

Poetic language: a Minimalist theory

School of Humanities, University of Strathclyde

PhD

2010

Gary Stewart Thoms

Declaration of authenticity and Author's Rights

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

The copyright of this thesis belongs to the author under the terms of the United Kingdom Copyright Acts as qualified by University of Strathclyde Regulation 3.50. Due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

Abstract

This dissertation proposes a new linguistic theory of poetic language, working within the framework of generative linguistics. I provide a substantial and systematic description of linguistic deviation in the language of English poetry, and I argue that previous linguistic theories that generate deviant forms by the application of an extra set of poetic rules are untenable both for theoretical and empirical reasons. I thus argue that we must break from this traditional approach and instead allow for poetic language to be generated by different means. I propose that poetic language sentences are generated by what I call ‘PF-uneconomical’ derivations; that is, the major classes of poetic language deviation, which I call ‘displacement’ (unlicensed movement) and ‘erasure’ (unlicensed ellipsis), are produced by derivations which invoke extra derivational steps in the PF-branch of the derivation relative to their well-formed counterparts. I show that this theory can model the ways in which the poetic language operations can affect both phrases and non-constituents, explain the non-occurrence of some forms, and account for the distribution of the different kinds of deviation attested in the poetic texts, and I discuss the ways in which the operations of the poetic language syntax may interact with grammar-external formal conditions such as metricality. I argue that the proposed theory follows entirely from Minimalist assumptions about the role of syntax and the interfaces, and that as such the theory presented here provides evidence for the importance of economy in the computations of the language faculty.

Contents

1	Introduction	1
1.1	Defining poetic language	2
1.2	What should the theory aim to achieve?	10
1.2.1	Descriptive adequacy	10
1.2.2	Explanatory adequacy	13
1.3	The data from poetry	15
1.4	Summary and dissertation outline	19
2	Linguistic deviation in poetry	21
2.1	Displacement phenomena	22
2.1.1	PP displacement	22
2.1.2	NP displacement	35
2.1.3	AP displacement	49
2.1.4	Other kinds of displacement	55
2.1.5	Summary	64
2.2	Erasure phenomena	66
2.2.1	Missing out words in English	67
2.2.2	Erasure of functional elements	72
2.2.3	Failed ellipsis	85
2.2.4	Article erasure	99
2.2.5	Summary	102
2.3	Overall summary	103

3	Previous approaches	105
3.1	Transformational poetic grammar	106
3.1.1	Empirical problems	111
3.1.2	Theoretical problems	124
3.1.3	Summary	129
3.2	Exploiting UG: O’Neil (2001)	130
3.2.1	Empirical problems	136
3.2.2	Theoretical problems	141
3.2.3	Summary	143
3.3	An OT approach: Fitzgerald (2007)	144
3.3.1	Empirical problems	149
3.3.2	Theoretical problems	163
3.3.3	Summary	167
3.4	Overall summary	168
4	A new theory	170
4.1	The Non-Uniformity Hypothesis	171
4.1.1	The Development Hypothesis and its problems	172
4.1.2	The Pragmatics Hypothesis and its problems	178
4.2	The initial proposal: Bypass Theory	183
4.2.1	Theoretical context	184
4.2.2	Proposal	185
4.2.3	Pieces	189
4.2.4	PF Concatenation	194
4.2.5	LF Combination	199
4.2.6	Blocking poetic language and blocking ordinary language	210
4.2.7	Poetic language blocking poetic language?	217
4.2.8	Summary	229
4.3	Deriving the data	229
4.3.1	Different forms of displacement	230

4.3.2	Uniting erasure and displacement?	232
4.3.3	Erasure, ellipsis and economy	234
4.3.4	The revised theory	251
4.4	Summary	256
5	Conclusion	258
5.1	The formal character of poetic language	258
5.2	PF Economy, possible derivations and the medium of language .	261
5.3	Poetic language and linguistic theory	267

Key for references to primary texts

Frequently quoted texts are referred to with abbreviated references. The references for verse texts appear as TITLEABBREVIATION:LINES, and the references for unlineated texts (marked with a \$ below) appear as TITLEABBREVIATION.PAGENO. Below are the relevant abbreviations and the matching full references.

- TRX** ‘The Rape of the Lock’ by Alexander Pope, in Pope (1966: 88-109).
‘X’ denotes the different Canto numbers within the poem (1-5).
- EL** ‘Eloisa to Abelard’ by Alexander Pope, in Pope (1966: 110-119).
- ESX** ‘An Essay on Man’ by Alexander Pope, in Pope (1966: 241-279).
‘X’ denotes the different epistle numbers within the poem (1-4).
- TB^{\$}** Gertrude Stein, *Tender Buttons* (Stein 1997 [1914]).
All examples taken from the ‘Objects’ section of the text.
- WH^{\$}** Samuel Beckett, *Worstward Ho* (Beckett 1999 [1983]).
- HI^{\$}** Samuel Beckett, *How It Is* (Beckett 1994 [1965]).
- LIX^{\$}** John Ashbery, ‘Litany,’ in *As We Know* (Ashbery 1979: 1-68).
‘Litany’ is written as two parallel columns, so ‘X’ denotes which column (1 or 2). Page references are given rather than line numbers for convenience, since the text does not provide line numbers.

Below are full references for other individual poems that are quoted in the text.

The page references are those that are quoted.

- John Ashbery:
 - ‘Album Leaf’: Ashbery (1997: 17) (originally appeared in *Some Trees*).
 - ‘The Thinnest Shadow’: Ashbery (1997: 30) (originally appeared in *Some Trees*).
 - ‘A Pastoral’: Ashbery (1997: 54) (originally appeared in *Some Trees*).
 - ‘Two Sonnets’: Ashbery (1997: 71) (originally appeared in *The Tennis Court Oath*).

- ‘Night’: Ashbery (1997: 74, 77) (originally appeared in *The Tennis Court Oath*).
 - ‘Rain’: Ashbery (1997: 84, 85) (originally appeared in *The Tennis Court Oath*).
 - ‘The Lozenges’: Ashbery (1997: 106) (originally appeared in *The Tennis Court Oath*).
 - ‘The Ascetic Sensualists’: Ashbery (1997: 111) (originally appeared in *The Tennis Court Oath*).
 - ‘Europe’: Ashbery (1997: 132, 134) (originally appeared in *The Tennis Court Oath*).
 - ‘To The Same Degree’: Ashbery (1997: 152) (originally appeared in *The Tennis Court Oath*).
 - ‘The Thousand Islands’: Ashbery (1997: 182) (originally appeared in *Rivers and Mountains*).
 - ‘The Skaters’: Ashbery (1997: 194, 203) (originally appeared in *Rivers and Mountains*).
 - ‘It Was Raining In The Capital’: Ashbery 1997 [1956]: 236) (originally appeared in *The Double Dream of Spring*).
 - ‘As You Came From The Holy Land’: Ashbery (2008: 430) (originally appeared in *Self Portrait In A Convex Mirror*).
 - ‘Grand Galop’: Ashbery (2008: 436) (originally appeared in *Self Portrait In A Convex Mirror*).
 - ‘Saying It To Keep It From Happening’: Ashbery (2008: 509) (originally appeared in *Houseboat Days*).
 - ‘The Thief Of Poetry’: Ashbery (2008: 527) (originally appeared in *Houseboat Days*).
 - ‘Tapestry’: Ashbery (1979: 90).
 - ‘A Box And Its Contents’: Ashbery (1979: 92).
- Robert Creeley:
 - ‘A Step’: Creeley (1970: 47) (originally appeared in *Pieces Two*).
 - ‘Gemini’: Creeley (1970: 56) (originally appeared in *Pieces Two*).
 - “‘Follow the Drinking Gourd’”: Creeley (1970: 58, 59) (originally appeared in *Pieces Two*).
 - ‘Echo of’: Creeley (1970: 90) (originally appeared in *Pieces Two*).
 - ‘Ice Cream’: Creeley (1970: 97) (originally appeared in *Pieces Two*).
 - ‘In London 3’: Creeley (1970: 123).
 - ‘Wellington, New Zealand’: Creeley (2006: 6) (originally appeared in *Hello: A Nournal, February 29–May 3, 1976*).
 - ‘Soup’: Creeley (2006: 14) (originally appeared in *Hello: A Nournal, February 29–May 3, 1976*).
 - ‘So There’: Creeley (2006: 24) (originally appeared in *Hello: A Nournal, February 29–May 3, 1976*).

- ‘Say Something’: Creeley (2006: 86) (originally appeared in *Hello: A Nournal, February 29–May 3, 1976*).
 - ‘Later’: Creeley (2006: 90) (originally appeared in *Hello: A Nournal, February 29–May 3, 1976*).
 - ‘Beach’: Creeley (2006: 104) (originally appeared in *Hello: A Nournal, February 29–May 3, 1976*).
 - ‘For Pen’: Creeley (2006: 105) (originally appeared in *Hello: A Nournal, February 29–May 3, 1976*).
 - ‘Desultory Days’: Creeley (2006: 166) (originally appeared in ‘*Hello: A Nournal, February 29–May 3, 1976*’).
 - ‘Blues’: Creeley (2006: 125) (originally appeared in ‘*Hello: A Nournal, February 29–May 3, 1976*’).
- Charles Olson:
 - ‘The K’: Olson (1987: 14).
 - ‘The Fool’: Olson (1987: 52).
 - Jackson Mac Low, ‘A(ce)’: Mac Low (2008: 48).

Chapter 1

Introduction

This dissertation is a linguist's attempt to discover what can be learned about language from the experiments of poets. I look at the ways in which poets deviate from the ordinary rules of language and, working within the framework of generative linguistics, I develop a theory that aims to account for the kinds of linguistic deviation we see in poetic language, to explain how these forms are interpreted and to explain why certain kinds of variation occur but others do not. I argue that previous linguistic theories that generate deviant forms by the application of an extra set of poetic rules are untenable both for theoretical and empirical reasons, and I thus argue that we must break from this traditional approach and instead allow for poetic language to be generated by different means. Instead I propose that poetic language sentences are generated by what I call 'PF-uneconomical' derivations; that is, the major classes of poetic language deviation, which I call 'displacement' (unlicensed movement) and 'erasure' (unlicensed ellipsis), are produced by derivations which invoke extra derivational steps in the PF-branch of the derivation relative to their well-formed counterparts. I show how this theory can model the ways in which the poetic language operations can affect non-constituents and account for the distribution of the different kinds of deviation attested in the poetic texts, and I discuss the ways in which the operations of the poetic language syntax may interact with

grammar-external formal conditions such as metricality. Finally I argue that the proposed model follows entirely from Minimalist assumptions about the role of syntax and the interfaces, and that as such the theory presented here provides evidence for such an approach to syntactic theory.

In this chapter I outline the relevant issues for developing a linguistic theory of poetic language and explain the kind of approach advocated here. The chapter is structured as follows. In 1.1 I define the terms of the theory to be developed, identifying the broad scope of the study and outlining its relation to literary-theoretic views of poetic language. In 1.2 I describe what the theory of poetic language should aim to achieve as a theory developed in the framework of generative linguistics. In 1.3 I discuss the data focus of the study and some specific issues that arise in the linguistic analysis of poetic language. Finally 1.4 summarises and outlines the structure of the rest of the dissertation.

1.1 Defining poetic language

It is evident that all poetic texts are made up of language to some extent, and in the majority of cases the language in literary texts is continuous with the language used in all other walks of life, such as everyday speech or other genres of written texts. But in the long history of the study of literature, the main point of interest in poetic language (or ‘literary language’ as it is also called) has typically been understood as the ways in which it differs from ‘ordinary language,’ typically with the aim of understanding what makes the language of poetry special or distinctive. This is the aspect of poetic language that I concentrate on in this dissertation.

The idea that the language of literature is somehow different from ‘ordinary language’ is contested by literary theorists and writers alike. Derek Attridge’s *Peculiar Language* (1988) provides a Derridean critique of this issue, discussing the ways in which a number of important literary figures have tried to rationalise the deviant nature of poetic language. He sums up the problem succinctly:

[...] if literature is a distinctive use of language with its own very special and unusual codes and practices, it is accessible only to the few who who are in a position to acquire familiarity with those codes and practices. It is at best elitist and at worst solipsistic. But writers have traditionally, and understandably, made a different claim for their work: that it speaks beyond the small circle of those with a professional commitments to literature, that it can engage with the language and the thoughts of everyone who speaks the same tongue, [...] To push this claim too far, however, is to endanger the existence of literature itself as a distinct entity, for if literature does *not* employ a special language, from what does it derive its appeal and strength? There is nothing in its armory but language, and if all its linguistic weapons are borrowed, they are likely to be more forceful on their own terrain.

(Attridge 1988: 1)

Analysing the different ways in which artists and theorists have engaged with this issue, Attridge concludes that “the domain of literature and of literary theory cannot provide its own self-sufficient and lasting answers to the question of the distinctiveness of literary language” (Attridge 1988: 16). He concludes that, while the discoveries made in engaging with this issue of distinctiveness may be valuable, there is a fundamental problem with identifying its source with the tools of literary theory. Thus the question of what makes poetic language different from ordinary language is ineffable in literary studies, at least from the perspective of post-structuralist literary theory.

Attridge’s conclusions are sound, and in keeping with those of many other literary theorists working with literature as a cultural entity, but this should not deter the generative linguist, who seeks to understand language as part of the natural world. The metaphor of language as a “weapon” highlights the fact that Attridge identifies the problem with pinning down the functional contribution of poetic language: that is, how it contributes to the creation of nuanced poetic meaning, the distinctive effect of literariness, or the sociological aspect of a text (all of which contribute to what he later describes as ‘the singularity of literature’ in Attridge 2004). However many linguists do not consider these issues to be subjects of linguistic research, and in analysing poetic language, they would be concerned with a different set of questions that concentrate on

the formal analysis of the language found in literature.¹ Thus, instead of trying to identify how poetic language contributes functionally to distinctiveness, we should aim to identify the formal properties which distinguish poetic language from ordinary language. To theorise these descriptive findings, one must use the tools of theoretical linguistics, which are fundamentally different from the tools of literary theory and thus resistant to the kinds of problems Attridge identifies.² This is the approach adopted in this dissertation.

The initial task is to define the formal characteristic which distinguishes poetic language from ordinary language, as this will define the focus of the theory. The majority of previous works that have proposed linguistic theories of poetic language have focused on linguistic deviation in the syntax of poetry, often focusing on unusual inversions in metrical poetry and related genres (Thorne 1965; Levin 1967; Dillon 1975; Austin 1984; Youmans 1982, 1983, 1986; O'Neil 2001; Fitzgerald 2007; see chapter 3 for a critical review of this literature). This is in keeping with the trend in literary studies, as deviation has been taken as a fundamental characteristic of literature by many theorists over the years, from Aristotle's *Poetics*, through to the Russian Formalists and Attridge's work cited above. Therefore, following the assumptions in the literature, the formal characteristic which I identify as distinguishing poetic language and ordinary language is syntactic deviation.³ The theory developed here aims to account for the kinds of syntactic deviation found in poetic language. I adopt this narrow definition of poetic language from hereon, so the reader should be aware that

¹This is typical of generative approaches to linguistics, which are formalist and distinct from functional approaches like that of M.A.K. Halliday and his followers (e.g. Halliday and Matthiessen 2004). For discussion of the differences between formal and functional approaches to linguistics, see Newmeyer (1983, 2003); for discussion of the difference between formal and functional approaches to literary linguistics, see Fabb (1997: ch.1).

²Note that, while formal approaches may not answer questions of literary theory, they may provide evidence of formal properties hitherto unknown to literary studies, providing a different basis for the conclusions of literary theory. See chapter 5 for discussion of this general issue, and some particular feeding-back into literary theory based on the conclusions of previous chapters.

³Such a study may be extended to look at various other kinds of deviation, such as deviation in morphological form, but I ignore this for the remainder of the dissertation for the sake of restricting the scope of the theory. Note that this is a non-trivial point if one assumes that morphology is syntactic in nature, as in the framework of Distributed Morphology (Halle and Marantz 1993). I return to this issue briefly in the discussion in chapter 5.

this is intended whenever I use the term.

The precise nature of the object of study needs to be clarified further, as ‘syntactic deviation’ may be construed in different ways. Here I take cases of syntactic deviation to be cases where a given string of words can be assigned a meaning as a sentence of English but it clearly violates a grammatical rule that is normally obeyed in ordinary English. Below is a simple example from Alexander Pope’s ‘The Rape Of The Lock’ that is discussed later:

- (1) Her lively looks a sprightly mind disclose,
Quick as her eyes, and as unfix’d as those; *TR2: 9-10*

The word order of the first line of this couplet is unlike that of ordinary English as the object *a sprightly mind* appears to the left of the verb *disclose*, rather than in its normal position to the right of the verb. While English has a set of standard rules for shifting objects to positions other than the standard one (such as topicalization), it does not have a rule which places the object in the position immediately to the left of the verb; the sentence sounds ungrammatical to a native speaker of English as a result. Given this, we can say that Pope has disobeyed the rules of English grammar that dictate where objects can appear, and therefore the sentence can be described as syntactically deviant. As such it is part of the empirical basis for the theory to be constructed here.

