deliberate and even laborious interpretive effort.

While I see my proposed notion of interpreting as a natural outcome of the ostensive-inferential model of communication, some fine-tuning of the model is needed to formulate this notion explicitly. Particularly, I argue that intention should be treated symmetrically between the acts of the speaker and addressee. Some speaker's acts are not manifestly intentional. At the same time, invoking only the speaker's intentions and not those of the addressee in the situations that require the addressee's effort of imagination and even responsibility for the resulting interpretation is inconsistent. Furthermore, interpretive intentions explain the differences among various interpretations by different addressees of the same speaker meaning that cannot be attributed to different constraints imposed by the speaker. I also propose that if we posit addressee's intentions, we should recognize that success conditions of a communicative act are not uniform for both the speaker and the addressee, but relative: a communicative act can be successful for the speaker but not the addressee, and vice versa. Such are the cases when the addressee recognizes certain content in an utterance that the speaker aims to conceal, as well as cases when the addressee "twists the speaker's words" or "puts words in the speaker's mouth."

To be sure, the acts of the speaker and the addressee should still be understood as distinct and complementary. The first difference is the directionality of intentions: the speaker's communicative intention is directed at the addressee, while the addressee's intentions are self-directed. However, such directionality of intentions is contingent. A speaker could articulate an internal monologue in an act of speaking with self-directed intentions. An addressee could form a representation of what the original speaker meant and direct it at a secondary audience in an act of retelling original speech to someone else. The second difference emerges in connection with the modularity of cognition: it is the difference in the sequencing of when the modules of input and production become active and when more general-purpose processes kick in. However, what happens in between the modules of input and production, the heart of the interpretive process, is remarkably similar whether one is interpreting an utterance into an internal representation, a thought into an utterance, or one mental representation into another (in case of an internal monologue).

I view my project as expanding the scope and proposals of the ostensive-inferential model of communication. Just as Sperber and Wilson's proposals pushed the Gricean account of communication beyond speaker meaning to include a broader range of ostensive communication, I believe my proposals help push Sperber and Wilson's account even further into embracing the acts of creative expression, or the interactions between artists and the audience, authors and readers, musicians and listeners.



### 1.3 Some crucial terms and assumptions

At this point, I must address the use of the term “content” in this work. In the philosophy of language, it is traditional (but not uncontroversial) to think of the content of our thoughts, beliefs, and communicative intentions as truth-evaluable propositions, or things designated by “that” clauses. For example: by saying, “It’s white,” the speaker meant that the snow is white. What “that” clause “the snow is white” specifies is a proposition.

There are several debates about the nature of propositions. Are propositions structured entities whose constituents and structure determine their identities?<sup>2</sup> If so, what is the nature of these constituents? The position associated with Russell is that propositional constituents are objects and properties. The position associated with Frege is that propositional constituents are “modes of presentation” of objects and properties.<sup>3</sup> My discussion aims to be as neutral as possible on this matter. However, I incline toward the Fregean position. (If at any point in the discussion it is no longer possible to remain neutral on these issues, I indicate as much.) I also remain open to the idea that the content of some thoughts and communicative intentions is not propositional but qualitative (e.g., when the speaker intends to induce a sensory state, sentiment, or emotion in the addressee).

Another important preliminary note is that on the ostensive-inferential account, it may not make sense to speak of the contents of utterances as propositions. If the content of a communicative intention is a proposition and an utterance is evidence for content, then the utterance is evidence

---

<sup>2</sup> For discussion of structured propositions, see King (2019).

<sup>3</sup> For discussion of Russellianism and Fregeanism in understanding propositions, see Fitch and Nelson (2013).

for a proposition, not a vehicle for it.<sup>4</sup> An utterance, rather, may provide only a propositional template, or a blueprint for creating a proposition in the mind of the addressee.<sup>5</sup>

## 1.4 The roadmap

In Chapter 2, I contrast two general approaches to communication, preservationist and non-preservationist, and two models of communication, code and ostensive-inferential, best representing each approach. The preservationist approach constitutes the idea that in communication, content is mostly preserved, perhaps even duplicated from the mind of the speaker to the mind of the addressee. Contrastingly, on the non-preservationist approach, the speaker only provides the audience with evidence to the speaker's communicative intentions and the audience must infer the speaker's intentions from such evidence. I also describe the discontinuity between our conceptions of everyday human communication and that of creative expression within artistic sphere (i.e., visual arts, music, literature, poetry, and so on). I argue that the perception of the discontinuity between the two spheres is due to the fact that in comparing regular human communication with creative expression, we essentially compare the preservationist conception of regular communication with the non-preservationist conception of artistic communication. The ostensive-inferential model can bridge our conceptions of communication across many domains, including creative.

---

<sup>4</sup> This point emerged in the discussion with my advisor Stephen Neale.

<sup>5</sup> A precedent for such a view appears in (Buchanan, 2010). Buchanan suggests that the addressee constructs a proposition using the *character* or the *context-invariant meaning* of a sentence uttered, along with contextual considerations, as a blueprint.

In Chapter 3, I argue that the acts of creative expression in arts, music, poetry, and literature fit the account of *weak communication* developed by Sperber and Wilson. The continuum of ostensive-inferential communication should therefore include the interactions between artists and viewers, musicians and listeners, authors and readers, and so on.

In Chapter 4, I first develop my notion of interpretation as straddling the divide between creating and deriving content using the act of a musical performer. Then I argue that everyone involved in a musical performance, including the composer and the audience, are interpreters in the same sense of the term. Finally, I develop the notion of an interpretive chain: a musical thought in the composer's mind is interpreted into the score, which is then interpreted into a performance, which is in turn interpreted into a mental representation in the listener's mind, which constitutes the listener's understanding of the music. It can be further interpreted into other performances and even works of art. I argue that every participant in this chain of interpretation both expresses their own creative intentions and is simultaneously constrained by the intentions of those who stand prior to them in the interpretive chain.

Chapter 5 transfers my notion of interpreting to the realm of regular human communication. First I use my notion of interpreting to describe the act of an intermediary reporting on the speech of another person. Next, I argue that all participants in an act of regular communication, even speakers and addressees, are also interpreters in the same sense—between the public utterances and their mental states and vice versa. The notion of interpretive chains describes how news, information, and beliefs propagate throughout the society.

Finally, the conclusive 6<sup>th</sup> Chapter summarizes my argument and illuminates its implications.



## CHAPTER 2: APPROACHES TO COMMUNICATION

### 2.1 The preservationist approach

When we speak of the traditional or common understanding of communication—the way a layperson would answer the question, “How does communication work?”, and how communication used to be described in philosophical literature prior to the second half of the twentieth century—what is typically described is something as follows. A speaker has a thought, wishes to share it, and expresses it through an utterance. An addressee interprets the utterance for content. Ideally, a copy of the content from the speaker’s mind is now in the mind of the addressee. This picture of communication can be termed the *preservationist* or *copying* approach to communication. Even the word *content* presupposes this view: it comes from the medieval Latin word *contentum* or “thing contained,” as if people exchange packages containing messages.

It is traditional to think of content as propositions. Therefore, on this preservationist approach, the same proposition “*the snow is white*” first inheres in the speaker’s brain state, then in an utterance, and finally, in the addressee’s interpreting brain state. This way, the content relates to the medium in the same way that the sculpture relates to its physical material. How does the content inhere in the medium, and how is it preserved in transmission? Different preservationist models of communication may give different answers to this question. The most prominent of such models is *code model*. On this model, there is a stable association, or code, between the content and how it is represented in its medium. In the most familiar example of code

communication, a red light is stably associated with the requirement to stop at a traffic light according to the code established by the traffic rules. That is, a red light encodes and, in this model, *means* “stop.” Someone who believes that the code model adequately describes how people communicate linguistically would say that the speaker encodes the content in the sounds of the words and the addressee decodes the content using the dictionary definitions of the words. Both the preservationist approach and the code model are simple and intuitive. They can undoubtedly describe many forms of communication, such as traffic signs, smoke signals, maritime flags, communication by primitive organisms, and, perhaps, animal calls (Scott-Phillips, 2014).

The purely preservationist accounts of everyday human communication figure in the writings of Locke, Frege,<sup>6</sup> and much of the twentieth-century philosophy of language. It is entrenched in both common understanding and philosophical tradition. In his account of testimony, Tyler Burge states that duplicating content is a paradigmatic case of verbal communication: “In interlocution, perception of utterances makes possible the passage of propositional content from one mind to another rather as purely preservative memory makes possible the preservation of propositional content from one time to another” (Burge, 1993, p. 481). The need for more active decision-heavy interpreting, he says, arises only in exceptional instances when there is doubt about whether a communication is successful: “This situation [of taking the speaker as a source of information and not as an object of interpretation] is basic for communication. Radical

---

<sup>6</sup>Frege was also one of the first philosophers who began to depart from the purely preservationist account. For example, he argued that thoughts about ourselves could not be communicated (Frege, 1956, p. 298). Each person is presented to themselves in a private and unique way. These private and unique modes of presentation cannot be shared with others, and therefore, thoughts involving them cannot be communicated.

interpretation is not, I think, the paradigmatic situation for theorizing about linguistic interchange” (Burge, 1993, p. 487).

The contrast between the producers and consumers of content is stark under the preservationist approach. The speaker is in control of *what* and *how* they communicate, while the addressees are powerless to affect what they receive. Everyday intuitions about speaking and interpreting seemingly confirm this contrast. Try as one might, they cannot hear “It’s raining” and *will* themselves to understand, “I love you.” At best, such understanding would be insincere.

While not directly associated with the preservationist approach to communication or the code model, Fodor’s modularity of mind hypothesis in its early version also associated utterance comprehension with a cluster of properties such as automatic, fast, mandatory, and informationally encapsulated (Fodor, 1983). In the early version of the modularity thesis, the addressee’s interpreting consisted of the operations of the special-purpose linguistic module and the general-purpose central systems. The module is specialized for a particular evolutionary purpose – turning the sound of the words into linguistic output that can be processed by the central systems. Therefore, at least the first stage of the addressee’s comprehending, one performed by the linguistic module, is fast, automatic, mandatory, and characterized by the lack of access to information from other parts of the mind. On the other hand, the operations of the central systems are general-purpose, deliberate, conscious, and informationally unencapsulated.

## 2.2 The anti-preservationist approach

In the second half of the twentieth century, the understanding of everyday human communication began to change. Philosophers, and in particular, Paul Grice, realized that much of human communication happens not by coding and decoding but by offering and interpreting the evidence for a certain kind of intention.

Grice thought that such intention had to fulfill three characteristics (Grice, 1969). Firstly, the speaker must intend to produce a particular response in the audience (e.g., a representation in the addressee's mind that someone is at the door). Secondly, the speaker must intend for the addressee to recognize that the speaker intends to produce a response from the addressee. If the speaker achieved the intended representation in the addressee's mind by letting them see for themselves, this would not qualify as Grice's communicative intention. Finally, the speaker must intend to achieve this particular response in the addressee *in virtue* of the hearer recognizing the speaker's intention. If the speaker opened the door to make the visitor visible to the addressee, this would still not fulfill Grice's requirements for a communicative intention, as the representation would be achieved in virtue of the addressee grasping the evidence for the visitor, not due to grasping the evidence for the speaker's intention. However, if the speaker gestured at the door to indicate that someone was there or uttered, "Someone is at the door," those situations would fulfill all three of Grice's requirements for a communicative intention.

All three requirements must also be fulfilled for the speaker to *mean* something. Grice distinguished *speaker meaning* from *natural meaning* (i.e., the way smoke *means* fire and bird migration *means* seasonal change). While in the code model, meaning is in virtue of the rules and



conventions, in Grice's description of verbal communication, speaker meaning is in virtue of the speaker's communicative intentions.

While the spirit of Grice's definitions of communicative intentions and speaker meaning are widely embraced, their exact forms have been widely debated. Many judged the three-part definition as too restrictive. For them, deliberately opening the door and letting the addressee see the visitor qualifies as a form of communication that should be explained by a theory of communication. Sperber and Wilson in particular argued that speaker meaning that fulfills all three of Grice's requirements is but a segment on a wider continuum of ostensive-inferential communication (Sperber & Wilson, 2015).

Grice described how the speaker and addressee are mutually constrained to assure communicative success. The speaker is constrained to form an utterance in the way that would provide the addressee the best evidence of the speaker's communicative intentions. Virtually anything can serve as such evidence: the conventional linguistic meaning of the words and phrases, the conversation topic, shared assumptions and beliefs, features of the immediate environment, pointing, gestures, and facial expressions. For their part, the addressee is constrained to form an interpretation that best reflects the evidence given by the speaker.

Grice's account of communication was still largely preservationist, since he identified intention demonstration and recognition as playing a role mostly in implicit communication, while he believed that explicit communication was largely code-like. Still, inspired by Grice's account, pragmatists began to recognize that demonstration and recognition of intention is pervasive and plays a critical role even in explicit and routine communication (Carston, 2000; Sperber & Wilson, 1986/95). It is easy to see that communication by offering evidence can at best

*approximate* content but is unlikely to *transfer, copy, or preserve* it. Therefore, if we communicate by offering evidence rather than coding, what enters the addressee's mind is not identical to what was originally in the mind of the speaker, even in the case of straightforward and mundane remarks. Different hearers may produce different but equally valid interpretations.

Different positions on whether content is preserved in communication should be thought of as occupying a range: it is possible to conceive of communication as duplicating content, nearly duplicating, as well as varying extents to which the content in the speaker's mind and the mind of the addressee overlap or to which extent they are similar. To be sure, the opposite extreme of preservationism, the idea that *nothing* is preserved between the speaker and the addressee is not viable: it would not explain how people can agree, disagree, and coordinate their actions based on the acts of communication. Most pragmatist philosophers today hold positions consistent with the middle of the spectrum.

Buchanan (2010), for example, argues that speaker intentions are often characterized by a certain generality and indifference toward the exact way the audience may interpret their utterance. Such generality allows for multiple, equally suitable interpretations. He proposes that the content of the speaker's communicative intention is a proposition-type or a *blueprint* from which the addressee constructs their own proposition.

Cappelen (2008) argues that an addressee occasionally plays a *content-producing* role and offers examples from other theorists' works. The content of certain terms varies with the context of interpretation: the indexical "you" in "Jesus loves you" pronounced by a televangelist acquires its content within the context of each addressee (Egan, 2009, as cited in Cappelen, 2008).

Epistemic modals (e.g., "Billy *might* be in Boston") create an *asymmetric dynamic* between

speaker and addressee. Addressees tend to interpret such statements in the strongest possible sense (“Billy is probably in Boston”), while speakers, upon questioning, may retreat to a weaker interpretation (“I only said that he *could* be in Boston, but I was not sure”) (Von Fintel & Gilles, 2008, as cited in Cappelen, 2008).

Bezuidenhout (1997) and Carston (2002) assert that there is only an *unspecified* relationship of similarity between the speaker’s mental state and the addressee’s interpretive mental state:

Since utterance interpretation is always in the first place colored by one’s own cognitive perspective, I think we should reject the idea that there is an intermediate stage in communication which involves the recovery of some content shared by speaker and listener and which is attributed by the listener to the utterance. [...] We need recognize only speaker-relative utterance content and listener-relative utterance content and a relationship of similarity holding between the two contents (where the speaker-relative utterance content is identical to the content of the thought the speaker expresses and the listener-relative utterance content is identical to the content of the thought the listener comes to entertain. (Bezuidenhout, 1997, p. 212)

The most explicit and formalized departure from the preservationist paradigm is the ostensive-inferential model of communication presented by Sperber and Wilson in their book *Relevance:*

*Communication and Cognition:*

It seems to us neither paradoxical nor counterintuitive to say that there are thoughts that we cannot exactly share, and that communication can be successful without resulting in an exact duplication of thoughts in communicator and audience. We see communication as a matter of enlarging mutual cognitive environments, not of duplicating thoughts. (Sperber & Wilson, 1986/95, pp. 192–193)

While it was not intended to replace the code model (Sperber and Wilson take coding and decoding as one of the many means by which people display and infer intentions), the ostensive-inferential account is a far departure from the code model and the preservationist approach in general. It expands the scope of Grice’s proposals and the scope of what a theory of communication should explain. Sperber and Wilson show that demonstration and recognition of intentions play a role in communication in general and not only in implicature, that speaker

meaning in the strict Gricean sense is only a limiting case of many rich ways people communicate, that *weak communication* (i.e., the kind that is not meant to cause a specific representation in the addressee's mind but only pushes the addressee's thoughts in a certain, if vague, direction) is equally as important as *strong communication*. Finally, they argue against glossing over and idealizing away the effects of intonation, gesturing, and other kinds of ostension in communication.

In the account of communication that Sperber and Wilson presented, the dichotomy between producers and consumers of content becomes more nuanced. The ostensive-inferential model no longer presupposes the speaker's absolute autonomy or the interpreter's complete dependency. Instead, the speaker is constrained in *what* and *how* they communicate by their expectation of what the addressee will infer. At the same time, the addressee must show a degree of autonomy and imagination in connecting the evidence to possible interpretations.

*Relevance* occasionally swaps the descriptions of the speaker's and the addressee's actions. It describes an act of forming an utterance as *interpreting* the speaker's thoughts:

We see verbal communication as involving a speaker producing *an utterance as a public interpretation of one of her thoughts*, and the hearer constructing a mental interpretation of this utterance, and hence of the original thought. [...] It follows from our general account of inferential communication that an utterance should be an interpretive expression of a thought of the speaker's. However, we see no reason to postulate a convention, presumption, maxim or rule of literalness to the effect that this interpretation must be a literal reproduction. How close the interpretation is, and in particular when it is literal, can be determined on the basis of the principle of relevance. (Sperber & Wilson, 1986/95, pp. 230–231)

For Sperber and Wilson, a thought constrains but does not determine an utterance. To go from a thought to a concrete utterance is to construct a surrogate or proxy for the thought within the relevant constraints (which is how the addressee is more frequently described). As for the

addressee, they are described as constructing, exercising imagination, and bearing personal responsibility for the interpretive representations they achieve:

[Certain acts of interpreting] require[s] *an effort of imagination* [emphasis added]: the hearer has to bring together relatively unrelated encyclopaedic entries and *construct* [emphasis added] non-stereotypical assumptions. (Sperber & Wilson, 1986/95, p. 222)

In general, the wider the range of potential implicatures and *the greater the hearer's responsibility for constructing them* [emphasis added], the more poetic the effect, the more creative the metaphor. [...] In the richest and most successful cases, *the hearer or reader can go beyond just exploring the immediate context* [emphasis added] and the entries for concepts involved in it, accessing a wide area of knowledge, *adding metaphors of his own* [emphasis added] as interpretations of possible developments he is not ready to go into, and getting more and more very weak implicatures, with suggestions for still further processing. The result is a quite complex picture, for which *the hearer has to take a large part of the responsibility* [emphasis added], but the discovery of which has been triggered by the writer. (Sperber & Wilson, 1986/95, p. 236)

While the contrast between a producer and a consumer of content has been blurred, in some ways the dichotomy persists within the ostensive-inferential model. Only the speaker's intentions play a role in explaining communicative behavior. While much has been said by Sperber and Wilson about the hearer's unique perspective, imagination, and taking responsibility for the resulting interpretation, they do not connect such psychological states with having intentions.

Additionally, while their version of the modularity thesis is more nuanced, they still associate the addressee's comprehending with its own special-purpose module, the pragmatic one. To be sure, under their version of the thesis, the term "module" is used in a broader sense of the term: it is still specialized, which allows it to yield fast and effective inferences, but no longer informationally encapsulated. While Fodor saw pragmatic processes as performed by the general-purpose central systems, Wilson describes pragmatics as a sub-module within the greater inferential mind-reading module (Wilson, 2005).

The non-preservationist positions, including the ostensive-inferential model, face a special challenge, one that does not confront the purely preservationist accounts. It is the challenge of precisely describing the relationship of similarity between the content in the speaker's mind and the content in the addressee's mind that should hold for the communicative act to be successful. For Buchanan (2010), the relationship is one between a proposition-type (the content of the speaker's communicative intention) and a proposition-token of the type (the content of the addressee's interpretation). For Cappelen (2008), to say that content is partially preserved between two token representations is to say that they represent overlapping sets of propositions. Finally, Sperber and Wilson describe a successful communicative act (focusing on the most complex case – one of non-literal meaning) in the following way. The speaker intends to communicate a complex thought consisting of many atomic propositions, some of which are more salient to the speaker than the others. The hearer does not entertain all the same atomic propositions, but only the most salient one, and then “constructs around it a complex thought which merely bears some similarity to [the speaker's complex thought]” (Sperber & Wilson, 1985, p. 156).

