



YALE UNIVERSITY
DEPARTMENT OF LINGUISTICS

**La sintassi è mobile: The Syntax of Italian
Opera Librettos**

Isobel Anthony

Advisor: Raffaella Zanuttini

May 2020

Submitted to the faculty of the Department of Linguistics in partial
fulfillment of the requirements for the degree of Bachelor of Arts

Acknowledgements

This thesis would not have been possible without the support of many people. Thank you, first and foremost, to my wonderful advisor Raffaella Zanuttini. Your endless patience, kindness, and genuine curiosity for this project were vital to its success, and I am grateful to have had the chance to learn from you these past four years. I look forward to many more years of discussing syntax, music, and baking recipes. Grazie di cuore, Raffaella.

Thank you to the Linguistics Department for being my academic home at Yale. To Jim Wood for your saintly patience and sincere interest in each of your students; to Kayleigh Bohemier for the countless LaTeX questions you answered; to Prof. Shaw, Chris Geissler, Prof. Aboh, Prof. Weber, Prof. Bower, and all the other professors, PhD students, and lecturers for your terrific classes and contributions to the department. I am thankful to have spent four years with this amazing group of individuals, and I cherish everything that I've learned from you all.

Thank you to my Saybrook family and my Grandma Suite ladies for being my loving extended family; to my fellow Linguistics majors for being supportive and enthusiastic colleagues and friends; to my OTYC family for being my opera home on campus; and to my music professors for their support in my musical journey.

Lastly, thank you to my family and to Elisabeth Howard. Mom, Dad, Ev and Nick, you all are my rocks and I could not have done this without your love and support. Thank you Liz for introducing me to opera. You were the start of all this, and I cannot thank you enough for believing so strongly in 15-year-old me.

La sintassi è mobile: The Syntax of Italian Opera Librettos

Isobel Anthony

May 2020

Abstract

I investigate the syntax of Italian opera librettos. I will present a case study of *Madama Butterfly*, written by Luigi Illica and Giuseppe Giacosa and set to music by Giacomo Puccini. My broad research question is: what are the syntactic traits of late-Romantic Italian opera librettos? Looking at this iconic opera from the turn of the 20th century, I provide analyses for some of the unusual syntactic traits of the libretto. For each trait, I provide the relevant background from the linguistic literature on spoken Italian. I then extend the intuitions of other linguists to the libretto as best as can be done. My findings are that the libretto has frequently overt subject pronouns, which are often inserted to fit the meter of the musical phrase. Secondly, the libretto has additional extraction rules from noun phrases. I found that prepositional phrases of a certain type (genitive) can be extracted from the DP for a poetic effect, which differs from spoken Italian. Lastly, I found that the ordering of constituents within the noun phrases in the libretto is often different from spoken Italian, which I argue can be explained via a Split DP. Ultimately, I explore the relationship between natural language and artistic forms, such as music and poetry.

1 Introduction

Truly linguistic analyses of operatic librettos are few and far between. This is in part due to an academic disinterest in librettos; in literary circles, they are considered lower quality than other literature and in linguistic circles they simply suffer from lack of attention. I will begin with a definition of a libretto: a libretto is the text of an opera. An opera, in turn, is a long musical form with a narrative. My senior essay is a in-depth look at one opera

libretto in particular: *Madama Butterfly* by Giacomo Puccini. I chose this opera because it is one of the most-performed Italian operas and firmly a part of the classical canon. Any analysis of Italian opera librettos would be remiss not to include an opera by Puccini, who remains one of the most popular opera composers to this day. My essay will explore various syntactic traits found in the libretto and I will compare these traits to both spoken Italian and a contemporary theatrical work. My goal is to outline the syntax of the librettos, focusing on a few specific characteristics, and determine if these characteristics are unique to opera and also what motivates them. To guide my discussion, I propose five hypotheses as to why each syntactic trait appears. They are as follows:

- (1) a. It is musically motivated.
- b. It is poetically motivated.
- c. It is dramatically (semantically) motivated.
- d. It is a syntactically unique trait that only occurs in opera librettos.
- e. It is a trait of early-20th century Italian.

The structure of this senior essay is as follows: in Section 2 I will outline the history of the Italian opera libretto; in Section 3 I will provide an account of the Italian language at the time *Butterfly* was written and give a short comparative syntax to a contemporary Italian play; in Section 4 I will present my research on pronouns in the librettos; in Section 5 I will describe the unusual syntax of Noun Phrases in the librettos; and Section 6 I will conclude and point to further study.

1.1 Useful Terms

I will be writing using musical terms as well as linguistic terms. As such, I will assume basic knowledge of linguistic terms and little-to-no knowledge of musical terms. This section will provide the basic definitions for the musical terms I will use.

- An **aria** is a portion of the opera that is the most ‘song-like.’ Typically, it has a recognizable melody, deals with a single emotion or action, and is a moment of dramatic reflection. The harmonic turnover is slower, meaning the arc of the section is compact and musically can stand alone outside of the context of the opera.

- **Recitative**, in this era of opera¹, is the portion of speech-like singing between arias, duets, trios, chorus numbers, etc. It has the fastest harmonic turnover with more minimal melodic material and is typically the section of music where the plot line is advanced. In early opera, this was mostly *recitativo secco*, which is voice accompanied by a solo instrument such as a harpsichord. By the late 1800s, recitative (often shortened to ‘recit’) was more often accompanied by full orchestra, which is the case in *Madama Butterfly*.
- A **measure** or **bar** is a structural unit within a musical piece. Measures contain a certain number of beats, specified by a *musical meter*; most commonly there are 2, 3, or 4 beats in a measure.
- The **time signature** specifies the musical meter. A waltz, for example, typically has a time signature of 3/4. This means that there are three (3) beats in the measure and the quarter note (4) gets the beat.
- A **quarter note** is the basic beat for many time signatures.
- A **half note** is equivalent to two quarter notes.
- An **eighth note** is half of a quarter note.
- A **cadence** is the end of a musical sentence. There are many types of cadences depending on what the end of a musical phrase is conveying. A **final cadence** ends the music and has a feeling of finality. A **deceptive cadence** is basically an open door to continue through the musical line.
- A **musical repair strategy** is the phrase I am using to talk about an instance where the syntax of the libretto is non-canonical in order to serve the musical line. Examples of this are the overt realization of a pronoun when the musical line needs an extra beat, or the reordering of a phrase in order to suit the meter of the music.
- A **metrical repair strategy** is the phrase I am using to describe an instance where the syntax is non-canonical in order to fit the meter of the libretto. Unlike a musical repair

¹As the 20th century goes along, recitative becomes more and more rare, giving over to **through-singing** in which arias and dialogues blend together.

strategy, the idea is that a metrical repair strategy is not necessarily unique to opera and may be used more generally in poetic texts.

1.2 *Madama Butterfly* Synopsis and Characters

Throughout this thesis, I will refer to various plot-points and characters as necessary for describing the pragmatic-syntactic interface. As such, I provide a complete character list and a short synopsis.

- **Butterfly** (also known as **Cio-Cio San**): a 15-year old Japanese ex-geisha
- **Pinkerton**: a lieutenant in the U.S. Navy on tour in Japan
- **Sharpless**: the United States Consul in Nagasaki, Japan
- **Suzuki**: Butterfly's maid and trusted friend
- **Goro**: the marriage broker
- **The Bonzo**: Butterfly's uncle and a priest
- **Kate Pinkerton**: Pinkerton's American wife, who he marries after leaving Butterfly in Japan

The story begins with Pinkerton's arrival in Japan. He has heard of the flexible marriage laws in Japan that permit the dissolution of a marriage after an extended period of separation, and has decided to take a Japanese wife for the length of his stay. The marriage broker arranges for him to marry the lovely young Butterfly, who has been left destitute by her father's debts and subsequent suicide. She has worked as a geisha for a few years but her dream has been to marry an American and thus raise herself and her mother out of poverty. She and Pinkerton wed in the first act, and mutually infatuated, they seem set to live out their fairy-tale together. In order to prove her devotion to him, she converts to Christianity, thereby isolating herself from all of her family and friends except for the ever-loyal Suzuki. Pinkerton's tour ends after a few months, however, and he returns to the United States having unknowingly conceived a child with Butterfly. She remains convinced of his loyalty and raises the child with the expectation that Pinkerton will return to bring them both back to the United States. Three years pass in which the money that Pinkerton left for her dwindles to nearly nothing. One

winter day, she receives news of his return and stays awake all night to watch the ship come into port. This scene is the famous Humming Chorus. When he arrives, however, he is with his new, permanent American wife Kate. Heart-broken, isolated, and disgraced, Butterfly bequeaths her son to the new Mrs. Pinkerton and takes the same course as her father.

2 History of the Italian Libretto

Opera began in the late 1500s as an Italian art form meant to refer to sung dramas of Ancient Greece. It spread soon after throughout Western Europe, with a continually robust representation in its homeland. One of the first considerations in investigating librettos is that as a form, they are a hybrid between poetry and prose. [Bonomi and Buroni \(2017\)](#) claim that the libretto is primarily a poetic form, although it becomes more and more prose-like throughout history. The author of the libretto, called the librettist, has held varying roles in the formation of opera since opera's advent in the late 1500s in Italy ([Bonomi and Buroni 2017](#)). Where at first the librettist and composer were often the same person, leading to a clear synergy between text and song, text became subservient as the role of librettist was outsourced ([Herzog 1996](#)). The result is that many works from the 1630s onward violate the three cardinal rules of text-setting, which are according to the 17th century Florentine opera-style manifesto called *Camerata*: “1) Text must be clearly understood; 2) Words must be sung with correct and natural declamation; 3) Melody must depict not graphic details of text, but the feeling of the whole passage, by intensifying accents of a person speaking the words in a highly charged emotional state” ([Herzog 1996](#)). While no one opera can claim perfect adherence to all three of these rules, the issue of the libretto's subservience has gradually righted itself over time.

In the early 1900s, which is where I begin my analysis of librettos, librettists and composers often worked in tandem from the onset. Puccini and his collaborator, Luigi Illica, began dramatic sketches of *Madama Butterfly* around the year 1900 after seeing the play of the same name by John Luther Long and David Belasco ([Groos 2016](#)). They then brought in the well-known Giuseppe Giacosa to fill out the poetry for the main solos and duets ([Senici 2016](#)). The result is a text that is pointedly tied to the music and music that celebrates the text. This will become relevant for each of the hypotheses that I make. There are instances where this proximity of composer and librettist motivates me to choose one hypothesis over the other.

3 Background on early 20th century Italian

Before I begin an earnest discussion of syntax, I must first describe the linguistic context in which *Butterfly* was written. It was first conceived at the turn of the 20th century. What we know as Italian is in fact one dialect that was then molded into the official language. Many other dialects existed and still exist, with widely varying syntactic structures (see [D'Alessandro et al. 2010](#) and [Benincà and Tortora 2003](#)). The ancestor of Modern Italian is the Florentine dialect, which gained its privileged status via its rich literary history, including contributions from Dante, Petrarch, and Machiavelli ([Migliorini and Griffith 1984](#)).

Butterfly came into existence at a pivotal time in the Florentine dialect's history. Just four decades before, Italy was unified under a central government with the capital located first in Turin and then in Florence from 1865 to 1871. Upon unification, the new government set out on the daunting task of spreading a national language.² For several years both the dialect of Florence and the dialect of Rome (where the capital was relocated in 1871) were serious contenders. Florentine eventually won out over Roman, in large part due to the support of the great authors and poets of the day. These 19th century writers mostly came from a literary tradition that was monopolized by Florence and the Florentine greats (Dante, Petrarch, Machiavelli, etc.), and so they advocated for their familiar written language as the best candidate for the new nation's unifying tongue.

It is impossible to divorce the history of spoken Italian at this time from the history of concurrent prose and poetry, as writers enjoyed significant propriety over language. Well-known author and poet Alessandro Manzoni was appointed by the Minister of Public Instruction to be the compiler of a Florentine dictionary. Manzoni worked with a team of fellow writers to complete the work and Florentine dictionaries became widespread. They were translated into several dialects and sent across the new nation. To aid the spread of literacy and thus the spread of Florentine, as of 1877, children over the age of 6 were required to attend school. The plan worked to some extent as illiteracy dropped during the decade between unification and the First World War, from 78% in 1861 to 50% in 1910. The task of establishing the national language was far from complete, though, with half the would-be speakers still unable to read. It was a top-down endeavor controlled by the educated elite, and it lacked the horizontal

²The *questione della lingua* or the 'language issue,' is a fraught subject in Italy's history. Both historians and linguists alike have written widely on the turbulent history of language and national identity. See [Migliorini and Griffith \(1984\)](#) for one such example.

participation that would have sped up the process.