This basic definition of the characteristics of poetic language has a few important features that should be highlighted, as they indicate how the scope of the theory is delimited. First, for a given sentence to count as data for the theory, its deviation needs to be describable in terms of grammatical rules that have been disobeyed; thus if a given text seems deviant according to intuitions but it is not clear what kind of rule has been disobeyed, its deviation cannot be formally characterized. Second, these rules are indeed *grammatical*, and as such they pertain to rules that reflect linguistic competence (in Chomsky’s terms), rather than performance. This means that deviation in style (like the use of overly complex sentences or fragmentation) is not considered to be relevant, since style is not described in terms of grammatical rules; style is an aspect of

performance, rather than competence.⁴ The same applies to non-grammatical discourse factors like cohesion (Halliday and Hasan 1976), which are also taken to be a matter for performance in the view of modern generative linguistics and, as such, outside of its remit (cf. Kolaiti 2005). Other processing factors, such as the cases of “centre embedding” famously discussed by Chomsky and Miller (1963), are also put to one side.⁵

Third, the study is restricted to texts that have a determinate meaning as a sentence of English describable as a truth-conditional proposition, because otherwise what we are dealing with cannot be described as linguistic data that is relevant to any sort of theory of language. It is a necessary limitation of the analysis of existing linguistic data that one must be able to attribute to a given string a sentential interpretation,⁶ as otherwise one cannot know the intended well-formed string in order to determine the ways in which the string has deviated from the ordinary rules of the language. To demonstrate the point, let us consider the following example, the first stanza from Jackson Mac Low’s ‘Mark Twain Life on the Mississippi Illustrated Harpers’ (Mac Low 2008: 51):⁷

- (2) Mississippi about. Reading keels.
 The well about. Is not
 Longest is four England,
 On not
 The hundred England.
 Mississippi is seems seems is seems seems is part is
 Is longest longest up seems the reading about. The England discharges
 Hundred about. Reading part England, reading seems

The lines of the poem cannot be interpreted as sentences of English, and only a few isolated sequences of words can be parsed as phrases (*the reading, the*

⁴This claim may be refuted if we accept Culler’s (1975, 1981) proposal that readers acquire a ‘literary competence’ that is analogous to linguistic competence, and that this literary competence includes rules that adequately characterize intuitions about style.

⁵Such examples are seldom found in poetic language, though see section 3.2 for discussion of a rare example.

⁶Technically this is not an adequate characterization, as one could also analyse an isolated phrase like *the car blue* as a deviant version of *the blue car*. Thus the criteria might be widened to cover any case where a phrase or sentence can be given a meaningful interpretation as a linguistic object, rather than just a sentential one. I will continue to refer to ‘sentential interpretations’ as a cover term including isolated phrases, for the sake of terminological convenience.

⁷The text was composed by an arbitrary algorithm which involved selecting words from a Mark Twain book according to the first letter (see Mac Low 2008: 49).

England, possibly *reading keels*). As a result, it is impossible to analyse what deviation there might be in the linguistic form; for example, with the first orthographic sentence, *Mississippi about*, one cannot be sure of the intended meaning of the string (if there is one) and hence one cannot be sure if words have been omitted or moved around in an unconventional fashion. Such examples are unusable as linguistic data, so they must be put to one side. I do this for the rest of the dissertation, concentrating on pieces of text that can be referred to as ‘sentences.’

Finally, the proposed definition does not consider the conceptual semantic meaning of poetic sentences, and as a result I do not consider cases of apparent deviation in this dimension to be within the remit of a theory of poetic language. That is, I do not consider the unusualness of a poetic sentence like (3), which is taken from John Ashbery’s ‘*Vetiver*’ (Ashbery 1990: 1):

- (3) Ages passed slowly, like a load of hay,
 As the flowers recited their lines
 And pike stirred at the bottom of the pond.

This sentence would be strange in ordinary usage, as we know that flowers cannot recite lines, since they are non-human and (largely) inanimate. Yet this kind of strangeness is common in the language of poetry, and as such it may be described as a formal characteristic that distinguishes poetic language from ordinary language, and thus part of the remit of the theory to be developed here. Given that the nuances of unusual sentences like (3) are one of the main concerns of many studies of poetry, this may appear to be a serious oversight for the proposed theory.

The nature of this kind of deviation has been the subject of a great deal of discussion over the years. As any speaker will verify, examples like (3) may be deviant in the strictest sense, but it still nevertheless receives an interpretation as a metaphor.⁸ A linguistic explanation of the phenomena displayed in (3),

⁸This metaphorical interpretation is subject to a great degree of variation: one might interpret *the flowers* as denoting some unidentified character within the narrative, or alternatively one may propose that the described action of *recited their lines* is metaphorical, describing the sound made by the wind passing through the flowers. The nature of this variation is itself

then, would essentially constitute a linguistic theory of metaphor interpretation. The problem with this, however, is that it is clear that the relationship between metaphor interpretation and semantic deviation is not direct (as is assumed by many authors, i.e. Cohen 1993). For example, Stern (1983) points out that a metaphorical interpretation can often be available in the absence of any semantic deviation, as demonstrated by the ambiguous (4):

- (4) Mary has a heart.

This sentence has both a literal interpretation, where *a heart* denotes an organ in Mary's body, and a metaphorical interpretation, where *a heart* refers to Mary's generosity of spirit; however, despite the fact that one is a metaphor and the other literal, both have the same semantic relation to the predicate possessive *have* and thus both are semantically well-formed.⁹ This shows that there can be metaphor without semantic deviation, and that a theory of metaphor must extend beyond a characterization of deviation in semantic structure. Interestingly, the opposite situation is also attested: Chomsky (1965: 148ff) shows that certain cases of semantic deviation do not allow for metaphorical interpretation:

- (5) a. John found sad.
 b. John persuaded great authority to Bill.
 c. Howard elapsed that Bill will come.

Chomsky noted that while examples like (3) involve mismatches in “selectional feature rules,” which describe semantic conditions on argument selection like [HUMAN] or [ANIMATE], the examples in (5) involve violations of “subcategorization rules” that dictate what kind of grammatical categories are selected by the predicates.¹⁰ Violation of subcategorization clearly leads to complete unin-

an important matter, one that is discussed greatly in the literature on metaphor, but it is not one I will consider here.

⁹The fact that the second meaning of *a heart* is abstract should not make a difference for selection, since possessive *have* can also take abstract nouns as complements, as in the directly analogous *Mary has compassion*.

¹⁰In section 4.2.3 I argue (following Chomsky (2001 *et seq*) and others that subcategorization is ultimately a semantic manner, that is, a well-formedness condition that applies at the semantic interface: subcategorization failure precipitates a failure in assigning a semantic interpretation to a given structure. This begs the question of how these two can be teased apart, given that they are both explained as semantic phenomena of essentially the same kind.

interpretability, as the sentences in (5) cannot be interpreted, metaphorically or otherwise; indeed such examples seem to have more in common with examples like (2) than (3).

The fact that one class of semantic deviations allows for metaphorical interpretation while another doesn't is clearly an interesting matter that warrants further attention, but what these cases and the opposite cases like (4) demonstrate is that developing a linguistic theory of metaphor interpretation is a difficult issue that takes us well beyond the remit of a theory of poetic language. This is perhaps to be expected, given that metaphor is clearly not a definitional characteristic of poetic language; metaphor is attested widely in ordinary language use, as attested by examples like (4), which are just as likely to be used in ordinary language as in a poetic text. The impossibility of separating poetic and ordinary uses of metaphor has been one of the biggest issues in recent discussions of metaphor in the massive literature on the subject,¹¹ so we may assume that the conceptual deviation seen in (3) cannot be identified as a characteristic of poetic language in the same way that the kind of deviation seen in (1) can be. Thus I conclude that while the relationship between metaphorical interpretation and linguistic deviation is an interesting issue, it is not one that is wholly relevant to the theory of poetic language. I put the issue of conceptual semantic deviation of the kind in (3) to one side for the rest of the dissertation.¹²

¹¹For discussion, see Sperber and Wilson (1986), Carston (2002), Wilson and Carston (2007), Recanati (2004), Lakoff and Johnson (1980), as well as many of the papers collected in Ortony (1993) and Gibbs (2008). For a particularly intriguing example of a recent linguistic theory of metaphor interpretation that is compatible with the present study, I refer the reader to the work of Josef Stern (2000, 2006, 2009), who proposes a compositional semantic model of metaphor interpretation that captures its context-sensitivity by appealing to an analogy with the Kaplanian treatment of indexicals and demonstratives.

¹²Fabb (2010: 1228) draws a similar conclusion, arguing that although poetic language may often be much more "difficult" due to its high use of complex and sometimes indeterminate metaphors, but nevertheless these aspects of interpretation "can be explained in terms of any ordinary pragmatics."

1.2 What should a theory of poetic language aim to achieve?

With this narrow description of poetic language in place, we may now consider the criteria that such a theory would need to satisfy. Since Chomsky (1965), the two main criteria of linguistic theory have been *descriptive adequacy* and *explanatory adequacy*. I will explain these criteria in turn and outline how they will be pursued in the present study.

1.2.1 Descriptive adequacy

Chomsky (1965: 24) defines descriptive adequacy in the following manner:

A grammar can be regarded as a theory of language; it is *descriptively adequate* to the extent that it correctly describes the intrinsic competence of the idealized native speaker. The structural descriptions assigned to sentences by the grammar, the distinctions that it makes between well-formed and deviant, and so on, must, for descriptively adequacy, correspond to the linguistic intuition of the native speaker (whether or not he may be immediately aware of this) in a substantial and significant class of crucial cases.

The quote indicates that there are a few different dimensions to descriptive adequacy. That the theory must confer the correct distinctions “in a substantial and significant class of crucial cases” is the requirement that a given theory should identify and describe the important data. A classic example of amendment of the theory in service to descriptive adequacy is provided by Chomsky (1957), where he shows that while phrase structure grammars may describe the structure of simple declarative sentences, they cannot describe speakers’ intuitions about the meaning of wh-questions; this led to the addition of a “transformational component” to the grammars, thus providing descriptive coverage of a wider range of data. A second important point is that it is ultimately the theory that decides what the “crucial cases” are, and that descriptive adequacy does not necessitate a data-led, empiricist approach to description; indeed it is the theory that decides what the data is. All description is rooted in theoretical

presumptions to some extent, as all data needs to be understood with respect to some set of criteria; without this there is no meaningful description, just data.

To understand how fundamental theoretical assumptions lead us to the “crucial cases”, consider the logic of Chomsky’s argument for the transformational component of the grammar. Chomsky’s theoretical assumption is that constituency – the fact that some words can be grouped together into meaningful formal objects known as phrases, and combinations of phrases form sentences – is a fundamental characteristic of language. The description of phrases leads to phrase structure grammars for simple sentences; *wh*-questions do not fit into these grammars in an orderly fashion, but they still display the basic characteristics of constituency, and their meanings seem to be systematically related to declarative sentences which are produced by the phrase structure grammars, in that *wh*-questions seem to be rearranged versions of declarative sentences (with some important differences relating to focus and scope). Since the theory values evidence for constituency and its relation to meaning, the data from *wh*-questions is taken to be a “crucial case” which should lead changes to the theory; thus Chomsky’s solution is to propose a set of transformations which map one phrase structure to another in accordance with evidence for constituency and meaning. Since the transformational rules also allow for the description of numerous other reorderings of this kind (such as passivization and topicalization), the addition of the transformational component increases the descriptive adequacy of the grammar.

The theory of poetic language proposed here aims to achieve descriptive adequacy in the same manner. Chapter 2 is concerned with identifying a wide range of different kinds of examples of poetic language data and describing the linguistic intuitions of speakers in response to them and why they are ungrammatical. The categorization of the data into two large classes of examples takes the work in standard linguistic theory as its guide, taking the kinds of linguistic forms which are assumed to be “crucial cases” in theoretical syntax (in particular, movement and ellipsis) and grouping the deviation data around them. I

demonstrate that the data is much more diverse than is sometimes assumed in the literature, and in chapter 3 I show that many of the crucial cases within the large categories cannot be handled by existing theories of poetic language. I argue that the difficulties in theorizing these crucial cases dictates that a radical rethink of how poetic language relates to ordinary language is required, and in chapter 4 I set about this task of providing a theory that can account for the wide range of data adequately.

Ensuring the coverage of a wide range of data is particularly important for this particular project, as a theory of poetic language must be data-driven. While regular work in theoretical linguistics can probe the grammar by constructing example sentences and judging them for acceptability, this study only takes ungrammatical forms as its data, and cannot make reference to any kind of non-occurring poetic language by inventing examples and judging them for grammaticality. Nevertheless, the theory should still strive to address the issue of what does and does not occur in poetic language as part of its descriptive coverage, so the only practical way of providing this is to describe gaps in the data which is gathered; that is, we might expect deviation in some kinds of linguistic form which are taken to be significant in linguistic theory but find no examples of this deviation in the data from poetic language. In this respect, the methodology of the present work is modeled on that of other data-driven subtypes of generative linguistics, such as diachronic syntax or typological work.¹³ For an example of this in diachronic syntax, consider the case of verb movement in Early Modern English, as studied by Roberts (1993) among others. Roberts examines representative EME corpus and find that lexical verbs occur to the left of negation in a significant number of cases; in contrast, the lexical verb is seldom found in positions to the right of negation without the presence of auxiliaries. The latter fact is taken to be an indication that non-movement of the verb in a sentence without auxiliaries would be ungrammatical for speakers of EME, just as it is for speakers of modern French and, taken together, these

¹³For discussion of the use of negative data in linguistic typology, see Newmeyer (2005) and the references cited therein.

facts lead to the conclusion that EME had obligatory verb movement to a higher position ('Inflection' or 'Tense') like that in French.

A similar approach is pursued in this dissertation: I take the non-existence of certain poetic language forms that are in principle plausible to indicate that they should be 'ungrammatical' in the relevant sense; that is, the theory of poetic language should predict that the relevant forms will not be attested. In this respect, the theory proposed here differs from most previous theories of poetic language, as it is strongly predictive and restrictive.

1.2.2 Explanatory adequacy

In one way or another, descriptive adequacy has always been a part of research in linguistics, but the requirement of explanatory adequacy was not explicitly formulated until Chomsky's *Aspects of the theory of syntax* (1965). The shift in focus brought about by this requirement pushed linguistics in a new direction, placing at its heart the issue of learnability. The fact that children can acquire the complex set of rules of a given grammar based on the relatively small input received during acquisition (which Chomsky later called 'Plato's problem,' or the argument from the poverty of the stimulus) remains the most remarkable feature of language, and Chomsky takes it to indicate that humans must be innately endowed with some set of linguistic capabilities which guide acquisition and predetermine the possible forms of acquired grammars.¹⁴ This is the theory of Universal Grammar (UG), and it has been fundamental to work in generative linguistics since *Aspects*. Explanatory adequacy thus places tight restrictions on the proliferation of theoretical technology of linguistic theory, placing a premium on generalization of various phenomena to a limited set of principles and general interface conditions.

Since the theory of poetic language to be developed here is within the

¹⁴Specifically, the nature of the argument is that it would be wholly impossible to explain the acquisition of language in terms of non-language-specific, domain-general learning processes, and that domain-general approaches would also fail to predict the strong statistical patterns observed in acquisition. For a compelling survey of this argument, see Pinker (1995), and for an up-to-date survey of this and related arguments from Specific Language Impairment, see Wexler (2002).

purview of linguistic theory more generally, the requirements placed upon it for explanatory adequacy are the same. This means that the theory should not proliferate descriptive technology where it is not required, and it should favour unified explanations of linguistic phenomena where possible. Regarding the theory of poetic language, explanatory adequacy dictates that the theory of poetic language should be as unburdensome as possible, and it discourages us from proposing a whole new set of linguistic mechanisms that do not receive independent support. Thus we should be discouraged from introducing a whole new set of rules and operations for producing poetic language forms, if they can be derived by independently mechanisms. I discuss specific implications of this condition in chapter 3 for existing theories of poetic language.

The theory proposed in chapter 4 aims to satisfy explanatory adequacy by providing an account of poetic language that is derived entirely from linguistic operations and constraints that can be motivated independently. The proposed explanation follows naturally in a framework that assumes that syntax is an optimal solution to the sound-meaning pairing in language, such as Chomsky's (1993, 1995a,b, 2000, 2001, 2004, 2008) Minimalist program, and it identifies poetic language as a species of 'suboptimal' language that is allowed to surface under specific conditions imposed upon the output of the grammar. The explanation given in chapter 4 requires a set of specific assumptions about the form of the grammar and its operations – in particular, I assume the Y-model of Chomsky and Lasnik (1977), Chomsky (1993, 1995b, 2000), Richards (2001) and others, and a decomposition of the basic operation Merge in a similar manner to that of Hornstein (2009) – and if successful, the theory proposed here would therefore constitute empirical evidence for this set of assumptions. The details of the different parts of the argument are discussed in great detail in chapter 4, and the implications of such an analysis of poetic language are discussed in chapter 5.

1.3 The data from poetry

With the empirical scope of the study defined, I now turn to some specific questions about the data that forms the empirical basis of the theory to be developed; that is, the poetic texts that are analysed in the following chapters. As mentioned above, one of the core commitments of this study is to provide a broad a description of the character of syntactic deviation in poetry. This means analysing a large cross-section of poetic texts and extracting from them examples of a wide array of different kinds of deviation. Nevertheless, certain kinds of texts need to be excluded for methodological reasons. The focus of this study is deviant linguistic form in poetry, and the basic methodology for identifying this data is reading the relevant texts and picking out sentences that are intuitively judged by native speakers as deviant. Given this, I have chosen to focus on poetry that is written in standard dialects of Modern English, restricting the analysis to texts from the 18th century onwards. Previous studies of poetic language, such as Dillon (1975), have included poetry from the Renaissance period, but the English of the time, Early Modern English (EME), is syntactically distinct from the contemporary English in many ways (see e.g. Roberts 1993 and references cited therein), as with the case of verb movement mentioned above. It follows that speakers of Modern English (the speakers I have access to) cannot reliably judge EME texts for well-formedness in all cases, and therefore that such data cannot be used as part of the empirical basis of the present theory.¹⁵

Naturally it also follows that I cannot study poetry from non-standard dialects of English, at least those for which I do not have access to speakers, so I

¹⁵It is worth noting at this point that many previous studies, such as Dillon (1975) and many others reviewed in section 3.1, have put this issue to one side, since they largely concentrate on EME texts as part of their empirical basis. This might be justified in a number of ways; for example, Emma (1964) conducts a comprehensive analysis of both the prose and poetry of John Milton, and he shows that the S-O-V order that appears regularly in his poetry is never found in the prose, thus showing that it is a distinctive characteristic of the poetic language and not part of the ordinary ('prosaic') language use of the time for Milton. Dillon (1975) supplements some of his empirical generalizations with comments by writers from the time who have described the linguistic properties of a given form of deviation; for example, he cites a note by Henry Peacham in 1577 on an "abuse" by poets involving the inversion of prepositions and their NP complements.

have also excluded such texts from the study too, focusing on standard variants of English.¹⁶ Note that this puts to one side a number of interesting questions that may have been taken into the scope of the study; for example, how do speakers of a dialect A read and interpret structures from another dialect B that would be ungrammatical in A? It is possible that the same mechanisms are involved for the interpretation of deviant structures in poetry as for the interpretation of structures that are deviant relative to one's own dialect, and thus we might expect that the theory proposed here may be able to extend its empirical remit to cover such cases, thus creating a general theory for the interpretation of ungrammatical sentences. This is a significant claim, however, one that I cannot evaluate in the space afforded by this dissertation. For now I restrict the remit of the theory to producing an explanation of poetic language, and therefore I exclude dialectal variation for the methodological reasons cited above.