It is not clear, however, to what extent these all are simply restatements of the idea that the act of communication is successful if the resulting content is sufficiently similar to the original.

Arguably, the challenge of describing the right relationship between the content in the mind of the speaker and the content in the mind of the addressee for the communicative act to be successful is a Gordian knot still waiting to be undone. Alternatively, the problem is unsolvable, if Fodor was right in saying that the systems of the mind responsible for pragmatic interpreting are too disorganized to make good candidates for precise analysis (Fodor, 1983, p. 127). I will

have more to say about this special challenge and how my proposals may advance it in the concluding remarks.

### **2.3 Approaches to communication implied in art and music**

The literary and aesthetic scholars have also long theorized about the nature of the expressive and interpretive processes, and about what makes it possible to achieve an understanding between the artists and the audiences, even though not all such scholars would be on board to term what happens in the art sphere as communication. In this section, I highlight interesting connections and disconnects between the conceptions of regular communication and artistic expression, both in common, or “folk,” understanding and within the theories constructed by philosophers of language and literary and aesthetic scholars. While many have sought insights for art interpretation in the theoretical knowledge about linguistic communication, I believe the other direction—searching for insights among the theories of artistic expression to explain regular communication—will prove just as prolific. Particularly, my notion of interpreting as straddling the divide between creating and deriving content springs from applying the ostensive inferential model of communication to the act of a musical performer.

There are obvious commonalities between our conception of everyday human communication and that of creative expression in the sphere of arts. The most salient one is the division between creators of content and its consumers. While we divide the participants of a regular communicative act into speakers and addressees, in the creative fields, the pairs of counterparties are artists and viewers, performers and the audience, authors and readers. The creative sphere also makes use of the meaning and interpretation duality: meaning is what the creator is trying to

get across, while an interpretation is what the audience obtains. The perception of similarities between the two spheres sanctions occasional borrowing of insights across the borders. Often the defense of various positions on how objects of interpretation in literature and art *should* be interpreted turns on the extent to which a given expressive activity is like a conversation, or, at least, like an extended utterance. For example, in defending moderate intentionalism in literature and art, Noël Carroll says, “I rely on the claim that in [ordinary] conversations we typically aim at understanding the intentions of our interlocutors. I further argue that I see no principled reasons to suppose that things stand differently with our ‘conversations’ with artworks” (Carroll, 1997).

Despite general commonalities, there are also notable differences between our common, or folk, understanding of regular communication and our understanding of creative expression and interpretation. As addressees in a conversation, we usually assume that there is a correct way to interpret the speaker. Contrastingly, we do not typically assume that there is a correct interpretation of a novel or a musical composition, although we recognize that there can be wrong or bad interpretations. As interlocutors in a regular everyday conversation, we do not act on the assumption that an addressee co-creates what the speaker means. In arts, however, the distinction between meaning and interpretation is sometimes blurred, and the idea of the audience contributing to the production of meaning along with the artist is not preposterous but time-honored. Finally, in arts, interpreting is often considered an active and creative pursuit, while the addressee’s interpreting, in common understanding, is a passive process.

The relationship between the theories of communication and the theories of interpretation in arts, music and literature is even more complicated than the relationship between our “folk theories.”



Theories of aesthetic and literary interpretation typically *prescribe*, rather than describe the relationship between meaning and interpretation. However, their premises imply (and sometimes explicitly appeal to) a descriptive picture of communication. Such theories are countless and difficult to organize concisely. At best, we can sort them into three broad categories. First, there are those theories that espouse that the work itself is the source of meaning. Literary texts and works of art are to be analyzed only in terms of their internal structure (for texts) and compositional elements (for artworks). This formalist view is associated with the “close reading” practice, promoted by I.A. Richards (1917), and the New Criticism movement, dominant in the middle 20th century, including works by T.S. Elliot (1921) and essays “The Intentional Fallacy” (Wimsatt & Beardsley, 1946) and “The Affective Fallacy (Wimsatt & Beardsley, 1949). The second set of theories associate meaning with the authorial intent. Intentionalism advocates that texts and artworks should be interpreted by inferring the author’s intentions. This view is most radically expressed by E.D. Hirsch (1967) and more moderately by Carroll (1992), Wollheim (1993), and Livingston (1998). Finally, the third, constructivist, view focuses on the reader and the audience and sees them as active agents playing a role in creating meaning. It is represented by various reader- and audience-response theories by Fish (1970), Barthes (1977), Krausz (1993), Margolis (1989), and others. The related idea that a meaning of a work is in the eye of the beholder is often shared by non-scholars.

Despite this muddled and complicated picture, we can trace some tentative connections between the theories of communication constructed by the philosophers of language and the theories of creative expression and interpretation. For example, there are reverberations between the code model and the formalist position in aesthetics and literature. However, few if any contemporary philosophers of language believe that regular communication is exclusively code-like. And few

if any literary scholars, even those espousing formalism, would agree that the message of a literary or an artwork could be precisely encoded and duplicated throughout the mind of the author, the artwork, and the mind of the audience. The ostensive-inferential model, on the other hand, has tentative connections to both, intentionalism and constructivism in art and literary interpretation. As it relates to intentionalism, the ostensive-inferential model may describe literary and art interpretation as calling for uncovering the intentions of the author or the artist, whether precise (rephrasable as a proposition) or nebulous (putting the audience in a particular frame of mind or setting a particular direction to their thoughts). As it relates to constructivism, the model can describe the aesthetic and literary interpretation as impossible without the creative input of the audience. Once the art or literary work sets the direction to the audience's thoughts, the unique and creative path they take is the interpreter's doing.

In the next chapter, I argue that the ostensive-inferential model can both bridge our understanding of communication across the linguistic and the creative spheres but also gain insights from the creative sphere. In the next chapter, I argue that Sperber and Wilson's description of weak communication applies to the interactions between artists and audiences. Moreover, many nuances of expression and interpretation that have only been described in the context of art are nonetheless a big part of the regular day-to-day communication. Art and music should be seen in continuity with ostensive communication, with the flow of theoretical insights going in both directions.

## CHAPTER 3: CONNECTING MUSIC AND OSTENSIVE COMMUNICATION

Not many philosophers of language and philosophers of music believe that music should inform our general account of communication. However, as often happens in philosophy, the two groups of philosophers are somewhat insular. Both possibly have outmoded and idealized visions of each other's subjects. In this chapter, I argue that the comparisons between music and everyday communication that lead to the conclusion that the two spheres are distinct rely on the no longer popular preservationist conception of communication and on the idealized concept of *absolute music*.

The first section of this chapter details the commonly offered differences between music and everyday communication. The second section argues that such differences imply the view that reduces regular communication to only one kind of content (propositional, as opposed to non-propositional) and one kind of method (*meaning*, as opposed to *showing*). Such a view is no longer dominant among the philosophers of language, thanks to the influence of Sperber and Wilson's ostensive-inferential model of communication. Those influenced by Sperber and Wilson treat regular communication as a continuum between strong and weak kinds of communication. With *strong communication*, a speaker makes their intentions obvious and concrete, conveying near propositional content by the way of meaning or telling. With *weak communication*, a speaker merely influences the addressee's state of mind, conveying non-propositional content by the way of showing. I argue that how one communicates through music can also vary from strong to weak. Therefore, one should not compare regular communication at its strongest with music at its weakest.

The third section highlights another mistake of those who argue that music and communication are categorically distinct – contrasting absolute music, or music regarded separately from its title, context, and occasion, with utterances, which are always considered within a context.

Understanding the intentions of a musician, just like understanding a speaker, is always bootstrapped by many factors – the facts manifest at the time of the occurrence, the participants' shared assumptions and beliefs, the non-musical (or non-linguistic in the case of the utterance) cues. The proper musical equivalent to an utterance is a performance on a particular occasion, which always occurs in *some* context. Unlike in the code model, where a single kind of signal is sufficient to convey the content, in the ostensive-inferential model of communication, a communicative act depicted in terms of a single signal is an artifice – whether it is absolute music or purely linguistic communication. When compared on equal terms and through the lens of the ostensive-inferential account, music and regular communication appear to have similar resources for making meaning and understanding possible.

The course of uniting everyday communication with the arts under the umbrella of ostension has been pursued by others, although, perhaps, still only as pioneering forays. In this approach, I follow McCallum, Mitchell and Scott-Phillips (2020) and Pignocchi (2019), although these authors did not focus on music.

### 3.1 Perceived differences

It is helpful to divide the frequently mentioned differences between music and regular communication into three types: the differences in (a) *what* they convey, (b) *how* they convey, and (c) what it means to *understand* what is conveyed. The differences in *what* music and regular communication convey boil down to the distinction between propositional and non-propositional content: regular communication is often viewed as the sphere of propositions by those who hold them as separate, while music is viewed as the sphere of non-propositional content. The differences in *how* they convey boil down to the distinction between *meaning* and *showing*, as outlined by Grice. Those who view music and regular communication as separate typically see regular communication as the mode by which speakers *mean*, state, explain, tell, and so on, while they see music as a mode by which speakers show, demonstrate, and display, similar to how people communicate with smiles and salutes. The perceived differences on the side of content production are mirrored on the side of reception: while understanding a speaker is viewed to be a matter of making the correct inference, understanding music is often understood to be a matter of having the right experience or reaction.

The schism appears even deeper because what is frequently compared with regular communication is not just any music but absolute music or the kind considered without a title, occasion, or any associated non-musical clues. Juxtaposed against utterances, which are always considered within a context, absolute music appears to be a different animal indeed. In the following sections, I detail the three aforementioned categories of perceived differences, as well as absolute music, in turn.

### 3.1.1 What is conveyed: propositional versus non-propositional content

It seems obvious to many linguists and philosophers that music conveys a different type of content from the one we exchange in regular conversations. In their view, regular communication consists mainly of utterances conveying propositional content, or the type of content that can be expressed with a truth-evaluable “that” clause. On the other hand, music conveys non-propositional content, according to them, particularly emotional or affective kind. For example, Ray Jackendoff, a prominent linguist and a contributor to the study of musical cognition, says that verbal communication specializes in propositional thought while music in emotion or affect:

Language is essentially a mapping between sound and “propositional” or “conceptual” thought. The messages it conveys can be about people, objects, places, actions, or any manner of abstraction. Language can convey information about past, future, visible things, invisible things, and what is not the case. Linguistic utterances can be used to offer information, make requests for action, ask questions, give instructions and orders, negotiate, undertake obligations (including promises), assert authority, and construct arguments about the differences between language and music. [...]

What [different uses of music] seem to me to all have in common is the enhancement of affect associated with an activity. If one wishes to call this “musical meaning” [...], I have no objection, as long as it is clear that it bears little or no relation to linguistic meaning. [...] Lullabies convey a sense of soothing intimacy; love songs convey affection and passion; ballads convey the emotional impact of a story. [...] Work songs convey the rhythm of work, often coordinated rhythm, and they convey as well the affect of coordinated action. Marches convey the coordinated action of walking, often militaristically or ceremonially. Religious music conveys transcendence and spirituality, with associated affect anywhere from meditative to frenzied. (Jackendoff, 2009, pp. 197-198)

Jackendoff lists several exceptions to this view. First, he concedes that “the distinction between propositional and affective content is actually not quite this cut and dried” (Jackendoff, 2009, p. 198). He views propositional content as foreign to music but recognizes that an utterance can convey affect. Next, he admits the existence of “hybrids” of music and language – poetry, song,

rap, and chant. Finally, he acknowledges that both music and verbal communication use tone to convey affect yet this similarity alone is insufficient to sway him. None of the concessions make him reconsider his main point that “the specialization of language to conceptual information and music to affect is actually quite extreme” (Jackendoff, 2009, p. 198).

Boghossian (2007) agrees that music and regular communication tend to different kinds of content – regular communication tends to propositional content while music to non-propositional. Interestingly, his concession is in the opposite direction of the one Jackendoff makes, allowing a limited role for propositionality in music:

[Musical meaning is] not so much, perhaps, representational meaning – truth-evaluable propositions that tell us how things are, though some, myself included, would be prepared to allow a limited role for such meanings; but, rather, expressive meanings – states of mind, most prominently. (Boghossian, 2007, p. 120)

### 3.1.2 How the content is conveyed: *meaning* vs. *showing*

Another common reason to view music and regular communication as separate is the perceived differences in *how* they convey content. Grice distinguished two ways of conveying – *showing* and *meaning* (Grice, 1989, p. 218). When a communicator *shows* something, they intend that the addressee’s understanding is produced by recognizing direct evidence. Communicators *show* with the help of facial expressions, gestures, and pictures. Contrastingly, when a speaker *means* something by an act or an utterance, they intend that an addressee’s understanding is produced by inferring the speaker’s communicative intentions. Speakers *mean* with the help of utterances and written sentences. In Grice’s famous example, if Herod presents Salome with the head of St. John the Baptist, Herod *shows* that John the Baptist is dead. If Herod says to Salome, “JB is dead,” he *means* it.

Grice conceded that utterances could sometimes *show* and that wordless communication acts could *mean*. If someone utters in a hoarse voice, “Listen to how I sound,” they *show* that they have a sore throat. If someone presents a bandaged leg as a response to an invitation to physical activity (Grice, 1989, p. 109), they *mean* that they cannot participate. Nevertheless, Grice dismissed the idea that ostension plays an important part in regular communication. He dedicated a large part of his work to analyzing how speakers *mean* and how hearers infer, largely ignoring ostension.

Many scholars who view music and regular communication as intrinsically different often associate regular communication with *meaning* (or saying, stating, explaining, etc.) while music with *showing*. In his entry to the *Stanford Encyclopedia of Philosophy*, Kania (2017) summarizes: “The sentence “I am sad” is about the emotions, but it is not expressive of sadness in the way a sad face is, though I could use either to express my sadness. Most people agree that music’s relation to emotion is more like that of a sad face than that of a sentence.”

Once popular expression theory of music is a good example of describing music the way Grice depicted cases of *showing*.<sup>7</sup> For example, Dewey (1934) believed that any expressive act (including a musical one) is a direct consequence of the author’s own emotional experience. While Dewey’s concept of artistic expression is substantially more complex than Grice’s concept of *showing*, Dewey set expressive acts apart from regular communication based on aspects similar to those captured by Grice’s descriptions of the cases of *showing*: “The act that *expresses* welcome uses the smile, the outreached hand, the lighting up of the face as media” (Dewey,

---

<sup>7</sup> Dewey (1934) and Collingwood (1938) espoused views consistent with the expression theory of art.



1934, p. 66). Dewey even offered his own distinction between *expression* and *statement*, similar to Grice's distinction between *showing* and *meaning*: "The problem at hand may be approached by drawing a distinction between expression and statement. Science states meanings; art expresses them" (Dewey, 1934, p. 87). He even noted the difference on the side of the addressee, implying that expressions are experienced, while statements are understood in some other way (the idea that I will address in the next section):

[...] an expressive object [is] different from the statement of a gazetteer, however full and correct they might be, as Wordsworth's poem is different from the account of Tintern Abbey given by an antiquarian. The poem, or painting, does not operate in the dimension of correct descriptive statement but in that of experience itself. (Dewey, 1934, p. 89)

More recently, Levinson arrived at a distinction also highly suggestive of the one between *showing* and *meaning* in his argument for separating performative interpretation (PI) and critical interpretation (CI; essentially, performing music and speaking about it):

[...] a CI typically aims to *explain* [emphasis added] (or elucidate) a work's meaning or structure – 'what is going on in it', in a common phrase – whereas a PI can at most highlight (or effectively *display* [emphasis added]) that meaning or structure. [...] if good the PI allows such structure and meaning to emerge clearly, but does not provide a determinate comment upon them. (Levinson, 1993, pp. 38-39)

### 3.1.3 What it means to understand: inferring vs. experiencing

The differences in how speakers and performers convey content must be mirrored by the differences in how the addressees and audiences grasp it. Indeed, those who view music as separate from regular communication believe that understanding music categorically differs from understanding speech. In this view, understanding a speaker is making the correct inference (as evidenced by the addressee's ability to provide a suitable paraphrase), whereas understanding music is having the right experience or proper emotional or aesthetic reaction (as evidenced by the listener's ability to describe their reaction):

We take the presence of the right emotional reaction [to music] to be indicative of understanding. We recommend music appreciation classes to those who profess not to see what the fuss is about. (Boghossian, 2007, p. 118)

...I never need to infer that some music I hear is expressive of sadness, let alone an inference that proceeds from some complicated premise about how others are likely to react. Rather, I hear the expression in the music as I taste the sweetness in the wine. (Boghossian, 2007, p. 128)

If I asked you, Do you understand German? it is clear what kind of question I would be asking, and that the evidence directly bearing upon an affirmative answer would be your ability to provide paraphrases, in other languages, for German sentences. It is equally clear that scarcely anyone who has thought seriously about music is prepared to take this as a satisfactory model for musical understanding. (Kivy, 1990, p. 93)

I am now suggesting that we customarily take a person's musical understanding to be evidenced by, to be constituted, really, by his or her ability to describe the musical happenings perceived, thought about, enjoyed. (Kivy, 1990, p. 98)

#### 3.1.4 Absolute music

The rift between music and regular communication appears especially dramatic when philosophers insist on using absolute music for comparison. The term “absolute music” was coined in the second half of the 19<sup>th</sup> century to refer to Western-style instrumental music written to be performed without a text or a program, such as Bach's *The Well-Tempered Clavier* and Brahms' *Third Symphony*. However, in the philosophy of music, the term seems to refer to something beyond a unique genre. Fascinated with the problems posed by absolute music, Peter Kivy (1990) dedicated a book to the idea – *Music Alone: Philosophical Reflections on the Purely Musical Experience*. However, he struggled to pinpoint the exact notion: “How, then, shall we “define” music alone? If we are wise, we will not try to define it at all; for our attempts, doubtless, will exclude either too much or too little and raise other philosophical problems as well” (Kivy, 1990, p. 15). Roughly, however, absolute music, or, as Kivy sometimes refers to it, “contentless music,” is the kind that allows a listener to respond *only* to music and not any other

clues to the composer's intentions. A "purely musical experience" is self-sufficient and unconnected to any extra-musical information or prior beliefs. I deduce that in Kivy's view, an instrumental piece written for a specific purpose, such as dancing or a ceremony, could still be enjoyed as absolute music by a listener who does not have any knowledge of its purposes but would no longer be "music alone" for a listener with such knowledge.

A comparison with utterances, always considered in context, makes the two modes seem as different as apples and oranges. It also makes the question of how people communicate through absolute music seem more problematic than the question of how people communicate regularly:

I start with the observation that we often respond to a musical performance with emotion – even if it is just the performance of a piece of absolute music, unaccompanied by text, title or programme. [...] These emotions feel like the real thing to me – or anyway very close to the real thing. [...] Like many others, I am drawn to the philosophy of music by a need to understand how such emotional responses are possible. How can absolute music move us in the way that it does, and to the extent that it does? [...] How is it possible for mere sound, lacking speaker intention, or any of the other resources which make linguistic meaning possible, to express meanings? (Boghossian, 2007, p. 117)

[...] the presence of language, with all its potential for conveying concepts, and the presence of delineated characters, such as Violetta and Alfredo, Romeo and Juliet and Friar Laurence provide materials for arousal of garden variety emotions far exceeding anything that can reasonably be postulated in absolute music. And that is why absolute music poses a problem far beyond opera, oratorio, song and programme music to those who wish to claim that it arouses the garden variety emotions. (Kivy, 2001, pp. 101-102)

Boghossian and Kivy both believe that absolute music lacks the resources that regular communication has. Kivy then denies that absolute music conveys meaning at all. For him, understanding music is aesthetically appreciating its sounds in the way that the composer intended for the listener to appreciate them (Kivy, 2004, p. 217). In contrast, Boghossian argues that the meaning is evidently present in absolute music – since people consider it rational to be

emotionally moved by absolute music – but it requires a distinct explanation from the one provided in the case of speaker meaning.

### **3.2 Communication as a matter of degree**

In the past, all regular human communication used to be seen as *strong communication*: strongly propositional and of the kind Grice described as *meaning* rather than *showing*. However, modern pragmatics is increasingly recognizing that everyday communication is vastly richer.