In addition to the difficulty of spreading a language through writing to an illiterate population, grammarians and writers faced several other roadblocks. Despite their valiant efforts to codify this national language, the actual dialect spoken by Florentines was a living, breathing being that was undergoing phonological changes even in the short period between unification and the First World War. There was a strict divide between Manzoni and language historian Graziadio Isaia Ascoli regarding whether to revert back to the literary Italian crystallized by Dante and his successors or to capture the spoken language of Florence as it was. Ascoli held the progressive view that normative interference with a language was futile and the process of language codification should be left to natural selection. This view, held by linguists today, was generally unpopular and prescriptivism dominated the conversation. Some non-trivial debates between the Manzonians and Ascolians were whether to include the recent monophthongization in Florence of words such as *nuovo* to *novo*, and vocabulary innovations such as *anello* from the prior *ditale* for 'thimble'. Ultimately, Florentines themselves rejected monophthongization in written language and so the diphthong won out in the new national Florentine. This debate ties into another difficulty of the process, which was that criticism of any given grammarian or writers' work on encoding Florentine almost always boiled down to an individual criticism of their style. The notion of 'good' language stalemated debates on competing forms, for example, with pronouns. *Lei* and *ella* co-existed as third-person singular feminine pronouns; some writers (such as Manzoni) preferred *lei* while others ridiculed *lei* and favored *ella* (such as poet Giosuè Carducci). While I have no data on the percentage of either usage, [Migliorini and Griffith \(1984\)](#) suggest that they were equally viable in spoken Florentine and coexisted in writing through the 20th century. Eventually *lei* became the naturally-selected subject pronoun and it is used today in modern spoken Italian. Regarding the process of crafting a national language, though, there was no consensus on *lei* versus *ella* because instead of a debate about natural linguistic tendencies, it was a debate about individual writing styles.

Adding to the maelstrom of difficulties in codifying the Italian language was the fact that the world was increasingly globalized. As people on the Italian peninsula came into contact with new technologies and cultures, the languages associated with them became a part of their language. The two most significant influences on the new national language were French and English. French literature and journalism remained important models for Italian writers and

French was taught as the second language in schools. During this period, there was a proliferation of borrowed French vocabulary, such as the color *marron* which became phonologized in Italian as *marrone*. Some French words were adopted in their original form, some phonologized like *marron* to Italian phonology, and others (especially idioms) were translated directly into Florentine. Italian language purists could not control the rapid expansion of French into the new national language. English began to slip into the new Italian as well with new technologies reaching Italy from abroad. Trolleys in the 1870s, for example, maintained its English name as it entered into the Italian lexicon (Migliorini and Griffith 1984). As such, the new Italian language emerging was not simply the heritage of Dante but also the heritage of a new and changing world.

In short, the initial steps towards making Florentine the national language were heavily mediated by prescriptivism and the notion that language can have an inherent good or bad quality. While this line of thinking is not extinct today, it was even more prevalent in the process of codifying and prescribing a national language. The success of this daunting endeavor is evident in the fact that Florentine mostly became, in one form or another, the national language used to this day. Those who wished to strong-arm the language of Dante were stymied by the tides of natural language development and by global shifts. As such, the dialect of Florence is not perfectly preserved in modern Italian. As more people spoke it and more generations learned it, it underwent natural linguistic evolution and has changed even since the first grammars were written. Finally, the establishment of a national language did not extinguish the kaleidoscope of Italian dialects. In fact, a literary movement towards realism³ in the late 19th century led authors to embrace their native dialects in their work, incorporating dialectal phrases and vocabulary to lend it a more “real” tone.

In the context of *Butterfly*, this linguistic background is important because the librettists lived through this linguistic turmoil. As writers, their linguistic heritage is the Florentine tradition, but as Italians, their heritage is a mix of dialects and ideologies. They wrote at a time when Florentine was expanding to fit the new nation and changing accordingly, but also when elements of dialect were welcome in drama via the movement towards realism.

³The idea of realism in literature, art, and drama (including opera) was that it should more closely mirror the experiences and stories of real people. Realists, including Puccini, sought to portray contemporary people with truth and accuracy instead of exaggerating or idealizing them.

3.1 A comparison to *La figlia di Iorio*

In this section, I will compare the syntactic peculiarities of *Madama Butterfly* to a 1904 play called *La figlia di Iorio* by Gabriele D'Annunzio.⁴ I find that noun phrases have similar word orders but the peculiar extractions from noun phrases in *Butterfly* do not extend to *La figlia*. In addition, D'Annunzio does not use as many overt subject pronouns in instances where there can be a covert pronominal.

D'Annunzio was a younger contemporary of Illica, Giacosa, and Puccini, and was hugely influential in Italian literature at the beginning of the 20th century. The reason I have chosen D'Annunzio from the many great writers of the time is that he was educated in Tuscany and Milan, which is where the *Butterfly* writers were educated, and would have a sense of the changing tides of Florentine. His work is more strictly metered than the libretto for *Butterfly* but contains some telling similarities and differences that are worth exploring. To begin, the word order in Noun Phrases is comparable to that in *Butterfly*. One example that occurs in both is the non-canonical word order of Noun Phrases with respect to the possessive pronoun. The following examples come from *La figlia di Iorio* and similar examples can be found in *Madama Butterfly*.

- (2) a. ...tu hai nel Paradiso **le nozze tue nuove** ...
...you have in.the paradise the wedding your new ...
'You (always) have your new wedding in paradise...'
- b. E screpolato t'è **il labbro tuo caro** dalla secchezza.
And cracked you.is the lip your dear from.the dryness
'And your dear lip is cracked from the dryness.'
- c. E sempre rinasca allo strazio **la carne sua maledetta!**
And always reborn.SBJV to.the torture the flesh her damned
'And may her damned flesh be always reborn to torture!'

As will be described in 5.3, both canonically-ordered and non-canonically-ordered noun phrases co-exist in both *La figlia* and *Butterfly*. However, they are marked in spoken Italian. The noun in spoken Italian always comes after the possessive pronoun, whereas in these examples, the noun comes before the possessive pronoun. From a brief survey of *La figlia*, it

⁴A project for future study would be to compare D'Annunzio's play to the operatic version of the same work, for which D'Annunzio wrote the libretto in 1906 and Alberto Franchetti wrote the music.

is notable that the examples are all inalienable nouns. *Nozze* 'wedding', *labbra* 'lip', and *carne* 'flesh' are all inextricable parts of a being. In both *Butterfly* and *La figlia*, *faccia* 'face' is often in this non-canonical order with respect to the possessive pronoun. More will be said on this in 5.3.

Another useful comparison to make is that the pronouns in use differ significantly. The strong, third person singular masculine subject pronoun in *La figlia* is always *egli*, which is supported in all of the stage directions. *Lui*, in contrast, is used as the accusative third person singular masculine object pronoun. This is not the case in **Butterfly**, where *lui* is present only twice. The following examples are from *La figlia*, showing the contrastive usage of *egli* and *lui* within the text of the play.

- (3) a. **Egli** è santo.
 He is saint
 'He is a saint.'
- b. E vivere con **lui** Mila non può!
 And live.INF with him Mila NEG can
 'And Mila cannot live with him!'

Likewise, in *La figlia d'Iorio* the strong, third person feminine pronoun is always *ella* and the accusative form is always *lei*. In *Butterfly*, *lei* is used only once and in the context of a pointedly polite second-person pronoun. The following example from *La figlia* shows the contrasting uses of *ella* versus *lei*.

- (4) a. **Ella** s'arresterà per qualche attimo...
 She stop.REFL.FUT for some moment
 'She will stop for a moment...'
- b. Meglio per **lei**, che ha perso conoscenza.
 Better for her who has lost consciousness
 'Better for her, who lost consciousness.'

Furthermore, there are fewer examples in the play of overt pronouns that do not classify as contrastive, new, or emphatic. That is to say that overt pronouns in *La figlia* follow the rules of spoken Italian, which is that overt pronouns constitute new, contrastive, or emphatic information.

Lastly, *Butterfly* makes use of a particular kind of Focus/Topic structure that does not appear in *La figlia*. This phenomenon, which will be described in 5.2, is not a part of the play at all and leads me to the idea that perhaps what can be found in *Butterfly* is unique to opera librettos.

4 Pronouns

A well-studied area of Italian syntax is the null subject parameter, whereby a subject is syntactically present but can be phonologically null or covert. In Italian librettos, subject pronouns are overt more often than in spoken Italian. To approach a theory of why this occurs in librettos, I will first outline the current thinking on null subjects in spoken Italian. Much literature has postulated that null subjects are allowed via the licensing role of a rich inflectional system (Chomsky 1993, Rizzi 1986, among others). According to this idea, Italian verbal morphology makes a six-way morphological distinction for all persons and number so it follows that a known subject would not need overt expression, as it could be inferred from the verbal morphology.

Most linguists assume that this syntactically-present but phonologically-null subject is a covert pronominal element (*pro*) that is in some kind of predictable distribution with an overt pronoun, and many linguists assume that this covert version has its own set of licensing restrictions. Frascarelli (2007) draws on Rizzi's 1986 theory of *pro* to outline the cases in which this null subject is licensed. She quotes the following from Rizzi (1986) (pg. 519-520, from Frascarelli 2007):

1. *pro* is governed by X°
2. let X° be the licensing Head of an occurrence of *pro*: then *pro* has the grammatical specification of the features on X° co-indexed with it. Frascarelli (2007)

What this means is that *pro* can occur if it has a governing head X° , and when it does, it takes on the features of X° . Rizzi assumes that the governing head X° is the head of AgrP, below what he called InflP which is analogous to TP (Rizzi 1986).

In slight contrast, Frascarelli (2007) argues that X° is also the head AgrP, but that for most cases where the subject is null AgrP is a projection in the C domain, responsible for agree relations and more generally responsible for discourse functions. Many prior schools of thought

held this pronominal element to be on par with overt pronouns but [Frascarelli \(2007\)](#) problematizes this assumption. *Pro* and overt pronouns occasionally have the same distribution, as shown in 5a. However, there are times when they are in different distribution. We can see this in 5b-d, where a pronoun in an embedded clause that is co-referential with a pronoun in the matrix clause cannot be overt, but can be null.

(5) a. **Lei_i*/**pro_i* andrà se Lucia_i troverà le chiavi.
 She will.go.3SG if Lucia will.find.3SG the keys
Intended: ‘She_i will go if Lucia_i find the keys.’

b. Lucia_i andrà se **lei_i*/*pro_i* troverà le chiavi.
 Lucia will.go.3SG if she finds the keys.
 Lucia_i will go if she_i finds the keys.

c. Se Lucia_i troverà le chiavi, **lei_i*/*pro_i* andrà.
 If Lucia will.find.3SG the keys, she will.go.3SG
 ‘If Lucia_i finds the keys, she_i will go.’

d. Se **lei_i*/*pro_i* troverà le chiavi, Lucia andrà.
 If will.find.3SG the keys, Lucia will.go.3SG

‘If she finds the keys, Lucia will go.’

The distribution is such that if a matrix and embedded clause have the same subject, only one can be overt. This is why *lei* is ungrammatical in all of the examples if it is meant to be co-referential with *Lucia*. In contrast, *pro* is a grammatical co-referent to *Lucia* but only when it is in the embedded clause.

[Frascarelli](#) posits that *pro* does not refer on its own and must be co-indexed with a Topic. [Frascarelli and Hinterhölzl \(2007\)](#) propose three kinds of topics, in a specific hierarchy, which is as follows.

(6) Aboutness-Shift Topic > Contrastive Topic > Familiar Topic

According to [Frascarelli and Hinterhölzl](#), the Aboutness-Shift Topic "has the discourse function of introducing a new topic (or proposing a topic-shift) in the discourse" ([Frascarelli \(2007\)](#), pg. 693). When there is a null subject in a matrix clause, [Frascarelli \(2007\)](#) assumes that the *pro* is in fact bound by a silent Topic, as in 7. This ambiguity would only arise in the case of third person pronouns because first and second involve the interlocutors, leaving little room for referential uncertainty.

- (7) Canta notte e giorno.
 Sings.3SG day and night
 ‘He/she/it sings day and night.’

In essence, *pro* can be bound by a Topic, or more specifically by an Aboutness-Shift Topic ([Frascarelli 2007](#)). This Topic can be a covert or overt full DP or a strong pronoun.⁵ Generally speaking, the assumption is that when a subject is phonologically present it is because it is new information, contrasting information, or emphasized information ([Frascarelli and Hinterhölzl 2007](#)).

Connected to the issue of null versus overt pronouns, [Chomsky \(1993\)](#) discusses a theory to explain the general minimization of structure in null subject languages. The thought is that there must be some rule governing the general silencing of pronominal subjects. [Chomsky \(1993\)](#) first proposed the principle, which he called the ‘Avoid Pronoun Principle’ 4.

- (8) *Avoid Pronoun Principle*
 Avoid pronouns, whenever possible.

The Avoid Pronoun Principle seems to capture the intuition about the systematic minimization of structure. If a pronoun does not convey new, emphasized, or contrasting information, then it is left null.