Beyond these methodologically-driven exclusions, this study takes a pluralistic and inclusive approach to data gathering. Two characteristics of this approach need to be discussed here. First, although the study is concerned with developing a theory of "poetic" language, the data focus is not restricted to genres of literature that are strictly defined as "poetry," that is, standard verse genres. Rather, I have also included poetic prose texts in the survey, in particular the late prose works of Samuel Beckett, since these texts display as much of the characteristic deviant language as many other verse texts.¹⁷ Second, as is indicated by the preceding discussion, the data includes examples from both traditional verse traditions and from experimental texts from the 20th century. This sets the present study apart from the others reviewed in chapter 3, as most previous studies have concentrated on more traditional poetries such as

¹⁶I also exclude some dialects that I do have access to, such as my own local dialect of Scottish English, since it would be difficult for the majority of readers to verify such judgments for themselves.

¹⁷Indeed some have argued that Beckett's prose is actually a form of verse that has not been broken up into lines: Perloff (1982) argues that Beckett's novella *Company* (1979) is actually written in a standard metre, and that its appearance as prose is nothing more than an artifact of its presentation.

the works of Shakespeare, Milton, Wordsworth, Shelley, Pope and Dryden; indeed the only 20th century author whose work is considered in previous works is typically e.e. cummings. I will not speculate on the possible motivations for this skewing of the empirical focus of previous theories of poetic language, but I will note that there are no compelling reasons from either linguistic or literary theory for separating the experimental and traditional in the formation of a theory of poetic language.

From the perspective of literary theory, any distinctions between ‘traditional’ and ‘experimental’ are artificial, at best a product of the cultural-historical biases of a given critical model; the literary theorist may point to the fact that many of the so-called traditional poets (like Shelley, for instance) had at one point in history been identified as experimental.¹⁸ The perspective from linguistic theory/literary linguistics is similar: there are no linguistic reasons to separate the traditional and experimental poetries when it comes to theorization of poetic language. It is true that much of the deviation in traditional poetry is brought about by inversions for the sake of rhyme and metre, and it is also true that most (but not all) experimental poetry lacks that motivation, since these texts are often written in free verse. However, in chapter 3 I show that it is not possible to model a causal relation between these inversions and their apparent motivations from rhyme and metre within a plausible linguistic theory of poetic language; rather, these aspects of form must be outside of the core operations of the grammar, and as such they are no different from non-linguistic formal conditions on the organization of a text. This is discussed in more depth in chapter 4.

On the other hand, there are a number of arguments *for* including experimental texts in the empirical coverage of the theory. Experimental authors are often concerned with creating linguistically deviant texts as a way of testing the boundaries of poetic expression. For example, in a famous letter to his

¹⁸The term ‘experimental’ might in some usages describe some aspect of the compositional practice of a given author, as with the aleatoric practice of Jackson Mac Low and other authors. Here and throughout I use the term in the more generally descriptive sense, to identify a group of works associated with the literary *avant garde*.

friend Axel Kaun, Samuel Beckett made the most explicit statement of his own conception of poetic language when he said that

language is most efficiently used where it is being most efficiently misused. As we cannot eliminate language all at once, we should at least leave nothing undone that might contribute to its falling into disrepute. To bore one hole after another in it, until what lurks behind – be it something or nothing – begins to seep through; I cannot imagine a higher goal for a writer today.

(Beckett 2001 [1937]: 171-172)

This reflects the attitudes of many other experimental authors writing in the early 20th century, as well as those of many other poets before and since, who have seen their experiments with the forms of language as a way of exploring the potential of language for creative expression. As a result, the texts written by Beckett and some of his contemporaries are rich resources of deviant linguistic forms, providing a wide range of different kinds of examples that are relevant to the present study.¹⁹ Most of the kinds of deviation found in the experimental texts are similar to some found in the traditional texts, but covering the experimental texts allows us to provide a more comprehensive picture of the kinds of deviation found in poetic language. Indeed one of the most important findings of the description in chapter 2 is that the kinds of deviation in traditional and experimental poetry are shown to be broadly similar, falling within the purview of the unified theory of poetic language without need for stipulation. This may be taken to be empirical evidence against the view that the language of experimental poetry is fundamentally different from that of more traditional literature, and thus it may form part of a counterargument to the ‘natural classicist’ position of Turner (1985), which argues that the avant garde art of the 20th century has a different relation to cognition than classical art.²⁰

The data reviewed in chapter 2 is taken from a comprehensive study of Alexander Pope’s poetry, in particular ‘The Rape of the Lock,’ and a selection

¹⁹For more discussion of theoretical motivations for the linguistic analysis of experimental literature, see Thoms (2008).

²⁰Pinker (2002) argues for a similar position, although his arguments for this position are less substantial.

from texts by various 20th century experimental authors, such as John Ashbery, Gertrude Stein, Robert Creeley, Charles Olson, Samuel Beckett and Jackson Mac Low. This only reflects a selection of the texts analysed; most of the specific phenomena that are identified in chapter 2 can be found in texts by other authors as well. Shortened references (ABBREVIATION.LINEREF) are used for frequently cited poems in accordance with a key that is provided in the frontmatter. Otherwise a longer reference (poet, poem/text title) is used, with the key in the front providing standard references to the containing volumes and page numbers for the poems/texts.

1.4 Summary and dissertation outline

In this chapter I have clarified the aims of the theory of poetic language to be developed in the following chapters. I have proposed that the theory should account for syntactic deviation in the language of poetry, and I have discussed a number of methodological issues relating to such a study. I have outlined the descriptive and explanatory goals of the theory and the data focus of the study.

The rest of the dissertation is organized as follows. Chapter 2 provides an extensive description of the kinds of syntactic deviation found in poetic texts. It is divided into two broad categories, displacement phenomena and erasure phenomena, which are defined in relation to the standard linguistic phenomena of movement and ellipsis respectively; in the chapter summary I discuss the nature of the distribution of displacement and ellipsis phenomena, noting facts that need to be explained by the theory to be developed. Chapter 3 provides an extensive critical overview of previous approaches to the theory of poetic language. I discuss variants of the dominant ‘poetic grammar’ theories and describe their theoretical and empirical inadequacies in detail. Chapter 4 provides the details of the theory of poetic language. I discuss the theoretical background and detail the ways in which the proposal deals with the empirical and theoretical challenges identified in chapters 2 and 3. Chapter 5 concludes

by discussing some additional implications of the theory.

Chapter 2

Linguistic deviation in poetry: a descriptive overview

As mentioned in the previous chapter, deviation has been taken to be one of the defining characteristics of poetic language in linguistic and literary theoretic discussions for a long time. And yet despite this emphasis, the phenomenon of deviation in poetic language has remained unsystematic and patchy. Literary theoretic descriptions have tended to avoid any serious detail in the discussion of the linguistic character of poetic deviation, while linguistic studies have typically focused upon very narrow data sets, focusing upon individual authors or specific recurrent ‘constructions’ (like poetic inversion). These studies have been unable to adequately describe the variety of poetic deviation found across literature.

In this chapter I attempt to provide a wider descriptive overview of the kinds of syntactic deviation found in poetic language in English. The data presented here is from a set of texts which provide numerous examples of different kinds of deviation. The data has been sorted into two large descriptive categories: displacement phenomena and erasure phenomena, which are defined in what follows. In the summary I describe the important generalizations that can be garnered from the description of the data, noting gaps in the paradigms and

differences between displacement and erasure in terms of their distribution. This overview then provides the empirical basis for the discussion and theorization to follow.

2.1 Displacement phenomena

This category gathers together examples in which words or phrases appear to have been displaced from their normal positions in sentences of English by non-standard means. The majority of examples in this section comes from a study of Alexander Pope's 'The Rape of the Lock,' since this particular poem is a concentrated source for a number of different kinds of relevant examples, although most of the kinds of deviation identified can be found in many other texts. Each subsection deals with the displacement of different syntactic elements, describing the distribution of that element in standard English and then showing the ways in which poetic language displacement phenomena depart from these distributional norms. I conclude the section with a summary of the overall behaviour of displacement in poetic language.

2.1.1 PP displacement

Preposition phrases (PPs) have a relatively wide distribution in English. PPs can appear within noun phrases (NPs) as in (6), and adjective phrases (APs) as in (7):

(6) The picture on John's table is beautiful.

(7) John is green with envy.

The place where PPs are found most commonly is after the main verb, within the verb phrase (VP). In some cases, like (8a), they are obligatory because they are arguments of the verb, cf. ungrammatical (8b). However, in other cases, like (9), the PPs appear optionally, in which case we refer to them as adjuncts:

(8) a. I put a book on the table.

b. *I put a book.

- (9) a. I kissed Mary.
 b. I kissed Mary on the cheek *in the nightclub* **at 2am**.

PPs like **at 2am** in (9b) can also occur sentence-initially, as in (10).

- (10) *At 2am* John will come in here for a cup of tea.

These PPs are known as ‘time adverbials’, and they differ in their distribution from the other kinds of PP-adjuncts in (9b): whereas the time adverbials modify the entire sentence, the other PP-adjuncts modify just the verb phrase. We call the former sentence- or TP-adjuncts, and the latter VP-adjuncts. VP adjuncts typically only occur to the right of the verb, just like argument PPs.

However, English has a set of well-attested operations that can move these PPs from their canonical positions at the right of the verb to sentence-initial positions. The first two are very similar and are typically grouped together in the study of movement operations in syntax: wh-movement, in (11) and topicalization, in (12):¹

- (11) To whom did you send the letter *t*?

- (12) To Mary, I sent a letter *t*.

Wh-movement is a question formation strategy² where a questioned wh-phrase like *who* is moved to the front of a sentence, and in cases where the wh-phrase originates in a PP, the wh-movement operation can take the whole PP with it; this is called ‘pied-piping.’ Pied-piping is optional in English wh-movement, as shown by (13). However, pied-piping of prepositions is obligatory in many other languages, such as German and Serbo-Croatian; thus the German and Serbo-Croatian versions of (13b) are ungrammatical (see Ross 1967):

¹The *t* in these examples marks the original position of the moved phrase, which appears underlined here; this is known as the ‘trace’ of movement.

²Wh-movement can occur in other non-question constructions, such as relative clauses. I will not go into the derivation of relative clauses here, but will simply point out that wh-containing PPs can also occur as the heads of relative clauses:

- a. The man to whom you referred earlier is at the door.

- (13) a. To whom did you send the letter *t*?
 b. Who did you send the letter to *t*?

Topicalization is a sub-type of what is known as ‘A-bar movement’, and it involves movement of a constituent to a sentence-initial position in order to make it clear that the sentence is about this particular constituent. Thus topicalization typically occurs with some degree of stress and an intonational break before the rest of the sentence (as indicated by the punctuation in (12); this is known as ‘comma intonation’). Topicalization is similar to wh-movement, in that it can also optionally pied-pipe the whole containing PP in some languages but not others.

Both of these types of movement are subject to a number of constraints; although the nature of these restrictions is a massive topic of research in contemporary linguistics, I will only survey a few here for illustration. First, these movement operations only apply to full constituents or phrases; they cannot apply partially, and they cannot strand sub-parts:

- (14) a. *To which did you go *t* party?
 b. To which party did you go?
 c. *About whose were you talking *t* mother earlier on?
 d. About whose mother were you talking *t* earlier on?
- (15) a. *To John’s, we went *t* party.
 b. To John’s party, we went *t*.
 c. *About that Bergman, I can talk *t* film all day.
 d. About that Bergman film, I can talk *t* all day.

Second, although it may seem that almost any wh-words can be moved, in fact movement can only occur from certain syntactic environments. These non-movement environments are known as ‘islands’ (Ross 1967), and two such environments are given below; I give the well-formed non-movement equivalents as (c)-sentences for clarity:³

- (16) a. *Who did you send a letter to Mary and *t*?

³The examples here involve movement of non-PP wh-phrases since this data is clearer, but the same holds for wh-movement (with or without pied-piping) targeting PPs.

- b. *John, I sent a letter to Mary and *t*.
 - c. I sent a letter to Mary and someone else. *Island: coordinate structure*
- (17)
- a. *Who did you deny the claim that John likes *t*?
 - b. *Mary, I denied the claim that John likes *t*.
 - c. I denied the claim that John likes Mary's sister. *Island: NP*

Third, these operations can only occur once in a given sentence in English: we do not get multiple instances of wh-movement or topicalization. Some speakers allow for the co-occurrence of topicalization and wh-movement (Culicover 2009: 361), as in examples like (20b)⁴, but topicalization below the wh-phrase, like in (20a) is never allowed:

- (18)
- a. *Who what did you give *t t*?
 - b. Who did you give *t* what?
- (19)
- a. ?*To John, a book, I gave.
 - b. To John, I gave a book *t*.
 - c. A book, I gave *t* to John.
- (20)
- a. *What, to John, did you give *t t*?
 - b. ?/??To John, what did you give *t t*?

Thus we can see that both topicalization and wh-movement are restricted in a number of ways, and that they only apply in certain situations. This means that PPs can only be found in the non-canonical sentence-initial position in a restricted set of situations.

One final situation in which a PP can appear in a position to the left of the main verb is a particular construction which is typically referred to as ‘locative inversion’. This is demonstrated by (21a); (21b) shows the ‘uninverted’ form, from which we assume the inverted form is derived:

- (21)
- a. On the stage stood an actress.
 - b. An actress stood on the stage.

⁴Judgments vary on this kind of data. Culicover reports examples like this as fully grammatical, but I find them very poor, and most speakers I have consulted have agreed. Nevertheless I will ignore examples of this form in what follows to avoid handling controversial data.

The movement operation that derives (21a) from (21b) is different from that seen in *wh*-movement and topicalization, as it is much more restricted: PPs can only move to this subject position with certain verbs, within a single clause, and the inverted structures are subject to a number of restrictions that are not experienced by the non-inverted structures; for example, *wh*-movement is not allowed out of a locative inversion structure:

- (22) a. *Who does on the stage stand *t*?
 b. *Why does on the stage stand an actress *t*?

I will not go into the specifics of locative inversions, but will simply point out that they are another example of a situation where a PP can be found in a non-canonical position.

In standard English PPs can occur in this limited set of configurations. However, in poetic texts, this distribution is widened significantly. For example, in ‘The Rape of the Lock’ PPs that would normally occur after the verb often appear in a position immediately to the left of the verb. Here is a small sample:

- (23) What dire offence from am’rous causes springs,
 What mighty contests rise from trivial things, *TR1: 1-2*
- (24) Know then, unnumber’d Spirits round thee fly,
 The light Militia of the the lower Sky; *TR1: 41-42*
- (25) A heav’nly Image in the glass appears,
 To that she bends, to that her eyes she rears; *TR1: 125-126*
- (26) Know farther yet; whoever fair and chaste
 Rejects Mankind, is by some Sylph embrac’d: *TR1: 67-68*
- (27) For lo! the board with cups and spoons is crown’d,
 The berries crackle, and the mill turns round; *TR3: 105-106*
- (28) Fair nymphs, and well-drest Youths around her shone,
 But ev’ry eye was fixed on her alone. *TR2: 5-6*
- (29) He summons strait his Denizens of air;
 The lucid squadrons round the sails repair: *TR2: 55-56*

In (23), the PP *from am'rous causes* is interpreted as a modifier of the verb *springs*, and as such it would normally occur within the VP, in a post-verbal position in the surface string; yet in (23), the PP has been displaced to a position between the subject and verb. Yet this does not resemble topicalization or any of the other typical displacement operations, since these operations always move the PP to a sentence-initial position. We see the same again in (24)-(29), and this is seen throughout Pope's poetry and in many other poetic texts of his period and others that have followed; (30) is a simple example from John Ashbery's poetry, and (31)-(32) are from Robert Creeley:

- (30) Children in the street
 Watch him go by.
 "Is that the thinnest shadow?"
 They to one another cry.

John Ashbery, 'The Thinnest Shadow'

- (31) "When they were
 young in Kentucky
 a man to freedom
 took them in a cave..."

Robert Creeley, "Follow The Drinking Gourd"

- (32) one, the smallest,
to the water goes.

Robert Creeley, 'Beach'

What is interesting about these examples is they are not all obviously motivated by metre or rhyme like the Pope examples; rather, displacement of the PP *to freedom* is arbitrary. This kind of arbitrary reordering is found throughout a great deal of more experimental modern poetry.