Additionally, the neatness of the distinctions between propositional and non-propositional content and between *showing* and *meaning* has come into doubt within modern pragmatics.

Finally, the role of intonation and facial expressions is no longer seen as one that can be overlooked in analyzing speaker meaning.

In their book *Relevance: Communication and Cognition*, Sperber and Wilson (1986/95) emphasized the idea of communication “as a matter of degree.” Sometimes a communicator makes their communicative intention obvious. However, just as often, they seek to only nudge the audience’s understanding in a particular direction – the phenomenon that used to be idealized away in older discussions. Sperber and Wilson refer to such cases as *weak communication*:

In the case of strong communication, the communicator can have fairly precise expectations about some of the thoughts that the audience will actually entertain. With weaker forms of communication, the communicator can merely expect to steer the thoughts of the audience in a certain direction (Sperber & Wilson, 1986/95, p. 60).

If musical communication also involves a spectrum between strong and weak (and I argue that it does), we should take care not to juxtapose musical and interlocutory examples from the different ends of this spectrum.

Furthermore, if there is a spectrum of *what* and *how* a communicator conveys, there must be a corresponding spectrum in how the audience understands what is conveyed – the continuity between inferring the intended import and experiencing it. Sperber and Wilson did not explicitly argue for such continuity, but it seems a natural conclusion of their spectrum views on propositionality and the meaning–showing distinction. Then, in comparing how a hearer in a conversation and the musical audience understand, we should take care not to mix the non-analogous ends of the spectrum.

In Sections 3.2.1-3.2.2 I elaborate on the idea of regular ostensive communication as a matter of degree to build a case for putting it in the same category as musical communication.

### 3.2.1 Degrees of propositionality

In theory, propositional content is the kind that matches a specific “that” clause. In reality, examples of utterances produced to convey a single concrete “that” clause are difficult to find. Even uttering as simple a request as “Pass the salt” on a particular occasion could be done to convey any number of the following nonequivalent “that” clauses:

- (1) that the speaker wants a particular addressee to pass the salt to the speaker
- (2) that the speaker wants any addressee in the immediate vicinity to pass the salt to the speaker
- (3) that the dish lacks salt
- (4) that the dish lacks flavor
- (5) that the cook is an amateur
- (6) that the speaker is dissatisfied with the dish
- (7) that the speaker is superior to the addressee (allowing the speaker not to say “please”)
- (8) that the speaker is on informal terms with the addressee.

As Buchanan (2010) describes, the speaker is often indifferent to which particular combination of “that” clauses the addressee will infer. In the above example, the speaker might be indifferent to whether the addressee will infer (1) or (2), (3), (4), (5) or (6), and (7) or (8).

Sperber and Wilson (2015) have argued that there is no clear border between propositional and non-propositional content. The way people ordinarily communicate includes the continuity of cases between the two extremes. On the high propositionality end of the spectrum, the authors place cases in which what a speaker means can be expressed as a fairly specific “that” clause. For example, a train station cashier impassively informs a passenger, “The next train to Oxford leaves at 12:00.” To the right of the high propositionality extreme are instances of less specific content. For example, if someone urgently shouts the same phrase at a train station, in addition to a specific “that” clause, some content is less specific – that the addressee must hurry or that the departure is imminent.

Farther still on the propositionality spectrum are vague implicatures that convey potentially endless lists of nonspecific “that” clauses. Suppose Mary, who hails from Oxford, lives in London with her husband, Paul. Lately, their marriage has hit a rough patch, and Mary expressed her dissatisfaction to Paul. “The next train to Oxford leaves at 12:00,” replies Paul. It is utterly non-specific what Paul meant, and the list of possibilities is open-ended:

- (9) that Paul wants Mary to leave
- (10) that Paul thinks Mary is free to leave
- (11) that Paul no longer loves Mary
- (12) that Paul thinks Mary no longer loves him
- (13) that Paul is upset about Mary’s complaint
- (14) that Paul refuses to address Mary’s complaint

Finally, at the opposite extreme are communicative acts conveying something so unspecific that even a list of possible “that” clauses seems inappropriate for describing such content:

Mary and Peter are newly arrived at the seaside. She opens the window overlooking the sea and sniffs appreciatively and ostensively. When Peter follows suit, there is no one particular good thing that comes to his attention: the air smells fresh, fresher than it did in town, it reminds him of their previous holidays, he can smell the sea, seaweed, ozone, fish; all sorts of pleasant things come to mind, and while, because her sniff was so appreciative, he is reasonably safe in assuming that she must have intended him to notice at least some of them, he is unlikely to be able to pin her intentions down any further. Is there any reason to assume that her intentions were more specific? Is there a plausible answer, in the form of an explicit linguistic paraphrase, to the question, what does she mean? Could she have achieved the same communicative effect by speaking? Clearly not. (Sperber & Wilson, 1986/95, pp. 55-56)

Poetic effects, hyperboles, metaphors, exclamations, and evocative non-verbal acts convey this type of content. Sperber and Wilson term this type of content *impressions*:

What is an impression? Is it a type of mental representation? Can it be reduced to propositions and propositional attitudes? What we are suggesting is that an impression might be better described as a noticeable change in one’s cognitive environment, a change resulting from relatively small alterations in the manifestness of many assumptions, rather than from the fact that a single assumption or a few new assumptions have all of a sudden become very manifest. (Sperber & Wilson, 1986/95, p. 59)

No sharp distinctions can be drawn between where propositions end and impressions begin. The contents of everyday communication can best be described as a continuum between the two extremes.

### 3.2.2 Degrees of *meaning* and *showing*

Sperber and Wilson (2015) also argue that regular communication involves both *meaning* and *showing* without clear borders between the two categories (in that they diverged from Grice).

They took Grice’s famous examples of *meaning* and *showing* meant to demonstrate a sharp distinction between the two and filled in several intermediate cases. For example:

- Herod lets Salome see the head of John the Baptist as if by accident.
- Herod shows the head of John the Baptist in response to Salome’s inquiry.
- Herod shows Salome a blood-stained sword in response to Salome’s inquiry.
- Herod makes a throat-slashing gesture in response to Salome’s inquiry.
- Herod explicitly tells Salome, “John the Baptist is dead.”

As one progresses through these cases, the evidence for John the Baptist’s death becomes less and less direct, but it is impossible to say at what point *showing* that John the Baptist is dead becomes *meaning* that John the Baptist is dead.

Ostension is pervasive in regular communication and should not be idealized away within our accounts of communication. For example, we routinely point or raise our hands in response to “Who is so-and-so?” Facial expressions, intonation, and body language are a regular part of our communicative repertoire. They do not always involve pure *showing* but fit on the continuum from direct to less direct evidence. For example, I may manifest my terrified face to an addressee to emphasize how scared I am (in the case of direct evidence), or I can *make* my facial expression exaggeratedly frightened (in the case of less direct evidence). The intonation with which I say “Ouch!” when I witness my friend stubbing her toe might be direct evidence of the discomfort I feel when I witness her pain, or it may be play-acting designed to communicate my empathy without a clear border between the two cases.

### 3.2.3 Understanding as a matter of degree

If we adopt a continuity view of *what* and *how* speakers convey, we must also consider a continuity view of *what* and *how* addressees understand, even though Sperber and Wilson did not



explicitly argue for such continuity. I now outline how communication as a matter of degree corresponds to understanding as a matter of degree.

On one end of the communication spectrum is *speaker meaning* – the narrowly defined situation in which the speaker intends to achieve a specific response in the addressee in virtue of the addressee recognizing the speaker’s intention. This end of the communication spectrum corresponds to the end of the understanding spectrum in which the addressee *infers* the intended import. On the opposite end of the communication spectrum, the speaker does not have a specific intention. The import is not a proposition but a vague impression. The speaker does not *mean* anything in Grice’s strict sense but instead *shows* the addressee the intended import. Impressions are not inferred but rather experienced. The result of being *shown* something (rather than told) is not an inference but a perception. Therefore, on the opposite end of the understanding spectrum, the addressee understands the speaker as a result of having the right experience, emotion, or affect.

For example, when Mary demonstratively opens the window and inhales the seaside air, she expects Peter not only to infer her feelings but also partake in her joy. When someone urgently shouts, “The next train to Oxford leaves at 12:00!” they are trying to get the addressee to not only infer a specific proposition but also *experience* the urgency. If, as a result, the addressee continues to shuffle barely, the shouter may consider their utterance failed even if the addressee confirms that they inferred the correct “that” clause. Similarly, Paul does not merely aim for Mary to arrive at one of the propositions (9)–(14). His utterance is meant to hurt her more than if he clearly stated, “Mary, you should leave.” If Mary’s understanding of Paul is not as visceral as a reaction to a sting or bitter taste, then, on some level, Paul’s communication has failed.

Speakers do not always aim to maximize clarity but to incite a reaction – to insult, hurt, inspire, and soothe. “Enhancement of affect” is not reserved only for musical uses, contrary to Jackendoff (2009, p. 198).

### **3.3 Erasing distinctions between music and regular communication**

When we do not deny the full range of regular communication, it becomes apparent that it involves just the kind of resources that, many believe, are the specialty of music. Mary could also achieve what she achieved by evocatively opening the window and sniffing the seaside air by musical means – by flipping open the piano and reproducing Beethoven’s “Ode to Joy” while gazing through the window. Paul may as well execute Mozart’s Requiem to communicate the end of his relationship with Mary. Many examples of musical performances fit well with the account of *weak communication*. Steering the listeners’ thoughts in a certain direction is often the intention behind elegies, nocturnes, lullabies, marches, anthems, and caprices.

However, assuming that regular and musical communication are comparable only at the weaker end of the communication spectrum would be a mistake. To continue comparing them across the spectrum, I need to address another mistake – the insistence on comparing speaker meaning with musical meaning apart from the context and non-musical cues accompanying musical meaning. This will show that Boghossian and Kivy are wrong in asserting that music lacks the resources of the regular communication and that it poses a special problem for explaining how meaning and understanding arise.

### 3.3.1 Absolute music versus utterances

Pragmatics makes a salient distinction whose analogy is currently lacking in the philosophy of music: one between the timeless meaning of words, phrases, and sentences and speaker meaning. The timeless meaning of a word or a sentence does not vary from context to context; it can be found in a dictionary and is not dependent on the speaker. On the other hand, speaker meaning on a particular occasion is determined by the speaker's intentions and is inextricable from its context and non-verbal cues. For example, if I mockingly say, "You poor thing!" to a friend wincing from a paper cut, the timeless meaning of the sentence I use is entirely different from the message of sarcasm I intend to convey.

In a conversation, the timeless linguistic meaning of the words and phrases is only one type of clue for understanding speaker meaning. Moreover, we can never understand a speaker by considering the timeless linguistic meaning alone. Understanding a speaker always requires taking into account factors outside of the purely linguistic medium – the situational cues, the prior history between the speaker and the addressee, the background assumptions, and a variety of encyclopedic knowledge. The sarcasm behind "You poor thing!" comes through the intonation and the background assumption that a paper cut is not a severe injury.

Absolute music is not a fitting analogy for an utterance on a particular occasion. A better analogy is a particular musical performance – always embedded in a context and rich with non-musical attributes, even when not noticeable. Knowing the title of the musical piece, the occasion of its performance (a wedding, a funeral, a military parade, or a party), and the composer's name, era, and biographical facts undeniably contribute to the listener's experience of the performance.

Despite Kivy citing *The Well-Tempered Clavier* as a paradigm of absolute music (Kivy, 1990, p.

14), just knowing that it is composed by Bach already ensures that a listener's understanding of the piece relies on factors beyond the sound alone. The performers' intentions come through not only the musical sounds but also their posture, movements, and facial expressions. Their clothing and stage decorations align with the intended effect on the audience.<sup>8</sup> For example, the recent review of a choreographed performance set to the electronic transcription of Mozart's Requiem at Lincoln Center describes how the stage set tips off the audience that they should expect a different kind of Mozart experience:

Abraham's work doesn't look like a traditional Requiem either. Dan Scully's lighting and scenic design are futuristic chic, an ever-changing light show of neon bars and circles of illumination — now red, now turquoise, now Easter purple. An oculus high on the rear wall bubbles with lava-lamp ooze. (Seibert, 2022)

The more context and non-musical cues to which the audience is privy, the stronger the case of musical communication. Perhaps, the reader can recall their first experience hearing "Flight of the Bumblebee" by Rimsky-Korsakov without knowing the title of the piece. The melody may have evoked a vague expressive effect—one of commotion or anxiety. Similarly, the sounds of "Morning Mood" by Grieg loosely suggest quiet joy to a listener who does not know its title. However, as soon as the listener learns the titles, the vagueness cedes to representational meaning—the hectic flight of an insect in the first case and the rising sun in the second. From then on, every time the listener hears these pieces, the knowledge of their titles always ensures a more concrete or representational understanding of these musical pieces.

---

<sup>8</sup> Sometimes composers include extra-musical cues in the notes to the score. For example, Terry Riley specified in the original performing instructions for his *In C* composition: "The 1/8 note pulse is traditionally played by a beautiful girl..." (Riley, 1964, as cited in Wilson, 2021).

Still, Boghossian and Kivy could take the listener's vague reactions to the pieces before knowing their titles as evidence of the listener's ability to respond to seemingly contextless music. They could cite other instances when the music seems truly "alone," such as an unknown instrumental piece played on the radio. Listeners can still respond to, enjoy, and think about such music and describe their reactions in common or technical terms. Nonetheless, the effect is not unique to music. We often encounter utterances without known context, for example, overhearing random strangers on the street or turning on a television in the middle of a show. In such cases, we still do not perceive the utterances as without context: our minds automatically build contexts, similar to how they fill in the missing visual information in a blind spot.

Sperber and Wilson discuss how readers of philosophy works consider artificial examples, such as many presented in this dissertation. An author typically provides a context or asks the audience to imagine one. Additionally, the words themselves contribute to creating a context by opening access to the "encyclopedic entries" about the concepts to which they refer (Sperber & Wilson, 1986/95, p. 185). For example, when I invited the reader to consider the utterance "Pass the salt" earlier in this dissertation, the reader must have inevitably pictured a scene around a dining table. They may have even imagined the speaker's intonation and the shape of the saltshaker. The key to creating this imaginary scene is the reader's knowledge of salt as a flavor enhancer. If an additional cue indicated that a different part of the reader's knowledge about salt was at play (for example, that salt can be used as a cleaner or a nonlethal shotgun round), the reader would come up with an entirely different imaginary scene. Correspondingly, the range of the propositions that would come to the reader's mind would also be quite different from (1)–(8) if they imagined the speaker uttering "Pass the salt!" while holding a shotgun.

Usually, we fill in the missing contexts on the fly without noticing. However, a remark can occasionally be too obscure to assume context easily. In such cases, an utterance can have a mysteriously vague expressive effect, similar to the effect of evocative unfamiliar music.

Imagine coming across a communication not prompted by a conversation or a social situation, not a part of a public sign, an art project, a book, or a publication. Suppose it is a scribble on a rock at a desolate location: “I did not find you here.” You may wonder if this is someone’s expression of loneliness, longing for a loved one, or the futility of a spiritual search. With only a glimmer of the author’s intentions, one may plunge into a wistful frame of mind pondering possibilities. However, just as learning the title of “Flight of the Bumblebee” turns its vague jittery effect into the representation of a flying insect, learning the context of the scribble will turn its vague evocative effect into propositional meaning. If you learn that two friends agreed to meet at this location, but only one turned up and eventually left, the wistful effect of the communication will disappear.<sup>9</sup>

Music is similarly never completely alone. If unaccompanied by a program, it may force the mind to create an artificial context – a desolate landscape, a battle scene, or a smoky cellar of a jazz club – the same way as hearing an unembedded utterance of “Pass the salt” forces the mind to imagine a dining scene. While musical sounds, unlike words, do not refer to objects or events,

---

<sup>9</sup> There is a real-life illustration of how utterances lacking easily constructible context can have a vague evocative effect. An automated twitter account *Horse\_ebooks* attracted thousands of followers, inspired performance art, and became a subject of exposés by the likes of *The New Yorker*. The bot simply broadcasted textual fragments taken from random sources. Yet, their effect has been called “strangely poetic,” “seemingly knowing and even portentous,” and compared to Zen koans (Orlean, 2013):

Everything happens so much.

I will make sure you never buy knives again.

Unfortunately, as you probably already know, people.

they still invoke previously stored beliefs. For example, the sound resembling a coyote howl in the beginning Ennio Morricone's "The Ecstasy of Gold" sets the listener's mind to creating a desolate scene in the American Southwest. The sounds of banjo may open access to even more complex encyclopedic entries, involving country and old-time culture, Appalachian region, and African American history. The greater someone's expertise in music, its various genres, and history, the more contextual information they will receive from the musical sounds. However, even a neophyte can tell whether a piece sounds modern or classical, was composed for children or a solemn occasion, is country or rock and roll, Latin or Native American, or played on a harp or a shofar. From these bits of information, a listener constructs context that may involve a mood, historical period, geography, setting, ritual, and even characters in costumes participating in a particular action. A more sophisticated listener would be privy to even richer information, able to discern the influences of the culture, traditions, social attitudes, and philosophies of the era and the composer.

I am not suggesting that listeners construct artificial contexts that *always* assign representational meaning to nonprogram music but that, in interpreting music, our minds inevitably invoke something beyond sound alone. The associations can be representational but also abstract, emotional, kinesthetic, tactile, and so on. Even minimalist and atonal music, such as Terry Riley's *In C* and music by Arnold Schoenberg, is not entirely "contentless." The critics have described the effect of *In C* in representational, perceptual, and emotional terms, as well as blending many musical and cultural traditions:

The effect is of a sparkling, glinting crystal which, as it slowly rotates, changes almost imperceptibly in color from a clear C major to a bright, yet more slowly pulsating, E minor then back to C rather triumphantly, and finally takes on the cast of a much more somber and enigmatic G minor. (Leedy, 1992)

Certainly “In C” touches on pre-baroque, jazz, and Cageian traditions, yet some of its motives also rub up against Riley’s own “String Trio of 1960. [...] ...the net result is an energized Balinese music unsprung from its hinges. A communal joyous cacophony of secular yet spiritual ecstasy. (Alburger, 1998, as cited in Force, 2004)

However, with sufficient context and non-musical cues, music *can* communicate complete propositional meaning. For example, if, in response to Salome’s question about the whereabouts of John the Baptist, Herod performed a dirge, his performance would make an example of propositional speaker meaning as good as any, specifically, *that* John the Baptist is dead.

Another example is “Chrissie’s Death” by John Williams from the *Jaws* soundtrack. If a listener hears the piece without knowing its origin or title and without seeing the accompanying movie scene, they may recognize only a vague sense of disquiet. However, when the sound is combined with the visuals of a woman enjoying a midnight swim, the listener cannot escape the realization *that* something is lurking underneath her and is about to strike.

There are many other examples of musical propositional meaning, as long as we do not excise the associated contexts and the non-musical cues. Songs can convey everything that Jackendoff (2009) expects only verbal communication can: messages about people and objects, places and actions, past and future, fiction and nonfiction, and concrete and abstract things. Melodic *adhan* gives instructions, calling Muslims to prayer. Christian cantons teach moral lessons. While I am unfamiliar with a song that contains a philosophical account of communication, many modern songs tackle social, political, and cultural issues. Sure, they do so with the help of words, but words are also not self-sufficient in fulfilling the speaker’s communicative intentions. Songs, chants, rap, and poetry are “hybrids” of different communicative modes (as Jackendoff insists) in the same way as *all* communicative acts are. One may say that a sarcastic remark, “You poor thing,” is a hybrid of words, intonation, and the assumptions about the severity of paper cuts.



Paul's exquisitely cruel reply to Mary is a hybrid of words, intonation, and his beliefs about their relationship.

No communicative act contains everything necessary for its understanding without recourse to context, previously stored beliefs and assumptions, and accompanying cues expressed through auxiliary mediums (for example, intonation and the setting for an utterance and the title and stage decorations for a musical performance). Absolute music is as impossible as absolute words or texts. Even if extraterrestrial beings received a musical or linguistic message from humans, they would not take such content as absolute. The recipients would either ignore the messages as noise or place them in their own artificial contexts, using alien history and experience.

### 3.3.2 Timeless meaning: musical and linguistic

It is worth returning to the special problem that Boghossian believes absolute music poses – the problem of being able to express meaning without the resources of regular communication. If absolute music is taken as something analogous to an utterance, or a vehicle for speaker meaning on a particular occasion, Boghossian's question is about what makes musical communication possible. It is not answerable, because, as I argued, absolute music and utterances are not analogous, and performances on particular occasions have the *same* resources as utterances for making communication possible.