Returning to *Madama Butterfly*, we see that the libretto consistently breaks with the Avoid Pronoun Principle. There are many examples of overt subject pronouns even when they have the necessary conditions to be null, i.e. they have a co-indexed DP or strong pronoun in AgrP

⁵This assumes a three-way distinction of pronouns where weak and clitic pronouns do not have enough structure to license *pro*. The idea, according to [Cardinaletti and Starke \(1999\)](#), is that weak pronouns are a subset of strong pronouns in regards to the amount of projections they have and the number of features, and clitic pronouns are likewise a subset of weak pronouns. See [Cardinaletti and Starke \(1999\)](#) for a full typology of pronouns.

and are not contrasting, emphatic, or new. However, for reasons that I will explore throughout the rest of this section, subjects are often overt.

First, we must generally consider that the language in a libretto is subject to a musical meter. The words fit within the confines of the music, with each syllable mapping to a musical beat. Secondly, the libretto is a poetic form and was at least originally conceived with metered text and rhymes. Take the example from the Bonzo's speech in Act I.

- (9) ...hai tu gli occhi asciutti?
 Son dunque questi i frutti?
 Ci ha rinnegato tutti!

In 9 there are eight syllables per line, each with three strong syllables. The final words rhyme in all three lines. While not all lines in *Butterfly* are so carefully metered, the metrical and musical environments do occasionally require the addition of an extra syllable or beat in order to keep with the rhythm of previous or proceeding lines. In this case, a musical or metrical repair strategy is necessary. I propose that the overt realization of subject pronouns that could otherwise be null are often cases of this musical or metrical repair strategy within the libretto. The reason that I do not choose between either a musical repair or a metrical repair is that [Bonomi and Buroni \(2017\)](#) claim that Puccini was notorious for changing the librettists' carefully metered text to fit the meter of a musical line he had already written. When we find an unexpected or superfluous overt subject pronoun, then, we cannot unilaterally assume that it will always be due to musical meter or in contrast that it will always be due to poetic meter. In determining which rule describes the overt realization of pronouns that would otherwise be null, I will consider [10a](#) and [10b](#). My five hypotheses are reprinted below.

- (10) a. It is musically motivated.
 b. It is poetically motivated.
 c. It is dramatically (semantically) motivated.
 d. It is a syntactically unique trait that only occurs in opera librettos.
 e. It is a trait of early-20th century Italian.

I will begin looking at unexpected overt subject pronouns with the third person masculine subject. There are only three instances of the overt third person feminine subject, and none

of them represent an unusual instance of an overt pronoun. This is likely because the story focuses on Butterfly and her loyalty to Pinkerton; statistically, she and other female characters are not spoken about in the third person so there is not significant data on third person feminine pronouns. I will then look at second person pronouns and then briefly first person pronouns. I will show that 10c does not describe the pattern and 10e explains which pronouns were in use but does not provide enough backing to explain the pattern. 10d can be shown to be false by a simple comparison with *La figlia di Iorio*. This leaves the possibility of 10a, 10b and 10d.

4.1 Embedded Subjects: A Tangent

Something that has fascinated me is the seemingly privileged nature of subject pronouns in embedded clauses. They tend to be overt more often than matrix subjects. Overt embedded subjects occur systematically in an informal register of English as well. English can have null subjects in a sort of 'diary speak' (Radford 2004).

- (11) a. I can't find my pen.
b. Can't find my pen.
- (12) a. I think I left it at home.
b. Think I left it at home.
c. *Think left it at home.
d. *I think left it at home.

The null matrix subjects can be dropped to achieve the diary speak, but embedded subjects cannot be dropped for most dialects (Haegeman 2013). I bring this up because I have noticed that in the libretto, it seems that subjects are most often overt in embedded clauses. I will touch on this again in 4.3 because in particular, the first person pronoun *io* seems to be overt most often in embedded clauses.

4.2 Egli

The subject pronoun *egli* is the third person singular nominative pronoun from Old Florentine, the ancestor of what we consider Modern Italian (Poletto 2014). This form is no longer present in Modern Italian except possibly in a formal written register. Modern Italian would use the

pronoun *lui*. In the period between the unification of Italy, in 1861, to the outbreak of the First World War, recall from 3 that Italian went through a standardizing process. According to [Migliorini and Griffith \(1984\)](#), *ei*, *eglino*, and *elleno* had become rare as subject pronouns, and *egli* was more literary, meaning it was present in prose but less so in spoken Italian.

What is more, [Cardinaletti and Starke \(1999\)](#) claim that *egli*, *ella*, *esso*, and *essa* belong to the weak class of pronouns in Modern Italian, which means that they are unable to co-refer with a *pro*. According to [Cardinaletti and Starke \(1999\)](#) and [Renzi \(2001\)](#), what distinguishes a strong from a weak pronoun is the number of features it carries, its animacy, whether or not it can coordinate with a lexical Noun Phrase, and, according to [Frascarelli \(2007\)](#), the prosody of the pronoun. In terms of the amount of structure it carries, [Cardinaletti and Starke \(1999\)](#) argue that weak pronouns are a structurally impoverished subset of strong pronouns. The prosody distinction that [Frascarelli \(2007\)](#) points out is hard to speak to in the libretto because each note has a prescribed pitch, but if I assume that the musical line is in fact the ‘natural’ prosody of the opera’s language, then I could perhaps come up with rules for prosody and overt pronouns in *Madama Butterfly*.

It is clear from the libretto that Illica and Giacosa use *egli* as the strong third person singular male pronoun. *Lui*, which is the only modern third person singular masculine pronoun, is only used twice in the libretto. Furthermore, *esso* and *essa* occur as weak pronouns in the libretto. With the role of weak pronouns already filled, the only other option for *egli* is to be the strong subject pronoun. There is not enough data to show that *egli* can coordinate with a lexical Noun Phrase, but it is evident that at least the animacy and features restrictions show *egli* to be the strong pronoun throughout the opera.

More generally, it seems from looking at *La figlia di Iorio* and other prose from the time that subject pronouns were in a degree of flux and often the choice fell to an author’s personal style. Evidence from a letter from the librettist Illica to the editor, Giulio Ricordi, shows that at least Illica used *egli* as a subject pronoun.

Puccini ha confidato ad un amico suo che de’ miei libretti ne fa anche senza [...] e che *del resto* nessuno sa capirlo, perché **egli** vagheggia una cosa...

‘Puccini confided to one of his friends that he writes even without my librettos and after all no one understands him, because he yearns for something...’

[Bonomi and Buroni \(2017\)](#)

This letter brings up two points worth discussing. First, the clause *perché egli vagheggia una cosa* 'He yearns for something' would be eligible for a null subject. There is a full DP topic present: *Puccini*. The pronoun is not contrastive or new information. The only possibility is that *egli* is emphatic in this clause, but this is difficult to distinguish in written language because emphasis is often relayed through prosody (Frascarelli 2007). This particular property will be easier to distinguish in a musical context when musical pitch can imitate prosody. Secondly, it is worth noting that in the Northern Dialect Continuum, the frequency of overt subject pronouns varies by region, with the highest percentage of usage in Bologna and a medium amount of usage in Milan (Heap 1997). Illica was born in Castell'Arquato, which is geographically between Milan and Bologna and he was schooled in Milan. I will not make any conjectures regarding Illica's linguistic background here, but it is again worthwhile to know that he used *egli* in his written language.

I will here discount my hypothesis that the unexpected presence of overt pronouns is a phenomenon syntactically unique to opera. In *La figlia*, which is distinctly not an opera, there are many overt pronouns that could in spoken Italian be covert. Take this instance.

- (13) *Aligi figliuolo di Lazaro | è innocente. Ma egli non sa.*
 Aligi son of Lazaro | is innocent. But he NEG know.3SG
 'Aligi, the son of Lazaro, is innocent. But he doesn't know.'

In this instance, the subject of the second sentence would be eligible for a null pronoun because it has a clearly defined antecedent, *Aligi*. However, *La figlia* is a strongly metered text: there are ten syllables per line consistently throughout the text. This passage follows that same meter: the two lines (separated by |) are both ten syllables in length. A covert pronoun would change the meter, reducing the second line to eight syllables. As such, this is strong evidence of the overt realization of a pronoun acting as a metrical repair strategy.

Returning to *Butterfly*, there are 17 examples of *egli* as a subject pronoun⁶. Of these 17 examples, *egli* is the subject of a matrix clause 10 times and the subject of an embedded clause the other 7 times. There is one instance of *ei* used as a possessive pronoun (*la vita ei*, 'his life').

I will now return to the background literature on null subjects to explore the relationship between the overt subject pronouns in the libretto and the research on spoken Italian. I will begin with a look at one particular instance of the matrix subject pronoun from Act I.

⁶I count *ei* as the shortened form of *egli*, so the distinction between them is merely phonological and not syntactic

- (14) **Egli** è bel, mi pare un re.
 He is.3.SG. beautiful, me appears a king
 ‘He is beautiful, he looks to me like a king.’ Act I

Egli occurs as the subject pronoun of a tensed phrase, with a clear co-reference established in the context of the whole first act. This co-referent is Pinkerton, whom the chorus sings about for approximately ten minutes as they evaluate what they think of him. *Egli* is not a Focus because it is not new material or *wh*-material. Likewise, it is not an Aboutness-Shift Topic as per Frascarelli (2007) or a contrastive Topic because the discourse of the scene has already established Pinkerton as the subject. The question, then, is whether this constitutes a unique syntactic rule of opera or if this is a musical or metrical repair strategy.

There is evidence that this overt subject is a case of metrical repair. This phrase musically mirrors a subsequent phrase, in 15.

- (15) Ma risposi non lo vo’
 But responded.1SG. NEG it want.1.SG.
 ‘But I responded I don’t want to.’ Act I



Figure 1: ‘Ma risposi non lo vo’...’

The syllable structures of 15 and 14 contain four trochees, which are strong beats followed by weak beats. This is supported by the distribution of the text along four strong beats in the music. In other instances, the pronoun is omitted as is normal, 16.

- (16) Bello è così che non si può sognar ...
 Beautiful is.3.SG. like-this that NEG CL can dream
 ‘It’s so beautiful like this that one couldn’t dream it ...’

At least preliminarily, we can say that this is an instance of the need for metric alignment in the music and poetry. This need overrides the natural language rule that would call for a null subject.

Within the same scene, the female members of the chorus say that they had been offered Pinkerton’s hand but refused it. They sing the following line, which has the same syllable and musical structure as 14 and 15:

- (17) **Ei** l'offrì pur anco a me!
 He it-offered rather also to me
 'He also offered his hand to me!'



Figure 2: 'Ei l'offrì pur anco a me...'

Ei clearly refers to Pinkerton, and the pronoun is not contrastive or new information. There has already been a pronoun referring to Pinkerton with *egli* in 14, so any doubt as to the subject of the clause should be gone. It is most likely not intended emphatically because the pronoun follows on the second beat, the weakest beat of the musical measure.

This line is meant to rhyme with 14, so that it has four strong beats. Musically and metrically, the overt pronoun repairs the line so that it does in fact rhyme and fit within the meter.

All of this points towards a particular trait of operatic syntax. There are still several examples of overt subject pronouns that follow the patterns of natural spoken Italian, where the pronoun is contrastive, new, or emphatic. 18 is one such example. Sharpless returns to Japan and comes to see Butterfly. He realizes she has had a son by Pinkerton, and points to the son saying:

- (18) **Egli** è suo?
 He is his
 'He is his?'

The pronoun is presenting pragmatically new information, so this constitutes an example of an overt pronoun that would also be overt in spoken Italian. The point is that in addition to allowing overt pronouns in the same contexts as spoken Italian, the libretto allows these non-contrastive, non-emphatic, and known subjects to be overt. This constitutes a trait of the libretto, not present in spoken Italian. It is a sort of caveat to the Avoid Pronoun Principle. It can still be said that pronouns are avoided when possible, but there are certain musical and metrical environments that require the insertion of a pronoun. As such, overt subject pronouns are not only controlled by information structure, but also by the music and the poetry. We can now return to 5, where we saw the difference in distribution between *pro* and pronouns, reprinted here as 19.

- (19) a. **Lei_i*/**pro_i* andrà se Lucia_i troverà le chiavi.
 She will.go.3SG if Lucia will.find.3SG the keys
Intended: ‘She_i will go if Lucia_i find the keys.’
- b. Lucia_i andrà se **lei_i*/*pro_i* troverà le chiavi.
 Lucia will.go.3SG if she finds the keys.
 Lucia_i will go if she_i finds the keys.
- c. Se Lucia_i troverà le chiavi, **lei_i*/*pro_i* andrà.
 If Lucia will.find.3SG the keys, she will.go.3SG
 ‘If Lucia_i finds the keys, she_i will go.’