All of these examples have a S(subject)-PP-V(erb) configuration, apparently derived by displacement of a PP from a standard S-V-PP one. This PP displacement occurs in many other configurations in poetic texts. For example, sometimes one PP is fronted to the pre-verbal position while another VP-internal PP stays in situ; this is a S-PP-V-PP configuration, and it is demonstrated by (33):

- (33) The merchant from th'Exchange returned in peace,
And the long labours of the Toilet cease. TR3: 23-24

In (33), the PP *from th'Exchange* can be interpreted as a modifier of the verb *returned*. What is interesting about this example is that this is in fact ambiguous: the PP could also be interpreted as a modifier of the subject, *The merchant*; that is, the line could mean that a certain merchant who was from the Exchange could have returned in peace from some unspecified place. An example like (33) would not be ambiguous in a non-poetic context, since this kind of PP-displacement does not occur in standard English usage; however in Pope's text, where this PP-displacement is extremely common, the second reading where the PP modifies the verb is available, and in the local context of that part of the poem, this reading is even preferred.

PP-displacement to a pre-verbal position also happens when there are other VP-internal elements too. In (34) the PP is shifted while an *as*-adverbial stays within the VP:

- (34) Boast not my Fall (he cry'd) insulting Foe!
Thou by some other shalt be laid as low. TR5: 97-98

PP-displacement to the pre-verbal position also co-occurs with VP-internal objects, thus deriving S-PP-V-O configurations:

- (35) Sol thro' white curtains shot a tim'rous ray,
And ope'd those eyes that must eclipse the day: TR1: 13-14
- (36) Close by those meads, for ev'r crown'd with flow'rs,
Where Thames with pride surveys his rising tow'rs, TR3: 1-2
- (37) So Ladies in Romance assist their Knight,
Present the spear, and arm him for the fight. TR3: 129-130
- (38) What Time would spare, from steel receives its date,
And monuments, like men, submit to fate! TR3: 171-172
- (39) Triumphant Umbriel on a Sconce's Height
Clapp'd his glad Wings, and sate to view the Fight, TR5: 53-54

In (35) the PP *thro' white curtains* has been displaced from the VP to a position following the subject *Sol* (here, the Sun personified), where it would normally have occurred after the object *a tim'rous way*, and similarly in the other examples. (37) resembles (33) in a way, since the displacement of the PP seems to introduce an ambiguity, in that the PP *in Romance* could also modify the subject NP *Ladies*, although this reading may be somewhat unusual in terms of the real-world meaning of the sentence in the context of the text. Note however that this reading would be the only possible reading in a non-poetic situation, since this kind of displacement is not found in normal usage.

PP-displacement occurs from constituents other than VPs. We saw earlier that PPs can occur within APs and NPs; the following examples show that PPs can be displaced from these constituents too. (40)-(46) involve PPs being displaced from Past Participle Phrases (PartPs), which have similar syntactic properties to standard APs:

- (40) If e'er one Vision touch'd thy infant thought,
Of all the Nurse and all the Priest have taught;
Of airy Elves by moonlight shadows seen, *TR1: 29-30*
- (41) And now, unveil'd, the Toilet stands display'd,
Each silver Vase in mystic order laid. *TR1: 121-122*
- (42) Some secret truths, from Learned Pride conceal'd,
To Maids alone and Children are reveal'd: *TR1: 37-38*
- (43) What boots the regal circle on his head,
His giant limbs, in state unwieldy spread; *TR3: 71-72*
- (44) Clubs, Diamonds, Hearts, in wild disorder seen,
With throngs of promiscuous strow the level green. *TR3: 79-80*
- (45) But this bold Lord, with manly Strength endu'd,
She with one Finger and a Thumb subdu'd: *TR5: 79-80*
- (46) There broken Vows, and Death-bed Alms are found,
And Lovers' Hearts with Ends of Riband bound; *TR5: 117-118*

In (40), the PP *by moonlight shadow* is an argument of the deverbal adjective *seen*, which is modifying the noun *Elves*, and the PP has been displaced from within the PartP to a position before the deverbal adjective. The same occurs in the other examples, too; interestingly, there is also reordering within the shifted PP in (43), in which the noun *state* and its adjectival modifier *unwieldy* are reordered (see section 2.1.3 for discussion of related examples).

Importantly, PPs are displaced to a position between the subject and the main verb in situations where some other VP-internal argument has been displaced to a sentence-initial position. In (47), the object *this* appears in the sentence-initial position while the PP *from the Mall* is fronted to a pre-verbal position:

- (47) *This* the Beau-monde shall from the Mall survey,
 And hail with Musick its propitious Ray.
 This the blest Lover shall for Venus take,
 And send up Vows from Rosamonda's Lake. *TR5: 133-136*

The fronting of the object to the sentence-initial position in this example may be analysed as instances of topicalization, though it need not be analysed as such, since it is not clear that the example has the same semantic and prosodic properties normally assigned to topicalization sentences (see section 3.1.1). What is important is that we know already that English does not normally allow for double topicalization, so it seems unlikely that the PP-fronting to a pre-verbal position is probably not a sub-type of topicalization.

One thing that becomes apparent from these examples is that the PP-fronting seen in these examples is not to a single designated pre-verbal position, as we can see from the above examples that the PP can appear before or after auxiliary verbs: in (26) and (47), the PPs occur after the auxiliary, in a position immediately before the main verb in the surface string, whereas in (27) and (34) the PPs occur before all the auxiliaries. The following demonstrate this variation further:

- (48) See the sole bliss Heav'n could on all bestow!
 Which who but feels can taste, but thinks can know: *ES4: 327-328*

- (49) No common Weapons in their Hands are found,
Like Gods they fight, nor dread a mortal Wound. *TR5: 43-44*
- (50) Unnumber'd throngs on ev'ry side are seen,
Of bodies chang'd to various forms by Spleen. *TR4: 47-48*
- (51) Amid the circle, on the gilded mast,
Superior by the head, was Ariel plac'd; *TR2: 69-70*

(48) is similar to (26) and (47), with *on all* being fronted below the modal auxiliary *could*. In (49)-(50), on the other hand, the PPs are fronted over the auxiliary *are*.

(51) is particularly interesting since it involves PP-fronting in a locative inversion construction, where the subject is the PP *amid the circle*. The PP *on the gilded mask* is fronted over the auxiliary *may* from within the VP, and this co-occurs with the parenthetical *superior by the head*. We saw earlier that locative inversion typically does not allow for extraction from within the VP in standard English (for reasons that are not fully understood⁵). As with the examples of fronting alongside object topicalization, this example also seems to indicate that the PP-fronting mechanism here is unlike other standard fronting devices used in English. Interestingly, there are other situations in Pope's poetry where the restrictions that normally apply to locative inversion are not observed; for example, in (52) we see that a wh-question is formed with the locative inversion construction, even though this typically disallows for such question formation:

- (52) Why round our Coaches crowd the white-gloved Beaux,
Why bows the Side-box from its inmost Rows? *TR5: 13-14*

Finally, we can also see that in (51) the NP argument *Ariel* has been fronted from within the VP to a position between the auxiliary and the main verb *placed*, much like the PP-fronting in the other examples discussed here. In the next section we will see that NP-fronting of this kind is just as common

⁵For discussion see Bresnan (1994) and Collins (1997).

and diverse as PP-fronting, even though NPs have a slightly more restricted distribution than NPs in standard English.

Before moving on, it is worth pointing out that PP-fronting is not just restricted to finite clauses like those seen above; rather, we also get PP-fronting from within infinitival clauses. (53) demonstrates this:

- (53) *Thy voice* I seem in ev'ry hymn to hear,
With ev'ry bead I drop too soft a tear. *EL: 269-270*

Here both the object NP *thy voice* the PP *in ev'ry hymn* are fronted from within the embedded infinitival clause *to hear*. Topicalization from embedded infinitives is typically allowed in English, but PP-fronting into intermediate positions is just as unusual in infinitives as it is in finite clauses. There are numerous examples of fronting of PPs from within infinitives to intermediate positions in the clause:

- (54) The graver Prude sinks downward to a Gnome,
In search of mischief still on Earth to roam. *TR1: 63-64*
- (55) Resolv'd to win, he meditates the way,
By force to ravish, or by fraud betray; *TR2: 31-32*
- (56) Strait the three bands prepare in arms to join,
Each ban the number of the sacred nine. *TR3: 29-30*
- (57) Nor fear'd the Chief th' unequal Fight to try,
Who sought no more than on his Foe to die. *TR5: 77-78*

In (54) the PP *on Earth* is fronted to a position between the infinitival auxiliary *to* and the adverbial *still* within the NP headed by *mischief*, and (55) is similar in that the PP is fronted from an infinitive that is modifying a nominal. In (56) the PP is fronted to a position between infinitival *to* and the embedding verb *prepare*, and in (57) the PP is fronted in a *than*-complement in a comparative clause. What this shows is that PP-fronting occurs with various different kinds of infinitival complements, and that it is typically to a position between the embedding element and the infinitival auxiliary *to*.

There are a number of examples where PPs are displaced to the sentence-initial position in ways that would not normally be allowed in standard English. Perhaps the most unusual examples are when the PP arguments of nominals are fronted out of the NPs:

- (58) Th' embroider'd King who shows but half his face,
 And his refulgent Queen, with pow'rs combined,
Of broken troops an easy conquest find. *TR3: 76-78*

- (59) of western New York state
 were the graves all right in their bushings
 was there a note of panic in the August air [...]
John Ashbery, 'As You Came From The Holy Land'

In (58) the PP *of broken troops* is interpreted as an argument of the nominal *an easy conquest*, yet it has been displaced from its normal position following the head noun *conquest*; note that this larger NP has itself been displaced to a pre-verbal position, since the NP is interpreted as the object of the verb *find*; this kind of double displacement is discussed further below. In (59), the PP *of western New York state* occurs at the beginning of the sentence (and the poem), and it seems to be interpreted as an argument of the nominal *the graves*, since it does not modify any other NP. NPs are typically understood to be islands for extraction (as shown by (17) earlier), so these cannot be examples of standard topicalization. Both of these examples are unusual for other reasons too: the nominal in (58) has already been displaced to a pre-verbal position below the subject and parenthetical; the nominal in (59) appears after the auxiliary *were* for reasons that are not clear (there are inversion structures that allow for this but (59) does not seem to be such a structure).

Note that there are also similar cases of argument PPs being separated from the nominals that they modify by apparent rightward displacement:

- (60) Words drip from the wound
 Spring mounts in me
of dandelion– lots of it
John Ashbery, 'Rain'

Although the meaning here isn't clear, it seems that the PP *of dandelion* is a PP modifier of *Spring*, yet it appears in a sentence-final position, separated from the nominal in the subject position. While there are standard movement operations that can move phrases rightwards to a sentence-final position, these 'extraposition' operations typically only apply to sentence adjuncts, and not to PP arguments moved from within subject NPs. Therefore the displacement in (60) is similar to those in (58) and (59), but instead the PP is shifted in the other direction.

There are other, less unusual examples of PP-displacement to the sentence-initial position which nevertheless could not be analysed as standard topicalization. There are also many examples of PPs being fronted in *wh*-questions and relative clauses which, as was pointed out earlier, also typically exclude topicalization:

(61) There stands a structure of majestic frame,
Which from the neighb'ring Hampton takes its name. *TR3: 3-4*

(62) But when to mischief mortals bend their will,
How soon they find fit instruments of ill? *TR3: 125-126*

(63) When what t'oblivion better were resign'd
Is hung on high to poison half mankind. *ES4: 251-252*

Similarly in *yes-no* questions:

(64) In tasks so bold, can little men engage,
And in soft bosoms dwell such mighty rage? *TR1: 11-12*

There are also examples in *if*-clauses:

(65) If to her share some female errors fall,
Look on her face, and you'll forget 'em all. *TR2: 17-18*

(66) But by your fathers' worth if yours you rate,
Count me those only who were good and great. *ES4: 209-210*

We can see throughout the examples that there is variation in where PPs are fronted to, as in some situations it is to a sentence-initial position above the

highest element in the complementizer area (the *wh*-phrase, raised auxiliary, *if* complementizer), but in others it is to a position just below these elements. What is important is that all these examples display fronting of PPs to positions in environments where topicalization would not be allowed. These are thus further examples of the kind of nonstandard displacement operations that have been documented so far in this chapter.

Finally, we can see that it is not just full PPs that can be displaced, but intransitive PPs too. This is demonstrated by the following examples from Gertrude Stein:

- (67) If there is no dirt in a pin and there can be none scarcely, if there is not then the place is the same as up standing.

TB: 9

- (68) That is spread, it shuts and it lifts and awkwardly not awkwardly the centre is in standing.

TB: 9

In (67), the prepositional particle *up* has inverted with the verb *standing*, and similarly with the particle *in* in (68): these particles standardly occur to the right of the verbs, so in these examples they seem to have been displaced leftward. These two examples appear within the same short prose poem, displaying a sort of parallelism, and this parallelism helps to guide the reader into interpreting the constructions in this way, since it may be easy for these interpretations to be lost in the mess of the rest of the prose's oddities, not least since the displacement of verb particles is unlike any other sort of phrasal displacement operations; that is, these examples cannot be plausibly analysed as sub-species of topicalization.

2.1.2 NP displacement

The distribution of NPs is different from that of PPs, and the specifics of NP-distribution are tied to certain aspects of their semantic interpretation in a given sentence. Consider the following sentence:

(69) The man is kissing the woman.

The two NPs in (69) have different grammatical roles: *the man* is the subject, the person who does the kissing, and *the woman* is the object, the person who was kissed. Speakers of English know that the subject typically occurs at the beginning of the sentence, to the extent that we would typically call this the ‘subject position’; this is the case even though the logical subject of a sentence need not always occur in this position, as evidenced by (70):

(70) The woman is being kissed by the man.

We also know that the definition of terms like ‘subject’ and ‘object’ in terms like ‘the agent who performs the action’ are also only contingent, given sentences like (71)-(72):

(71) Bill worries John.

(72) Bill amuses John.

These examples involve predicates that are known as ‘psych verbs,’ since they describe the psychological states of participants, and they pose immediate problems for simple accounts of subjecthood in terms of agency or action. With these verbs, no actions are performed and no agents are involved, but rather the predicates describe the states of experiencers of psychological states, as caused by their experience of certain actors or objects in their environment; in each case, *Bill* could be replaced by *the economy* or *darkness*. In both examples the post-verbal argument *John* is the experiencer of the psychological state, and the argument in the subject position performs no action and does nothing to contribute to the situation other than exist. As such, these kinds of predicates cause immediate problems for simple definitions of what makes a subject argument.⁶

Yet despite any apparent problems in defining subjecthood etc, we know that the lexical meaning of these verbs all bring with them knowledge about which kind of argument should appear in which position; for example, we know

⁶For discussion of psych predicates and their implications for theories of subjecthood and argument linking, see Belletti and Rizzi (1988), Pesetsky (1995) and references cited therein.

that the object of *worries* in a sentence like (71) needs to be an animate human, and that it cannot be something like *the padlock*, except in some (Disney-like) situation where *the padlock* has been contextually assigned these properties (for example, *the padlock* is the name of some committee who decide on what Bill should do); on the other hand, the subject of *worries* can be anything capable of causing worry, be it *Bill*, *the economy* or *the frying pan*. Thus the lexical meaning of verbs typically tells us a lot about which arguments are to be associated with which positions, and hence we are provided with information about which arguments should be where from interpretation.

As with PPs, NPs can also be moved from their standard positions by wh-movement and topicalization, as we saw with some of the other examples in the previous section, and they are subject to the same restrictions, with some minor exceptions. There are a number of other movement operations that involve NPs in English, the passivization in (70) being one of them, but I will not survey their details here, accepting instead the fairly uncontroversial assumption that the lexical meaning of the relevant verbs, their verbal morphology and auxiliaries will thus determine which NPs occur where in the relevant surface strings. That is, we know from the morphological form of the verb and the presence of the auxiliary *be* in (70) that the argument in the subject position is a passivized object.⁷

We will see in this section that NPs are displaced from their standard positions just like PPs in literary examples. Although the distribution of NPs and PPs in standard English differ in some ways – PPs can only appear in the subject position in a limited number of cases, NPs cannot appear as TP-adjuncts – their distributions with respect to displacement in poetic language are very similar. For example, there is a significant amount of NP-displacement from VP-internal positions (i.e. of objects) to the pre-verbal position, as demonstrated

⁷Which is to say I put A-movement to one side in this discussion, since it belongs the more technical realms of syntactic discussion: most non-linguist speakers would not say that an argument that has in fact undergone A-movement has been displaced, since A-moved arguments are only every pronounced in their post-A-movement positions. This is unlike A-bar movement, which involves clear and detectable displacement of arguments from their standard pronunciation positions.

by the following examples:

- (73) Belinda still her downy pillow prest,
Her guardian SYLPH prolong'd the balmy rest TR1: 19-20
- (74) What tender maid but must a victim fall
To one man's Treat, but for another's Ball? TR1: 95-96
- (75) This casket India's glowing gems unlocks,
And all Arabia breathes from yonder box. TR1: 133-134
- (76) Her lively looks a sprightly mind disclose,
Quick as her eyes, and as unfix'd as those; TR2: 9-10
- (77) Fair tresses man's imperial race ensnare,
And beauty draws us with a single hair. TR2: 27-28
- (78) Th'advent'rous Baron the bright locks admir'd;
He saw, he wish'd, and to the prize aspir'd. TR2: 29-30

In (73) the first line is clearly interpreted as *Belinda still prest her downy pillow*, perhaps with some ambiguity about the modification by the adverb *still*; given this, it seems that the object of the verb *prest* is *her downy pillow*, and since this occurs to the left of the verb, we can say that it has been displaced. The same occurs in the other examples, where an NP that is unambiguously the object of the verb is displaced from its standard postverbal position to a position preceding the verb. Just as with PP-shifting, this cannot be analysed as an instance of topicalization, since NP-topicalization is always to a sentence-initial position.