However, it is possible to take absolute music as something analogous to utterance *types*, or a vessel for *timeless* musical meaning. Then the second interpretation of Boghossian's question is, "What makes timeless musical meaning possible?" and it is a valid question. However, this question no longer demonstrates the exceptionality of music and is analogous to the following

question about ordinary communication: in virtue of what do words, intonational patterns, and facial expressions have timeless meaning? In comparing the sources of timeless musical and timeless ordinary meaning, one still finds more similarities than disconnects between music and everyday communication.

The timeless meaning of a sentence is a type of evidence with which the speaker makes their intentions overt to the addressee. Grice (1989) suggested that words, phrases, and sentences acquire their timeless linguistic meanings through the history of using them to express particular speaker meanings – an explanation by convention. Similarly, certain musical sound structures are prone to invoking musical intuitions in a listener, giving the listener clues to the overall impact that the performer is seeking to produce.<sup>10</sup> Such structures are less obvious and defined than words, linguistic phrases, and sentences. Musical theorist Fred Lerdahl and cognitive scientist and linguist Ray Jackendoff teamed up to describe how the brain segments “the surface of the music” into such structures in their book *A Generative Theory of Tonal Music* (Lerdahl & Jackendoff, 1996) according to innate rules. It is possible that with regular use of such structures and their combinations within certain contexts, they may acquire timeless musical meaning just as words and sentences acquire timeless linguistic meaning. Then, in each case of a concrete performance, such timeless musical meanings serve as the evidence for constructing the listener’s overall interpretation of the performance, just as the linguistic meanings of the words serve as the evidence for constructing the addressee’s pragmatic interpretation of what the

---

<sup>10</sup> There is another sense of timeless musical meaning. Two kinds of musical understanding are involved in a performance: one by the audience (their experience of the music) and one by the performers (their understanding of the score and how it translates into a performance). The timeless meaning that contributes to the understanding by the audience is conveyed by the structural elements of the melody, whereas the timeless meaning that contributes to the understanding by the performers is conveyed by the score.

speaker means. With enough non-musical clues, such interpretation can be even close to propositional.

For example, a musical piece consisting of strong steady rhythm played on brass instruments is associated with marching for a typical Western listener. Most likely this association has been established by convention – using this type of musical pattern in military parades. When a pattern of this type appears in the performance of Prokofiev’s ballet *Romeo and Juliet* (in the beginning of the part known as “Dance of the Knights”), together with many non-musical cues given by the performance of the ballet, it contributes to the listener’s constructing the representation of two powerful family clans engulfed in a bitter conflict. Some musical pieces become so connected with certain contexts that they become carriers of the corresponding timeless meaning as clearly as words. For example, one cannot hear the opening fanfare of *Also sprach Zarathustra* by Strauss without getting an impression that a significant announcement is coming or “Carol of the Bells” by Leontovych without thinking of Christmas.

It is also possible that musical patterns acquire timeless musical meaning by imitation, for example due to their resemblance to the emotional tone of a person, the rhythm of human action, or natural sounds. This answer is consistent with a popular theory of musical expression, the resemblance theory (Davies, 1994). According to this theory, a listener may recognize music as sad because it resembles the pattern of the voice of a person overcome with sadness. Music can also imitate naturally occurring sounds. Vivaldi’s *The Four Seasons* opens with violin trills resembling bird songs. Honegger’s *Pacific 231* sounds not just like any train gaining speed but like a particular class of steam locomotives. Neither the explanation by convention nor one by imitation need be exclusionary – both doubtlessly play a role in explaining how musical meaning

is constructed. Furthermore, neither explanation for the timeless musical meaning makes music discontinuous from everyday communication. After all, imitation also plays a role in constructing meaning in regular communication. Consider onomatopoeic and phonesthetic words. The sound of the former suggests the sound of the thing that the word names (for example, “Bang!” or “Shush!”). The sound of the latter suggests other attributes of the object they name. For example, the sounds -gl and -kl tend to evoke associations with flickering light (“twinkle,” “glisten,” “glitter”). The combination with -mor appears in many words associated with evil and menace, such as “Mordor,” “mortal,” “murder” (Harbeck, 2014). It would not be surprising if speakers consciously or subconsciously selected words that sounded in line with the reaction they wanted to produce in addressees – for example, avoid harsh-sounding words when trying to soothe. However, it is important to note that the concept of timeless meaning – whether linguistic or musical – is also an abstraction. The demarcation between the *timeless* meaning and the *occasion* meaning of a word or a melodic element is subject to a sliding scale.

In summary, we should not conclude that music is separate from regular communication by comparing musical communication at its weakest with regular communication at its strongest. Musical communication can often be strong, conveying nearly propositional content, while regular communication can often be weak, conveying vague impressions and affect. Both employ ostension and inference. It is also wrong to compare music separately from its context and non-musical cues to utterances, which are always considered within given or imagined contexts. When we do not make such comparison mistakes, music and ostensive communication appear to have similar resources for making *occasion* meaning possible. They also have similar sources of *timeless* meaning.

If music and ostensive communication are continuous, then the ostensive-inferential model of communication should be able to give an account of a musical performance. In the next chapter, I show that it is, but with certain modifications. I propose that in describing the act of a musical performer, we need to be able to describe an action that constitutes both deriving and creating content. I also propose positing interpretive intentions of the musical performer to explain why there can be many faithful performative interpretations of the same score. In Chapter 5 of this dissertation, I argue that such modifications are also useful in theorizing about regular communication.

## CHAPTER 4: MUSIC

### 4.1 The failure of traditional models to describe music

Musicologists and philosophers of music have occasionally used models of communication to describe music. When doing so, they usually designate composers as creators of content and audiences as addressees. However, there has not been a uniform treatment of the score, the performance, and the role of the musical performer. Some, like Davies, treat the score as an utterance and the performers as addressees:

[Scores] are to be read as instructions addressed to the work's potential performers, and it is by following these instructions that players generate instances of the work. The proper interpretation of a score requires knowledge both of conventions for the notation and of the performance practice shared by the composer with the musicians to whom the score is directed. (Davies, 2001)

Others, like Cross, treat the performance as an utterance and the performers as the creators of content:

Like language, music can also be conceived of as a communicative medium; indeed, both seem to fit equally well within a widely used theoretical model of communication, the 'information theory' model of Shannon and Weaver (1949). [...] In a musical context, one can think of the sender as the performer, the receiver as the listener, the channel as the air and the information transmitted is the sonic patterns that constitute the music. (Cross, 2005)

In common understanding and the writings of music critics, musical performer is neither purely a creator of content, nor purely its consumer. The role of the performer requires both – fidelity to the composer's intentions and the creativity to realize those performative effects that are not specified by the score. Even the most detailed score does not contain all the information

necessary to perform the piece, leaving many decisions to the performer. For example, tempo is never specified precisely: it is sometimes marked as a *range* of beats per minute, sometimes described in Italian words (such as *largo*, *adagio*, *andante*), and sometimes not indicated at all. Depending on the work, the performer might be tasked to determine even “the larger structure of the performance by deciding on the number of repeats, variations, elaborations, links, and the like. Moreover, her choices settle many aspects of the performances micro-detail, including intonation and melodic inflections, timbre, and texture, attack and decay, emphasis and phrasing” (Davies, 2001, p. 20). This is why performers can realize the same musical score in countless ways that can all be considered faithful renditions of the same work and triumphs in their own right.

Neither the ostensive-inferential nor the code communication models can appropriately describe the relationship that unites the composer, score, performer, performance, and audience. The code model may be helpful in reflecting some elements of this complex relationship, just as it describes one way a speaker can give evidence of their communicative intentions to the addressee. However, the code model cannot capture the entire musical chain that begins with the composer and ends with the audience, because a musical performance is not the case of specific content being duplicated throughout a chain of transmission in a coding-decoding manner.

The ostensive-inferential model has a higher chance of describing a musical performance adequately, as it holds important truths transferrable from regular communication into the realm of music. Like an utterance, a musical performance can be described as an interaction between two parties (or two sets of parties) in which one party (the speaker or performer) intends to achieve a particular effect on the other party (the addressee or audience) and uses the utterance

(the performance) as ostensive evidence or as a cue for what they intend to achieve. The intended effect need not be a propositional representation but a general cognitive effect. Both parties are constrained by the mutual beliefs about the practice and its conventions, and by the mutual expectations of each other's thought processes. However, unlike utterances, musical performances are themselves interpretations of something else (namely, the score), the product of derivation rather than creation. The ostensive-inferential model currently lacks the resources to describe an intermediary segment within a communicative chain that is simultaneously an interpretation and an expression of creative intentions. It also lacks the role of an interpreter whose task is not purely to derive content but derive-cum-create it. Finally, the model currently does not recognize that an interpreter may have intentions directed at an audience of their own – something we intuitively attribute to musical performers.

It is possible that the ostensive-inferential model can accommodate the act of the performer without modifications by describing it as a two-step procedure: as an addressee, the performer interprets what is specified by the score and then, as a creator in their own right, creatively supplements what is underspecified. However, the next section of this chapter explains why there is no non-arbitrary way to tease these two steps apart. I suggest that the performers' creative resolution of what is underspecified by the score *is* the process of inferring the melody. A new notion of interpreting is required to describe a musical performer appropriately, one that does not involve a division between the properly derivational and creative parts of the interpreting process.

The third and fourth sections of this chapter argue that the proposed notion of interpreting also describes listeners and composers. A listener is an interpreter in the sense I am advocating



because the listener's internal interpretation of the music is constructed with their creative input. A composer is an interpreter in the sense advocated here because the score they produce is simultaneously an expression of their creative intentions and an interpretation. Sometimes, the score is an interpretation of lyrics or movie scenes. However, it is always an interpretation of the composer's inner music.

The last section of this chapter develops the idea of interpretive chains, where an expressive act or object is an interpretation of other expressive acts or objects. The notion has not been addressed by traditional models of communication but is familiar from the art world (artworks are frequently interpretations of other artworks) and oral lore (ideas, traditions, and legends are passed down and reinterpreted by each generation).

#### **4.2 Musical performers: prying creativity from inferring**

Orchestras require conductors, even though the players are proficient at reading notation. The notation leaves many parameters of the performance unspecified, and if each player made their own choices for the underdetermined elements, the orchestra would sound confused and disjointed. The conductor's role is to develop a unified plan for specifying what is unspecified in the score. They communicate the unified plan to the performers with a variety of signs, including baton and hand motions, facial expressions, posture, and even inhalations:

...baton can [...] shape the sound. The nature of the downbeat — how abrupt, how delicate — tells the orchestra what kind of sound character to produce. The baton can smooth out choppy phrases by moving through the beat in a more sweeping way. A more horizontal motion can create a more lyrical quality... [...] A squint [...] can convey a distant quality to the music... [A conductor] pushes forward to achieve more intensity

from the orchestra. Sometimes she leans back to have the musicians play softer. (Walkin, 2012)

Hence, the conductor acts simultaneously as the composer's addressee and as a speaker with intentions directed at the orchestra.

The score does not *determine* the performance but rather *constrains* or outlines a field of possibilities of how a performance may proceed. To put it figuratively, the score is not like the rails that keep a train on a defined path but more akin to a river down which the captain must chart a course. Here is what this means in the words of a conductor:

Every gesture we make in performance eliminates a myriad of other possibilities. No performance of a piece of music can ever present all that it has inside it. Think of it this way: We are standing together in a darkened Academia Gallery in Florence before the statue of David. The conductor has the opportunity to light it with one light. It will be recognized as David, but it will only illuminate some of it. [...] Only through multiple performances can an audience begin to perceive the infinite possibilities of a piece of classical music. (Mauceri, 2017)

The conductor makes interpretive choices according to their creative intentions toward the audience. Ideally, the composer's and the conductor's intentions should not contradict each other. Then the audience conforms their interpreting to both sets of intentions, and, if perceptive, will differentiate between the two. For example, even though Herbert von Karajan and Bruno Weil are conducting the same composition, a critic attributes the unique quality of each performance to the conductors' styles:

Herbert von Karajan's 1975 Mozart Requiem is of the grand ceremonial type with large choral forces and full orchestra accompaniment—all of which sounds quite massive compared to more recent period-style performances—particularly Bruno Weil's powerfully lean and incisive version. (Carr Jr, 2015)

The performers also express their creative intentions toward the listeners within the constraints imposed by the score and the conductor's directions. Like the score, conducting conveys content only as a range of possibilities. Many nuances of the playing remain underdetermined even after

the conductor has weeded out some choices: for example, the strength of the keystrokes, the angle of the airstream blown into the mouthpiece, or the precise section of a string that brings out the note remain up to the performer. Some of what is underdetermined is resolved randomly without the performer's conscious control, some is the result of the performer's choice and conscious effort, and some is the result of the performer's choice and effort when the piece is learned but becomes automated through regular practice. Orchestras often respond differently to the same conductors: a study involving hundreds of listeners, professionals, and non-musicians has confirmed that listeners can identify orchestras by their distinctive styles, even while controlling for the conducting effect (Bertsch, 2002).

Without the conductor's directions, the scope of the performer's choices is even broader. Subtle shades in the attack, phrasing, balance, and pitch (for example, choosing to play older classical pieces with slightly lower pitches in line with historical practices) will create unmistakably different performances of the same work while staying within the confines of the original score:

While melodies and harmonies may be specified in advance, the precise realization of dynamics, rhythmic subtleties, timbre, intonation, and articulation arises at the moment of performance, and will vary (often considerably) from performance to performance, even when the piece is played by the same musician. One evening, a cellist may linger a bit longer at cadences, may relax her or his tempi somewhat, or choose a delicate, graceful tone for a peak phrase. On the next, she or he may choose to play with more rhythmic drive, propulsion, and rigidity, and strive for an intense, penetrating tone to wrest the greatest degree of excitement from that same peak phrase. (Gould & Keaton, 2000)

In calling for creative input by the performer, the performance of an existing work is unmistakably similar to the act of creating an original composition. After all, a composer, as I argue in section 4.4, is also constrained in crafting the melody. The similarities between composing (or improvising) and performing an existing work are at times so striking that a

question is raised about whether there is a principled difference between the two categories.

Gould and Keaton see no difference:<sup>11</sup>

What this all shows is (i) classical performance [of an existing work] can be improvised in a way quite similar to a jazz performance. That is, it might stray from the score according to the performer's unique ideas [...] (ii) it by necessity goes beyond the score in given ways even if it reflects the score more closely than a jazz performance reflects a formula; (iii) both involve an immediacy of expression that we find even in carefully crafted linguistic utterances. Both sorts of performances, that is, involve intentional acts that are performed according to rules, but have an element of fluency and immediacy about them. (Gould & Keaton, 2000, p. 146)

One may object to putting the composer's crafting of a melody and the performer's creative interpretation of an existing work into the same category by saying that the choices available to the performers are still limited compared to the options open to the composer. According to this objection, changing the tempo or deciding on the intensity of the keystrokes are not on equal footing with crafting the structure of a melody. However, some performances of an existing work involve more creative decisions from the performer than some improvisations (those that contain few novel musical phrases). Davies has noted a distinction between "thin" and "thick" works:

If it is thin, the work's determinative properties are comparatively few in number and most of the qualities of a performance are aspects of the performer's interpretation, not of the work as such. The thinner they are, the freer is the performer to control aspects of the performance. [...] ...if the work is thick, a great many of the properties heard in a performance are crucial to its identity and must be reproduced in a fully faithful rendition of the work. The thicker the work, the more the composer controls the sonic detail of its accurate instances. (Davies, 2001, p. 20)

There are no rigid rules about the extent of the performer's freedom or which decisions are their prerogative: the answer varies from work to work, but also by genre and historical period.

Classical works are typically thick (some thicker than others), while jazz works are thin. Works

---

<sup>11</sup> Davies (1987) used to agree but has later changed his mind (Davies, 2001, p. 16).

specified by a melody or a chord progression give the performer almost as much discretion over the sound as the composer typically enjoys: “for [extremely thin works], the player creates the larger structure of the performance by deciding on the number of repeats, variations, elaborations, links and the like” (Davies, 2001, p. 20). Additionally, what seems like a modest opening for creative freedom can make a big difference in the impact on the audience. This report by Cooper (2014) shows how a slight nuance in the performance by the Vienna Philharmonic (choosing to play according to the conventions of an earlier time rather than strictly following the notation) impacts the listening experience:

To demonstrate the Philharmonic’s way with a waltz, [its lead violinist] Rainer Honeck [...] played an excerpt from Joseph Lanner’s “Steyrische Tänze.” First, he played it in strict 3/4 time, as written. Then he played it in the Viennese style, achieving the trademark lilt by rushing the second beat a tad and hesitating before the third. “This is a very good example of playing not 1-2-3, but 1-2 ... 3,” he said. [...] That slight uncertainty, and unpredictability, can create a pleasing tension for the listener, who, consciously or not, cannot anticipate quite when the next beat will come. (Cooper, 2014)

How can the ostensive-inferential model describe the act of the performer that involves both deriving an interpretation and creating certain content of their own? The most obvious solution is framing the performer’s act can be framed as a two-step process. First, the performer acts as an addressee and infers the instructions intended by the composer from the score. This step creates the first mental representation of how the performance must progress in the performer’s mind, one that contains gaps for what is underspecified by the score. Second, the performer acts in the role of a creator and produces a new representation, creatively filling in the gaps in a way that expresses the performer’s intentions toward the audience. This new representation is what determines the performance. In this way, the dichotomy between deriving and producing content remains unmuddled, while the communicative intentions remain associated strictly with the

producer's act, as is traditional. Davies seems to suggest something similar to the two-step proposal I just outlined:

One performs the given work by following the instructions determinative of it. The discovery of those instructions requires an appropriate interpretation of the score. But when this is achieved, many details of the given performance remain unspecified. The performer must decide what to do about them. This requires a further act of interpretation. Whereas the first concerns the interpretation – that is, the reading – of the score, the second focuses on the interpretation – that is, the rendition – of the work. (Davies, 2001, p. 110)

However, in Davies' thinking, the gaps are filled not within the second mental representation, as I propose, but at the time of the physical realization of the sound. I will consider both possibilities to see how such separation between driving and creating might work.

The above description of the performer's two-step act assumes that one can effectively separate elements in the composer's instructions that determine the performer's actions and those that allow a choice as if the latter comes in the form of gaps requiring specific parameters. The problem with this description is that *everything* in musical notation offers a range of possibilities, from a basic musical note onwards.<sup>12</sup> The duration of a note is not an absolute but a relative quantity (relative to the speed of beats). Even though it is sometimes described as corresponding to a frequency, the pitch conveyed by a note is a psychoacoustic, subjective property: it is the *sensation* of a frequency. Timbre and volume affect the perception of pitch even in highly trained musicians (McDermott & Oxenham, 2008). Before the invention of frequency measuring and tuning instruments, pitch varied greatly. Composers and conductors had to frequently transpose their music—or play the melody higher or lower—due to the different tuning of the musical

---

<sup>12</sup> See Ostertag (2012) on a related issue – the difficulty of specifying the exact acoustical properties of a work.

instruments. The standard A440 pitch was adopted by the International Organization of Standardization in 1955, but not all orchestras tune to it even now (Gribenski, 2021). The same note sounds differently on the same instrument depending on the timbre or tone color—an elusive quality that a skillful musician can vary with different playing techniques. No part of musical notation is realizable in sound without the performer’s input (or, to some extent, chance). Separating sound elements that are the effects of notation from those that are the performer’s choices is like trying to imagine the Cheshire cat’s grin without the cat.

Davies’ idea of separating the performer’s deriving the instructions determinative of the performance from the performer’s choosing along the mental and physical lines also fails. Davies’s two stages of performer’s interpretation are the mental “reading” of the score and the physical “rendition,” or striking-the-keys. As I understand, he believes that the underdeterminacies of the score are settled in the striking-the-keys stage. However, it is likely that at least some of the underdeterminacies are settled at the reading stage, just as a person who is silently reading a poem can provide details not specified by the text—such as pausing, rhythm, and emotional intonation—in their imagination. If musical mental representations are like “music in our heads,” it is natural to assume that multiple readings can form based on the same notation. Studies in musicology have shown that the same brain areas are activated when perceiving music and imagining it. These studies also reveal that imagined music contains such attributes as tempo, timbre, and absolute pitch (Zatorre et al., 1996). Musicians also attest to a “trial and error” feedback loop between the imagined reading and the physical execution: an idea about

how something should be played emerges during the reading, gets tested in practice, and a decision is made to either keep playing it that way or mentally revise it and try again.<sup>13</sup>

This shows that it is impossible to divide the act of the performer into the deriving-of-the-determinate and filling-in-of-the-underdeterminate parts in a non-arbitrary way. The performer's creative resolution of what is underspecified by the score, expressive of the performer's intentions toward the audience, just *is* the process of deriving the melody. What about the interpretive process of the listeners?