Above, *pro* is allowed in more contexts than the pronoun *lei*. In an embedded context where the embedded subject co-refers to the matrix subject, *pro* is always licensed while *lei* is not. In a matrix subject, *pro* is not allowed if it is intended to co-refer to the embedded subject because there is no co-referential TopP above the *pro* to impart its features to *pro* (as per [Frascarelli \(2007\)](#)).

There is not enough data from the libretto of *Butterfly* to determine if *pro* and subject pronouns have a different distribution from spoken Italian.

As I mentioned, *egli* appears 7 times as the subject of an embedded clause, both in subjunctive clauses, [21](#), and indicative clauses, [20](#).

- (20) S’egli non torna e presto, siamo male in arnese.
 If-he NEG returns and soon, are.3PL bad in attire.
 ‘If he doesn’t return soon, we are in a bad state.’
- (21) Ma temo ch’egli ignori.
 But fear.1SG that-he ignores.SBJV
 I fear that he ignores us.

In [21](#) it could be argued that the pronoun is necessary to disambiguate the subject because the inflectional morphology of a subjunctive clause does not distinguish between persons when the subject is singular. This is because, for first, second, and third person singular subjects, the verb has the same ending. While some might claim that the lack of differing morphology forces the presence of a pronoun, there are instances of an embedded subjunctive clause with no pronoun, as in [22](#) and [23](#).

(22) Non c'è vagabondo che a sentirlo non sia di gran prosapia.
 NEG there-is vagrant that to hear-it NEG be.3SG.SBJV of great lineage
 'There isn't a vagrant that won't tell you they're of great lineage.'

(23) Vo' che mi veda indosso il vel del primo dì.
 Want.1SG that me see.3SG.SBJV dressed the veil of.the first day
 'I want him to see me dressed in the veil from our first day.'

As such, we need a different or more refined theory for overt subjects in embedded clauses. Before concluding, I will look at the first person singular overt subject pronouns in *Butterfly*, which is often overt in embedded clauses. I will then note the second person subject pronouns, although they do not fit with the previous discussion for various reasons.

4.3 Io

In *Butterfly*, *io* is the most commonly used overt pronoun, with twenty-eight instances of overt realization in the libretto. It occurs six times as the subject of embedded clauses and the other 22 times as the subject of the matrix clause. The proportionally high usage of *io* is likely due to the nature of the narrative, which is told through the first person for all of the characters.

As with *egli*, there are instances of an overt *io* that come with the subjunctive form. I have shown that for *egli*, it is not enough to say that the lack of inflectional distinctions between persons accounts for the overtness of the pronoun. Rather, there is another hypothesis at work for these cases.

(24) E volete ch'io chieda ad una madre...
 And want.2PL that-I ask.1SG.SBJV to a mother
 'And you want that I ask a mother...'

(25) Dammi ch'io baci le tue mani care.
 give.me that-I kiss.1SG.SBJV the your hands dear
 'Give me your dear hands so I can kiss them.'

(26) Reggimi la mano ch'io ne discerna il nome.
 Hold.out.me the hand that-I of.it discern.1SG.SBJV the name
 'Hold out your hand to me so that I can discern the name.'

(27) Ma bisogna ch'io le sia sola accanto...
 but is.necessary that-I her be.1SG.SBJV next.to
 'But it's necessary that I be alone with her...'

(28) Non c'è gran male s'io vo' quell'ale drizzare ...
 NEG there-is great harm if-I want those straighten.INF
 'It's no great harm if I guide those wings ...'

As I mentioned in 4.1, embedded subjects seem to have a privileged status. They are the most frequently overt subjects in the libretto at least. Poletto (2014) points out a similar phenomenon in Old Italian, in which embedded subjects seem to be most often overt, although she does not offer an analysis. Examples 25 and 26 are an imperative structure in Italian, in which the embedded subject is typically overt.⁷ However, for the other examples, I do not have sufficient evidence to give a compelling analysis of why they are overt.

I will first explore whether the music or meter necessitates them being overt. Beginning with 24, we see a distinct lack of musical repair.



Figure 3: 'E volete ch'io chiedo ad una madre...'

In this instance, the overt pronoun is on the tail end of a weak beat and elided with the complementizer *che*, so it is not at all placed in an accented or focused rhythmic position. It is also set on a musically unimportant note. As such, there is not musical repair at work here. Looking at other examples, we see the same thing.



Figure 4: 'Reggimi la mano ch'io ne discerna il nome.'

The pronoun *io* in the embedded clause, if anything, impedes the metrics of the phrase by adding an extra syllable on a very short note. Musically, it is neither emphasized nor treated

⁷This kind of imperative can be found in spoken varieties of Italian with fairly high frequency. According to native speaker judgements, the acceptability of the sentence decreases if the embedded subject is null.

with any sort of special consideration. It seems that both in the poetry of the phrase and the musical line, the pronoun does not add anything. Syntactically, the only possible role it has is to disambiguate the subject of the embedded clause due to the ambiguous subjunctive morphology.

Likewise, the instances of overt pronouns in non-subjunctive clauses are not metrical or musical repairs, as in the example below, where the verb *volere* is already conjugated for the first person singular, albeit in a common abbreviation, so the overt subject is not necessary for morphological disambiguation.



Figure 5: ‘Non c’è gran male s’io vo quell’ale...’

There is no special musical treatment for the overt pronoun; it is written on a single beat with the complementizer *se* (abbreviated to *s’*). As before, the overt pronoun does not help the meter or the music. My remaining hypothesis is that it could be explained via syntactic theory.

Returning to 4.1, in English, the Diary Speak register has been well-researched by Liliane Haegeman. For most dialects of Diary Speak, the dropped subject is a root phenomenon, so embedded subjects remain overt. The idea that Haegeman (2019) proposes, which she calls the Truncation Hypothesis, is that the availability of a dropped subject depends on the ability of the clause to terminate at the level of SubjP. Essentially, if there is nothing above SubjP, then one can omit everything from SubjP up in the clause. While this works neatly for this register of English, there is no analogy between the dropped subject of Diary Speak and null-subject languages like Italian. As pointed out by Samek-Lodovici (1996), they do not have the same distribution. The dropped subject in English can co-refer to an agentive by-phrase but *pro* in Italian cannot co-refer to an agentive by-phrase. The following is from Samek-Lodovici (1996).

- (29) a. Questa mattina, la mostra è stata visitata **da Gianni**_i.
 This morning the show is been visited by Gianni
 ‘The morning, the show was visited by Gianni.’
- b. Più tardi, lui_i/**pro*_i ha visitato l’università.
 More later he has visited the-university

‘Later, he (Gianni) visited the university.’

In addition, Diary Speak is a root phenomenon, with the exception of one British English dialect⁸. As we have seen in the libretto, embedded subjects are often covert, which can be explained by the Avoid Pronoun Principle. In searching for an explanation as to why they are overt, the research on Diary Speak is not helpful. The basic issue is that Haegeman seeks to explain the covertness of the root subject while I seek to explain the overtiness of an embedded subject. I believe, however, that future study could unite these phenomena.

4.4 Second Person Pronouns

This subsection has a slightly different aim than the previous subsections. As an English-speaker, I have always been fascinated with the politics of formal grammatical ‘you’ vs informal ‘you’. In Italian, the present-day formal ‘you’ is *Lei* and it is conjugated the same way as singular third person. At least before the 19th century, the polite ‘you’ was consistently *Voi* and was conjugated in the same way as the second person plural. There is admittedly little formal research on the transition from *voi* to *lei* as a second person formal subject pronoun. However, an informal publication by a grammarian of the 1920s shows that *lei* and *voi* coexisted as subject pronouns during the time period (Roselli 1926) with subtle social variance.

Throughout *Madama Butterfly*, *voi* is the preferred polite pronoun. Take this interaction from after their wedding ceremony.

(30) **Butterfly:**

Adesso **voi** siete per me l’occhio del firmamento.
Now you.POL are.2PL for me the.eye of.the heavens

‘Now you are for me the eye of the heavens.’

Even before this, Pinkerton speaks to Butterfly in the form of *tu*, but she continues to address him using the polite *voi*. The imbalance is pointed and pragmatically important as it is meant to solidify the power imbalance between Butterfly and Pinkerton.

(31) **Pinkerton:**

⁸See Haegeman and Ihsane (2001) for an analysis of this dialect. They do not provide a definitive analysis of why this occurs but posit that the dropped embedded subjects are instances of pronoun ellipsis. This theory would not help to explain why pronouns are overt in Italian embedded clauses, since the pronominals are already covert in Italian unless contrastive, emphatic, or new.

Sei tutta vestita di giglio.
COP.2SG all dressed of lily

‘You are all dressed in white.’

There is one pointed instance of *lei*, which comes in the second act. Butterfly is imitating what she imagines to be an American judge, and in her imagined dialogue, the judge uses *lei* instead of *voi*.

- (32) un bravo giudice ...dice al marito: ‘**Lei** vuol andarsene? Sentiam perché?’
A skilled judge ...says to.the husband: ‘You want leave.INF? Hear.1.PL. why?’
‘A skilled judge says to the husband: ‘You want to leave? Can we hear why?’

This pointed use of *lei* does not have to do with any kind of musical repair. It is a curious cultural commentary, for which no formal research exists. The only thing I can extrapolate from this is that perhaps *lei* was a more modern innovation at the turn of the 20th century, and to imitate the modernity of the American legal system compared to the Japanese system, Butterfly uses a different pronoun.

There are few examples in the libretto of *tu* being overt when it could be covert. Most of the examples, such as the one below, are in fact contrastive or emphatic.

- (33) Tu sei con Dio ed io col mio dolor.
You are with God and I with.the my pain
‘You are with God and I with my pain.’

There is a similar distribution with *voi*. Most of the cases where *voi* is overt, it is syntactically or semantically necessary. As such, I have no data to offer from second person pronouns on the curious debate of why pronouns are more often overt in librettos.

4.5 Conclusion on Pronouns

The ability of pronouns to become overt in matrix clauses when the music or meter requires it constitutes a trait of the language of the libretto, not present in spoken Italian. It is a sort of caveat to the Avoid Pronoun Principle. It can still be said that pronouns are avoided when possible, but there are certain musical and metrical environments that require the insertion of a pronoun. As such, overt subject pronouns are not only controlled by information structure, but also by the music and the poetry.

With regards to overt embedded subjects, a slightly different theory is in order. Any theory of overt pronouns in embedded clauses would have to capture the fact that they seem to be entirely optional within this form of Italian. They do not necessarily syntactically disambiguate the person, as we can see by their absence in many ambiguous embedded clauses. They also do not fix the meter of the poetry nor receive musical treatment. What's more, they are not contrastive, emphatic, or new, so they do not follow the criterion of overt pronouns in spoken Italian. As far as I can tell, it seems to be purely volitional when they occur or do not occur. This reveals something about overt pronouns in this seemingly null subject language; at least in librettos, embedded subjects are not subject to the Avoid Pronoun Principle and do not require any kind of emphatic, new, or contrastive feature in the phrase to be overtly realized. When they do or do not occur is subject to some other realm, not explicable by pure syntax, musical meter, poetic meter, or semantics.

Linguistically, the overt realization of the pronoun for poetic/dramatic purposes contains an insight into the nature of pronouns. In summary, the Avoid Pronoun Principle, which governs spoken Italian, seems to be mostly present in this form of Italian but has an additional environment in which pronouns must be overt. This additional environment is the music or meter, which utilizes the added syllables of a pronoun to repair. As for the overt realization of pronouns in embedded clauses, future study must be done to better understand their distribution and to develop a theory for their existence.

5 Noun Phrases

There have been several studies of Italian Noun Phrases, most famously by Guglielmo Cinque and Giuliana Giusti. Generally speaking, these studies have posited that Italian noun phrases are more flexible with regards to movement than English noun phrases. This section will draw heavily from Cinque's (1980) and also from Delmonte's (2018) analysis of early 20th-century Italian poetry. I will present data on two types of noun phrases found in opera that are marked or ungrammatical in modern spoken Italian. I will argue that noun phrases in Italian opera librettos have different movement and ordering rules than spoken Italian, and that this is primarily syntactically- and semantically-motivated rather than musically or metrically required. The word order of noun phrases could be an exception, however, in that the re-ordered noun phrases receive special musical treatment, which I will discuss in 5.3.

5.1 Hyperbaton and Italian opera DPs

Delmonte (2018), investigating early 20th century Italian poetry, points out several cases of a construction called *hyperbaton*, which he defines as follows:

Hyperbaton: Focusing by dislocation of [a] portion of a constituent, typically an adjective from a noun phrase or a main verb from a verbal complex. (Delmonte (2018), pg. 37)

Delmonte (2018) claims that the hyperbaton in the early 20th century texts is directly inherited from written Latin, which makes robust use of the phenomenon. According to this definition of hyperbaton, the *Butterfly* libretto makes robust use of the phenomenon. I will extend research from Cinque and Giusti to this phenomenon, bringing Delmonte's work into the world of Minimalism. The results of this analysis show two things: 1) the dislocation of constituents is a Topic and Focus phenomenon, and 2) the phenomenon of hyperbaton can be accounted for by the rules that govern movement in Italian syntax.