As with the PPs, we can see that this displacement is not to a designated pre-verbal position, since we see variation in the order of the displaced objects and auxiliaries. In (74) we see the object *a victim* occurring between the auxiliary *must* and the main verb *fall*, and we saw a similar pattern in (51) earlier. There are further examples of this order:

- (79) Gums and Pomatums shall his flight restrain,
While clog'd he beats his silken wings in vain; TR2: 129-130

- (80) Steel could the works of mortal pride confound,
And hew triumphal arches to the ground. *TR3: 173-174*
- (81) In God's, one single can its end produce,
Yet serve to second too some other use: *ES1: 55-56*
- (82) His safety must his liberty restrain:
All join to guard what each desires to gain. *ES3: 277-278*
- (83) Here thou, great ANNA! whom three realms obey,
Dost sometimes counsel take—and sometimes Tea. *TR3: 7-8*

The following is a similar example from Robert Creeley:

- (84) Can't myself
let off this
fiction. "You
don't exist,

baby, you're
dead."
[*italics in original*] *Robert Creeley, 'Echo of'*

(79)-(82) and (84) demonstrate examples with modal verbs again, and (83) is an example from Pope of the object appearing after the dummy auxiliary *do* in an older morphological form. This form was out of use in the modern usage of Pope's era, but it is used here apparently in imitation of older forms of English. The distribution of dummy *do* is similar to that of modal verbs, so we may surmise that the position of the shifted object in (83) is similar to that in the others. Note that there are no obvious regularities in the properties of the NPs that are displaced, with respect to phonological 'weight,' (in)definiteness, pronominal weakness or any other linguistic property that sometimes condition movement operations in syntax.⁸

We also see the S-O-aux-V order as well in the following examples:

⁸Such conditions do apply with other syntactic operations: Heavy NP Shift (discussed later in this section) is conditioned by phonological weight in English; Object Shift (which these examples resemble superficially) in the Germanic languages is sometimes conditioned by (in)definiteness; weakness vs strongness conditions object cliticization in Romance.

- (85) Self-love, the spring of motion, acts the soul;
Reason's comparing balance rules the whole.
Man but for that no action could attend,
And but for this were active to no end: *ES2: 59-62*
- (86) In vain thy Reason finer webs shall draw,
Entangle justice in her net of law, *ES3: 191-192*
- (87) And sure, if fate some future bard shall join
In sad similitude of griefs to mine,
Condemn'd whole years in absence to deplore,
And image charms he must behold no more;
Such if there be, who loves so long, so well;
Let him our sad, our tender story tell; *EL: 359-364*

The NP-displacement in (85) is to a position in front of the auxiliary *could*, and it co-occurs with the parenthetical *but for that* between the subject and auxiliary. The end result is somewhat chaotic but still interpretable, and what examples like this demonstrate is that there can be multiple disruptions to word order without complete distortion of the meaning of the sentence.

In (86) the NP-displacement over the auxiliary *shall* co-occurs with the PP *in vain* in the sentence-initial position. This does not seem to be an example of double displacement from the VP, as PPs like *in vain* can typically occur in such positions without topicalization-type intonation or discourse context; these are TP-adjuncts (like the time adverbials discussed in section 2.1.1), and they do not get a movement analysis. However, there are numerous other examples of double displacement where this kind of non-movement analysis is not available. We saw already with (51) that PP-fronting to the pre-verbal position can co-occur with displacement of the object, and it seems that almost all the possible permutations of VP-argument reordering are found in the sample texts. The following example demonstrates double fronting to a pre-verbal position:

- (88) Beauties *in vain* their pretty Eyes may roll;
Charms strike the Sight, but Merit wins the Soul. *TR5: 33-34*

Here the PP *in vain* must have been moved from some other position like the VP, since it is not occurring in a sentence-initial TP-adjunct but in the pre-

verbal position. It is also of interest that, if the two arguments were to occur within the VP, the most natural order would be the reverse of that seen in the pre-verbal position i.e. V-O-PP, *roll their pretty Eyes in vain*. The two shifted phrases also precede the auxiliary *may*, again demonstrating the proliferation of positions for displacement.

The same S-PP-O-V reordering is also seen in the following examples:

- (89) Love *in these labyrinths* his slaves detains,
And mighty hearts are held in slender chains. *TR2: 23-24*
- (90) Some *to the sun* their insect-wings unfold,
Waft on the breeze, or sink in clouds of gold; *TR2: 59-60*
- (91) Some *o'er her lap* their careful plumes display'd,
Trembling, and conscious of the rich brocade. *TR3: 115-116*

And here is an example from Creeley:

- (92) My mind
to me a nightmare is—
Robert Creeley, 'Desultory Days'

(89) and (92) are similar to (88) in that the two displaced phrases have changed order when they have been displaced, since the most natural order for the standard VP would be *detains his slaves in these labyrinths*. This is not necessarily the case in (90)-(91), as they may be reordered in the VP more freely.

There are also examples of double displacement with the S-O-PP-V order:

- (93) Remembrance and reflection how allied!
What thin partitions *Sense* from Thought divide! *ES1: 225-226*
- (94) A nymph there is, that all thy pow'r disdains,
And *thousands more* in equal mirth maintains. *TR4: 65-66*

(93) preserves the order of the two fronted phrase *Sense* and *from Thought*, and we might assume that, in such situations, the two have been fronted as one unit, although it should be noted that such a process would be implausible with a syntactic operation like topicalization, since they do not form a constituent

within the VP. In (94) we see double fronting within the second of a pair of coordinated VPs within the relative clause. There is object fronting in both VPs, since *all thy pow'r* is interpreted as the object of *disdains* in the first constituent, but there is an additional fronting of the PP *in equal mirth* to the pre-verbal position below the coordinating conjunction in the second conjunct. There are numerous other examples of displacement within coordinated VPs, but no real consistency in the data; that is, in some examples, both conjuncts have parallel displacement, but in many others this is not the case, in that one conjunct may have displacement but the other does not.

There are further permutations of double displacement in the sample texts. We have already seen cases where displacement of a PP to the pre-verbal position co-occurs with (apparent) object topicalization; as noted above, this makes it seem unlikely that the PP displacement is a sub-species of topicalization, since English typically lacks double topicalization. Yet we see in the texts that there are numerous instances of double displacement. For example, the following examples show two phrases being displaced to a sentence-initial position, with the PP preceding the object:⁹

- (95) A heav'nly Image in the glass appears,
To that she bends, *to that her eyes she rears;* *TR1: 125-126*
- (96) *On her white breast a sparkling Cross she wore,*
Which Jews might kiss, and Infidels adore. *TR2: 7-8*
- (97) Nay oft, *in dreams, invention* we bestow,
To change a Flounce, or add a Furbelow. *TR2: 99-100*
- (98) *With anxious beating hearts the dire event they wait,*
Anxious, and trembling for the birth of Fate. *TR2: 141-142*
- (99) *In various talk th' instructive hours they past,*
Who gave the ball, or paid the visit last: *TR3: 11-12*

⁹Note that (95) is the same example as (25) above, but it has been renumbered since it is used to make a separate data point, as indicated by the underlining. The same practice is used throughout the rest of the dissertation.

- (100) Full *o'er their Heads* the swelling Bag he rent,
And all the Furies issu'd at the Vent. TR4: 91-92

In the following examples, the object precedes the PP in the sentence-initial position:

- (101) *The drops to thee*, *Brilliante*, we consign;
And *Momentilla*, let the watch be thine; TR2: 113-114

- (102) *This just behind Belinda's neck* he spread,
As *o'er the fragment* she bends her head. TR3: 133-134

Again there is no real consistency in the displacement with respect to relative order: in (25)-(100) the displaced phrases seem to be reversed in their relative order, yet in (101)-(102) the relative order has been preserved. An interesting aspect of (101) is that the displaced phrase *the drops* seems to be modified by the parenthetical *Brilliante*, which occurs after the displaced PP *to thee*; this seems to indicate that the parenthetical has been displaced with the NP, and that the PP has been subsequently displaced to a position between the two. However, there is no way to know whether this is how such a word order would be produced; rather, all that such an example shows is that the ways in which parenthetical modifiers are attached to the phrases they modify can be interrupted by displaced phrases, and this is unlike how standard movement operations work in natural language.

There are even examples where both an AP complement and an object are displaced to the sentence-initial position:

- (103) *Sunk in Thalestris' Arms* the Nymph he found,
Her Eyes dejected, and her Hair unbound. TR4: 89-90

In this example, the AP *Sunk in Thalestris' Arms* forms a small clause with the NP *the Nymph*, and both are displaced, in reverse order, to the beginning of the sentence.

At this point it is worth noting that, in all of these examples of double displacement, the subject of the sentences that contain the displacement are

nominative pronouns, like *we*, *she* and *they*. They are thus clearly identifiable as the subjects of the sentences, since nominative case only shows up on pronoun subjects in English. This helps to make the sentences interpretable, since in the absence of overt subject marking it may become increasingly difficult to identify which phrase bears which grammatical role in the presence of such significant reordering of the arguments.

As with PP displacement, we also see NP displacement in a number of contexts other than finite clauses. For example, there are numerous examples of displacement in infinitival clauses, where an NP is moved to an intermediate position between the embedding element and infinitival *to*:

- (104) 'Tis these that early taint the female soul,
 Instruct the eyes of young Coquettes to roll,
 Teach Infant-cheeks a bidden blush to know,
 And little hearts to flutter at a Beau. TR1: 87-90
- (105) Coffee, (which makes the politicians wise,
 And see thro' all things with his half-shut eyes),
 Sent up in vapours to the Baron's brain
 New stratagem, the radiant Lock to gain. TR3: 117-120
- (106) Was it for this you took such constant Care
The Bodkin, Comb and Essence to prepare? TR4: 97-98
- (107) Nor fear'd the Chief th' unequal Fight to try,
 Who sought no more than on his Foe to die. TR5: 77-78
- (108) The same, his ancient Personage to deck,
 Her great great Grandsire wore about his Neck TR5: 89-90

In (104), the NP object *a bidden blush* is displaced from its position within the embedded VP *to know* to a position between the infinitival auxiliary and *Infant-cheeks*, the object of the matrix embedding verb *teach*. In (105) and (106) the infinitive is within an NP and the displaced objects occur immediately following the embedding NP. (106) is particularly interesting since not only one NP but a set of three coordinated NPs are shifted to this position; depending on one's

syntax of coordination, this may be seen as NP displacement or &P displacement. (107) and (108) are more unusual still: in (107) the NP *th' unequal fight* seems to be interpreted as the object of *to try*, but the embedding construction *Nor fear'd the Chief* is itself somewhat disordered; in (108) it seems that *his ancient Personage* should be interpreted as the object of *to deck*, but there is no obvious embedding element for the infinitive. While these examples do get interpretations, we can see that the displacement of NPs can become increasingly disruptive when it co-occurs with other sorts of disruptions.

NP-displacement is seen in numerous different kinds of infinitives. There are cases where the NP is fronted within an infinitival subject, such as (109):

- (109) Warn'd by the Sylph, oh pious maid, beware!
 This to disclose is all thy guardian can: TR1: 112-113

Here the sentence is interpreted as *to disclose this is all thy guardian can do*, where the infinitive is the subject, but the object of the embedded infinitival verb, *this*, has been displaced from within the infinitive to a position preceding the infinitival auxiliary. This cannot be analysed as an example of topicalization, since topicalization or any other kind of extracted is not normally allowed from within subjects. Subjects are typically described as “derived islands,” part of the class of extraction-disallowing environments described at the beginning of the previous section.

The following example also demonstrates displacement from within a bare *to*-less infinitive, one that follows the verb *dare*:

- (110) The rebel Knave, who dares his prince engage,
 Proves the just victim of his royal rage. TR3: 59:60

This displacement occurs within a relative clause to some position between the main verb and the infinitival form *engage*. Note that an example of this sort might appear ambiguous in some other contexts, since *dare* can also take an object and an infinitival complement, as in a sentence like *John dared Bill to try the water*. In such a sentence, *Bill* would be the object of the embedding verb *dare* and it would also control the subject of the infinitive *to try*, and it would

contrast with the interpretation of (110), where the displaced NP *his prince* is interpreted as the object of the embedded verb *engage*, i.e. as the person who is engaged with. (110) is unambiguous because the infinitive appears without infinitival *to*, since *dare* with a bare infinitive complement does not occur with an object; thus *John dared Bill try the water* is ungrammatical. We will see later that grammatical particles like infinitival *to* are often missed out in poetic texts, so while (110) may be unambiguous in its own context, similar examples may have different interpretations in other contexts.

We saw earlier in this section that NPs that are displaced to a pre-verbal position can either precede or follow an auxiliary verb, and in section 2.1.1 we saw that, while the same holds for PP displacement in finite clauses, PPs displaced from within infinitives only tend to occur before the infinitival auxiliary *to*. However, it turns out that there are also some examples of NP-displacement to a position between the infinitival auxiliary *to* and the non-finite form of the verb, such as (111) below:

- (111) Some nymphs there are, too conscious of their face,
 For life predestin'd to the Gnomes embrace. *TR1: 79-80*

Infinitival *to* here would typically occur immediately adjacent to the verb *embrace*, but here the object *the Gnomes* has been displaced to a position between the auxiliary and verb. This shows that, even with infinitives, there is no consistent position to which phrases are displaced in the intermediate field between the VP and the subject; rather, NP displacement seems to occur freely to almost any position.

Unusual NP displacement is found in a number of other contexts too. For example, in the following examples the NP objects and PPs are reordered after the verbs:

- (112) The nymph exulting fills *with shouts* the sky;
 The walls, the woods, and long canals reply. *TR3: 99-100*

- (113) The rising tempest puts *in act* the soul,
 Parts it may ravage, but preserves the whole. *ES2: 105-106*

Similar examples are seen in Ashbery's poetry:

- (114) The dog ran *over us*
The ball with all his might.
 We might escape, in the daylight
 The barn of his personal loss.
John Ashbery, 'The Ascetic Sensualists'

- (115) Like a long room
 Monsignor
 pushed *away it*
 studio artificially small
 pine rounds
John Ashbery, 'Europe'

Examples like this could be analysed in two ways: rightward displacement of the NPs or leftward displacement of the PPs. While there is no way to know exactly which reordering has taken place, it is natural to assume that this is rightward displacement of the NPs, since English standardly has a similar kind of movement operation known as Heavy NP Shift. In HNPS, a long or 'heavy' NP is shifted over another VP-internal argument to the end of the sentence:

- (116) a. I gave the ball to John.
 b. *I gave to John the ball.
 c. (?)I gave the ball that Mary had asked for three weeks ago to John.
 d. I gave to John the ball that Mary asked for three weeks ago.

(112)-(115) are unusual because the NPs that have been shifted – *the sky*, *the soul* and *the ball* – are not 'heavy' in the relevant sense (i.e. phonetically); rather, the first are only disyllabic indefinites without any further embedding within them, and *it* is a monosyllabic weak pronoun (definitionally non-heavy). In (114), the final PP *with all his might* modifies the VP rather than the shifted NP, so it does not provide the standard motivation for HNPS. Whether or not they are illicit applications of HNPS or something else altogether is a technical question, but for now all that needs to be observed is that they present another kind of example of displacement in poetry that would not normally occur in standard English.

There are also examples where it seems that NPs have been moved leftward within a VP, such as the following example from Robert Creeley:

- (117) The car
moving
the hill
down

Robert Creeley, 'A Step'

In this example, the NP object *the hill* and the particle *down* are rearranged, perhaps by leftward NP-displacement, since the standard word order within the VP is *moving down the hill*; there is a possible interpretation where the car is the thing that is making the hill move down, in which case the word order is perfect, but given that this is entirely implausible and the context of the poem indicates that displacement may be a more likely interpretation. As with the previous examples, it is unclear which of the two elements has actually been displaced, but nevertheless we know that there has been some degree of VP-internal displacement that does not normally occur in English, but unlike those there is not a legitimate English movement rule that might be applied (albeit erroneously) to derive the word order here. Rather, it is simply the product of some other displacement process.

Finally, we can see that, just like with PP-displacement, we also find NP-displacement to sentence-initial positions in environments where we would not normally get topicalization in English. We see object fronting in wh-questions, relative clauses and yes-no questions:

- (118) Honour forbid! at whose unrivall'd Shrine
Ease, Pleasure, Virtue, All, our Sex resign. *TR4: 105-106*
- (119) Just where the Breath of Life his Nostrils drew,
A charge of Snuff the wily Virgin threw; *TR5: 81-82*
- (120) But still this world, so fitted for the knave,
Contents us not. A better shall we have? *ES4: 131-132*
- (121) Truths would you teach, or save a sinking land?
All fear, none aid you, and few understand. *ES4: 265-266*

This example is from Charles Olson:

- (122) And here is significance, lost to those
 Who scholastics are, and follow their nose.
Charles Olson, 'The Fool'

As with PP-displacement, we also see NP-displacement in Pope's poetry with *if*-clauses, here to a position below the complementizer *if*:

- (123) Happy! ah ten times happy had I been,
 If Hampton-Court these Eyes had never seen! *TR4: 149-150*

Ashbery's poetry provides a similar example within an *as*-clause, where the object *kyrie eleison* is fronted to a position immediately below *as*:

- (124) Yet when the doorbell rang
 It reduced all that living to air
 As "kyrie eleison" it sang.
John Ashbery, 'It Was Raining In The Capital'

As with the examples of PP-displacement, what these examples show is that NP-displacement in poetic texts is extremely liberal, occurring in many contexts and to positions which do not normally allow such displacements.

2.1.3 AP displacement

Adjectival phrases (APs) have a more restricted distribution than NPs and PPs, and there are fewer operations for moving APs than for NPs and PPs. Adjectives can occur as prenominal or post-nominal modifiers:

- (125) a. I saw the sick man.
 b. I saw a man sick with the flu.

They can also occur as predicates in copula constructions and small clauses:

- (126) a. John is sick.
 b. The cake made John sick.

APs can also appear as arguments of verbs like *feel*:

- (127) I feel sick.