### **4.3 The musical audience**

The stereotype of listeners as passive receivers is persistent. However, the musical chain does not always end with the listener, as when the listener is also a musical performer who subsequently turns the memory of the performance into a performance of their own. The dynamic between a music teacher and a student is like that: the teacher's playing is like a set of instructions for how the student must play. As the student listens to the teacher, their interpreting cannot be any less active than a performer's interpretation when reading the score. Similar considerations to those described in the previous section must now apply to the student: the teacher's instructions in the form of sound rather than notation constrain but do not determine the student's subsequent performance. To be sure, the sound instructions are more determinative of the performance than

---

<sup>13</sup> I am grateful to Gary Ostertag for helpful suggestions about mental readings and the feedback loop between the mental reading and the physical execution.



the notational instructions. Nonetheless, they too leave some room to chance (and the student's skill) and some to the student's expression of their state of mind and creative freedom:

The arranging of all these elements to create a coherent structure and performance is also a matter of the moment, and is influenced not only by the artist's preparation, but also by her or his mood. One might have a feeling of conquest or victory, pain or loneliness, longing or despair, or perhaps pure joy. This may arise from the performer's personal situation, or from her or his imagination or memory. This may be expressed in the timbre or mood the player elicits from her or his instrument. (Gould & Keaton, 2000, p. 146)

The example of the student listening to the teacher and turning their listening experience into a performance of their own should begin to undermine the intuition that a listener is always a passive consumer.

Another example of an active listener is an interpretive dancer, facing an audience of their own. Just as a musical performer interprets the score "into" their performance, the dancer interprets musical sounds "into" the dance. The music, the genre, the conventions of the dance art, and the choreographer's instructions will constrain but not determine the dancer's creative choices. The process guiding the dancer's movements is the process of interpreting that I have been advocating – neither purely inferential nor purely creative. It is a process of creating a surrogate for the content in one medium (musical) within a different medium (kinesthetic). The process involves obeying constraints but also expressing creativity and intention to affect the dancer's audience, just like the process underlying the actions of the musical performers who precede the dancer in the performance chain.

Furthermore, there is no reason to treat listeners seated in the audience who do not present *overt* evidence of their active interpreting as passive receivers. Even if they do not present such evidence at the time, they may do so in the future. For example, a listener may transform the memory of their musical experience into a musical performance of their own (like the student in

the example above) or into creation in a different medium—for example, a critical essay. But even simply listening to music might be a process of “active construction,” with intentions aimed at oneself. The idea of “creative listening” is receiving wider acceptance in musical psychology (Dunn, 1997; Hargreaves et al., 2012; Kratus, 2017). There is even psychological and neurological evidence that all three main musical activities—composition, performance, and listening—are linked by the same cognitive processes and structures in which imagination and creativity play an essential role:

Listening to music is an active creative process that exists at different levels of engagement: people display different listening styles, and they consciously use music in different situations in order to produce different psychological states in themselves. All music processing involves centrally-stored personal networks of association: our minds contain internal networks of mental representations, or schemata, which mediate all musical activities, and not just the act of listening itself: these include composition, improvisation, and performance. (Hargreaves et al., 2012)

What are the inputs and the output of such “creative listening”? The inputs are the most basic melodic structures—something analogous to the outputs of the linguistic module in regular communication—the outputs of the melodic module (Clark & Lindsey, 1990). Perhaps, they are the syntactic structures that the listener’s brain constructs according to innate rules in the way *A Generative Theory of Tonal Music* by Lerdahl and Jackendoff (1996) describes. A typical listener, they assert, hears music not as a uniform sound stream but as organized patterns. A piece of music, according to the authors,

...is a mentally constructed entity, of which scores and performances are partial representations by which the piece is transmitted. One commonly speaks of musical structure for which there is no direct correlate in the score or in the sound waves produced in performance. One speaks of music as segmented into units of all sizes, of patterns of strong and weak beats, of thematic relationships, of pitches as ornamental or structurally important, of tension and repose, and so forth. Insofar as one wishes to ascribe some sort of “reality” to these kinds of structure, one must ultimately treat them as mental products imposed on or inferred from the physical signal. (Lerdahl & Jackendoff, 1996, p. 2)

The syntactic structures that Lerdahl and Jackendoff describe serve as inputs to a more global, creative-inferential process that constructs the listener's overall reaction to the music *in context*. This creative-inferential process is "creative listening," while the listener's overall reaction is the musical analogy to an addressee's pragmatic interpretation of a speaker. The overall reaction can be mostly affective, just like a listener's interpretation of a speaker whose intention has mostly non-propositional content—a comforting utterance by a friend, "It's going to be ok!" or a passionate and inciting slogan at a political rally. It can also be nearly propositional in the right context or with enough accompanying non-musical clues, such as in the case of a performance of an opera or ballet. Many other inputs go into this creative-inferential process in the addition to Lerdahl and Jackendoff's syntactic structures: the listener's knowledge of the musical genre, the encyclopedic entries associated with certain sound patterns, the cues from the performance setting, and so on. The factors personal to each listener inevitably also reflect in the ultimate product: personal memories, emotions, and desires brought up by the performance.

What reason do I have to say that the process of "creative listening" involves self-directed intentions? The idea of a listener *choosing* to interpret music in a particular way, with intentions aimed at oneself, is unusual. However, consider a pianist playing from the notes for their own enjoyment. Suppose they are playing to put themselves into a relaxed state of mind. First, they will choose a piece appropriate for this purpose. However, beyond choosing the right kind of piece, they may also choose to execute it in a way that is most conducive to the production of the desired impact upon *themselves*. In other words, they express self-directed intentions. The audience in the concert hall may also seek to make an impact upon themselves (otherwise, they would not attend the performance). People commonly listen to music to achieve a particular psychological state, such as a better mood, sharper focus, or greater motivation for physical

activity. Just as the pianist in the above example may linger (in their performance) on the sound elements that promote the effect they are striving to achieve on themselves, the audience may linger (in their attention and focus) on the features of the performance that are in line with the effect they are seeking, whether consciously or subconsciously.

Finally, the word “choosing” should not be taken to imply that the creative process can only consist of active consideration and rejection of alternatives. Creation—by a composer, a performer, a listener, or anyone—can be influenced by many internal factors that the agent does not directly control, such as mood, emotions, or memories. A performer might produce a different take on the same musical piece, depending on whether they are calm or emotional. Similarly, “Moonlight Sonata” may normally elicit a sense of tranquility in a listener but arouse a profound sadness following a loss.

The intention always seamlessly drifts in and out of the creative process. Walton (1990) aptly describes this feature in the study of imagination, *Mimesis as Make-Believe*. There is a state of *spontaneous* imagining in which “our fantasizing minds stray, seemingly at random, without conscious direction.” There is also a state of *deliberate* imagining, in which “we form intentions to imagine this or that and carry them out.” The line between the two states is blurry. In one moment, the person is an observer of their thoughts, and in the next, an active director. This description also fits the experience of being consumed by music: drifting away with the sound and intermittently forcing attention back to particular aspects of the experience. Losing control over one’s creativity and surrendering to the moment is common to creators of all stripes.

## 4.4 The composer

Composing music is often viewed as a paradigm of pure artistry. Nevertheless, it also fits the pattern of making choices within specified constraints. In fact, constraints make composing possible.

Constraints on composing music begin with its building blocks: chord progressions, themes, and harmonic techniques. In tonal music—which encompasses the majority of classical Western, contemporary popular, and some tonal world folk music—chords do not follow each other randomly but are organized around a single note, called tonic, the tonal center toward which the harmony progresses. While there are infinite possible chord progressions, most music we hear contains a limited set of common, relatively short patterns. Classical and popular music sound different despite sharing the vocabulary of sound patterns, thanks to other attributes of musical complexity. The progressions, common musical themes, and harmonic techniques are often associated with the music sounding happy, sad, foreboding, and so on:

...in lockstep with the establishment in Western music of the major–minor tonal system, in the seventeenth century there evolved a whole arsenal of musical themes and harmonic techniques whose emotive character became instantly recognizable to the competent listener. Furthermore, this fundamental recognizability endures to this day. As long as the composer remains within the major–minor tonal system, he or she can still rely on the competent listener to recognize when the music is sad, when joyful, when ominous, and so on. These emotive building blocks are stock-in-trade of the film composer, the composer of popular songs, and are part and parcel of the whole classical music repertoire from the beginning of the seventeenth century to the breakdown of tonality in the twentieth. (Kivy, 2007, p. 220)

Just as a speaker cannot speak random words and expect the addressee to grasp what the speaker means, a composer cannot expect a listener to recognize a particular emotive theme in the music without relying on some mutually shared vocabulary of musical patterns. That is why the genres

with which we are most familiar sound the most meaningful to us, while the music of a distant tribe may not.

Only toward the beginning of the 20<sup>th</sup> century did classical composers begin to innovate and expand the constraints of tonality. This development culminated in atonal music, which completely rejects organization around a single tone. Nevertheless, even atonal music requires constraints:

...by removing the last vestige of traditional melodic, harmonic, and formal constraints, [atonality] presented the composer with a theoretically unlimited number of compositional choices. Nothing was impossible. [...] Faced with the awesome task of perhaps having to choose every single pitch from scratch, the composer had to invent constraints if he was to compose at all (Meyer, 2010).

One such newly devised constraint within atonal music was the twelve-tone technique, associated with Arnold Schoenberg, which ensures that all twelve notes appear within the composition without emphasis on any single note.

Another compositional constraint is the purpose for which the musical piece is commissioned. A composition is hardly ever the beginning of content transmission but is an interpretation of something else. In composing a song, the composer is constrained by the rhythm and the meaning of the lyrics. In constructing a film score, the composer must respect the scene cues, heighten the drama of climactic points, and defuse the tension in the moments of resolution.

I have already mentioned that the validity of the distinction between improvising (which, for my purposes, is the sister of composing) and interpreting an existing work has been questioned in the philosophy of music. Particularly, Gould and Keaton (2000) characterize the performance of an existing work as improvising within the constraints of the score. They also describe pure improvising as closely tracking a preexisting entity:

...how much of a truly new entity is created in an improvisation? Even an improvised performance is recognizably based on a preexisting framework: a jazz performer may improvise over “Giant Steps,” a singer may embellish a bel canto aria, or a harpsichordist may realize a continuo part from a figured bass in a trio sonata. In each case, the performance realization is based on a preexisting entity: the form and harmonic progression that identifies the piece. The player may remain very close to that framework, or may stray distantly, but he will never leave it to the extent that the identification is lost. (Gould & Keaton, 2000)

Others describe how improvising is constrained by the stock of preexisting phrases:

There is another respect in which improvising is structured. The musician usually acquires a repertoire of phrases and figurations. Some of these will be of her own invention and others belong more widely to the performing community. When she improvises, she draws on this stock. The difference between formulaic improvisation and excitingly original improvisation often lies in how the units are combined, contrasted, and developed, not in their novelty. (Davies, 2001)

The difference between performing a work and improvising is not the degree of freedom but the character of constraints: a score constrains the former, while a preexisting form and a repertoire of phrases constrain the latter.

Finally, the most important constraint on the composer’s written work is their mental representation of a melody—the composer’s “inner music” that they are aiming to express publicly. It would be an understatement to say that we do not understand well how composers might “think” musically or represent music internally. Perhaps, there is an innate system for internally representing music, a “language of musical thought” (Rawbone, 2021), organized according to the rules of musical generative grammar, such as those Lerdahl and Jackendoff (1996) attempted to describe. Or, perhaps, music only borrows the resources of other cognitive domains (Sperber, 1994). However, it is safe to say that “musical thought” also only constrains and does not determine its public expression: the composer may produce many sound and notational versions expressing a particular musical idea. There is undoubtedly a “trial and error” feedback mechanism between the composer’s mental representation they are trying to publicly

express and the physical sound, just as when the performer attempts different ways to realize the score. That is why the composer is also an interpreter in the sense I have been advocating: the act of satisfying various constraints and the pure creative impulse are intertwined in a way that is impossible to pull apart non-arbitrarily.

Furthermore, even the composer's musical thought might not be the ultimate beginning of the interpretive chain: the inner music may be an interpretation of an emotion, a memory, or a story. As the composer's musical thought turns into notation, notation turns into a reading in the performer's mind, the reading into the sound, and the sound into a new mental state within the listener—each instantiation of the content is underdetermined by those preceding it. Each stage of transformation requires creative input by the actor within the constraints specific to each stage. The actors are neither purely producers nor purely consumers of content but creative interpreters.

#### **4.5 Musical performance as a relay of interpretations**

I described a musical performance as a relay of interpretations without a determinate beginning or end: an emotion may be interpreted “into” a musical thought, then into a written score, into the sound produced by the performers, into representations in the mind of the audience, and into other works inspired by the performance. No definite content is transmitted from one nexus of this chain to another. Instead, the content is constantly inferred-cum-created according to the constraints specific to each nexus and the intentions specific to each actor. The traditional models of communication force one to see each actor in the chain of transmission either



exclusively as a producer of content or a consumer. I suggest instead that each actor is engaged in the process of interpreting that straddles the divide between deriving and creating.

This way of thinking about a musical performance helps make sense of the idea that the audience partially creates the meaning they interpret, which is prevalent in musical and common discourse but strange in traditional views of communication<sup>14</sup>. Pragmatic theory, if applied to music in its traditional form, would force a simplistic and awkward conclusion that the audience of a musical performance aims to interpret the *composer's* meaning, constituted by the composer's creative intentions. It also fails to acknowledge interpretive intentions, for example, those of the audience and the performers. In my view, however, each participant in the chain of transmission expresses their own intentions in addition to navigating constraints. The intentions of each participant are constrained by the intentions of those standing before them in the chain and constrain the intentions of those who come after. Hence, there is no straightforward way to define the meaning of a performance in terms of the intentions of any single participant or group of participants.

Listeners can *choose* to focus on either the performers' or the composer's intentions in their own interpreting. Those familiar with the composition may concentrate on the performers' intentions, while in the performance of a new work, the attention will be on the composer's ideas and not the ingenuity of the orchestra. While trained musicians often focus on the technical aspects of the playing (their fingers sometimes move along with the pianist's), non-musicians focus more on the internal affective experience (Meyer, 2008).

---

<sup>14</sup> Perhaps, one could argue that Donald Davidson's idea of radical interpretation (Davidson, 1973) is a precedent to the notion of interpreting I am advocating. An exploration of the connections to Davidson's idea cannot be pursued here due to the constraints of time and space.

Consequently, the audience contributes to the meaning of a performance in multiple ways. They choose *what* constitutes the meaning they interpret by choosing whose intentions constrain their interpreting (although they are rarely aware of this step and may even switch it midway through the experience). They also choose *how* to interpret by letting their personal preconceptions, memories, experiences, and associations into the creative-inferential process of crafting the interpretation of music in context.

Could an account of communication informed by a musical performance be helpful in explaining regular everyday communication? The following chapter posits that the notion of interpreting advocated here, as well as the idea of interpretive intentions, applies to how utterances are formed and comprehended. Relays of interpretations, I argue, where no participant is exclusively a producer or a consumer of definite content, is the right way to think about all human ostensive communication.

## CHAPTER 5: OSTENSIVE COMMUNICATION

In this chapter, I first argue that one can think of someone reporting on the speech of another speaker the same way we think of the act of a musical performer. In relaying what someone else has meant, the reporter strives for fidelity to the original source but also expresses their own communicative intention toward the secondary audience. Therefore, my notion of interpreting as deriving and creating simultaneously applies in the sphere of regular communication. In sections 5.2-5.3, I argue that it also describes regular speakers and addressees.

### 5.1 Creative intermediary: reporting on the speech of others

Relaying what someone else said happens routinely: a waiter delivers a customer's order to the chef; a child tattles on a peer's offensive remark; a spokesperson communicates the president's statement; a biographer retells a general's confession. While one's own interpreting may seem almost a deterministic process, we feel otherwise about interpreting by others. We recognize that different interpreters make different, equally acceptable interpretations. We usually accept reports as reliable even when they contain phrasing, intonation, and facial expressions distinct from those of the original speaker. Furthermore, we often realize that in expressing their interpretive hypotheses, interpreters may pursue their own intentions to impact the secondary audience. Just as we do not view a conductor as a pure middleman between the composer and the audience, we do not view a reporter as lacking independent preferences, purposes, feelings, and agendas, even while respecting all the constraints on the interpretation imposed by the primary

speaker. Finally, we realize that a report always occurs in a context different from that of the original utterance and can, therefore, be interpreted differently by the secondary audience. Despite all these realizations, we often consider acts of communication with three distinct participants – the original speaker, the reporter, and the secondary addressee – as successful. As a result, we often coordinate our actions and form our beliefs based on the reports of what someone has said.

I see three possibilities for how the ostensive-inferential model can describe the reporter's act. First, the reporter could be described primarily as the addressee, inferring the communicative intention of the original speaker and passing the original intention through to the secondary audience. Second, they could be described as expressing a communicative intention, but only of a primitive kind, one that would not change what the secondary audience would understand had they heard the original utterance firsthand. (Perhaps, it is an intention to convey whatever the original speaker meant). Alternatively, the reporter's act could be described as two-step: first they are acting as the addressee and then as the speaker, with unique communicative intentions toward the secondary audience that change what the secondary audience mentally represents had they heard the original utterance firsthand.

The first two possibilities can be dismissed. If the reporter is an intentional agent (as opposed to a mechanical recorder), they must have intentions distinct from the original speaker because each person has a unique cognitive history. Such intentions cannot be excluded from the speaker meaning at the time of the report. Even if the reporter uses the exact words, their own intentions will manifest through intonation, expressions, and connections to the context. A child relaying what the mother has said, word for word, may roll their eyes. The child's intentions must at least

contribute to the speaker meaning behind the entire act: their facial expression influences the understanding of the secondary hearer. A person's cognitive history is constantly in flux, ensuring that even the same speaker will not report their own speech with the same intentions on a new occasion. Therefore, our model needs to account for the *unique* intentions of the reporting speaker.

Let us now pursue the third possibility: the reporter first assumes the role of the addressee and then the role of the speaker. In the role of the addressee, the reporter constructs a *mental* representation,  $R_m$ , of the content of the original speaker's intentions. Next, in the role of the speaker, the reporter constructs a *public* representation, an utterance  $R_p$ , that provides the secondary audience with evidence for the content of the *reporter's* communicative intentions. However,  $R_p$  should still provide evidence for the *original* speaker's intentions. Otherwise, the communication will not be considered successful by all three parties. Therefore, the two tokens, mental  $R_m$  and public  $R_p$ , should both represent the original speaker's content.

This way, there are two sets of intentions determining speaker meaning on the occasion of a reporting utterance – the reporter's and the original speaker's – just as a musical performance reflects the intentions of both the performer and the composer. Sometimes (but not always), perceptive secondary audience may differentiate between the two sets. Fox News and CNN audiences sometimes find themselves in such a situation. If a commentator says, "The president said that Q" in a particular context, some listeners may take the utterance as is, while others may choose to discern a component of meaning attributable to the president's intentions and one attributable to the intentions of the news network.