Taking a look at the evidence from *Madama Butterfly*, we can divide the instances of Delmonte's *hyperbaton* into two groups. The first, most common type is the separation of possessive or subject prepositional phrases from head nouns. The second is the relative reordering of adjectives, possessive pronouns, and nouns within the NP. 5.2 will look at the former and 5.3 will look at the later.

5.2 Movement out of opera DPs

This first subsection will deal with articulated noun phrases, which are noun phrases with a head noun and a possessive/subject PP of the type [*di* NP]. The examples here from *Butterfly* contain an articulated noun phrase with its PP displaced from the canonical position. All of the following examples are taken from various parts of the libretto, with a gloss of the complete, canonically-ordered noun phrase below.

- (34) Questa è la cameriera che **della vostra sposa** fu già **serva amorosa**
This is the maid that of.the your wife was already servant lovely
'This is the maid that was already your wife's lovely servant.'

Complete noun phrase: serva amorosa della vostra sposa

- (35) Già **del femmineo sciame** qual di vento in fogliame s'ode **il brusìo**
 Already of.the feminine swarm like of wind in branches one.hears the bustle
 'Already one hears the bustle of the feminine swarm, like branches in the wind.'

*Complete noun phrase: **il brusìo del femmineo sciame***

- (36) ...guarda ben fiso, fiso **di tua madre la faccia!**
 watch.SG.IMP. well fixed, fixed of your mother the face
 'Watch closely your mother's face!'

*Complete noun phrase: **la faccia di tua madre***

- (37) **Di sua voce il mistero** l'anima mi colpì
 Of her voice the mystery the.soul me hit
 'The mystery of her voice hit my very soul.'

*Complete noun phrase: **il mistero di sua voce***

In modern spoken Italian, this word order is marked or ungrammatical. This phenomenon seems to violate rules of extraction from noun phrases. In spoken English, as well, this kind of extraction is extremely marked.

- (38) I can already hear **the bustle of the women**.

- (39) *Already **of the women**, I can hear **the bustle**.

According to the Principle of Subjacency, which was suggested in the 1980s as a means of explaining the constraints of movement, movement can only occur across one cyclic node in a language like English. These cyclic nodes in this theory are DP and CP. Italian, however, is thought to have fewer locality constraints on movement; movement can occur across both a DP and a CP without violating Subjacency. More contemporary accounts of movement constraints are not based on the idea of Subjacency (which is within Government and Binding Theory), but rather the Phase Impenetrability Condition (consistent with Minimalist Theory), by which only the highest head and the phrase in its specifier may extract from a phase (Radford 2004). Cinque (2014) contributes to the discussion with the intuition that the highest SpecDP is an A-position as opposed to an A'-position, analogous to SpecCP in the sense that it can serve as an escape hatch for movement. Thus anything moving out of a DP must first move through the SpecDP. Furthermore, movement out of a DP is acceptable only under certain conditions (Cinque 1980). Firstly, only a PP of the type [*di* NP] is acceptable

for movement. Secondly, Cinque (1980) posits that only possessive PPs⁹ or subject PPs are eligible for movement. To better understand what a subject PP is, take the following example. 40 is ambiguous between two readings, the first of which is an object-reading and the second of which is a subject-reading.

- (40) a. Il desiderio di te
 the desire of you
 ‘The desire of you’ *Somebody’s desire of you*
- b. Il desiderio di te
 the desire of you
 ‘Your desire’

In the object reading, the NP within the PP is the structural object of the head noun, so in Cinque’s words, ‘Somebody desires you’ 1980. In the subject reading, the NP within the PP is the structural subject of the head noun, so ‘You desire something/somebody.’ It depends entirely on the noun whether it can take an object PP, a subject PP, or both.¹⁰ Cinque (1980) divides Italian nouns into seven classes based on their argument structure. Event or action nominalizations, which denote an event or action in some sense, make up classes I through IV. Then ‘object nominalizations’ and ‘agent nominalizations’ make up classes V and VI and tangible nominalizations, which bare no relation to other categories such as verbs or adjectives, are class VII (such as *appartamento* ‘apartment’). The exact argument structure of each individual class is not relevant for my point here. What the reader needs to know is that there are certain classes of nouns, such as *desiderio* ‘desire,’ that can take both object or subject PPs. Extraction of the PP can only occur with the subject-reading of the PP, so only the PP in 40b can move (Cinque 1980). Likewise, any PP expressing a possessive relationship (which Cinque (1980) argues is in fact a subject PP) can move. In the following example, the PP introduces the possessor of the head noun and so is eligible for movement out of the complete noun phrase.

⁹Cinque (1980) in fact cites possessive PPs, in which the NP of the *di NP* is the possessor of the head NP, to be subject PPs.

¹⁰See the first chapter of Wood (2019) for a comprehensive summary of the various approaches to nominalizations and the derivation of argument structure.

- (41) L'appartamento di Musetta
 the-apartment of Musetta
 'Musetta's apartment'

I will use subject PP and possessive PP somewhat interchangeably, in keeping with Cinque (1980).

The third, most important consideration for movement in spoken Italian is that any extracted PP must contain *wh*-material, as in 42.

- (42) Una persona [_{PP_i} **di cui**] apprezziamo **la grande generosità** *t_i* è Giorgio.
 A person [of whom] appreciate.1PL the great generosity *t* is Giorgio
 'A person whose great generosity we appreciate is Giorgio.'
 (Cinque 1980:47)

Above, the PP *di cui* moves through the cyclic node SpecDP to the SpecCP of the clause. SpecDP is not a landing site, however, because, according to Giusti (1996), a SpecDP cannot check the *wh*-feature. In order to get this feature checked, it moves to SpecCP, as in 43. This is why it does not stop in SpecDP.

- (43) Non so [_{PP_i} **di che associazione**] Gianni sia **il presidente** *t_i*.
 NEG know.1SG [of which association] Gianni is.3SG.SBJV the president
 'I don't know which organization Gianni is the president of.'
 (Giusti 1996:107)

In short, the extraction rules of Italian DPs are such that a subject PP of the type [*di* NP] can move across the cyclic nodes DP and CP. They must have both an empty landing site, which would be an empty SpecDP or SpecCP, and also *wh*-material, which motivates the movement in the first place.

The difference between these examples from the linguistics literature and the examples from the libretto are that the latter do not have *wh*-material and yet the PP can still move across these phrase boundaries. There does not seem to be a *wh*-feature on the C° that needs to be checked in SpecCP. Some other feature must be driving movement. Otherwise, though, the pattern seems to be generally similar. Interestingly, this kind of movement occurs very widely in Italian librettos spanning centuries. Take these two examples from *Doriclea*, a 17th century Venetian opera with music by Francesco Cavalli and libretto by Giovanni Faustini.

- (44) Dell'innocenza mia senti le grida
 Of.the-innocence mine hear.SG.IMP the cries

‘Hear the cries of my innocence’

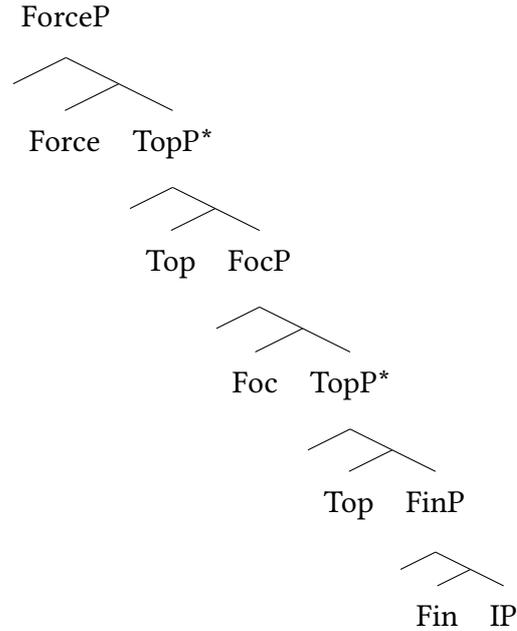
- (45) *Pria delle sfere arresterassi il moto*
Before of.the spheres halt.SG.SBJV the motion
‘Before the motion of the spheres halts’

Here, as in *Butterfly*, a subject PP without *wh*-material moves out of its DP to SpecCP. Working from Cinque (1980), I claim that this movement is not in fact a violation of Subjacency or the Phase Impenetrability Condition but rather it can be explained via the more generous rules of extraction in Italian DPs along with Italian’s robust left periphery. The significant difference between the librettos and spoken Italian is the absence of *wh*-features on the PP that moves. I posit the movement is motivated by some other feature in the left periphery. Working off of Rizzi’s (1997) seminal work, I will now look at where exactly within the CP layers, otherwise known as the Left Periphery, this extracted PP lands and also what might motivate this movement.

Within a different theoretical framework¹¹, Delmonte (2018) proposes that this phenomenon, which he terms hyperbaton, is a topicalization structure that refers back to Latin rhetoric, in which these kinds of exceptional structures are common. The underlying thought is that the syntactic irregularity is in fact part and parcel of the information structure, so when these phenomena occur they are entirely semantically and pragmatically motivated. Any noun or noun complement moved out of its phrase is a result of a Topic or Focus movement. This is in keeping with the idea proposed by Rizzi (1997), which is that the domain of the CP is responsible for the interface between syntax and pragmatics. He proposes that the CP layer is in fact split into several projections with specific functions. His full proposed split CP is reprinted below in 46.

¹¹Delmonte (2018) works within the theory of Lexical Functional Grammar (LFG) which builds off of psycholinguistic models to represent language based on the lexicon, with no transformations taking place.

(46)

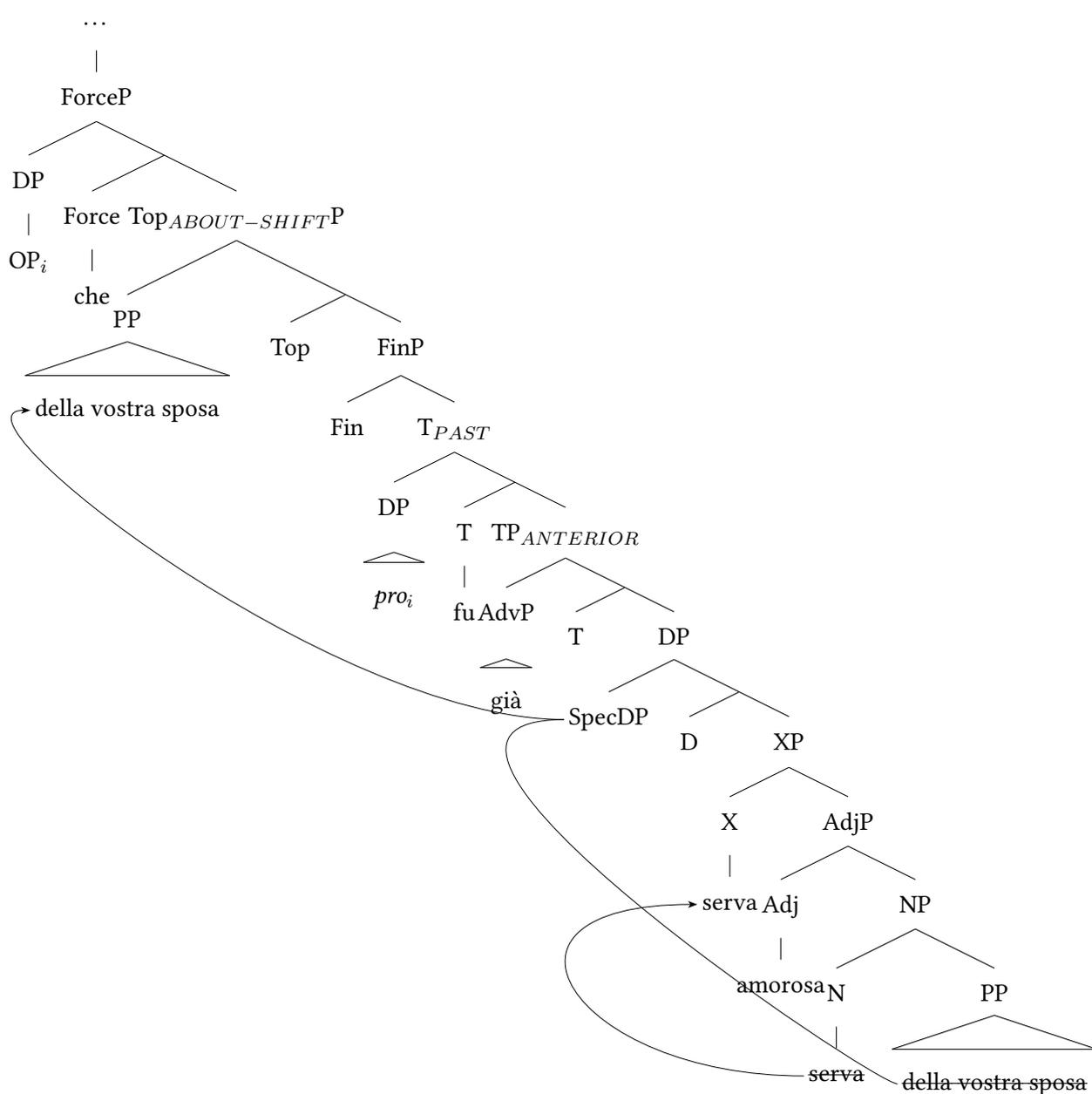


ForceP defines the clause type, TopP introduces a Topic, which is defined as old or known information, FocP introduces new information and *wh*-phrases, and FinP marks the finiteness of the clause. IP here is analogous to TP. There can only be one Focus ((Rizzi 1997)), but there are several types of Topics ¹² ((Frascarelli and Hinterhölzl 2007)).