Adjectival complements can be questioned in wh-questions with words like *how*, but these phrases are typically analysed as wh-adverbials, and AP modifiers in NPs cannot be wh-questioned anyway. AP complements can be topicalized in some situations, as in (128):

- (128) John said the game would be terrible, and terrible it was indeed.

However, AP-topicalization is typically more marked than with NPs and PPs.

Despite this limited distribution and general resistance to movement, there are many examples of different kinds of AP displacement in English poetic texts, although they are fewer and less diverse than the NP- and PP-displacement examples. The following examples display similar fronting of APs to the pre-verbal position familiar from the previous two sections:

- (129) With tender Billet-doux he lights the pyre,
And breathes three am'rous sighs to raise the fire.
Then prostrate falls, and begs with ardent eyes
Soon to obtain, and long possess the prize. *TR2: 41-44*
- (130) Oh had I rather unadmir'd remain'd
In some lone Isle, or distant Northern land; *TR4: 153-154*
- (131) Of various habits, and of various dye,
The pierc'd battalions dis-united fall, *TR3: 84-85*
- (132) Not tyrants fierce that unrepenting die,
Not Cynthia when her manteau's pinn'd awry, *TR4: 7-8*
- (133) Where all must fall or not coherent be,
And all that rises rise in due degree; *ES1:145-146*

There are also examples from Ashbery's poetry:

- (134) O how this sullen, careless world
Ignorant of me is!
John Ashbery, 'Two Sonnets'

In (129) the adjective *prostrate* has been displaced from its typical position after to the verb to the pre-verbal position; although the orthographic sentence

doesn't contain a subject, the interpretation is that conjunction *then* conjoins the VP headed by *prostrate* and the previous two. In (130) the full adjectival phrase *rather admired* (the adjective and its modifying adverbial) is inverted with the verb *remained* and it appears below the subject, and the auxiliary has been fronted in a conditional inversion construction. The other examples are similar, involving movement of verbal complements to positions before the verbs that they modify. All of these are superficially similar to the AP topicalization construction seen in (128), but since the displacements are to non-initial positions they cannot be analysed as normal topicalization.

There are also examples of displacement of AP-modifiers from within nominals:

- (135) In various talk th' instructive hours they past,
Who gave the ball, or paid the visit last: TR3: 11-12
- (136) What boots the regal circle on his head,
His giant limbs, in state unwieldy spread; TR3: 71-72
- (137) For this your Locks in Paper-Durance bound,
For this with tort'ring Irons wreath'd around! TR4: 99-100
- (138) Of systems possible, if 'tis confest
That wisdom infinite must form the best, ES1: 43-44

In (135), the adjective *last* is interpreted as a modifier of *visit*, but it appears to the right of the noun rather than its standard position (for English) to the left, between the noun and the determiner *the*; thus, it seems that the adjective has been displaced rightward over the noun. An interesting aspect of this example is that, if the determiner was *a* rather than *the*, the example may not have this interpretation where *last* modifies *visit*, since it could also be interpreted as an adverbial modifier of the VP *paid a visit*, and this interpretation of *or paid a visit last* would typically be preferred since it does not involve a nonstandard displacement operation. Quite why this would be the case is an issue that is discussed further in later chapters.

In (136), it seems that the adjective *unwieldy* has also been displaced rightward from its position modifying the noun *state*, although this example is stranger since it is missing a determiner where it would normally have one, i.e. *in a state unwieldy*. Note also that this AP-displacement has occurred within a PP that has itself been displaced, here to the pre-verbal position, showing that double displacement doesn't just need to involve displacement of two separate elements, but it can also involve displacement of one element from an already displaced element. It is worth noting that this is not possible with most other kinds of movement operations in natural language, at least according to standard sources in the theoretical literature.¹⁰

In (137) the adjective *bound* appears after the PP *in Paper-Durance*, whereas it would normally be in a position immediately following the nominal *Locks*. The interesting thing about an example like this is it is not clear whether the element that has been displaced is the PP *in Paper-Durance*, which is contained by the AP, or the AP, which normally appears post-nominally and not pre-nominally because it contains a modifier; the former situation would make this example like those in section 2.1.1. This does not introduce any ambiguity in meaning, however, as both ways of doing things produce the same word order and meaning. It seems that there are no clear criteria for determining which process has taken place, perhaps since both operations are as arbitrary as each other.

An interesting aspect of the kinds of inversion seen in the examples in (138), as well as (135) and (136), is that it is not always obvious that they should be analysed as displaced adjectives or as something else. There are numerous examples of these apparently displaced adjectives in experimental texts, like John Ashbery's poetry:

(139) Rust dark pouring
Over the body,

¹⁰See Rochemont and Culicover (1990) for discussion. There is one possible exception, and that is so-called 'smuggling' movement, which has been used by Collins (2005a,b) to explain passivization and raising, and by Hicks (2009) to explain tough-movement.

John Ashbery, 'Saying It To Keep It From Happening'

- (140) In the apartment fallen
The tree began to take root.

John Ashbery, 'To The Same Degree'

- (141) It is the surface black which attacks the shape,
Bending it to its present uses.

John Ashbery, 'The Thousand Islands'

However in cases like these it is less clear that the adjectives should be interpreted as displaced pre-nominal modifiers, or rather as what we might call 'parenthetical adjectives,' that is, adjectives that modify the nominal much in the same way as non-restrictive relative clauses do. Consider (142), which illustrates a more controlled example of a parenthetical adjective, with the relevant punctuation, a comparable example with a non-restrictive relative, and then compare these with (143), which illustrates a pre-nominal adjective modifying a noun:

- (142) a. The dog, black, jumped back into the pond.
b. The dog, which was black, jumped back into the pond.

- (143) The black dog jumped back into the pond.

The adjectival parenthetical *black* in (142a) contributes roughly the same meaning to the sentence as the non-restrictive relative in (142b), namely that there is a dog that happens to be black at the point just as it jumps back into the pond; this dog may be a blond labrador that has just rolled around in mud after getting out of the pond. These contrast with (143), which, without a significant amount of context, can only mean that a dog that is always black has just jumped into the pond. The different kinds of modification thus have different semantic interpretations, and whether or not a given adjective is interpreted as a displaced pre-nominal modifier like *black* in (143) or a parenthetical adjective like *black* in (142a) can give us an indication as to whether or not we will say it has been displaced or not.

Turning to the examples from Ashbery's text, it seems that, at least in (140) and (141), and perhaps (139) as well, the adjectives are interpreted as parenthetical adjectives rather than as displaced pre-nominal modifiers; for example, it is perfectly acceptable to interpret *fallen* in (140) as a modifier of *the tree* where the tree has just fallen, even though the reality of a tree taking root after it has fallen is somewhat strange. At the very least these examples are ambiguous between the two types of modification, and there may be a bias for the non-displacement example since the language in the surrounding area of the examples is relatively stable, displaying few unusual deviations and displacements. These are to be contrasted with (135)-(138) from Pope, where the high level of deviation in the surrounding texts makes it perfectly plausible that either meaning should be derived, and given the context of some individual examples the displaced meaning is to be preferred over the parenthetical meaning. What this shows is that context can play a significant part in whether or not deviation is even detected, since sometimes potentially displaced pieces of text may simply be interpreted as separate 'chunks' that are not related to the main sentence in the same way as a displaced constituent in a standard example. The role of contextual factors is something that will be discussed in some depth in the later sections, but for now we shall simply observe that there are many examples which are typically interpreted as involving unusual displacement of pre-nominal adjectives.

In the experimental texts, there are even examples where the order of sets of pre-nominal modifiers are reversed, by displacement of one of the adjectives:

(144) at
 that always vague edge is
 the public so-called condition,
 which nobody knows enough
 ever, even those are supposed to be it.

Robert Creeley, 'Blues'

The nominal in this example is interpreted as *the so-called public condition*, where *public condition* is a compound noun and *so-called* is an adjectival modifier.

Public is a nominal modifier since *condition* is the head of the compound, and in (144) the nominal modifier and the adjective have been reversed. It isn't clear whether it is *public* or *so-called* that has been displaced, but nevertheless it seems that some sort of displacement has affected these elements even within the nominal phrase.

To conclude the subsection, we can see that, just as with NPs and PPs, there is AP displacement to the sentence-initial position in situations where we would not normally get standard topicalization, such as in a relative clause in (145):

- (145) Then in a Bodkin grac'd her Mother's hairs,
Which long she wore, and now Belinda wears. TR5: 95-96

The AP *long* is displaced to a sentence-initial position below the relativizing wh-phrase. As with the examples of PP- and NP-displacement, (145) cannot be analysed as topicalization since this does not occur within relative clauses, so this is a further example of unusual displacement of an AP in a literary text.

2.1.4 Other kinds of displacement

This subsection gathers together other kinds of displacement from the texts that do not fall clearly into the previous subsections, or which should be separated from those kinds for important reasons. What this final subsection shows is that the displacement phenomena in literary texts are truly diverse, targeting constituents that seldom move, and sometimes even non-constituent elements, for displacement.

TP displacement

First, we see examples where entire embedded sentences are displaced. Sentences are known as Tense Phrases (TPs) in the literature (since at least Pollock 1989), and they can be displaced to a sentence-initial position in certain circumstances. For example, purpose-clauses like *to get a tan* can be fronted as in (146):

- (146) a. You need to go to Spain to get a tan.
 b. To get a tan, you need to go to Spain.

This kind of TP-fronting is limited, however, as dependent clauses with predicates like *seem* and *hate* resist fronting:

- (147) a. John seems to like cheese.
 b. *To like cheese, John seems.
- (148) a. John hates to talk about childhood.
 b. ?*To talk about childhood, John hates.

This is because the TP *to get a tan* in (146) is an adjunct that can be omitted optionally, whereas the TPs *to like cheese* and *to talk about childhood* in (147) and (148) respectively are obligatory complements of the verbs, and such TP complements cannot normally be fronted.¹¹

In the literary texts, there are a few examples where different kinds of TPs are displaced to positions other than the sentence-initial position. The following example shows the displacement of an adjunct TP:

- (149) Seas roll to waft me, suns to light me rise;
 My footstool earth, my canopy the skies.” *ES1: 139-140*

The phrase *to light me* is a purpose infinitive TP adjunct, and it modifies the verb *rise*; here it has been displaced to the familiar position between the matrix subject and the verb. This may be seen as an example of intermediate fronting, bearing resemblance to the examples of NP and PP displacement where a constituent that can normally be topicalized seems to have been shifted only so far up the sentence. Such examples are strange, but only so strange, in that they might be plausibly analysed as a sub-species of standard topicalization.

In the following example, on the other hand, we see displacement of an argument TP within a relative clause:

- (150) No bandit fierce, no tyrant mad with pride,
 No cavern'd hermit, rests self-satisfied;
 Who most to shun or hate mankind pretend,

¹¹See Abels (2003) for discussion of the relative immobility of TPs.

Seek an admirer, or would fix a friend.

ES4: 41-44

The infinitive *to shun or hate mankind* is interpreted as a complement of the verb *pretend*, and it has been displaced to a position between the verb and the adverbial *most*; *most* could be interpreted as the determiner in nominal phrase that has undergone nominal ellipsis (as in *some students turned up but most didn't bother*), but it seems that it is interpreted here like the modern *mostly*. The important thing is that constituents like the infinitive complement of a verb like *pretend* cannot usually be topicalized, so this example is doubly unusual since it cannot even be viewed as a special sub-species of topicalization.

Adverb displacement

It is well-known that adverbs are restricted to appearing in specific positions with respect to other elements in a sentence. For example, in English manner adverbs like *quickly* typically appears to the left of a main verb but to the right of the highest auxiliary, whereas epistemic adverbs like *probably* can also occur to the left of highest auxiliary but not to the right of lower ones or the main verb:

- (151) a. John has quickly learned French.
 b. ?* John quickly has learned French.
 c. *John has learned quickly French.
- (152) a. John probably has been fired for negligence.
 b. John has probably been fired for negligence.
 c. *John has been probably fired for negligence.
 d. *John has been fired probably for negligence.

Adverbs can also have different meanings when they appear in different positions, and they take scope with respect to other elements. It is typically assumed that adverbs do not undergo movement, unlike other elements like NPs and PPs. Therefore the interpretation of an adverb in a given sentence should diagnose quite accurately its structural position.

There are a number of examples in the texts where it seems that adverbs have been displaced from the position in which they are interpreted:

- (153) Some secret truths, from Learned Pride conceal'd,
To Maids alone and Children are reveal'd: *TR1: 37-38*
- (154) With tender Billet-doux he lights the pyre,
And breathes three am'rous sighs to raise the fire.
Then prostrate falls, and begs with ardent eyes
Soon to obtain, and long possess the prize. *TR2: 41-44*
- (155) But grant that those can conquer, these can cheat:
'Tis phrase absurd to call a villain great.
Who wickedly is wise, or madly brave,
Is but the more a fool, the more a knave. *ES4:229-232*
- (156) And next to him, a young girl
Seated on the pavement, sitting
Merely. *LI1: 39*
- (157) Where it is
was and
will be never
only here.
Robert Creeley, 'Gemini'
- (158) Did the young
couple come
only home
from London?
Robert Creeley, 'Wellington, New Zealand'
- (159) [...] all the dear or awful
passages apparently
I've gone through.
Robert Creeley, 'Soup'
- (160) can you keep
it ever
together,
Robert Creeley, 'So There'

In (153), the adverb *alone* appears immediately adjacent to the nominal *Maids*, yet the interpretation of the sentence is where *alone* modifies the full coordinate nominal *Maids and Children*, since the interpretation where *alone* modifies just one of the conjuncts is both contradictory and unnatural. Therefore it seems that the adverb has been displaced from its natural position to the right of *Children*. The interesting thing about this particular displacement is that it bears no resemblance to other kinds of natural language displacement, not only since it involves movement of an adverb, but also since there is no plausible structural position to which one could move an adverb to create this configuration. In structural terms, a movement analysis of this displacement would involve lowering of the adverb into a right adjoined position within a nominal contained by a coordinate structure, and such a movement process is unheard of elsewhere in natural language syntax. Rather, it seems that what we see in an example like this is simple word order rearrangement.

In (154) the adverb *soon* appears in a position preceding the infinitival auxiliary *to*, but the adverb is interpreted as if it is modifying the verb *obtain*, and in such cases it should appear between *to* and *obtain*. This is confused by the VP coordination in this final line, where both *obtain* and *possess* share the object *prize*, but it seems clear that *soon* only modifies *obtain* while the other adverb *long* modifies *long*, in which case it is clear that *soon* should appear on the edge of the VP rather than on the edge of the TP.

In all of the other examples, the majority from Robert Creeley's sparse poetry, it seems pretty clear that the adverbs have been displaced from their standard interpretation positions. In (155), the manner adverb *wickedly* should appear to the left of the copula, adjacent to *wise*, so it has been displaced leftwards. In (156) the adverb *merely* should appear to the left of the verb it modifies, *sitting*, but instead appears to the right. In (157), *never* should appear either between *will* and *been* or to the immediate left of *will*, but instead it seems to have been displaced to a position following *be*, where such adverbs do not normally occur. In (158), the focus adverbial *only* appears between *come*

and *home*, but since *come home* is an idiom *only* can only take scope outwith the construction, so it would typically occur in a position to the left of *come*. In (159), *apparently* would normally appear between *I've* and *gone*, even though, as an epistemic adverb it can sometimes appear in such a position; due to the nature of the construction, however, the adverb is most naturally interpreted lower, and as a result it seems that the adverb has been displaced. In (160), the adverb *ever* is interpreted as if it appears to the left of *keep*, yet it has been displaced to some position within the idiom.

The examples all get a similar explanation, in that it is clear that the adverbs should occur in one position according to interpretation but they are seen in another, and we may say on the basis of this that the adverbs have been displaced. Adverb displacement does not create significant problems for interpretation, but it is an unusual process, one that is largely unlike other movement operations in ordinary language syntax.

Subject displacement

Throughout the analysis in this section we have seen that the displacements that have been observed have typically been analysed with respect to the position of the subject, which has been assumed to be something of a constant. Yet there are also examples where we may say that the subject itself has been displaced from its standard sentence-initial position to some other position, since it occurs in a non-sentence-initial position with respect to one or more other elements. Consider:

- (161) Already see you a degraded Toast,
 And all your Honour in a Whisper lost! *TR4: 109-110*
- (162) But errs not Nature from this gracious end,
 From burning suns when livid deaths descend, *ES1: 141-142*
- (163) Why has not man a microscopic eye?
 For this plain reason, man is not a fly. *ES1:193-194*

- (164) And out over the park where crawled roadsters
 The apricot and purple clouds were
 And our blood flowed down the grating [...]
John Ashbery, 'Two Sonnets'

- (165) still, the rods, could not they take long
 More anthems until Dust
 flocks disguised machine. The stone
 the valentine couldn't save... Hooks
John Ashbery, 'Europe'

In (161) the subject *you* appears to the right of the verb *see*, indicating that either the verb has been displaced leftward or the subject has been displaced rightward. There is no clear indication of which analysis is correct; the presence of the sentence adverbial *already* seems to indicate that the verb has not shifted since, if it were shifted to a sentence-initial position, it would also precede the adverb; however, this is not hard and fast evidence, since both of these possible displacements are non-standard and their relation to adverbials like *already* are not clear or testable. The choice between calling (161) an example of verb displacement or subject displacement is merely one of terminology, at least at this stage. (164) is similar, in that the reversal of the subject *roadsters* and the verb *crawled* within the relative clause could be produced by one kind of displacement or the other.