In the traditional understanding, the two tokens,  $R_m$  and  $R_p$ , result from drastically different mental processes. The process resulting in  $R_m$ , a pragmatic process associated with the addressee's actions, is one of comprehension – a process of satisfying constraints and not involving intention. It is often described as at least partially automatic, not ostensive, not directed at an addressee (which would be the reporter themselves), and, therefore, not containing a pattern of satisfying an addressee's expectations. On the other hand, the process resulting in  $R_p$ , a formatic process associated with the speaker's actions, expresses an intention (in this case, the reporter's intention), is creative and directed at an audience (in this case, the secondary audience), and aims to satisfy the audience's needs. An easy side-by-side comparison of  $R_m$  and  $R_p$  is as follows:

*Table 1: Comparison of a reporter's mental representation of what the speaker means and the reporting utterance, as traditionally understood*

<b>Reporter's mental representation of what the original speaker means <math>R_m</math></b>	<b>Reporter's public representation of what the original speaker means <math>R_p</math>, or the reporting utterance</b>
<ul style="list-style-type: none"> <li>• Result of a <i>pragmatic</i> process</li> <li>• Not directed at an addressee</li> <li>• Context-insensitive</li>   <li>• Intention-independent</li> </ul>	<ul style="list-style-type: none"> <li>• Result of a <i>formatic</i> process</li> <li>• Directed at a secondary audience</li> <li>• Context-sensitive: contains a pattern of satisfying the secondary addressee's expectations (references objects and events consistent with the addressee's way of referencing objects and events and makes assumptions not obvious to the addressee explicit)</li> <li>• Reflects the reporter's intentions toward the secondary audience</li> </ul>

Here is how Carston contrasts the two mental processes and their resulting tokens:

Having thoughts is a strikingly different kind of mental activity from comprehending utterances. We do not have to undergo a process of comprehending occurrences (tokenings) of our thoughts (hence Mentalese sentences), as we must comprehend occurrences of linguistic utterances. Thoughts are, in effect, an end-product: they are the result of utterance comprehension processes, or are prompted by the deliverances of our sensory-perceptual systems, or are generated from within the central thought system itself (in ways that remain largely mysterious). A thought token is not ostensive, it is not a communicative act, it doesn't involve an interaction of speaker and addressee, it doesn't come with a presumption of optimal relevance, which warrants a particular expectation of effect and effort and so a particular pattern of processing geared to satisfying that expectation. There surely cannot be a 'pragmatics of thought' in the sense of pragmatics at issue here. (Carston, 2002, p. 76)

I will now argue that the two tokens –  $R_m$ , the result of pragmatics, and  $R_p$ , the result of formatics, are, in fact, incredibly similar and exhibit the same patterns, at least in the case of reporting the speech of another speaker. First, let's grant that the reporter is both the creator and the internal audience of  $R_m$  (I will shortly show that this line of thought is perfectly permissible). Now, both  $R_m$  and  $R_p$  satisfy certain constraints: the constraint of having to reliably represent the original utterance and the constraint of being accessible to their audiences (the reporter in the case of  $R_m$  and the secondary audience in the case of  $R_p$ ). Both, I will argue, also express the intentions of their creator, the reporter. Contrary to Carston, a thought does fulfill a threshold of optimal relevance for the thinker. It does contain a pattern of processing geared to satisfying the needs of the thinker. While the idea of 'pragmatics of thought' sounds strange initially, the patterns of pragmatic and formatic processes appear as such mirror images of each other, there is a sense in which they are a *single* collection of mental processes. A side-by-side comparison of how I advocate we view  $R_m$  and  $R_p$  is as follows:

Table 2: Proposed comparison of a reporter's mental representation of what the speaker means and the reporting utterance

<b>Reporter's mental representation of what the original speaker means <math>R_m</math></b>	<b>Reporter's public representation of what the original speaker means <math>R_p</math>, or the reporting utterance</b>
<ul style="list-style-type: none"> <li>• Result of a <i>pragmatic</i> process</li> <li>• Directed at the internal audience (the reporter)</li> <li>• Context-sensitive: references objects and events consistent with the reporter's internal way of referencing objects and events and "spells out" novel and relevant assumptions</li> <li>• Reflects the reporter's self-directed intentions</li> </ul>	<ul style="list-style-type: none"> <li>• Result of a <i>formatic</i> process</li> <li>• Directed at a secondary audience</li> <li>• Context-sensitive: contains a pattern of satisfying the secondary addressee's expectations (references objects and events consistent with the addressee's way of referencing objects and events and makes assumptions not obvious to the addressee explicit)</li> <li>• Reflects the reporter's intentions toward the secondary audience</li> </ul>

Admittedly, there is a certain awkwardness in speaking about being the audience of one's own thoughts and having intentions aimed at oneself. Let us make sense of both ideas before proceeding with the discussion. While it may feel odd to say that the thinker is the internal audience of a thought, there is no philosophical problem with the idea. There is hardly any other way to characterize one's role when attending to an internal monologue. We may even address ourselves using second-person pronouns and names: "Don't you even think of doing that, Sveta!" Similarly, there is no philosophical problem with self-targeting intentions. It is not uncommon to attempt to bring a psychological effect upon oneself: calm down, focus, or reassure. Trying to convince oneself to see an argument from a different perspective or to adopt a specific belief is not out of the range of normal mental activities. The idea of adopting one



interpretation over another at will has also been encountered in philosophy.<sup>15</sup> A duck and a rabbit are equally valid ways to read the outlines of the iconic image made famous by Ludwig Wittgenstein's *Philosophical Investigations* (Wittgenstein, 1953). A viewer "stuck" on one image can usually *will* themselves to switch to the alternative view.

Firstly, both  $R_m$  and  $R_p$  are constrained by the need to reliably represent the original utterance. The reliability standard is purely pragmatic: the involved parties must be able to coordinate their actions in a mutually satisfying way based on the achieved representations (as long as no one has the intention to deceive). That is, each party strives not to be called a liar or a dimwit or receive a slap for, say, an unwelcome kiss because of how the original utterance may have been represented.

At the same time, both  $R_m$  and  $R_p$  must satisfy the needs of their respective audiences. The audience of  $R_m$  is the thinker, while the audience of  $R_p$  is the secondary audience. Both must reference objects and events mentioned in the original utterance to make them accessible to the addressees. Both must also spell out the assumptions implicated in the original utterance if they are relevant but not obvious in the new context.

The audience of  $R_p$  is once or many times removed from the original speaker—a chef getting a customer's order through a server, a teacher getting an account of an insult through an aggrieved child, a journalist receiving the president's statement through a spokesperson. In all such cases, the reporter must reference objects and events differently from the original speaker if the second-

---

<sup>15</sup> The position that one can adopt an alternative interpretation at will is more moderate than the radical position of "doxastic voluntarism," a belief that we can directly adopt beliefs at will for non-epistemic reasons. It is merely a position that one can adopt an alternative interpretation at will for epistemic reasons since the same epistemic reasons may allow multiple interpretations.

order hearers were not privy to the context of the original utterance. “That tree” of the original utterance may have to be referenced as “the tree in the middle of so-and-so’s yard” in the reporting utterance, and “Joe” as “my brother’s friend.” Sometimes, a reporter may need to relay an utterance containing a term whose reference is not easily accessible to the secondary addressee even with a paraphrase. For example, a reporter may have to relay “The long-term interest rates are determined by the bond market” to a child or to someone who lives in a country with a state-run economy and, therefore, lacks the concept of a bond market. To make the reference of the term “bond market” accessible to such an addressee, the interpreter must put in a significant effort. Ideally, the reporter has to impart the whole encyclopedic entry containing information about the denotation of the mentioned concept. However, the task would often entail transferring such a vast amount of information that it would be impractical to do so as a part of the reporting act. The reporter would have to instead reference something rather distantly related to what the original speaker meant.

$R_m$  must also reference objects and events differently from the original speaker, but the referencing will be according to the *reporter’s* private ways of referencing them. We conform our internal representations to our private ways of individuating and conceptualizing things and to our private language the same way we conform our public utterances to what we expect are the addressee’s ways of individuating and conceptualizing things and to their vocabulary.

Sperber and Wilson (1986/95, p.1992) concur:

It seems plausible that in our internal language we often fix time and space references not in terms of universal coordinates, but in terms of a private logbook and an ego-centred map; furthermore, most kinds of reference - to people or events for instance - can be fixed in terms of these private time and space coordinates.

Therefore, “that tree” of the original utterance will be referenced with a unique placeholder (perhaps, even a stored image) within the reporter’s mental representation, individuating a particular tree only for the reporter. How would a hearer internalize a concept whose reference is not accessible to them, such as “bond-market” to someone not familiar with how market economies work? The experience reminds me of the early years of my philosophy studies, laboring to comprehend almost every utterance produced by my professors. Often, getting at the reference of a term would take days of research, and still result only in a vague outline of what the speaker referenced.

Next, the reporting utterance  $R_p$  must spell out the assumptions obvious in the initial conversation but not at the time of the report. Imagine a dialog between a host and a guest at a party:

(15) The chef is planning to serve swordfish.

(16) I eat only kosher fish.

In the context of the conversation, the assumption that

(17) swordfish is not kosher

is unspoken but manifest. Sperber and Wilson (1986/95, p. 195) call such unspoken but implicated assumptions *implicated premises*. Now, imagine that the host goes to the kitchen and utters the following to the chef:

(18) One of the guests said they eat only kosher fish.

In the context of the host's reporting, (17) may or may not be manifest, depending on what preceded (18) in the conversation and whether the chef is knowledgeable about kashrut. If the chef believes that swordfish *is* kosher and if the prior context did not make (17) manifest, the chef could take the host's utterance of (18) as the host's attempt at small talk or as sharing a curious fact about the guest. If so, uttering (18) without making (17) obvious will violate the guest's communicative intention. To report what the guest meant accurately (i.e., in a way that would achieve mutually satisfying coordination among all three parties), the host would need to rephrase the original utterance to either make (17) explicit or describe the elements of the original context that would make (17) manifest. For example, the host could say either of the following to the chef:

(19) One of the guests said that they eat only kosher fish, and swordfish is not kosher.

(20) When I told one of the guests that you would serve swordfish, they said they eat only kosher fish.

However, if the host knows that the chef already knows (17) or if (17) is manifest between the host and the chef via the preceding context, then (19) and (20) are not necessary, and uttering (18) alone will be sufficient.

There are many other examples in which reporting an utterance accurately requires including an explanation of the original context or additional premises. Often, what needs to be added is not just linguistic context but historical, cultural, and social. What is an innocent remark in one conversation can be perceived as offensive, threatening, or obscene if the original context is not explained. Suppose you are lightheartedly teasing a childhood friend, jokingly calling them insulting names and yelling, "I will kill you!" If this remark is relayed without explaining the

social context – namely, that it was made during horseplay between old friends – the police may want to ask you some questions. Furthermore, contemporary addressees often interpret communicative acts of the past in a way the original speaker would no longer endorse. This obvious fact often gets lost in modern discourse, when ink gets spilled criticizing various public figures for statements made decades ago in a different social and cultural environment.

If the original speech targets experts, while the report aims at non-specialists, the background assumptions possessed by the experts and necessary for understanding the speech must be explained to the lay audience. An explanation of the implicated premises is often needed when we relay something to an audience of a different generation than the original speaker, for example, when relaying parents' discussion to children or teenage-talk to grandparents.

Similarly, the reporter's mental representation,  $R_m$ , spells out an implicated premise, such as (17), depending on whether the premise is known and relevant to the reporter. (Of course, a mental representation cannot "spell out" things. What I mean is that the reporter *attends* to an implicated premise mentally.) An utterance can implicate an indefinite number of premises. However, it is unlikely they all would be represented in the hearer's understanding. For example, (16) tacitly implies many premises in addition to (17):

- (21) The speaker (the guest) believes the chef is planning to serve swordfish to *all* guests.
- (22) The speaker believes that the addressee (the host) influences what the chef serves.
- (23) The speaker believes that the addressee knows what "kosher" means.
- (24) The speaker believes that the addressee knows that the speaker is a guest (as opposed to a member of the waiting staff).

(25) The speaker believes that the addressee speaks English.

An implicated premise is represented in the hearer's understanding only if it is novel and relevant to the hearer. Imagine being the host in this situation. If you know that swordfish is not kosher, (17) will pass your attention without being flagged, and so will (21)–(25) unless they are novel and relevant. You will internally represent something similar to the following:

(26) The speaker implicated that they would not eat swordfish.

On the other hand, if you are unfamiliar with the rule that swordfish is not kosher, you will pay special attention to (17). Your mental representation of what the guest meant will be as follows:

(27) The speaker implicated that swordfish is not kosher and that they would not eat swordfish.

Under certain circumstances, you may also internally represent (21)–(25). For example, suppose you initially believed that the kosher-eating guest was staying only for cocktails and leaving before dinner. Then, you would internally represent (21) as follows:

(28) The speaker implicated that they were staying for dinner and would not eat swordfish.

If you are not the host in charge of relaying the guests' wishes to the chef, you will internally represent (22) as follows:

(29) The speaker implicated that they would not eat swordfish and that they thought the addressee was the host.

Finally,  $R_m$  and  $R_p$  differ from the original utterance because they reflect the reporter's unique perspective (i.e. unique background assumptions) within the given constraints. While most pragmatists would probably agree that such perspectives are already reflected in  $R_m$  and not only in  $R_p$ , they may not agree that invoking a unique perspective merits invoking an intention aimed at oneself (in this case, the reporter). They may say that  $R_m$  only reflects a perspective of the reporter and not a self-targeting intention, while  $R_p$  reflects the reporter's intention towards the secondary audience. I now demonstrate that there is no principled difference between the effect of a unique perspective and the effect of a self-targeting intention. Imagine a consultant lecturing on consumer habits:

(30) Now people have become more health-conscious.

A reporter may paraphrase the original utterance (30) in various ways that would typically pass most people's judgment of accurate reporting. Paraphrases could happen by choice or because memory is economical: by the time of the report, the short-term memory of the exact utterance has faded while its understanding remains. Consider possible reporting utterances  $R_p$  conveying what the original speaker meant by uttering (30):

(31) Recently, consumers have become focused on health.

(32) Today, more American adults diet and exercise.

(33) In the 21<sup>st</sup> century, we prioritize healthy living.

Most hearers should agree that (31)–(33) still constitute accurate reporting, despite the differences in wording. In most situations, (31)–(33) would equally allow the speaker, the reporter, and the secondary audience to coordinate their actions successfully. Each one of (31)–

(33) differs from (30) in its own way. Do the variations between the original utterance (30) and the reporting utterances (31)–(33) reflect slightly different internal interpretations  $R_m$  by various reporters (the result of pragmatics) or slightly different communicative intentions of different reporters toward the secondary audience (the result of formatics)?

The reporter's understanding  $R_m$  of what the original speaker meant by (30) cannot be formed based on the utterance alone – it must involve some preexisting assumptions. (This suggestion parallels the thesis in the philosophy of science known as the Duhem–Quine thesis: one cannot make a prediction based on a scientific hypothesis alone and must consult other background assumptions.) For example, to realize that “now” in (30) cannot refer to the timeline of the most recent thirty seconds, the reporter must rely on the preexisting belief that it takes years for a population to change health habits. To understand that “people” in (30) cannot include toddlers, the reporter must have the preconception that health considerations matter only to adults.

The preconceptions necessary to understand the speaker uttering (30) are acquired gradually through life experience. A young child may lack them and is more likely to misrepresent the speaker meaning. Most Western adults have similar ideas necessary to understand (30).

However, by their nature, these ideas cannot be the same. For instance, middle-aged hearers who witnessed the rise of the organic food movement may take the “now” in (30) as referring to the time after the turn of the century. For younger generations, the “now” may signify a shift towards the fad diets of the past decade. Moreover, the terms attached to highly complex encyclopedic entries, such as “bond market” or philosophical terms, like “consciousness,” “intention,” and “meaning,” most likely reference drastically different objects for each person, often depending on their experience in the particular field. Therefore, different reporters may build slightly



different mental representations  $R_m$  as a result of hearing (30) and, subsequently, construct different reporting utterances  $R_p$ . Thus, at least some of the variations between (30) and (31)-(33) are due to the reporters' different understandings (the result of pragmatics) and not due to the reporters' unique communicative intentions (the results of formatics).

However, it is not only the reporters' prior beliefs and assumptions that affect the resulting understanding – their values, preferences, goals, desires do as well. Between two equally suitable interpretations, the hearer will most likely choose the one that will preserve not only their web of beliefs but also their web of agendas. Those who embrace an active lifestyle may construct a mental interpretation of what the speaker meant by “health-conscious” in (30) emphasizing sports and exercising. A vegan will have more food-centric associations. Confirmation bias is not just frequent but intrinsic. Most of the time, it goes unnoticed. Sometimes, when a report reflects the reporter's values or preferences noticeably, we may say, “He heard what he wanted to hear,” not necessarily to declare a deception or a misunderstanding but to note the interpreter's bias. Can we confidently say whether such reporting bias is the result of the reporter's unique communicative intention toward the secondary hearers or a self-targeting intention, an intention to preserve the webs of beliefs, preferences, and goals for the reporter? On many occasions, there will not be a principled difference between the two possibilities.

To summarize, the reporter's act involves constructing two token representations – a *mental* representation  $R_m$  of what the original speaker meant and a *public* one  $R_p$  intended for the secondary audience. However, the two tokens are remarkably similar: both are constrained by the fidelity to the original speaker and both allow to express the reporter's personal perspective. In this way, both the pragmatic and the formatic processes are processes of interpretation. Like

the act of a musical performer, the reporter's act straddles the division between creating and inferring. Next, I demonstrate how the same description applies to the addressee of a communicative act.

## **5.2 Audience as a creator of content**

In the pragmatic works inspired by *Relevance*, the contrast between the speaker and addressee is less dramatic than in the past philosophy of language. The speaker is seen as equally constrained by the linguistic meaning of the words and considerations of the context as the addressee. The addressee is no longer a reflexive copier of content, showing a degree of freedom and even imagination in certain kinds of interpreting. The more nuanced version of the modularity hypothesis no longer takes the addressee's interpreting as informationally encapsulated.

However, in other respects, the contrast persists. Only the speaker's intentions are a part of the theory explaining communicative behavior. Pragmatics is still seen as occupying its own special purpose sub-module within the greater mind-reading module, usually yielding fast and mandatory inferences. I will now argue that the contrast between the speaker and the addressee should be reduced even further. The situations in which interpreting is neither fast nor mandatory are common. The freedom and imagination required for interpreting even mundane utterances justify postulating the addressee's interpretive intentions, and certain communicative situations even require it. Finally, the degree of similarity between the pragmatic and formatic processes implies a low likelihood that they evolved separately for different evolutionary purposes to occupy distinct modules of the mind.

Many actions in our daily lives require our conscious control when performed for the first time but become reflexive, fast, and mandatory through repetition without being associated with a specialized mental module. Examples of such actions are driving, navigating in a familiar neighborhood, playing an instrument or a sport, getting out at the right train stop, and even preparing dinner under certain conditions<sup>16</sup>. Even forming utterances can become automatic. For example, one may spontaneously utter “Good, thank you!” in response to “How are you?” or automatically recite lines frequently used in their line of work, such as questions for foreign visitors by border control officials and Miranda rights by police. Similarly, situations in which interpreting utterances is fast and mandatory tend to involve remarks processed by the dozen every day, such as “It’s raining” and “Where is the milk?”

On the other hand, comprehending utterances is not always spontaneous. For example, I took more than a moment to understand what the then Secretary of Defense Donald Rumsfeld meant by his famous utterance:

[...] as we know, there are known knowns; there are things we know we know. We also know there are known unknowns; that is to say, we know there are some things we do not know. But there are also unknown unknowns—the ones we don't know we don't know. (Rumsfeld, 2002, as cited in Graham, 2014)

However, the best examples of interpreting that is neither spontaneous nor mandatory involve figurative language. Even though Sperber and Wilson do not go as far as postulating interpretive

---

<sup>16</sup> Daniel Kahneman (2002) postulated two types of mental processes:

...judgments can be produced in two ways (and in various mixtures of the two): a rapid, associative, automatic, and effortless intuitive process (sometimes called System 1), and a slower, rule-governed, deliberate and effortful process (System 2).

System 1 and System 2 are not domain specific. The same task can be performed by either System 1 or System 2, depending on the circumstances.

intentions, they emphasize the addressee's imagination, creativity, and even responsibility for the resulting interpretation in such situations.

First, consider their examples involving zeugma, a figure of speech in which the same word repeats in different parts of the sentence but relates differently to each part:

(83) Mary went on holiday to the mountains, Joan to the sea, and Lily to the country.

(84) Mary lives in Oxford, Joan in York, and Lily in a skyscraper.

(85) Mary came with Peter, Joan with Bob, and Lily with a sad smile on her face.  
(Sperber & Wilson, 1986/95, p. 222)

The three examples have similar structures, but only (84) and (85) contain zeugma. This figure of speech creates an effect reminiscent of a grammatical mistake and signals to the hearer that interpreting will not proceed as usual. The odd construction requires a unique explanation. With (84), Sperber and Wilson explain, a hearer is expected to not only grasp that Mary, Joan, and Lily live in different kinds of towns and buildings but also envision the various ways in which their lives are affected by the choice of an address. With (85), the addressee is expected to describe the story behind Lily's sad smile in their own way.