The motivating feature for movement would be [+Topic] or [+Focus]. As we will see, the PPs that move are unusual candidates for Topic or Focus, but I will argue that this is due to a wider use of FocP and TopP in the language of the libretto. Combining Rizzi (1997) and Delmonte (2018), I propose that for 35 and 34, the explanation for the non-canonical ordering is a TopP in the embedded clause to which the possessive PPs can move. My proposed structure, based on a combination of Rizzi (1997) and Cinque (1999 and 1995) is in 47, with the full canonical ordering in 48.

¹²see Frascarelli and Hinterhölzl (2007) for a full typology of Topics in Italian and German.

(47)



(48) Questa è la cameriera che **della vostra sposa** fu già **serva amorosa**
 This is the maid that of.the your wife was already servant lovely
 ‘This is the maid that was already your wife’s lovely servant.’

Canonical sentence: **Questa è la cameriera che fu già serva amorosa della vostra sposa**

Che realizes the head of ForceP and takes TopP as its complement. Since this is the highest TopP, it is an Aboutness-Shift Topic with a [+aboutness] feature on the head (Frascarelli and Hinterhölzl 2007). The specifier of TopP hosts the topicalized PP *della vostra sposa*. This PP originates in the complement to the noun *serva*, and then raises across the DP layer, stopping in SpecDP before raising to SpecTopP. The NP remains *in-situ*, deriving the surface structure of 48.

One of the important features in distinguishing kinds of Topics and Focuses according to [Frascarelli and Hinterhölzl \(2007\)](#) is prosody. While the musical setting does not perfectly mirror spoken prosody, I will describe the language of the opera as though the prosody is the prescribed notes of the composer. With this in mind, there is additional evidence that *della vostra sposa* ‘of your wife’ has a [+aboutness] feature. [Frascarelli and Hinterhölzl \(2007\)](#) posit that an aboutness-shift triggers low to high intonation, which is supported by Puccini’s musical line.



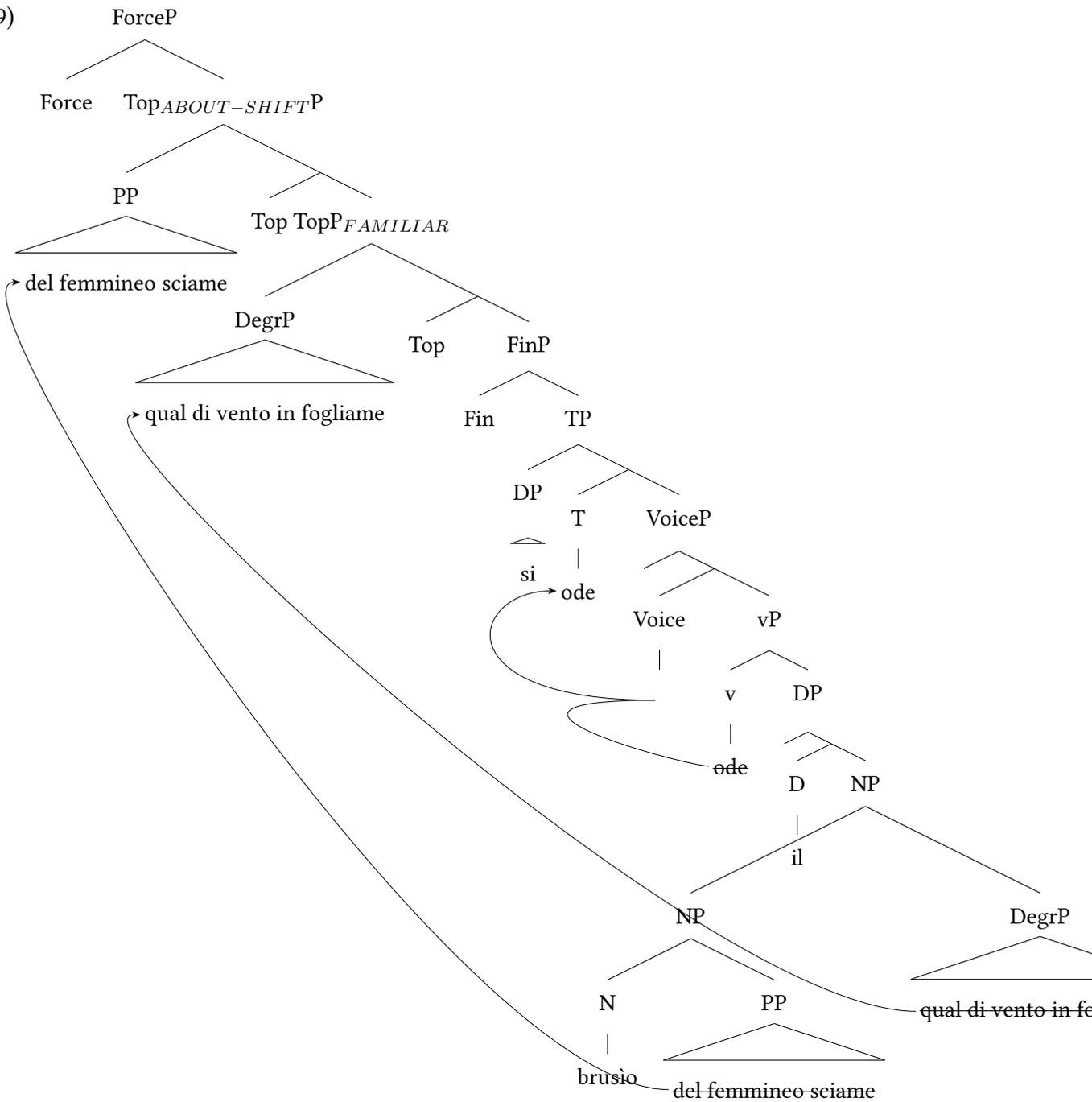
Figure 6: ‘Questa è la cameriera che della vostra sposa fu già serva amorosa.’

The intonation rises throughout the embedded clause, beginning with *della* ‘of the’ and peaking on the word *sposa* ‘wife’. Although it is an imperfect system, it at least another data point to show that the PP *della vostra sposa* is an Aboutness-Shift Topic.

For 35, I propose a similar structure to the one I have proposed for 48. The sentence is reprinted below tree 49 as example 50.¹³

¹³Note that in the tree, I have omitted the adverb *già* ‘already’ for brevity’s sake. I assume that it originates as a modifier of the vP *ode*. It would then move to some projection higher than the highest TopP.

(49)



- (50) Già **del femmineo sciame** qual di vento in fogliame s'ode **il brusio**
Already of.the feminine swarm like of wind in branches one.hears the bustle
'Already one hears the the bustle of the feminine swarm, like branches in the wind.'
Canonical sentence: S'ode già il brusio del femmineo sciame, qual di vento in fogliame.

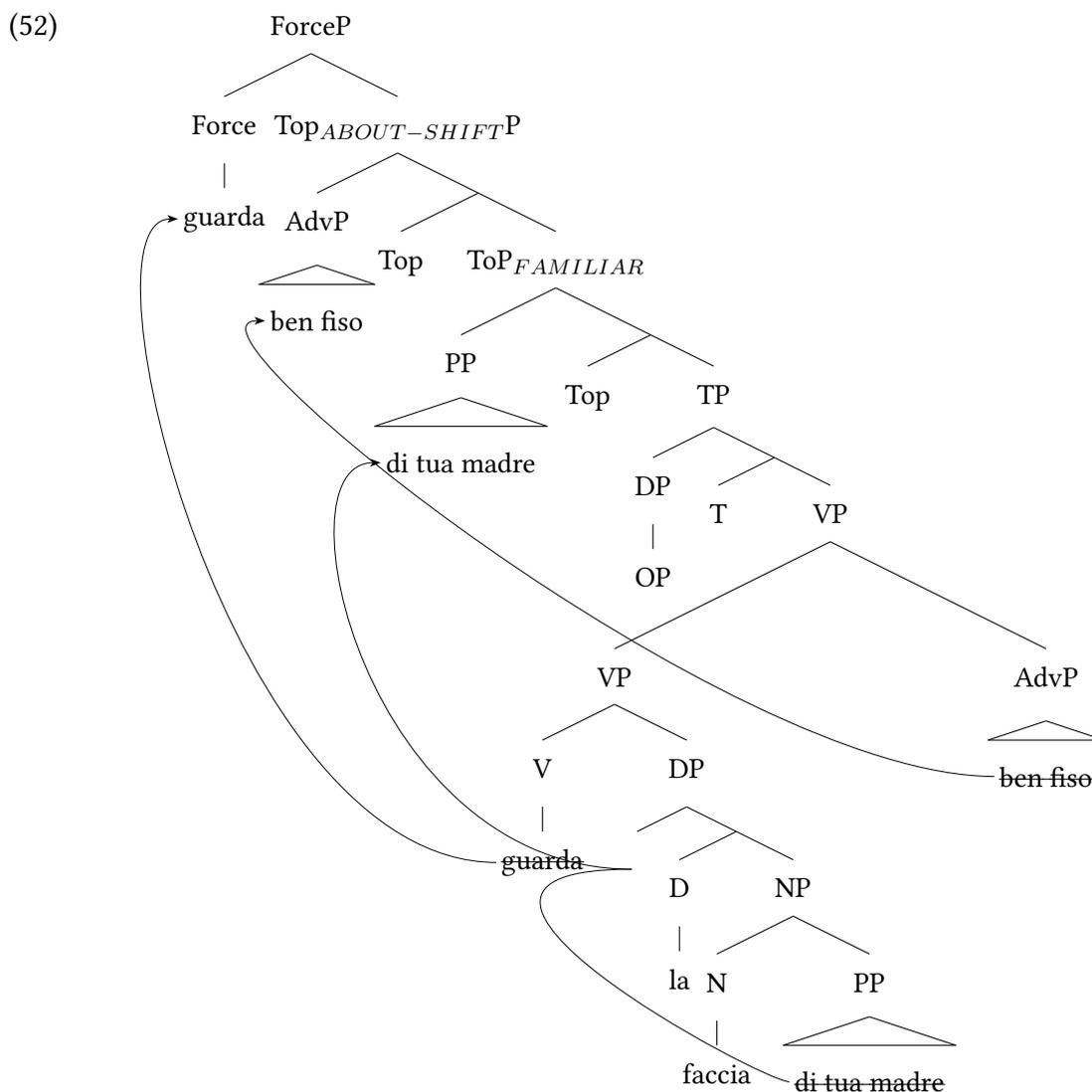
The possessive PP *del femmineo sciame* originates as the complement of the N *brusio* and moves to the specifier of the first TopP, which is an Aboutness-Shift Topic. The DP *il brusio* remains *in-situ*. The DegrP *qual di vento in fogliame* originates as a modifier to the NP then

moves to the Familiar TopP. Italian has V-to-T movement, and so presumably the verb *ode* head-moves to T°, passing through Voice° along the way and the impersonal *si* originates in SpecTP as per Cinque (1988).

Turning to example 36, reproduced below in 51, there is a similar explanation.

- (51) ...guarda ben fiso, fiso **di tua madre la faccia!**
 watch.SG.IMP. well fixed, fixed of your mother the face
 ‘Watch closely your mother’s face!’

Complete noun phrase: **la faccia di tua madre**



I propose that, as an imperative, the verb *guarda* ‘look’ A-moves to Force. This is consistent with Zanuttini’s (1997) proposal that the imperative in Italian moves to C. Given the split CP, it makes sense that the projection of C would be Force.

From there, the AdvP *ben fiso* is topicalized and the PP *di tua madre* is topicalized. *La faccia* ‘the face’ remains *in-situ*. Looking at the imposed prosody (the musical line that Puc-

cini wrote), it is clear that the adverb phrase *ben fiso* can be described as an Aboutness-Shift Topic. The musical line is ascending, with its intonation on *fiso* ‘fixed’. As with the previous examples, the rise in intonation seems to be as close as Puccini’s musical line gets to a low to high intonation that, according to [Frascarelli and Hinterhölzl \(2007\)](#), is evidence of an Aboutness-Shift Topic. The PP *di tua madre* ‘of your mother’ has the lowest intonation of the phrase, which according to [Frascarelli and Hinterhölzl \(2007\)](#) indicates a Familiar Topic.