However, in (162), (163) and (165) it seems clearer that what we are seeing is subject displacement. In (162) the subject *Nature* appears to the right of both the main verb *errs* and sentential negation, and these two elements do not form a constituent which could have been displaced as one unit. It is not implausible to suggest that both *errs* and *not* have both been displaced over the subject, of course, and the only reason to prefer one analysis over another in this situation is appealing to very vague guidelines of least effort or parsimony, but once more the two choices are effectively terminological variants of one another. (163) is similar due to the presence of negation to the left of the subject; while the verb is expected to appear to the left of the subject since the highest auxiliary always moves to such a position in wh-questions of this kind, negation only moves to

this position when it is attached to the verb as a clitic in a *hasn't* form. Thus it seems that either negation has been displaced leftward over the subject or the subject has been displaced rightward over negation.¹² (165) is similar to this too, involving a yes-no-question instead of a wh-question.

What these examples show us is that the word order rearrangements that we see in poetic texts can be so unlike standard movement operations that it becomes difficult even just to settle on a specific analysis for what kind of movement operation this could be, were it indeed to be a sub-species of standard movement operations like topicalization. It is also interesting to note that, in all of these examples, it is unclear which element would have been focused for topicalization, and that the word order rearrangement we see in these examples has no clear and direct influence on prosody or focus.

Non-constituent displacement

Throughout this section on displacement, it has been assumed implicitly (and sometimes explicitly, as in the discussion of (162) in the last section) that the displacement operations we see in the literary texts should be analysed as sharing at least some basic ingredients in common with standard movement operations in syntax. Perhaps the definitional characteristic of movement in syntax is its structure-preserving nature: that is, movement applies to constituents, not to simple sequences of words, and this means that when we move a constituent,

¹²An additional, more technical analysis could be to suggest that the subject hasn't been displaced at all, and that in these situations, standard subject movement from the base position within Spec,vP to to Spec,TP has been suspended. This analysis has been used in other situations where the subject appears lower than its standard position, in English for locative inversion (Collins 1997), in German for situations where the object is scrambled over the subject (Wurmbrand 2006), and in various other places to account for word order in VSO languages like Scottish Gaelic and Irish (i.e. McCloskey 1996). Such examples involve suspending/parametrizing the EPP or allowing other elements like verbs to check the EPP features on T that are normally checked by the subject; the technical merits of these different options continue to be disputed in the literature, so I will not go into their details here.

What is important is that, in these other linguistic phenomena, the 'subject movement suppression' analysis is motivated by empirical evidence from various sources, to cover core facts of word order; in this way, they are entirely unlike our examples from Pope, which are clearly deviant and only liable to occur in poetry where there are many other distortions in word order throughout. In the former, technical innovations in the grammar are clearly motivated by a desire to capture core facts, but in the latter this motivation is almost entirely lacking. For more discussion see section 4.2.6.

we move the structure contained within it. It has been assumed throughout that displacement in poetic texts shares these characteristics, and all of the examples seen thus far have displayed this basic characteristic. Indeed the fact that displacement in poetic texts shares this characteristic is perhaps one of the main reasons for viewing displacement operations as sub-species of movement operations.

However, there are also examples of displacement of non-constituent elements from throughout the poetic texts.

- (166) For this, ere Phoebus rose, he had implor'd
 Propitious heav'n, and ev'ry pow'r ador'd,
 But chiefly Love – to Love an Altar built,
 Of twelve vast French Romances, neatly guilt TR2: 35-38
- (167) Here Britain's statesmen oft the fall foredoom
 Of foreign tyrants, and of nymphs at home; TR3: 5-6
- (168) One self-approving hour whole years outweighs
 Of stupid starers and loud huzzas: ES4: 255-256

In each of these examples, pieces of NPs are displaced without carrying with them their full internal structures, stranding argument PPs (underlined). In (166), *an Altar* is modified by the PP *of twelve vast French Romances*, and the article-plus-noun part of the NP is displaced to a pre-verbal position to the left of the verb *built* (alongside displacement of the PP *to Love*). In (167) the same occurs with fronting of the article-plus-noun part of the NP *the fall of foreign tyrants*. In (168) it is the adjective-plus-noun part of the NP *whole years of stupid starers and loud huzzas* that is displaced to the pre-verbal position before *outweighs*.

All of these examples seem to resemble the NP displacement examples discussed previously, in that they involve the displacement of nominals to the pre-verbal position; however, these are not examples of displacement of phrases, since the strings that are displaced in these examples are not the full nominal phrases according to interpretation; if this were phrasal displacement, the

PP-modifiers that remain in the post-verbal position would have to have been displaced too. In this respect, the displacements in (166)-(168) are unlike any movement operations in natural language, although we can also note that they are in fact quite similar to other examples discussed in this chapter, where a nominal element has been moved from its VP-internal position to a position to the left of the verb.

2.1.5 Summary

In this section, we have seen a number of different kinds of displacement phenomena in poetic language. These phenomena have been shown to be very diverse: all kinds of phrases can be displaced; phrases are seldom displaced to designated positions, but rather to various different ones with no specific structural characteristics; multiple displacements can occur at once, alongside one another or sometimes even from one from another (i.e. displacement from a displaced constituent); displacement can occur in various environments to various position; both constituents and non-constituents are displaced.

The rearranged word orders seen in the poetic text do not exhaust the logical possibilities, however. There were no reported instances of displacement of heads (like auxiliaries or determiners), even though there are analogous natural language movement operations which allow for movement of these elements (i.e. head movement). Furthermore, certain combinations of reordering were unattested: for example there are no examples of S-O-V-PP order in the corpus reviewed, despite the fact that almost every other alternative reordering of these elements was attested. We saw that certain kinds of movement were much more common than others: object displacement was shown to be very common, as was PP complement and adjunct displacement, but subject displacement was only attested infrequently. Finally, it is also worth noting that although displacement is often rather arbitrary, it seldom intersperses phrases within other, smaller phrase; that is, displacement is typically to positions that mark boundaries between larger phrases that form phonological units. If we expected

displacement to be entirely arbitrary, we may expect it to displace NPs to positions within other NPs or into other such awkward positions, but this is not found. In chapter 4 I will discuss ways in which these gaps in the empirical paradigm may be explained by the proposed theory.

What all of this shows is that this displacement that we get in poetry is very unlike the highly constrained movement operations that are provided by natural language syntax; rather, it seems clear that what we see in these examples is a matter of simple word order rearrangement, where a sequence of words is displaced to a new position in the surface string. In some cases, such as the displacement of PPs, NPs or APs to sentence-initial positions in simple finite declarative sentences, I have used the terminology of syntactic movement, identifying such examples as ‘topicalization’ since they bear enough of a surface resemblance to standard topicalization that they could warrant such an analysis. However, we can see that many of these cases do not have the same effects on discourse or prosody as standard topicalization, in that they do not always create clear focus on the displaced element (this is discussed in more detail in section 3.1.1). This is because these ‘proper’ movements often occur in the context of numerous other non-standard displacements, and in such cases it is not clear what should or should not be called the ‘topic’, or what should or should not be given focal stress when reading the text. This is discussed more in the next chapter.

The case of pronouns is particularly instructive. It is interesting to note that there were few examples of pronoun displacement, even though pronouns may also be targeted by phrasal movement. Indeed pronouns are very often targeted by movement rules in both Germanic and Romance (such as Object Shift or cliticization), and they were also subject to such movements in earlier stages of the development of English (see Wallenberg 2009), so we might have expected the exact opposite – pronouns to move more than full NPs – if we were to assume that displacement was a development of standard movement rules. The fact that the opposite tendency was found seems to indicate that

displacement is a separate phenomenon that is constrained by different factors.

Given all this, it seems plausible to suggest that many of the other displacement operations that have been taken to be instances of standard movement (i.e. the cases of ‘topicalization’ that occur alongside other kinds of displacement, as in (96) and many other examples) may in fact be better analysed as further examples of the more arbitrary word order rearrangement process of displacement that is shifting the same elements to other positions in the same texts. Therefore we could say that displacement is even more widespread in poetic texts than has been described in this chapter.

The significance of this characterization of displacement in poetry is to be drawn out in the chapters to follow, but for now it is enough to observe that displacement is an operation that is distinct from standard movement that rearranges the word order of a sentence. What is needed is a theory that can explain how English speakers are able to produce linguistic meanings from these unusual sentences.

2.2 Erasure phenomena

This category gathers together examples from poetic texts in which words or phrases appear to have been omitted by non-standard means by a process which I call ‘erasure.’ The data discussed here comes from a variety of sources, but the majority comes from more modern experimental poetry, since this non-standard omission of words is a more prominent characteristic of this kind of poetry than older styles, although relevant examples also come from Pope and other older sources. Poetic erasure phenomena have not been systematically analysed anywhere else in the literature, and although the data is necessarily more messy than that which has been discussed so far in this chapter (since it is more difficult to analyse that which is missing than that which is simply displaced), the section sorts the evidence into individual subsections that identify specific kinds of ungrammatical deletion of syntactic elements. Throughout the section

I will refer to the operation of omitting words as deletion, and I will distinguish between two different kinds of application of this operation: ‘erasure,’ where deletion is ungrammatical, and ‘ellipsis,’ where deletion is grammatical. Erasure is the ill-formed deletion we see in poetic language, and it is the focus of this section, while ellipsis is the regular and well-formed deletion phenomena we find in ordinary English. First, I begin by laying out the general characteristics of ellipsis in standard English, describing its general properties and conditions for good use, before then going on to show describe erasure in poetic texts.

2.2.1 Missing out words in English

Natural language standardly allows for the omission of words of phrases when certain conditions are met, and this well-formed deletion of phonological material is called ellipsis.¹³ (169) demonstrates three standard ellipsis cases:

- (169) a. John brought three friends, although Mary said he should only bring two in case it’s too busy.
 b. A: I met one of your friends earlier today.
 B: Really? Which one?
 A: John.
 c. John really wants to eat the entire cake, but he knows that he really shouldn’t.

In these examples, strings of words are interpreted but not pronounced in a variety of syntactic positions. In (169a), the nominal object of the verb *bring* is pronounced as *two*, but it is clear from interpretation that the nominal has more structure than just *two*, since this would be a strange thing to say outwith a context where *two* referred to some specific thing designated with that as its name or label. *Two* in (169a) is clearly a nominal modifier of a head noun, and according to interpretation, that nominal is *frienda*: it is clear that Mary said John should only bring *two frienda*, not *two bananas*. We say, then, that the

¹³There are other kinds of phonetically unrepresented constituents of relevance in the literature, most notably the traces of overt movement under the Copy Theory of Movement (Chomsky 1993), but I will ignore this here since this technical issue does not relate to the discussion at hand. For a recent analysis of the relationship between ellipsis and the Copy Theory, see Thoms (to appear). I return to the significance of this connection in chapter 4.

word *friends* is indeed present in the structure of the nominal, but it has been omitted by a process called ‘NP ellipsis.’

In (169b), B’s short reply *which one?* encodes much more than the meaning of an isolated wh-nominal, whatever that might be.¹⁴ Rather, it has the same meaning as the full wh-question *which one (of my friends) did you meet earlier today?*; this meaning seems to be dependent upon A’s utterance, although it is not just a straightforward copying of that part of A’s utterance, since the pronoun switches from *your* to *my* between the two utterances. This full-sentence ellipsis in the presence of wh-movement is known as ‘IP ellipsis’ or more commonly ‘sluicing,’ and it can occur in embedded questions as well as question fragments like (169b). Some linguists (i.e. Merchant 2004) have argued that fragment answers, like A’s response, also have full elided structures, and that the fragment answers involve topicalization of the answer followed by ellipsis of the full sentence complement in a parallel fashion to sluicing. Thus A’s response, *John*, has a structure like *John, I met earlier today*, and the fragment is produced by ellipsis of the sentential part, just as in sluicing.

In (169c), the second sentence can be interpreted as *he knows that he really shouldn’t eat the entire cake*, where a full verb phrase is ‘filled in’ after the auxiliary *shouldn’t*. This is not the only option though, as the sentence can also be interpreted as *he knows that he really shouldn’t want to eat the entire cake*, where a larger verb phrase is filled in for the gap in the ellipsis sentence. However, these two options do exhaust the options for interpretation of the second sentence, as the verb phrase cannot be interpreted as any old predicate, like *he really shouldn’t eat a single slice*, even if the context allows for such an interpretation on reasonable assumptions. This shows that there are constraints on how we can interpret the constituent in the ellipsis, and that these constraints are imposed with respect to the linguistic context of the utterance; that is, some

¹⁴There are numerous proposals in the literature (such as Culicover and Jackendoff 2005) which suggest that the isolated wh-phrase could import the same meaning as the full sentence, without any elided structure; these accounts rely upon general pragmatic mechanisms for the construction of the full question meaning. See Merchant (2001) and Lasnik (2007) for extensive arguments against this.

relevant phrase (here, a verb phrase) must antecede the ellipsis site for it to be interpreted. The kind of ellipsis seen in (169c) is known as ‘VP ellipsis,’ but the antecedent requirement is a more general one that constrains the interpretation of ellipsis.

These examples demonstrate some specific kinds of ellipsis in English, and they demonstrate some particular constraints on ellipsis and a requirement for antecedence (see also Hankamer and Sag 1976). However, it should be noted that there are further constraints on ellipsis than just an antecedent requirement; consider the following variations on (169), where the same antecedence contexts are maintained:

- (170) a. *John brought some siblings, although Mary said he should have brought every to make sure the party looked busy.
 b. A: I met one of your friends earlier today.
 B: *Really? Which one did you?
 c. *John really wants to bring the entire cake, but he knows that he really shouldn’t bring.

In all of these examples, ellipsis is ungrammatical, even though there are clear antecedents for each example. In (170a) the nominal *some siblings* provides an antecedent for nominal ellipsis within *every sibling* just as it did with the nominals *one friend* in (169a) above, yet for some reason deletion of the noun *sibling* is not possible in the context of the quantificational determiner *every*. In (170b), it seems not to be possible to elide the constituent directly below the dummy verb *did*, even though the context is just as clear as it is in (169b). (170c) is similar, in that it does not seem to be possible to delete the constituent that follows *bring*, even though it was possible to delete the constituent that followed *shouldn’t* in (169c).

Although explanations are varied and the problem is still not fully understood, these general constraints on deletion are generally known as the ‘licensing conditions’ on ellipsis, where ellipsis is only allowed if it occurs adjacent or near-adjacent to a ‘licensing head’ of some sort. The standard accounts state that ellipsis is licensed only by a small set of elements in specific constructions, such

as those shown in (169) above: auxiliary verbs, quantificational determiners and numerals (but not *every*), and moved wh-phrases (or the syntactic elements associated with them) to name a few, but not by others, like main verbs and non-wh nominal phrases (like *you*). Though these categorizations seem somewhat stipulative (see Thoms to appear for a recent discussion), they capture clear intuitions about the differences between the data in (169) and (170), in that ellipsis is unacceptable in the latter but acceptable in the former.

It is implicit from the presentation of some of the examples above as ‘NP ellipsis,’ ‘IP ellipsis’ and ‘VP ellipsis’ that ellipsis involves the deletion of full phrasal constituents, and that it is not just the omission of individual words or discontinuous strings. This is generally assumed to be the case, given examples like (171). The text in strikethrough indicates words that are supposed to be interpreted in the ellipsis but not pronounced; the ungrammaticality of these examples indicates that the pronounced strings do not create the meanings indicated:

- (171) a. *I bought three green apples, and John bought four ~~green~~ tomatoes.
 b. *Bill has been meeting students all week, and John ~~has been meeting students all week~~, too.
 c. *I know some people who will vote for the Lib Dems, but I don’t know why ~~they will vote~~ for the Lib Dems.

The ellipses in each of each of these examples occur next to elements that are standardly taken to be licensors – a numeral in (171a), auxiliaries in (171b), a wh-phrase in (171c) – yet the partial deletions seen are not allowed. This shows that when ellipsis is licensed by a particular element, the whole complement constituent is deleted, and that ellipsis cannot just target individual words.¹⁵

¹⁵There are examples of situations where words can be ‘missed out,’ such as with complementizer deletion in the following:

- i. John knows that I am ill.
 ii. John knows I am ill.

However, examples like (ii) are not taken to involve ellipsis of the complementizer *that* but rather the use of a null complementizer, which is a different lexical item from the non-null element.

There are apparent exceptions to this rule, most notably in the area of VP-ellipsis. The two main cases are demonstrated below:

- (172) a. John will bring three guests, and Bill four.
b. John will bring the wine, and Bill will the beer.

(172a) is an example of what is known as ‘gapping,’ and it is so called because it seems that there is a ‘gap’ in the second sentence: it is interpreted as *Bill will bring four guests*, and it seems that the non-constituent string *will bring* is omitted by ellipsis (the deletion of *guests* would be a separate instance of NP ellipsis). (172b) is similar, involving omission of just the verb *bring*, and this is known as ‘pseudogapping.’ However, these superficially similar constructions are actually very different in their syntactic properties, and many have argued that, while pseudogapping does involve ellipsis, gapping does not (see Johnson 2009 and the references cited therein). Furthermore, a number of linguists have argued for subsuming pseudogapping to standard VP-ellipsis, arguing that the phrases that escape the ellipsis (*the beer* in (172b)) do so by moving out of the VP (see Jayaseelan 1990, 2001, Lasnik 1999, Takahashi 2004; though see Thoms to appear and section 4.3.4). Gapping is discussed in more detail in what follows, but for now it is enough to observe that examples like (172) do not present a real counter-argument to the generalization that ellipsis targets full constituents.

What these examples show is that ellipsis is subject to two distinct sorts of constraints: antecedence, which is often referred to as ‘recoverability,’ and licensing. Natural language only allows for ellipsis when these conditions are satisfied simultaneously, and we can see from some of the examples that ellipsis can be plainly deviant or ungrammatical when these conditions are not met. In what follows, we will see that there are many examples in the poetic texts of deviant deletion where these conditions have been disregarded. In what follows I will discuss different examples of this, introducing specific comparable ellipsis constructions and their characteristics along the way when necessary. We will see that the deletion phenomena we get in poetry are different from those that

we get in standard ellipsis in a number of ways.