Inventing a unique narrative as a hearer is not intrinsically different from doing so as a speaker. Different addressees will have unique visions of how Lily's life differs from that of her friends and why she appears unhappy. The differences in their interpretations cannot be attributed to the differences in the constraints within which each addressee constructs their own representation of meaning. I suggest that the differences are due to different interpretive intentions by the addressees.

Consider another example from Sperber and Wilson:

Compare the interpretation of (81) and (82):

(81) My childhood days are gone, gone.

(82) My childhood days are gone.

[...] To justify the repetition of 'gone', the hearer must think of all the implicatures that the speaker could reasonably have expected him to derive from (82), and then assume that there is a whole range of still further premises and conclusions which the speaker wants to back. For this, *the hearer must expand the context* [emphasis added]. As a result (81) might suggest, say, that the speaker is experiencing a torrent of memories which the hearer is *being trusted to imagine for himself* [emphasis added]. (Sperber & Wilson, 1986/95, p. 222)

The repetition of the word “gone” constitutes an unusual construction that takes the hearer out of the “autopilot” mode. An interpretive representation cannot simply occur to the hearer: they need to actively conjure a non-straightforward scenario in which the repetition makes sense. Such conjuring is not unlike the creative work of a speaker. Furthermore, there is nothing mandatory about such a representation. In contrast to Sperber and Wilson’s hearer who imagined the speaker experiencing a torrent of memories, I pictured the speaker repeating “gone, gone” because they suffered an irreversible loss, similar to how someone would echo “My vase! My vase!” if a valued heirloom was suddenly shattered. Explaining the differences between these two interpretations requires an appeal to more than the constraints that direct the addressee’s interpreting, since the constraints are the same for both interpreters in this case. It requires an appeal to unique *interpretive intentions* guiding the interpreters.

Postulating interpretive intentions is also difficult to avoid in a communicative act in which the speaker and hearer do not cooperate even though the former does not aim to deceive. Imagine a police officer on a domestic disturbance call, interrogating the spouse of a victim. The suspect says, “Officer, my wife and I had an altercation. I tried to leave the room, but she was in my

way. So, I moved her.” The skeptical officer confirms their understanding: “You mean you shoved her?”<sup>17</sup>

Here, the speaker’s communicative act is not equally successful for the speaker and hearer. The former is likely dissatisfied with the outcome as they failed to create the intended representation in the hearer’s mind. They also clearly intended to *avoid* a specific representation but failed at this objective as well. On the other hand, the hearer is likely happy with the outcome. They took the utterance as evidence of what led to the 911 call, not of the speaker’s communicative intentions. They probably reached the correct conclusions relevant to *their* agenda. The fact that the act satisfies the hearer but not the speaker suggests that the two agents have conflicting intentions in regard to the outcome of this communicative act.

Diverging intentions between the speaker and hearer can also occur in situations that are not hostile. For example, consider this simple conversation:

Speaker 1: “Are you ok?”

Speaker 2: “Yes, I am fine.”

Speaker 1: “You do not sound fine.”

Hearers often take utterances as evidence of more than the speaker’s intentions. They often use utterances for evidence in answering other questions they find relevant: the circumstances that prompted the utterances, intentions *behind* the communicative intentions, concealed thoughts and emotions, and deeper mental life of the speaker. Utterances should be considered not as parcels of intended content but as flashlights that briefly illuminate the inner reality of the speaker that is

---

<sup>17</sup> I have come across this example in literature but can no longer recall the source.

normally hidden. In those brief flashes, the hearer searches for evidence relevant to their own interests.

Routine remarks such as “Pass the salt” or “Shut the door!” can also present a picture in which the speaker’s and hearer’s goals for the communicative act differ. A hearer can take “Pass the salt!” as a request but also as a vague allusion that the dish is bland or the cook is an amateur. The hearer can interpret the tone of “Shut the door!” as a signal of the speaker’s irritation. They can also take the speaker’s failure to use terms of politeness as evidence for the speaker’s view on the group power structure. The distinction between the information that is intended by the speaker and that which is sought out by the hearer to confirm their biases or advance their agendas is not always clear. For example, the speaker may deny that they failed to use “please” to convey superiority, but whether their denial is sincere might not be clear even to the speaker themselves.

Real-life speakers recognize addressees’ interpretive intentions even though our theories do not. In constructing utterances, speakers often take greater care to avoid betraying certain content than to convey some other. Speakers *expect* their addressees to search for the answers to their own questions and to confirm their preexisting biases and often try to counteract such efforts. For example, in forming an answer to the question, “How are you?” the speaker might be focused on not betraying internal anguish through intonation, facial expressions, or a carelessly chosen word. Those of us who have ever received a letter of rejection may attest to how carefully worded such letters usually are, dressing a potentially hurtful content in the attire least likely to offend. Thus, speakers often treat a communicative act less like playing “pass the baton” and more like “hide-and-seek.”

Another evidence for the addressee's intentions is the case in which the addressee holds the speaker responsible for the attributed meaning even after the speaker denies the attribution (e.g., "I did not mean it in a sexist way!"). To be sure, in some such cases addressees have good reasons to continue holding the speaker responsible. The speaker may express themselves in a way that preserves the option of denying their communicative intention in case the addressee reacts hostilely to what was implied (i.e., the speaker maintains *plausible deniability*) (Pinker et al., 2008). However, in other cases, the speaker's denial is honest, but the addressee persists in reacting negatively to how they understand the speaker. The addressee's unique goals and agendas regarding the outcome of the communicative act play a role in their unwillingness to reconsider their reaction. Sometimes, hearers may not care for the speaker's intentions at all in deciding their reaction to their words. Did the speaker *intend* to imply that the meal was flavorless when they requested the salt or were they merely careless to leave the implication possible? The chef may be offended regardless. The granddaughter may be offended by the grandfather's comment that she perceives as sexist and patronizing, even if she wholeheartedly believes that his intention was not to offend her. In such cases, the speaker may accuse the addressee holding on to the wrongfully attributed meaning of "twisting the speaker's words" or "putting words in the speaker's mouth." All these examples would be difficult to explain without allowing that hearers have their own intentions to affect the outcomes of communicative acts independent of the speaker's intentions.

The hearer's intentions not only concern the outcomes of the entire act but might also be required to resolve the underdeterminacy arising from a single word or phrase. For example, there is no way to construct an understanding of "moved" in "I moved her" or "health-conscious" in "Now people are more health-conscious" which is unaffected by the hearer's personal values,



preferences, and goals. In the previously discussed example, a hearer perceiving “moved” may represent either “shoved” or “picked up and relocated” depending on their incentive to get the bad guy. A hearer perceiving “health-conscious” may represent “prioritizing an active lifestyle,” “avoiding processed foods” or “practicing self-care” depending on their personal values and preferences. Therefore, the linguistic meaning of a single word is a constraint within which the hearer is free to construct meaning according to their own choices. The extent of the constraint can vary: some words, linguistic constructions, and rhetorical effects place tight constraints on the hearer’s interpretive intentions while others – such as zeugma, unusual repetitions, and metaphors – leave a lot of flexibility.

The more novel and complex the object of interpretation, the more the process of interpreting it acquires creative aspects. Some metaphors and poetic language require as much imagination to interpret as they do to write, making poetry criticism a creative pursuit. An interpretation of an existing literary work can become a creative tour de force in its own right, such as the Talmud (a compilation of interpretive commentaries on the Hebrew scriptures), *The Aesthetics of Thomas Aquinas* by Umberto Eco (1988), and *Shakespeare: The Invention of the Human* by Harold Bloom (1999). In such cases, an interpreter who uses an original work as a constraint to create an entirely new literary achievement is not any less of a creator than the original author. Oscar Wilde expressed this idea in figurative terms in his book *The Critic as Artist*:

Ernest

But is Criticism really a creative art?

Gilbert

Why should it not be? It works with materials, and puts them into a form that is at once new and delightful. What more can one say of poetry? Indeed, I would call criticism a

creation within a creation. For just as the great artists, from Homer and Aeschylus, down to Shakespeare and Keats, did not go directly to life for their subject-matter, but sought it in myth, and legend, and ancient tale, so the critic deals with materials that others have, as it were, purified for him, and to which imaginative form and colour have been already added. (Wilde, 1982, p. 365)

If the addressee interprets an utterance by *constructing* content within the relevant constraints, what remains of the difference between the act of the addressee and that of the speaker? As I will demonstrate in the next chapter, the speaker is also an interpreter in this sense.

### 5.3 The speaker as an interpreter of thoughts

Why is it often difficult to publicly express a thought that “feels” fully formed?<sup>18</sup> Everybody struggles to articulate a complex thought or make an esoteric idea accessible to an uninitiated audience at some point. Therefore, everyone can appreciate that expressing thoughts reliably is not always straightforward. Similar to how an addressee’s interpretation may not reliably (in a pragmatic sense) represent what the speaker meant, there is no guarantee that an utterance reliably represents the speaker’s thought (in a way that would make the speaker happy with the resulting utterance). Speakers may struggle for words, take several attempts, and miscommunicate what is on their minds, especially when the subject is idiosyncratic. Sometimes, they signal trouble with “What I am trying to say...,” “I am sorry, I am not being very articulate,” or “Well, you *know* what I mean...” Some thoughts may take hours or even days to articulate and not result in a satisfactory outcome. Schopenhauer (1918, as cited in Alshanetsky, 2019) famously said of his work *The World as Will and Representation*, “What is to be imparted

---

<sup>18</sup> In his book “Articulating a Thought,” Eli Alshanetsky (2019) investigates the phenomenon of the difficulty of articulating thoughts.

by it is a single thought. Yet in spite of all my efforts, I have not been able to find a shorter way of imparting that thought than the whole of this book.” Certain subjects seem especially prone to generating problems for our verbal expressive abilities, including theoretical arguments within intellectual fields, moral and aesthetic judgments, thoughts describing mixed emotions, subtleties in social interactions, and unusual sensations (Alshanetsky, 2019, p. 132).

I argue that the notion of interpreting that straddles creativity and constraints provides the solution to the puzzle of why fully formed thoughts are sometimes difficult to articulate. If the speaker is seen as an interpreter, in the sense of the term that I have previously outlined, there is no mystery: any interpretive process is unreliable. The thought itself and the various contextual constraints still leave room for creativity in forming the utterance and for the possibility of failure.

In the past, philosophers did not seem to acknowledge the problem of articulating thoughts. The traditional view was that the content of an utterance is a *direct reproduction* of the speaker’s mental content. Some even believed that anything that could be thought could be said (Searle, 1969). At present, however, pragmatists and philosophers of language recognize that a thought and the utterance representing that thought may not correspond exactly. Consider this example by Bezuidenhout:

Suppose I am writing a guidebook to various American cities. I am in San Francisco, wandering around making notes about things to include in my projected book. I write in my notebook 'This city is hilly'. Of course, when I come to writing up my notes in a way suitable for publication I cannot simply write 'This city is hilly', unless I also make it clear in the context that it is San Francisco I'm writing about. So either I must write explicitly 'San Francisco is hilly', or I must make it clear that this section of the book is devoted to a description of San Francisco, say by having a chapter or section heading to this effect. In other words, a speaker or writer must take the listener's or reader's needs into account, and sometimes has to use an expression which is less context-sensitive than

the one she would have chosen to express her thought were she talking only to herself. (Bezuidenhout, 1997, p. 212)

The passage shows that the constraint of needing to represent a thought reliably must be balanced against the need to represent objects in a way that is most accessible to the audience; this results in deviations between the utterance and thought. Constraints of economy can result in situations where either the utterance or the thought behind it can be true, but not both, as this example by Sperber and Wilson shows:

For example, suppose I earn £797.32 pence a month. You, a friend I have not seen for some years, ask me over a drink how much I am earning now. If I remember the exact figure, I can choose between the strictly literal and truthful answer in (105a), and the less than literal (105b), which I know to be strictly speaking false:

(105) (a) I earn £797.32 pence a month.

(b) I earn £800 a month.

In the circumstances, there is no reason to think you need an exact figure. From either reply you will be able to derive exactly the same conclusions about my status, standard of living, purchasing power, life style, and whatever else you are planning to use my salary as an indicator of. Aiming at optimal relevance, I should therefore choose the reply which will convey these conclusions as economically as possible. In other words, I should choose the false but economical (105b) rather than the complex but strictly literal and truthful (105a). (Sperber & Wilson, 1986/95, p. 233)

Furthermore, considerations of speed and effectiveness of expression, politeness, and style cause less than literal reproduction of the thought. While working on this dissertation, I often had to remind people coming through my office to shut the door. The way I expressed this routine intention varied according to the situation. Could the addressee hear me? If their ears were covered with headphones, I would make a gesture resembling shutting the door. Did I have enough time for a multi-word utterance before my addressee was out of reach? I would yell, "The door!" Did they have a similar point of view? I would express my request verbally and point out the door I wanted shut. Was the addressee a senior? I would say, "please" or "do you mind." Different jargon, speaking styles, and vocabulary are used depending on the social circle:

I may address my request to shut the door to a “dude” or a “sir” depending on the addressee. Just as a reporter changes the original phrasing to suit the secondary context and the needs of the secondary audience, the speaker expresses their thoughts less than literally to suit the primary context and the needs of the primary audience.

However, when the best possible balance among all the constraints is achieved, various equally suitable ways to form the utterance remain, requiring a leap of creativity to finalize the result, the same way as a leap of creativity is required of the addressee to choose between many equally suitable interpretations. In other words, a speaker *interprets* their thoughts to others:

Reported speech or thought is thus not the only interpretive use of language. [...] we want to argue that there is an even more essential interpretive use of utterances: on a more fundamental level, every utterance is used to represent a thought of the speaker's. (Sperber & Wilson, 1986/95, p. 230)

This is because the kind of content that can be represented in one medium does not correspond exactly to the kind of content in a different medium. The content of mental representations does not correspond to the content of public utterances exactly, just like the content expressed in canvas and paint does not exactly correspond with the content of the artist's inner vision. To interpret between mismatching mediums is to create a surrogate or stand-in. The more dissimilar the mediums, the greater the leap of faith required to construct an approximate correspondence. For example, imagine having to reproduce Mona Lisa with Jellybeans or Michaelangelo's David with Legos. No close-to-literal approximation is possible in such cases. In contrast, the more similar the mediums, the easier it is to achieve a more literal correspondence. Reproductions of David in bronze, plaster, and fiberglass are more accurate but also less creative.

What is the content of thoughts and utterances? This has been a controversial subject. It has been suggested that the thought content is not always of the same kind. We may think in

propositions, mental maps, stored records of sensory states, or combinations of all of these on different occasions (Boroditsky & Prinz, 2008). However, the content of the utterance cannot be the same proposition, map, or record of a sensory state as the content of the thought it is meant to convey. After all, the utterance is *evidence* of thought content not a vehicle for it. Instead, the utterance may provide a *blueprint* for creating a proposition, a mental map, or a sensory state in the mind of the addressee. Perhaps, the most literal correspondence between thoughts and utterances is achieved when the thought is most language-like, or the closest to expressing a particular proposition.

That is why *weak communication* presents the best illustration of the speaker as a creative interpreter. With *weak communication*, the speaker does not intend the addressee to understand a specific proposition but only pushes their thoughts in a certain, vague direction. Consider this example of weak communication by Sperber and Wilson that I discussed previously:

Mary and Peter are newly arrived at the seaside. She opens the window overlooking the sea and sniffs appreciatively and ostensively. When Peter follows suit, there is no one particular good thing that comes to his attention: the air smells fresh, fresher than it did in town, it reminds him of their previous holidays, he can smell the sea, seaweed, ozone, fish; all sorts of pleasant things come to mind, and while, because her sniff was so appreciative, he is reasonably safe in assuming that she must have intended him to notice at least some of them, he is unlikely to be able to pin her intentions down any further. (Sperber & Wilson, 1986/95, pp. 55-56)

What kind of content does Mary wish to share with Paul? The answer is so uncertain that even Sperber and Wilson seemingly struggle to describe it:

If asked what she wanted to convey, one of the best answers Mary could give is that she wanted to share an impression with Peter. What is an impression? Is it a type of mental representation? Can it be reduced to propositions and propositional attitudes? What we are suggesting is that an impression might be better described as a noticeable change in one's cognitive environment, a change resulting from relatively small alterations in the manifestness of many assumptions, rather than from the fact that a single assumption or a

few new assumptions have all of a sudden become very manifest. (Sperber & Wilson, 1986/95, p. 59)

They appear to suggest that the content of an impression is not a proposition or a set of propositions. It is also the kind of content that does not have straightforward correspondence in a public language: “Is there a plausible answer, in the form of an explicit linguistic paraphrase, to the question, what does she mean? Could she have achieved the same communicative effect by speaking? Clearly not” (Sperber & Wilson, 1986/95, p. 59). When the thought content is the least propositional, expressing it publicly requires the most creativity. Such can be thoughts about perceptions, emotions, intricacies of social situations, aesthetic aspects, and topics that are not frequent subjects of discourse in one’s culture.

People are not uniformly good at the creative aspect of interpreting their thoughts, the same way not everyone is uniformly good at creative expression using canvas, music, poetry, or interpretive dance. While most people are competent speakers of at least one public language, some have a “way with words.” In describing the taste of lychee, the smell of durian, or the sensation of shin splints, each person’s success will depend on their skill with words and, perhaps, to some extent on luck. I searched for a fitting description of the durian smell for some time, though it was on the tip of my tongue. The memory of the smell was transparent and familiar but did not readily correspond to words. The best I could do was that durian smells like a mixture of sulfur and overripe bananas. Food writer Richard Sterling, much better at using sensory language than me, describes the smell as “turpentine and onions, garnished with a gym sock.” However, the ever-eloquent Anthony Bourdain captured it most creatively: after eating durian, he said, “Your breath will smell as if you’d been French-kissing your dead grandmother.”

## CHAPTER 6: CONCLUSION

### 6.1 Summary

The ostensive-inferential communication model of Sperber and Wilson is a substantial advance in our understanding of how everyday human communication works. In this dissertation, I have intended to contribute to the project by developing a notion of interpreting that does not draw a strict line between deriving and creating content, arguing that such a line cannot be drawn in a non-arbitrary way. I have also argued that this notion of interpreting applies equally to those traditionally viewed as exclusively producers of content (speakers) and those traditionally viewed as exclusively consumers of content (addressees). While the original ostensive-inferential model recognizes only the speaker's communicative intentions, I suggested that it should be modified to treat the intentions of everyone involved in a communicative act equally. That is, it should recognize that some speakers' acts do not involve obvious intentions, as they might be automated through repetition, while some addressee's acts of interpreting are creative, laborious, and intentional. Finally, the proposed notion of interpreting supports the idea that music and arts should be viewed as a part of the broad spectrum of ostensive communication, which also includes everyday conversing.

In Chapter 2, I first presented an account of the ostensive-inferential model and the non-preservationist approach to communication that it represents. I contrasted it with the code model, which illustrates the preservationist approach, according to which the content from the speaker's mind is completely or nearly duplicated in the addressee's mind. However, in the non-



preservationist approach, the speaker only gives evidence of their communicative intentions to the addressee. The contrast between the actions of the speaker and of the addressee is stark under the preservationist approach; that is, the speaker is in control of what and how they communicate while the addressee has no control over what and how they understand. In contrast, under the ostensive-inferential model, this difference is blurred: both the speaker and the addressee are mutually constrained and the latter's creative input is required to resolve what remains underdetermined by the utterance.

In Chapter 3, I further argued that the ostensive-inferential model could also apply to interactions between artists and viewers, musicians and listeners, and authors and readers and that music and arts should be seen as a part of the ostensive communication universe. Although my focus was on music, the argument equally applies to visual arts, architecture, poetry, and more. Those who view music and regular communication as discontinuous, I argued, make two crucial mistakes. First, they hold the no longer popular preservationist conception of communication. Second, they compare utterances to an idealized notion of absolute music (i.e., music taken separately from the context). This creates the perception that communication and music cater to different types of content and that the latter lacks the resources of regular communication. However, when compared on an equal basis (that is, utterances on a particular occasion compared to musical performances on a particular occasion) and seen through the lens of the ostensive-inferential model, music and regular communication transact in the same types of content and have the same resources for conveying meaning and ensuring understanding.