Figure 7: ‘Guarda ben fiso, fiso di tua madre la faccia.’

Sentence 37, repeated here as 53, requires a slightly different explanation because the PP that is topicalized is extracted from a subject DP, as opposed to an object DP as in the previous sentences. I propose, as such, that the PP moves to SpecDP and then the whole DP is topicalized. The DP *il mistero* cannot remain *in-situ* as per the previous examples because the object *l’anima* is also moved out of its base position into a position higher than the verb, but after movement it still ends up below the subject. This of course is a clear violation of the Subject Criterion. The Subject Criterion, which comes from the Extended Projection Principle, holds that the subject position is obligatorily filled within the clause structure. It is not a possible launching site for movement ([Rizzi 2016](#)).

- (53) **Di sua voce il mistero** l’anima mi colpì
 Of her voice the mystery the.soul me hit
 ‘The mystery of her voice hit my very soul.’

Complete noun phrase: il mistero di sua voce

A possible solution would be to say that instead of any NPs or PPs focusing or topicalizing, the VP *mi colpì* focuses to the right. Such a structure is not in fact impossible in Italian, which [Samek-Lodovici \(2015\)](#) argues in his 2015 paper. Samek-Lodovici in fact claims that a unified analysis of Focus phrases would hold all Contrastive Focuses to be *in-situ* and “discourse-given constituents generated lower than a contrastively focused constituent may optionally move above it and precede it” (2015, pg. 11). While I cannot add evidence from the libretto to support this, I can at least support the claim that *mi colpì* is the focused element with a

musical analysis. Samek-Lodovici (2015) claims, as do many others, that focused phrases are higher in pitch. Under the pretense that the musical line is a sort of constructed prosody, *colpi* is indeed the highest note and has a long note value compared to the rest of the phrase, and occurs on the strongest beat (the first beat of the 3/4 measure).



Figure 8: ‘Di sua voce il mistero l’anima mi colpi’

The whole phrase harmonically leads to the high note of the second syllable of *colpi*, which even someone without musical training might be able to hear in context. As such, I will go ahead with the analysis that the verb is the focused constituent, not the subject or the object. Samek-Lodovici (2015) proposes a two-part movement for right-dislocation, whereby a projection above TP called RP takes the focused element in its specifier. Then another, unspecified projection above RP (XP for now) will remnant-move everything else into its specifier. Since this is not important to my analysis of movement out of noun phrases, I will not show a tree for it. I simply will note that this kind of Focus can occur, and furthermore, that it can co-occur with the kind of left-dislocated Focus discussed previously in this section.

There is still a significant issue to address which is that the possessive PP *di sua voce* extracts from subject position, which is not compatible with any theory of syntax even if I assume that *il mistero* remains *in-situ*. In spoken Italian, even a PP with *wh*-material is not able to be extracted from subject position.

- (54) * $[_{PP_i}$ **Di quale associazione**] il president t_i è corrotto?
 [Of which organization] the president t is corrupt

I see no way around this other than stipulating that PPs can extract from subject noun phrases in this variety of Italian. I do not have enough data from the *Butterfly* libretto alone to determine if this is indeed a trait of Italian opera librettos.

5.3 Movement within opera DPs

This section looks at the word order *within* the Italian opera libretto noun phrase. In *Madama Butterfly*, the possessive pronoun is often in a non-canonical place. In spoken Italian, this

ordering only occurs in vocative phrases or with a particular subset of nouns. This non-canonical ordering is in fact quite regular. I will first present the canonical ordering, according to [Cinque \(1995\)](#).

(55) ARTICLE>POSSESSIVE PRONOUN>QUALITY>SIZE>SHAPE>COLOR>NATIONALITY>N

[Cinque \(1995\)](#) then argues that in the Italian NP, the N obligatorily raises above some of the lower adjectives to an unspecified projection XP, deriving the following order.

(56) la sua bella grande palla rossa
 the his/her beautiful big ball red
 ARTICLE > POSSESSIVE PRONOUN > QUALITY/EVALUATIVE > SIZE > N > COLOR

In the opera libretto, the order is often the one outlined by [Cinque \(1995\)](#). However, there are several instances of a different order. These instances can be simplified to two classes. In one class, the noun precedes the possessive pronoun and the adjectives, as in an Italian vocative. In other orders, the noun seems to stay lower in the clause and an adjective raises above the possessive pronoun even. I present these possibilities in the following:¹⁴

(57) ARTICLE > N₁ > QUALITY/POSSESSIVE PRONOUN > POSSESSIVE PRONOUN/QUALITY > N₂
 > COLOR/QUALITY

The following examples are taken from various places throughout the opera. Examples [58](#), [59](#), and [60](#) contain N in the first position and examples [61](#) and [62](#) contain N in the second position.

(58) **La vita ei** non appaga ...
 The life his NEG satisfy ...
 ‘His life isn’t worth it...’
Ordering: ARTICLE > N > POSSESSIVE PRONOUN

(59) Mi piace **la treccia tua bruna** fra candidi veli.
 Me pleases the locks your brown between snow_white veils
 ‘I like your brown locks, on your snow white veil.’
Ordering: ARTICLE > N > POSSESSIVE PRONOUN

¹⁴I omit several adjective classes because there are no examples in the libretto containing these classes.

- (60) Sempre **il mite suo sembiante** con strazio atroce vedrò.
 Always the face her gentle with guilt atrocious will.see.1SG
 ‘I will always see her gentle face with atrocious guilt.’

Ordering: ARTICLE > N > POSSESSIVE PRONOUN > QUALITY

- (61) **L’esotico suo odore** m’ha il cervello sconvolto...
 Her exotic sent me.has the brain confused...
 ‘Her exotic scent has confused my brain...’

Ordering: ARTICLE > QUALITY > POSSESSIVE PRONOUN > N

- (62) **Colla nuova mia vita** posso adottare nuova religione.
 With.the new my life can.1SG adopt new religion
 ‘With my new life, I can adopt a new religion.’

Ordering: ARTICLE > QUALITY > POSSESSIVE PRONOUN > N

In 58, 59, and 60, the noun has raised to a projection above the possessive pronoun. This phenomenon is not in fact unique to opera. This kind of movement, where the noun precedes the possessive pronoun, is actually valid in spoken Italian for a certain set of nouns, such as *casa* ‘house’ and *amore* ‘love’. *Casa mia* ‘house mine’ and *amor mio* ‘love mine’ are perfectly acceptable, although notably do not allow other post-nominal adjectives. Generally, as I have mentioned, this ordering is allowed in spoken Italian for vocatives such as *amore mio* ‘my love’ but aside from vocatives, the post-nominal possessive pronoun has a literary or poetic feeling.¹⁵ Moreover, in the contemporary play, *La figlia di Iorio*, there are several examples of this noun in a higher position. Likewise, there are even more examples of this in the 17th-century opera *Doriclea*. While I have not extensively studied the subset of noun phrases that contain this kind of music in *Doriclea* and *La Figlia di Iorio*, I can say conclusively that the libretto of *Madama Butterfly* extends the set of nouns that allow post-nominal possessors from the set in spoken Italian.

Before turning to a purely syntactic analysis, as I did in the previous section, I will first look at these particular instances in the musical context to investigate whether this non-canonical ordering is chiefly musically- or metrically-motivated, as per the hypotheses I have been working through.

Looking at an example of a post-nominal possessive pronoun, 58, there is evidence for a musical repair, although not a metrical repair, as we see in Figure 9.

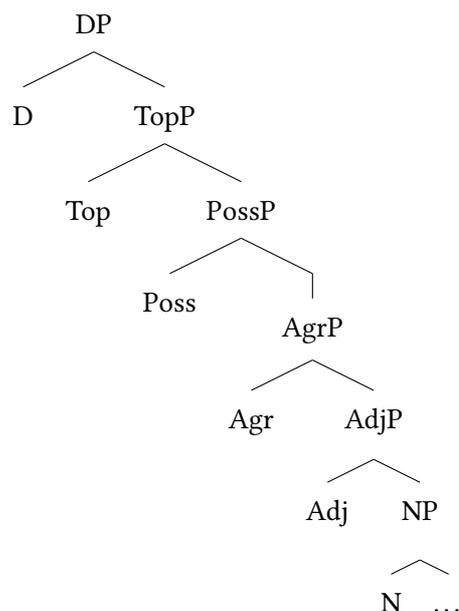
¹⁵Many thanks to Paula Benincà for her insights on post-nominal possessive pronouns.

significant point I am exploring. Perhaps this syntactic requirement came about as a necessity of the music, or perhaps it was firmly a part of the grammar before music was added. Either way, it is a common, systematic occurrence.

Turning back to [Giusti \(1996\)](#) and [Cinque \(1995\)](#), I will now look at how this phenomenon is informed by the syntax of spoken Italian. I will show how the second subset of noun phrases, where a quality adjective precedes the possessive pronoun, can be explained exclusively by Giusti's 1996 work.

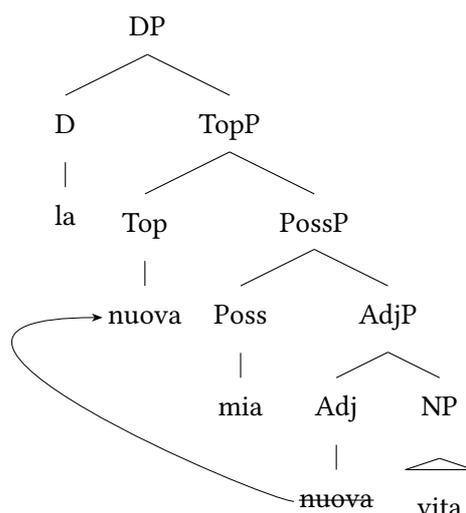
[Giusti \(1996\)](#), among others, argues that DPs are analogous in many ways to CPs in that both have an internal articulation. One instance of this claim is the ability of DPs to contain FocP and TopP. [Giusti \(1996\)](#) claims that Italian DPs contain TopPs but not FocP, arguing that the projections allowed in a DP are language-specific. Her proposed structure for the Italian DP is the following. Note that [Cinque \(1995\)](#) does not specify the projection to which N moves, but here Giusti specifies it as an AgrP above the AdjP.

(63)



Combining the insights of both [Cinque \(1995\)](#) and [Giusti \(1996\)](#), there is enough structure to explain the non-canonical movement within the noun phrases of the libretto. In [62](#), we can consider *nuova* 'new' to be a Topic within the DP. It is old information semantically but the main idea of what she is discussing, meaning it qualifies as a Topic but not as a Focus. The entire scene has emphasized the newness of Butterfly's experience and how this new marriage will change her fortune and she adds to this discussion with this sentence about adopting a new religion along with her new life. As such, it semantically checks out for *nuova* to head move to the TopP. My proposed structure is below.

(64)



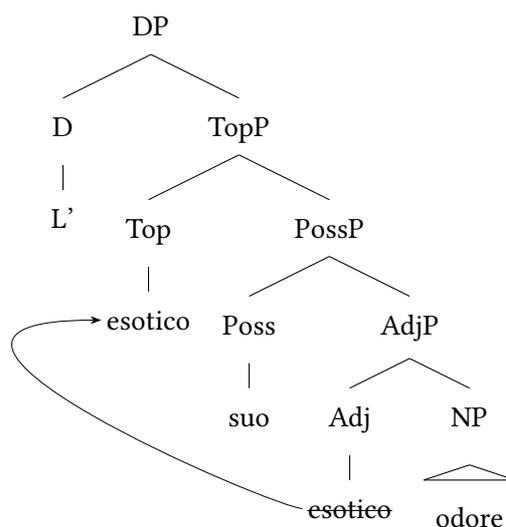
Here, the quality adjective *nuova* initiates below the PossP in AdjP. Some sort of Topic feature on the adjective motivates its movement to TopP, where that feature is checked. The noun, for all intents and purpose, remains *in-situ*.

This particular examples works well within Giusti's (1996) framework, and according to her paper, could even be found under the right semantic circumstances in spoken Italian. According to Giusti (1996), however, a prenominal qualitative adjective must have a significant prosodic break between the raised adjective and the possessive pronoun and the adjective would have to scope over the entire noun entity. Specifically, it must be known information because she argues that noun phrases in Italian do not host a FocP. The example she gives is in 65.

- (65) a. i suoi capelli bianchi
the his/her hair white
- b. i suoi bianchi capelli
the his/her white hair
- c. i bianchi, suoi capelli
the white his/her hair

61 also works within this framework. A similar thing happens, whereby the adjective *esotico* 'exotic' is topicalized within the NP.