In Chapter 4, I argued that applying the ostensive-inferential model to a musical performance requires certain changes to be made to the model that are also useful in addressing regular

communication. First, a new notion of interpreting is needed for describing the performer's act of interpreting a score, one that does not draw a line between inferring determinative and creatively filling in what is underspecified by the score. Second, the same notion of interpreting must apply to those traditionally viewed as the ultimate producers and consumers, the composers and listeners. Third, the model should equally recognize the creative intentions of everyone involved in a musical performance, that is, the composer, performers, and listeners. Finally, a musical performance should be considered as a relay in which each agent directs their intentions toward those standing ahead of them in the communicative chain and is simultaneously constrained by the intentions of those standing behind. Thus, the meaning of a musical performance cannot be constituted by the intentions of any single agent or set of agents.

The notion of interpreting developed in relation to a musical performer can also describe someone who reports a speech, as suggested in Chapter 5. A reporter, like a performer, is both constrained by the necessity to represent the original utterance accurately and driven by their own communicative intentions toward the secondary audience. When comparing the *pragmatic* process underlying the formation of the reporter's mental interpretation of the original speaker and the *formatic* process underlying the formation of the reporting utterance, the two processes are remarkably similar in a way that justifies seeing them as the same set of bi-directional processes. Moreover, this same set underlies the addressee's pragmatic interpretation and the speaker's articulation of a thought.

In addition, the intention should be treated symmetrically between the speaker and addressee. With quick everyday phrases, both speaking and interpreting can be fast and automated, like other repetitive actions. However, with novel and distinctive subjects, both speaking and

interpreting can be active, imaginative, and intentional. Postulating the interpreter's intentions accomplishes the following. First, it corrects the inconsistency with which the addressee's actions are currently treated within the ostensive-inferential model: the addressee's imaginative effort and responsibility for the resulting interpretation are currently justly acknowledged but not linked to the psychological state of having an intention. Second, it explains how different hearers may form different although equally suitable interpretations that cannot be explained by the differences in the constraints – when the addressees produce varying interpretations upon hearing the same utterance, in the same context, and under the same circumstances. Finally, postulating the interpreter's intentions creates a framework to discuss communicative acts that are not uniformly successful for both the speaker and addressee but relative to each agent: that is, an act can be successful for the speaker and not the addressee and vice versa. Such situations occur when the speaker is focused on concealing certain content from the addressee and those commonly described as “twisting the speaker's words” or “putting words into the speaker's mouth.”

## **6.2 Implications**

The most significant implication of this proposal is establishing continuity between regular everyday communication and creative expression and interpretation within art, literature, poetry, and music. Despite past arguments to the contrary, the psychological processes involved in understanding a speaker are similar to those involved in understanding an author, artist, or musical performer. Even *physical* interpreting by a performer, artist, or artisan, such as a particular performance of “Moonlight Sonata” or a sculptor's interpretation of Venus, is

continuous with an addressee's pragmatic interpretation of a speaker. This line of thinking follows McCallum, Mitchell and Scott-Phillips (2020) and Pignocchi (2019), but the uniqueness of this project lies in the focus on the continuity between regular communication and music.

In this context, I question if my proposals are relevant to the debates within the literature on aesthetic and literary interpretation on how we *should* interpret literary texts and works of art. Specifically, should the creator's intention take precedence over the audience's perspective in interpreting a work of art? However, in the spirit of Sperber and Wilson's *Relevance*, I only aimed to describe how the addressees, readers, and audiences already interpret utterances, texts, music, and works of art rather than how they should. The notion of the audience contributing to the creation of meaning within the constraints imposed by the artist/author's intentions, that is, the marriage of the normative prescriptions advocated by the intentionalist and constructivist theories of interpretation, is a descriptive upshot of the communication account advocated here.

Furthermore, I discussed the problem that challenges all non-preservationist accounts of communication – specifying the success conditions for a communicative act that does not preserve but only approximates the content. This issue has implications beyond the field of philosophy and academia. That content is preserved in communication is ingrained in social interactions (Cappelen & Lepore, 2006), and the idea that a speaker conveys specific content and the addressee receives it nearly intact upholds interpersonal relations. Society is built on the beliefs that people can be held responsible for the truth of their utterances, the contents of their minds can be shared reliably, they can base their knowledge about the world and reasons for action on the testimony of others, and they can understand and be understood. The worry is that questioning to what extent such beliefs are only approximate may be detrimental to most human

activities (Cappelen & Lepore, 2006). Nevertheless, the utility of these beliefs does not make them accurate.

Accepting that we only approximate and do not copy speakers' and even our own thought content should not throw society into a tailspin. Many people's practical purposes can be achieved with a quite distant approximation of content. We nod in agreement, add statements to the conversation, and coordinate our actions while noticing variations between what was in the speaker's and the addressee's minds only on occasion. Perhaps, the extent to which communicative acts are successful if measured by the achieved similarity of content is routinely overestimated.

Furthermore, modern science and technology already equip us to recognize that none of our representations ideally map onto what they are meant to represent. For example, consider the representations of a minute: a minute represented by a child's toy hourglass, a minute on a digital stopwatch, and a minute as measured by an atomic clock. All these devices strictly differ in their representation of a minute while fulfilling the practical purposes for which they were designed. A toy hourglass might be 20 seconds off the atomic clock, but it is good enough to measure each player's turn in a children's game. Similarly, a stopwatch determines who wins and loses a race without precision in terms of the millionths of a second. Atomic clocks might differ on some granular level, but if satellites launch and shuttles dock without a hitch, these deviations are irrelevant. The purposes for which the measurement representations are constructed dictate the level of synchrony demanded from these representations.

Similarly, the communicators' intentions dictate the extent of the similarity between their communicative representations. For example, when walking down the street, if I say, "That dog

is old,” my way of representing “that dog” and “old” internally will differ from that of my addressee, but the interchange should be satisfactory for both. However, this statement would not be sufficient in a vet’s office. In certain contexts, it is necessary to ensure that the communicator’s representations are similar to some granular degree, equivalent to the atomic clock synchrony. For example, when the vagueness of certain utterances contributed to several airline disasters, aviation authorities worldwide introduced greater standardization of the phrases to be used in air traffic communication (Estival et al., 2016).

Finally, the notion of interpreting I have advocated here calls for recognizing utterances and works of art and literature not as discrete wholes but as segments in the chains of interpretation. A paradigm of such meaning-cum-interpreting is oral lore – the way societies communicate legends, knowledge, law, history, and morality across generations, continually re-interpreting the message to make it relevant for each generation and reflect new facts, values, and preferences. Oral lore exemplifies the model of communication in which there is no primary producer, ultimate consumer, or definite message. Instead, endlessly mutating content is passed on from one agent to another, whose effect on the content is both creative and interpretive. Similarly, *every* thought, utterance, and creative object is an instance of such transformation. This work is an illustration of my argument: it is an attempt to interpret those who have previously described the ostensive-inferential communication and a humble improvisation all at once.

## BIBLIOGRAPHY

- Alshanetsky, E. (2019). *Articulating a thought*. Oxford University Press.
- Barthes, R. (1977). The Death of the Author. In R. Barthes, *Image / Music / Text* (S. Heath, Trans., pp. 142-147). New York: Hill & Wang.
- Bertsch, M. (2002). Can you identify the Vienna Philharmonic Orchestra compared with the Berlin or New York Philharmonics? *Proceedings of Forum Acusticum Sevilla*.
- Bezuidenhout, A. (1997). The Communication of De Re Thoughts. *Noûs*, 31(2), 197-225.
- Bloom, H. (1999). *Shakespear: the Invention of the Human*. Riverhead Books.
- Boghossian, P. (2007). Explaining musical experience. In K. Stock (Ed.), *Philosophers on Music* (pp. 117-129). Oxford: Oxford University Press.
- Boroditsky, L., & Prinz, J. (2008). What thoughts are made of. In G. R. Semin, & E. R. Smith (Eds.), *Embodied Grounding: Social, Cognitive, Affective, and Neuroscientific Approaches* (pp. 98-115). Cambridge University Press.
- Buchanan, R. (2010). A puzzle about meaning and communication. *Noûs*, 44(2), 340-371.
- Burge, T. (1993). Content Preservation. *The Philosophical Review*, 102(4), 457-488.
- Cappelen, H. (2008). The creative interpreter: Content relativism and assertion. *Philosophical Perspectives* 22, 23-46.
- Cappelen, H., & Lepore, E. (2006). Shared Content. In B. Smith, & E. Lepore (Eds.), *The Oxford Handbook of Philosophy of Language* (pp. 1020-1055). Oxford: Oxford University Press.
- Carr Jr, V. (2015). *Mozart: Requiem/Karajan*. Retrieved from [www.classicstoday.com](http://www.classicstoday.com): <https://www.classicstoday.com/review/review-14268/>
- Carroll, N. (1992). Art, Intention, and Conversation. In G. Iseminger (Ed.), *Intention and Interpretation* (pp. 97-131). Philadelphia, PA: Temple University Press.
- Carroll, N. (1997). The intentional fallacy: Defending myself. *The Journal of Aesthetics and Art Criticism*, 55(3), 305-309.
- Carston, R. (2000). Explicature and Semantics. *UCL Working Papers in Linguistics*, 12, 1-44.

- Carston, R. (2002). *Thoughts and utterances: The pragmatics of explicit communication*. Oxford: Blackwell.
- Clark, B., & Lindsey, G. (1990). Intonation, Grammar and Utterance-Interpretation: Evidence from English Exclamatory-Inversions. In *UCL Working Papers in Linguistics Volume 2*. University College London.
- Collingwood, R. (1938). *The Principles of Art*. Oxford: Oxford University Press.
- Cooper, M. (2014, July 30). *A Sound Shaped by Time and Tools*. Retrieved from New York Times: <https://www.nytimes.com/2014/08/03/arts/music/what-makes-the-vienna-philharmonic-so-distinctive.html>
- Cross, I. (2005). Music and meaning, ambiguity and evolution. In D. Miell, R. MacDonald, & D. Hargreaves (Eds.), *Musical Communication* (pp. 27-44). Cambridge: Oxford University Press.
- Davidson, D. (1973). Radical interpretation. *Dialectica*, 27, 314-28.
- Davies, S. (1987). Authenticity in Musical Performance. *British Journal of Aesthetics*, 27, 39-50.
- Davies, S. (1994). *Musical Meaning and Expression*. Ithaca, NY: Cornell University Press.
- Davies, S. (2001). *Musical Works and Performances: A Philosophical Exploration*. Clarendon Press.
- Dewey, J. (1934). *Art as Experience*. New York: Penguin Group.
- Dunn, R. (1997). Creative Thinking and Music Listening. *Research Studies in Music Education*, 8(1), 42-55. doi:10.1177/1321103X9700800105
- Eco, U. (1988). *The Aesthetics of Thomas Aquinas*. Cambridge, MA: Harvard University Press.
- Elliot, T. S. (1921). *The Sacred Wood: Essays on Poetry and Criticism*. Alfred A. Knopf.
- Estival, D., Farris, C., & Molesworth, B. (2016). *Aviation English*. London and New York: Routledge.
- Fish, S. (1970). Literature in the reader: Affective stylistics. *New literary history*, 2(1), 123-162. doi:10.2307/468593
- Fitch, G., & Nelson, M. (2013). *Singular Propositions*. Retrieved from The Stanford Encyclopedia of Philosophy: <https://plato.stanford.edu/entries/propositions-singular/>
- Fodor, J. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.



- Force, K. A. (2004). La Monte Young, Terry Riley, Steve Reich, and Philip Glass: The Evolution of Minimalism and Audience Response. [Unpublished master's thesis]. University of Ottawa.
- Frege, G. (1956). The Thought: A Logical Inquiry. *Mind*, 65(259), 289-311.
- Gould, C. S., & Keaton, K. (2000). The Essential Role of Improvisation in Musical Performance. *The Journal of Aesthetics and Art Criticism*, 58(2), 143-148. doi:10.2307/432093
- Graham, D. (2014, March). Rumsfeld's Knowns and Unknowns: The Intellectual History of a Quip. *The Atlantic*. Retrieved from <https://www.theatlantic.com/politics/archive/2014/03/rumsfelds-knowns-and-unknowns-the-intellectual-history-of-a-quip/359719/>
- Gribenski, F. (2021). Sounding standards: A history concert pitch, between musicology and STS. In *Rethinking Music through Science and Technology* (pp. 26-46). Routledge.
- Grice, H. P. (1989). *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.
- Grice, P. (1957). Meaning. *The Philosophical Review*, 66, 377-388.
- Grice, P. (1968). Utterer's Meaning, Sentence Meaning, and Word-Meaning. *Foundations of Language*, 4, 225-242.
- Grice, P. (1969). Utterer's Meaning and Intentions. *The Philosophical Review*, 78(2), 147-177.
- Harbeck, J. (2014, March 21). *Why Is the Mor in Voldemort (and Mordor and Dr. Moreau) So Evil-Sounding?* Retrieved from Slate: <https://theweek.com/articles/449613/why-mor-voldemort-evilsounding>
- Hargreaves, D., Hargreaves, J., & North, A. (2012). Imagination and creativity in music listening. In D. Hargreaves, D. Miell, & R. MacDonald (Eds.), *Musical imaginations: Multidisciplinary perspectives on creativity, performance and perception*. (pp. 156-170). New York: Oxford University Press.
- Hirsch, E. D. (1967). *Validity in Interpretation*. New Haven and London: Yale University Press.
- Jackendoff, R. (2009). Parallels and nonparallels between language and music. *Music perception*, 26(3), 195-204.
- Kahneman, D. (2002). *Biographical. NobelPrize.org. Nobel Prize Outreach AB 2002*. Retrieved from <<https://www.nobelprize.org/prizes/economic-sciences/2002/kahneman/diploma/>>
- Kania, A. (2017). *The Philosophy of Music*. (E. Zalta, Editor) Retrieved from The Stanford Encyclopedia of Philosophy: <https://plato.stanford.edu/archives/fall2017/entries/music/>

- King, J. C. (2019). *Structured Propositions*. Retrieved from The Stanford Encyclopedia of Philosophy: <https://plato.stanford.edu/entries/propositions-structured/>
- Kivy, P. (1990). *Music Alone: Philosophical Reflections on the Purely Musical Experience*. Cornell University Press.
- Kivy, P. (2001). Experiencing the musical emotions. In P. Kivy, *New Essays on Musical Understanding* (pp. 92-118). Oxford: Clarendon.
- Kivy, P. (2004). Music, language and cognition: which doesn't belong? In I. C. SC, K. Korta, & J. M. Larrazabal (Eds.), *Truth, Rationality, Cognition, and Music* (pp. 216-232). Springer Netherlands.
- Kivy, P. (2007). Music, Language, and Cognition: Which Doesn't Belong? In *Music, Language, and Cognition : And Other Essays in the Aesthetics of Music* (pp. 214-232). Oxford University Press.
- Kratus, J. (2017). Music Listening Is Creative. *Music Educators Journal*, 46-51. doi:10.1177/0027432116686843
- Krausz, M. (1993). *Rightness and Reasons: Interpretation in Cultural Practices*. Ithaca, NY: Cornell University Press.
- Leedy, D. (1992). *Liner notes to "In C: 25th Anniversary Concert"*. New Albion NA 017 CD. Retrieved from <https://www.discogs.com/release/1263216-Terry-Riley-In-C-25th-Anniversary-Concert/image/SW1hZ2U6NDc5NTg4OTA=>
- Lerdahl, F., & Jackendoff, R. S. (1996). *A Generative Theory of Tonal Music*. MIT Press.
- Levinson, J. (1993). Performative vs. Critical Interpretation in Music. In M. Krausz (Ed.), *The Interpretation of Music*. Oxford: Clarendon Press.
- Livingston, P. (1998). Intentionalism in Aesthetics. *New Literary History*, 29(4), 831-846. doi:10.1353/nlh.1998.0042
- Margolis, J. (1989). Reinterpreting Interpretation. *The Journal of Aesthetics and Art Criticism*, 47(3), 237-251.
- Mauceri, J. (2017). *Maestros and Their Music. The Art and Alchemy of Conducting*. . New York: Alfred A. Knopf.
- McCallum, K., Mitchell, S., & Scott-Phillip, T. (2020). Art Experience. *Review of Philosophy and Psychology*, 11, 21-35. doi:<https://doi.org/10.1007/s13164-019-00443-y>
- McDermott, J., & Oxenham, A. (2008). Music perception, pitch, and the auditory system. *Current opinion in neurobiology*, 18(4), 452-463. doi:10.1016/j.conb.2008.09.005

- Meyer, L. (2008). *Emotion and Meaning in Music*. Chicago and London: The University of Chicago Press.
- Meyer, L. (2010). *Music, the Arts, and Ideas: Patterns and Predictions in Twentieth-Century Culture*. Chicago and London: The University of Chicago Press.
- Orlean, S. (2013, September 24). Horse\_Ebooks is human after all. *New Yorker*. Retrieved from <https://www.newyorker.com/tech/annals-of-technology/horse-ebooks-is-human-after-all>
- Ostertag, G. (2012). Critical Study Julian Dodd. Works of Music: An Essay in Ontology. *Nous*, 46(2), 355-374.
- Pignocchi, A. (2019). The continuity between art and everyday communication. In *Advances in Experimental Philosophy of Aesthetics* (pp. 241-266). Bloomsbury Publishing.
- Pinker, S., Nowak, M., & Lee, J. (2008). The logic of indirect speech. *Proceedings of the National Academy of Sciences*, 105(3), 833-838.
- Rawbone, T. (2021). The Conceptual Structure of Music: Congruence, Modularity, and the Language of Musical Thought. *Worldwide Music Conference*.
- Richards, I. A. (1917). *Practical Criticism: A Study of Literary Judgement*. New York: Routledge. doi:<https://doi.org/10.4324/9781315127194>
- Schroeter , L. (2012). Bootstrapping our way to samesaying. *Synthese*, 189(1), 177-197.
- Scott-Phillips, T. (2014). *Speaking Our Minds: Why Human Communication is Different, and how Language Evolved to Make it Special*. Palgrave Macmillan.
- Searle, J. (1969). *Speech Acts*. Cambridge: Cambridge University Press.
- Seibert, B. (2022, August 12). Review: Kyle Abraham’s Out There ‘Requiem,’ With Nods to Mozart. *The New York Times*. Retrieved from <https://www.nytimes.com/2022/08/12/arts/dance/review-kyle-abraham-requiem.html>
- Sperber, D. (1994). The modularity of thought and the epidemiology of representations. In J. Tooby (Ed.), *Mapping the Mind: Domain Specificity in Cognition and Culture* (pp. 39-67). London: Cambridge University Press.
- Sperber, D., & Wilson , D. (1985). Loose talk. *Proceedings of the aristotelian society*, 86, 153-171.
- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and Cognition*. Oxford: Blackwell. (Second edition 1995).

- Sperber, D., & Wilson, D. (2015). Beyond speaker's meaning. *Croatian Journal of Philosophy*, 15(2), 117-149.
- Walkin, D. (2012, April 6). The Maestro's Mojo. *New York Times*. Retrieved from <https://www.nytimes.com/2012/04/08/arts/music/breaking-conductors-down-by-gesture-and-body-part.html>
- Walton, K. (1990). *Mimesis as Make-believe: On the Foundations of the Representational Arts*. Cambridge, MA and London, England: Harvard University Press.
- Wilde, O. (1982). *The Artist as Critic*. University of Chicago Press.
- Wilson, D. (2005). New directions for research on pragmatics and modularity. *Lingua*, 115.8, 1129-1146.
- Wilson, T. (2021). The Rougarou Concerto and Initial Observations of the Flex Ensemble. [Unpublished doctoral dissertation]. Louisiana State University.
- Wimsatt, W. K., & Beardsley, M. C. (1946). The Intentional Fallacy. *The Sewanee Review*, 54(3), 468-488.
- Wimsatt, W. K., & Beardsley, M. C. (1949). The Affective Fallacy. *The Sewanee Review*, 57(1), 31-55.
- Wittgenstein, L. (1953). *Philosophical Investigations*. (G. E. Anscombe, Trans.) Oxford.
- Wollheim, R. (1993). Art, Interpretation, and Perception. In *The Mind and Its Depths*. Cambridge, MA: Harvard University Press.
- Zatorre, R., Halpern, A., Perry, D., Meyer, E., & Evans, A. (1996). Hearing in the mind's ear: A PET investigation of musical imagery. *Journal of Cognitive Neurosciences*, 29-46.