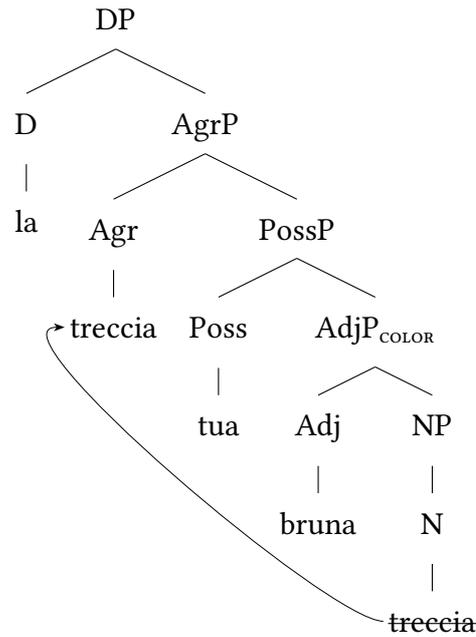
(66)



So these two particular word orders in noun phrases are not in fact unique to the libretto. Although on the surface they seem non-standard, they can be explained via the rules of standard spoken Italian. The difference is that without any significant prosodic break, these re-ordered noun phrases have an additional poetic feeling that could be interpreted as marked in spoken Italian.

The other examples where the noun precedes the possessive pronoun, though, are more distinct from spoken Italian. They have the ordering of Italian vocative phrases without being vocatives, and can also have other modifying adjectives which vocatives cannot have. My proposed explanation is derived entirely from [Cinque \(1995\)](#), and is quite simple. I propose that in the language of Italian opera or poetry, the noun moves to a projection higher in the phrase than that in spoken Italian. While the limit for spoken Italian on N-movement is just above a Color Adjective, the limit in the libretto is above the possessive pronoun. Sticking to [Giusti's \(1996\)](#) label for this projection, I propose the following structure for [59](#).

(67)



In 67, the structure is the same as it would be in spoken Italian. The difference is that the AgrP, to which the noun obligatorily moves, is optionally higher than the PossP, which differs from the rules of spoken Italian, where the AgrP is always below the PossP. As I have argued earlier in this section, this particular syntactic trait could be motivated by musical necessity. The other kind of movement within noun phrases to a TopP, is not necessarily motivated by music but there is some evidence that it fits better within the musical line. It also constitutes a case of movement of a topicalized adjective to the specifier of TopP, not unheard-of in spoken Italian. It is likely low-frequency in spoken Italian, which is why it has a bit of a marked feel within the libretto. Semantically, though, the movement is justified and Giusti (1996) has plenty of evidence to suggest that it is possible within Italian.

5.4 Conclusion on DPs

In this section, I have looked at movement both out of and inside of DPs in the *Madama Butterfly* libretto. For the former, I worked from Rizzi's (1997) split CP hypothesis to show how the movement out of DPs is a Focus/Topic construction. Unlike spoken Italian, movement out of DPs can occur without a *wh*-feature in the libretto. Instead, [+Topic] and [+Focus] features drive movement. Possessive PPs are unusual candidates for Focus or Topic, but I argue that the language of the libretto makes broader use of Topic and Focus.

I have then worked from Cinque (1995) and Giusti (1996) to look at movement within the DPs in the libretto. I have noted two supposed irregularities in the libretto, where the

order of heads seem to be different from the possible order in spoken Italian. I have explained that the re-ordering of adjectives with respect to the possessive pronoun is in fact a Topic construction within the domain of DP, as per [Giusti \(1996\)](#). This kind of re-ordering is not unique to the libretto and can occur with some prosodic work in spoken Italian. The other re-ordering of the noun with respect to the possessive pronoun I have argued is a syntactic trait of librettos and poetry more widely. The simple explanation for this trait is that the AgrP projection that hosts the moved N is in fact higher in poetic language than in spoken Italian. Both kinds of re-ordering fit within the musical line more smoothly than their canonically ordered counterparts. Whether Puccini wrote the music to fit the libretto or re-wrote the libretto to fit the music is unclear and may never be known. However, the fact that it occurs often shows that the relationship to the music is at least partial evidence that it could be a musical repair strategy. This section has shown how Italian opera librettos are not entirely distinct from spoken Italian, but have significant differences.

6 Conclusion

This has been the first truly syntactic analysis of an opera libretto. I have shown that *Madame Butterfly* has several unique syntactic phenomena that make it interesting to study. In particular, I have explored the ways in which the libretto makes interesting usage of FocP and TopP with regards to movement both out of noun phrases and within noun phrases. Whereas spoken Italian only allows movement out of a DP if the moved constituent has a *wh*-feature, the libretto allows movement out of the DP for possessive or subject PPs to a Topic or Focus Phrase. I have also found an instance of movement out of subject position, which violates the Subject Criterion. With only one instance of this, I only note that it occurs but do not offer firm theoretical explanations. It could be that Italian opera librettos violate the Subject Criterion, but I do not yet make this claim.

In addition, I show that the unusual word ordering within noun phrases can be explained via [Giusti's 1996](#) proposal for a Split DP, where there are FocP and TopP inside the DP, analogous to a Split CP. The libretto differs from spoken Italian with regards to what can be topicalized or focused within the DP.

Aside from noun phrases, I found that pronouns are often overt in contexts that would allow a covert pronoun. For the libretto, I argue that the Avoid Pronoun Principle still holds

with the additional caveat that the pronoun cannot be avoided in certain musical and metrical environments. There are times when a musical line requires an extra beat or a poetic line requires an extra syllable or two. In these instances, the pronoun must be overt in order to musically or metrically ‘repair’ the line.

On this brief but fascinating journey through the libretto, several points remained to be explored. I have not explored the possibility of hyperbaton in verb phrases (as per [Delmonte \(2018\)](#)). There are several instances of past participle displacement within the sentence, as in the following.

- (68) M’**ha** coll’ingenua arti **invescato**.
me’has with.the’naive charms ensnared
‘She has ensnared me with her naive charms.’

Canonically, the past participle should come directly after the auxiliary except when there is an adverb in the phrase. 68 has no adverbial elements. Either the prepositional phrase *coll’ingenua arti* is dislocating or the past participle is dislocating. This curious puzzle, which [Delmonte \(2018\)](#) alludes to with Romantic Italian Poetry, is a project for another time. A future study should explore whether this constitutes an unusual focus or topicalization structure or something else entirely.

Greater takeaways from this thesis are that natural language may be different for artistic forms. The libretto stems from spoken and written Italian, of course, but pointedly breaks from this language throughout the interactions with music and poetic meter. There is a sense in which this may be analyzed as a change in register. The movement of the PPs out of the noun phrases, for example, lends a particularly poetic tone to the language that deserves a closer look. Aside from Topic and Focus, there is another layer of more abstract meaning that takes place with the change in syntax. A convincing argument has been made by [Thoms \(2010\)](#) for a separate account of poetic language from spoken language. While I have stuck to pre-existing Minimalist accounts of Italian to explain the phenomena in the libretto, I would be interested to extend my analysis of Italian librettos to his work on English poetry. Overall, librettos deserve more study because the relationship of language and other art forms could illuminate both the pragmatics of artistic speak and also just how flexible natural language can be.

And with that, *la commedia è finita*.

References

- Benincà, P., and C. Tortora. 2003. The syntax of Italian dialects. Oxford studies in comparative syntax. Oxford: Oxford University Press. URL <https://books.google.com/books?id=DVTTtAEACAAJ>.
- Bonomi, Ilaria, and Edoardo Buroni. 2017. La lingua dell'opera lirica. Itinerari. Filologia e critica letteraria. Bologna: Il Mulino.
- Cardinaletti, Anna, and Michal Starke. 1999. The typology of structural deficiency: A case study of the three classes of pronouns. In Clitics in the languages of Europe, ed. Henk van Riemsdijk, 145–233. New York: Mouton de Gruyter.
- Chomsky, Noam. 1993. Specification of empty categories. In Lectures on government and binding: The Pisa lectures, 231–284. Berlin: Walter de Gruyter.
- Cinque, Guglielmo. 1980. On extraction from NP in Italian. Journal of Italian linguistics 5:47–99.
- Cinque, Guglielmo. 1988. On *si* constructions and the theory of arb. Linguistic Inquiry 19:521–581.
- Cinque, Guglielmo. 1995. On the evidence for partial N-movement in the Romance DP. In Italian syntax and universal grammar, 287–309. Cambridge: Cambridge University Press.
- Cinque, Guglielmo. 1999. Adverbs and functional heads; a cross-linguistic perspective. New York: Oxford University Press. URL <https://www.google.com/books/edition/Adverbs-and-Functional-Heads/sFd79vhd-n0C>
- Cinque, Guglielmo. 2014. Extraction from DP in Italian revisited. In Locality, ed. Enoch Olaté Aboh, Maria Teresa Guasti, and Ian Roberts, 86–103. Oxford: Oxford University Press. URL <http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199945269.001.0001>
- Delmonte, Rodolfo. 2018. Syntax and semantics of Italian poetry in the first half of the 20th century. Umanistica Digitale 2. URL <https://umanisticadigitale.unibo.it/article/view/8053>.

- D'Alessandro, R., A. Ledgeway, and I. Roberts. 2010. Syntactic variation: The dialects of Italy. Cambridge: Cambridge University Press. URL <https://books.google.com/books?id=JfdVPgAACAAJ>.
- Frascarelli, Mara. 2007. Subjects, topics and the interpretation of referential pro. Natural Language & Linguistic Theory 25:691–734.
- Frascarelli, Mara, and Roland Hinterhölzl. 2007. Types of topics in German and Italian. On Information Structure, Meaning and Form: Generalizations Across Languages 100:87–116.
- Giusti, Giuliana. 1996. Is there a FocusP and a TopicP in the noun phrase structure? University of Venice Working Papers in Linguistics 6:105–128.
- Groos, Arthur. 2016. Madama Butterfly between east and west. In Giacomo puccini and his world, ed. Emanuele Senici and Arman Schwartz, 49–84. Princeton: Princeton University Press. URL <https://muse.jhu.edu/book/46282/>.
- Haegeman, Liliane. 2013. The syntax of registers: Diary subject omission and the privilege of the root. Lingua 130:88–110.
- Haegeman, Liliane. 2019. Register-based subject omission in english and its implication for the syntax of adjuncts. Anglophonia 28. URL <http://journals.openedition.org/anglophonia/2873>.
- Haegeman, Liliane, and Tabea Ihsane. 2001. Adult null subjects in the non-pro-drop languages: Two diary dialects. Language Acquisition 9:329–346.
- Heap, David. 1997. Subject pronoun variation in the Northern Italian geolinguistic continuum. Quaderni d'italianistica 18:235–250.
- Herzog, Silvia. 1996. Strophic discourse and Stefano Landi's "La morte d'Orfeo": Syntax in music and poetry during the cinquecento and seicento. Doctoral Dissertation, University of Southern California, Los Angeles. URL <https://search.proquest.com/docview/304295646/abstract/531A46458B2E480EPQ/1>.
- Migliorini, Bruno, and T. Gwynfor Griffith. 1984. The Italian language. Boston: Faber and Faber.

- Poletto, Cecilia. 2014. Word order in Old Italian. Oxford studies in diachronic and historical linguistics. Oxford: Oxford University Press. URL <http://dx.doi.org/10.1093/acprof:oso/9780199660247.001.0001>.
- Radford, Andrew. 2004. Minimalist syntax: Exploring the structure of english. Cambridge: Cambridge University Press.
- Renzi, Lorenzo. 2001. I pronomi personali. In Granda grammatica italiana di consultazione, 535–594. Bologna, Italy: Il mulino.
- Rizzi, Luigi. 1986. Null objects in Italian and the theory of pro. Linguistic Inquiry 17:501–557.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In Elements of grammar, ed. Liliane Haegeman, 281–337. Dordrecht, Netherlands: Kluwer Academic Publishers.
- Rizzi, Luigi. 2016. EPP and ECP revisited: The role of labeling. Romance Languages and Linguistic Theory 10:211–232.
- Roselli, Bruno. 1926. Tu, lei, and voi. Italica 3:49–53.
- Samek-Lodovici, Vieri. 1996. Constraints on subjects: An optimality theoretic analysis. Doctoral Dissertation. URL <https://search.proquest.com/docview/85640775?pq-origsite=summon>.
- Samek-Lodovici, Vieri. 2015. The interaction of focus, givenness, and prosody: a study of Italian clause structure. New York: Oxford University Press.
- Senici, Emanuele. 2016. Introduction: Puccini, his world, and ours. In Giacomo puccini and his world, ed. Emanuele Senici and Arman Schwartz. Princeton: Princeton University Press. URL <https://muse.jhu.edu/book/46282/>.
- Thoms, Gary. 2010. Poetic language: A minimalist theory. Doctoral Dissertation, University of Strathclyde, United Kingdom.
- Wood, Jim. 2019. Icelandic nominalizations and allosemy.
- Zanuttini, Raffaella. 1997. Negative imperatives. In Negation and clausal structure: A comparative study of romance languages, Oxford studies in comparative syntax. Oxford: Oxford University Press.