2.2.2 Erasure of functional elements

One of the most common kinds of erasure in poetic texts involves omission of words that lack lexical semantic content. These are ‘function words,’ and they encode grammatical functions and relations between words in sentences. Four examples of these kinds of words are given below:

- (173) a. I do not know Japanese.
b. I hope to learn Japanese next year.
c. I think John is a speaker of Japanese.
d. There is a man in the garden.

The underlined words in these examples do not make any real contribution to lexical semantic meaning, in that they do not have any meaning in isolation, yet they are required in these environments, and missing them out in these examples leads to ungrammaticality. In (173a), the verb that appears to the left of the negation is known as *do*-support, and it is typically described as a ‘dummy’ or ‘pleonastic’ verb that is only inserted as a ‘last resort’ to make the sentence grammatical; this is due to the nature of negation and verbs in English, and this dummy *do* appears in a number of other environments which are sometimes assumed to be ‘last resort’ ones that can be occupied by auxiliary verbs like *be* and *have* but not main verbs like *know*.¹⁶ In (173b) the infinitival auxiliary *to* is required to mark the embedded verb *learn* as an infinitive, but this does not convey any lexical semantic meaning. (173c) is similar in that the finite copula *is* makes no contribution to semantic meaning; while some uses of *be* that do seem to carry some degree of semantic meaning (for example in the intransitive use *to be or not to be*), in general the copula serves only to predicate one noun of another, and most linguists agree that its contribution to semantic meaning

¹⁶For discussion of *do*-support see Chomsky (1957), Pollock (1989), Lasnik (2004) among others. For an alternative analysis that argues against the ‘last resort’ characterization of *do*-support, see Thoms (2010b).

is null.¹⁷ The subject nominal *there* in (173d) is known as an ‘expletive,’ and its only contribution to the sentence is to fill the subject position in the existential construction, as the sentence is identical in meaning to a sentence like *A man is in the garden*. Expletive *there* is not the same as the locative adverb *there* in a sentence like *I put the keys there*, and it is assumed in the literature that it is only required for formal reasons (relating to Case and other abstract aspects of syntax). There are numerous other examples of function words in different domains of syntax, and their morphosyntactic categorization is a large research project in contemporary work in syntax.¹⁸

All of these lexical items are in some way more ‘expendable’ than contentful lexical items, appearing only for formal syntactic reasons, and the reason for their appear in sentences like those in (173) is not obvious to the English speaker upon simple introspection; that is, a speaker may tell you why he or she uses the word *dog* or *cat* in a sentence, or the verb *introspect*, but they may be at a loss to explain why they use words like infinitival *to* or expletive *there*, at least in terms of what it contributes to what they are trying to say. Indeed it is interesting and perhaps instructive that, in standard ellipsis, the presence or absence of these meaningless words does not matter for the calculation of the antecedence relation. Recall that an ellipsis must always have an antecedent, and that the meaning of the ellipsis is always near-identical to that antecedent: in such examples, there cannot be a difference between the ellipsis and antecedent in terms of lexical content. However, there can be differences between the ellipsis and antecedent with respect to the use of these function words, as the following examples demonstrate:

- (174) a. Fixing a car is easy if you know how ~~to fix a car~~ (Merchant 2001: 22)
 b. I know that John DID arrive late, but I don’t know why ~~he arrived late~~
 c. There is someone at the door, but I don’t know who ~~is at the door~~

¹⁷For the semantic view, see Heim & Kratzer (1998); for a syntactic view, see Schütze (2004).

¹⁸See for example van Riemsdijk (2004), which collects together work on the borderlines of functional categories and their implications for syntactic theory.

In (174a), the ellipsis contains a verb with the infinitival auxiliary *to*, yet this auxiliary is not present in the antecedent *fixing a car*; in (174b) *do* appears in an emphatic form, yet the interpretation of the ellipsis does not require the use of *do* in this form; in (174c) an existential construction containing expletive *there* is the antecedent of the ellipsis, yet the expletive isn't in the ellipsis. We can see, then, that whatever the identity relation between the antecedent and the ellipsis may be, it is not sensitive to the presence or absence of these elements.

It is perhaps unsurprising, then, that there are many examples in the poetic texts of erasure of function words like these. Perhaps most common of these is omission of *do*-support in examples where we would normally expect to see it. In the following examples I mark the gaps left by the missing elements with an underscore (_):

- (175) ... rather, it is in the disrepair
Of these lives that we _ not find despair

LII: 53

- (176) And no more in our society _ living melodies
Break forth under the little or no shade.

John Ashbery, 'A Pastoral'

- (177) What have I seen,
_ now see?

Robert Creeley, 'Ice Cream'

- (178) Go no nor come again.

WH: 40

- (179) So skull _ not go. What left of skull _ not go.

WH: 46

- (180) But it _ not reproach us shall now recesses

Jackson Mac Low, 'A(ce)'

In (175) the main verb *find* in the embedded sentence is preceded by negation, just like (173a), yet the dummy verb *do* does not appear, and the same happens

in (179) and (180). (176) is similar in some ways in that we expect *do*-support in the position of the gap because of the fronting of the negated phrase.

(177) is slightly different from the other examples, in that it seems to be missing more than just *do*. The text is interpreted as a pair of wh-questions, roughly equivalent to something like *what have I seen, what do I now see?*; given this, it seems that dummy *do* has been erased along with the subject and the wh-phrase. It might be possible to propose that the ‘underlying sentence’ is not two separate questions but rather a coordinate structure like *what have I seen, do I now see?*, in which case the wh-phrase is not erased in the second phrase, but rather the first wh-phrase is shared by coordination (an instance of what is known as ‘across the board wh-movement’). However, in this case we are still missing both a subject and *do*-support. The morphological form of the verb *see* indicates clearly that it should be appearing alongside *do*-support in a separate question rather than being coordinated with the other verb phrase, since if it was coordinated it would appear in the *seen* form. Thus it seems that the subject and the *do* that would normally occur in the question have both been erased. I will return to erasure of subjects in what follows, but for now it is enough to show that this example involves a clear instance of erasure of functional *do* in a wh-question. Note that such an example could either be construed as erasure of the non-constituent string *do I* or as two separate instances of single-word erasure; it is impossible to tell which has taken place in this situation, just as it was impossible to tell which constituents had been displaced in which direction in the multiple displacement examples discussed earlier.

In all of these examples, we see the omission of the functional element *do* where we would normally see it in standard English.¹⁹ In addition to these, there are also related examples that involve the erasure of ‘main verb’ *do*. Main verb *do* is often called a proform verb or light verb, since its meaning is defined

¹⁹The issue of whether or not these examples could just represent examples of non-standard dialects is discussed in what follows; see footnote 58 in section 4.3.3 for a particularly instructive point. For now, it is enough to point out that the data in these examples do not demonstrate generalizations about verb movement that are consistent for these texts: for example, in most cases in Beckett’s text the verb occurs to the left of negation, unlike in (178) and (179).

anaphorically in the context of use; that is, it possesses little or no lexical semantic content in itself, as we can see with these sentences:

- (181) a. I said I would run a marathon this year, but I don't know if I will be able to do it.
 b. I have done a lot this evening: washed the dishes, cleaned the sink, irritated Molly. . .

In these examples the meaning of *do* is dependent upon the contextually salient verbal antecedents: in (181a), *do it* means *run a marathon this year*, while in (181b) *done* receives its meaning from the various verb phrases that follow it. While the 'anaphoric' nature of such verbs is not the same as that of anaphoric NPs, what we can see is that *do* has little semantic content in itself, similar in some respects to dummy *do*.

In the following poetic examples, it seems that light verb *do* or some equivalent verbal form should appear in the positions indicated by the gap, indicating again that *do* has been erased, just like with the previous examples:

- (182) What can the rain that fell
 All day on the grounds
 And on the bingo tables ___ ?

John Ashbery, 'Album Leaf'

- (183) [...] white all white and no head does that mean soap. It does not ___
 so.

TB: 17

- (184) Time to lose. Gain time to lose. As the soul ___ once. The world ___
 once.

WH: 20

In (182) the wh-question contains the modal verb *can* but no other lexical verb, and since this kind of construction is not allowed in English,²⁰ the interpretation of the sentence indicates that some light verb *do* has been erased. In (183) the light verb seems to have been omitted from the verbal proform *do so*; this *do* would normally co-occur with the dummy *do* that appears to the left of negation,

²⁰This is not the case in all languages though, as such constructions can appear in languages like Dutch and Italian.

but here it seems to have been erased. In (184), the verb seems to have been omitted from the *as*-parenthetical (the phrase is interpreted as a parenthetical, despite the punctuation), since *the soul* seems to be the subject and it co-occurs with an adverb which clearly modifies some missing verbal element. The following string *the world once* seems to be interpreted as a sentence that has been coordinated within the *as*-parenthetical, thus indicating that it involves parallel *do* erasure too.

In addition to erasure of dummy and light verb versions of *do*, there are also many instances of other erased auxiliary verbs and related main verbs. The most commonly omitted auxiliary is *be*, and its main verb equivalent copular *be* is also erased in numerous situations, and in some cases it seems that *have* has been erased too. Consider:

- (185) And after ___
 Taken out behind the stairs and ___ stood them
 In the kitchen... the flowers blowing in the wind
 Felt funny just the same...

John Ashbery, 'Night'

- (186) Where if not there ___ it too?

WH: 18

- (187) The same narrow void. Before the staring eyes. Where ___ it too if not there too?

WH: 19

- (188) No words for what when words ___ gone.

WH: 28

In (185) there are two gaps that indicate missing auxiliaries, both of which are indicated by the morphological form of the main verbs. The two VP-conjuncts are in the adverbial clause headed by *after*, so given the general syntactic properties of these clauses we would expect that both VPs should appear in non-finite forms. The first conjunct has as its main verb *take* in past participle form, indicating that there is a missing auxiliary in the *-ing* form; given that *taken* lacks a direct object in this context, the interpretation is that the missing auxiliary

is the passive auxiliary *be* inflected as *being*. The second conjunct contains the past participle form of *stand*, and this verb has a direct object, thus indicating that the missing auxiliary is *have* in the *-ing* form. Thus the sentence is interpreted as *and after being taken out behind the stairs and having stood them in the kitchen*, with the two *-ing* forms of the auxiliaries erased. This coordinated meaning is somewhat unusual, since it is not clear what the subject of the first conjunct should be, but the morphological form of the verbs and the context of the *after*-adverbial seems to make this interpretation the only one available. What this shows is that the interpretation of erased functional elements may be driven to a great extent by morphological evidence in the lack of clear semantic evidence from the context.

(186) and (187) are similar examples from the same text of wh-questions that seem to lack a clear main verb, and interpretation of these examples indicates that the missing verb should be copular *be* in the positions indicated by the gaps. In fact, the putative position of the erased verb is not always clear, since in (186) the *if*-clause *if not there* appears to have been displaced from its usual sentence-final place to a position immediately following the wh-phrase: in principle, the auxiliary could have been before or after the displaced *if*-clause in an unerased version of the sentence. It is also unclear where the auxiliaries would have appeared since, while we know that auxiliaries normally move to a position next to the wh-phrase in main wh-questions like these, we also know that there is a tendency in these texts for altering the normal behaviour of verbs with respect to these kinds of movement.

In (188), the copula has been erased from within the relative clause, since it is interpreted as *when words are gone*; in this respect, this example is like (186) and (187). However, it seems that there are additional missing elements in this example, since the matrix sentence is missing both a subject and a main verb; that is, the sentence is interpreted as *there are no words for what when words are gone*, rather than as a standalone nominal phrase. As it happens, there are arguments in the literature on ellipsis for this option being available in many

standard cases, and in such a case we could say that the missing elements *there* and *are* have been elided by standard mechanisms rather than erased in a non-standard way. This makes sense, since the example does not seem deviant or ungrammatical simply because of the absence of *there are*, and the NP seems to be perfectly interpretable as a fragment.

In the literature, which typically focuses on (but is not restricted to) fragment answers to questions, fragments are often analysed as consisting of topicalization phrases that have undergone ellipsis. Thus a version of (188), factoring out the erased copula, would look something like this:

(189) No words for what when words are gone, ~~there are t~~.

Merchant (2004) provides a number of arguments to back up this topicalization analysis of fragments, and as such these structures are similar to the sluicing cases discussed earlier, where the full sentential complements of moved wh-phrases are erased. Merchant also argues that fragment nominals can always be interpreted as having an elided *there is*, regardless of whether there is a viable linguistic antecedent, since this can always be supplied by general pragmatic conditions. The important thing about English fragments, however, is that they cannot be embedded, since English only has topicalization in matrix clauses. Thus *there is*-omission can only occur with sole fragments like (188) and (189); the string cannot just be deleted in any situation, as the ungrammaticality of (190) attests:

- (190) a. *I know that ~~there is~~ no reason to believe you.
 b. *I want to know if ~~there is~~ a man in the garden.
 c. *I spoke to John because ~~there is~~ a good chance he will come round.

These arguments can be extended to cover the other expletive constructive *it is*, since the two constructions can interchange in a number of the relevant environments.

In the poetic texts, there are numerous examples of *there is* and *it is* strings being erased in non-matrix contexts:

- (191) Where in
the world then ___ an-
other place?

Robert Creeley, "Follow the Drinking Gourd..."

- (192) this voice is truly changeable of which ___ so little left in me

HI: 15

- (193) The time to show a message is when ___ too late and later there is no
hanging in a blight.

TB: 4

- (194) Say yes that the bones may pain till ___ no choice but stand. Somehow
up and stand.

WH: 8-9

- (195) ... and not
Even aiming at the heavens far above it
Yet seemingly nearer, just because ___ so
Vague and pointless

LI2: 37

- (196) [...] a new kind
Of demand that stumps the absolute because ___ not new
In the sense of the next one in an infinite series

John Ashbery, 'The Skaters'

- (197) If it is not dangerous then ___ a pleasure and more than any other if it
is cheap it is not cheaper.

TB: 5

- (198) Dim white and hair so fair that in that dim light ___ dim white.

HI: 15

In (191) there seems to be an *is there* string missing, since the text is interpreted as a simple wh-question, and there is only one nominal argument, *another place*, indicating the wh-question has been derived from an existential *there is*-construction. Since this string is in a wh-question, inverted in the sentence-initial position following the wh-phrase and the adverbial *then*, it could not have been elided in the manner of a sentence fragment. The relative clause in

(192) is also missing a *there is* string, since the AP *so little left in me* must be predicated of the PP *of which* in an existential construction. This sentence is itself unusual, since the relative clause seems to have been extraposed from the subject position of the matrix clause: the relative clause modifies the nominal *this voice*, yet it appears in a sentence-final position after the predicated AP *truly changeable*. Nevertheless, the erasure interpretation for the relative clause is clear-cut, and the two kinds of disruption do not accumulate to make the sentence uninterpretable. (194) is a similar example that has undergone various other kinds of erasure (discussed below); the gap here indicates the place where a *there is* string has been erased in the sentential complement of an *until*-adverbial. (193) is also similar to (192), in that it involves erasure within an embedded clause, but in this case it is an *it is* expletive-verb string that is erased. This example contains another unerased *there is* string, indicating that the erasure of these strings is not systematic.

(195)-(198) are all superficially similar to (193), since they are all missing *it is* strings in the positions indicated: in embedded *because*-modifiers in (195) and (196), in (197) in a *then*-clause in a conditional, in (198) within an embedded clause. The difference between these examples and (193), however, is that the *it is* strings that have been omitted in these examples are not expletive-verb strings, but rather strings with anaphoric pronoun *it*. In (195) *it* is ambiguous, in that it could refer either to the action *aiming at the heavens* or to the same unidentified nominal that is the referent of the prior (unerased) pronoun *it* in the discourse. The other examples involve anaphoric *it* referring to nominal referents within the discourse. Anaphoric pronouns are similar to expletives in that they are functional nominals that encode minimal semantic content (the similarity extends to the fact that the same lexical item *it* can be either an expletive or anaphoric pronoun), so these examples of erased anaphoric pronouns may be grouped with the other kinds of erasure seen above. What is important is that all of the examples involve erasure that does not resemble the ellipsis that we see in sentence fragments like (189), and that these examples are in some way

deviant because they employ this non-standard strategy.

There are further examples of erasure of sentence-initial strings and auxiliaries in embedded clauses. The follow examples demonstrate erasure of the infinitival auxiliary *to* and additional elements in infinitives:

- (199) How ___ be
young and yet to be loved?
Robert Creeley, 'Ice Cream'
- (200) Say yes that the bones may pain till no choice but ___ stand. Somehow
up and stand.
WH: 8-9
- (201) An occasion for a plate, an occasional resource is in buying and how
soon does washing enable a selection of the same thing ___ neater.
TB: 7
- (202) All things seem ___ mention of themselves
John Ashbery, 'Grand Galop'
- (203) it seems ___
in your tracks
because it
was ending for the first time
John Ashbery, 'The Thief of Poetry'

We can see that in (199) the infinitival auxiliary in the first conjunct has been erased, since the other conjunct is an infinitive with the auxiliary; given that this kind of coordination typically involves coordination of likes, the clear interpretation of the example is *how to be young and yet to be loved?*, with a missing *to*. (200) is similar, where the auxiliary *to* is erased from the infinitive *to stand*. In (201)-(203) it is not just the infinitival auxiliary that has been erased, but also copular *be* which would normally follow it in these constructions; while *seem* can sometimes be followed by an adjectival complement in sentences like *John seems sick* (a small clause structure), this is not possible for nominals like *mention of themselves* in (202) or the PP *in your tracks*, so these examples must involve erasure. Note that these examples involve erasure of a non-constituent,

and that this makes these examples clearly different from standard kinds of ellipsis found in English.

(199)-(203) show erasure of functional elements that appear in embedded contexts in infinitives, but this sort of erasure is not restricted to embedded clauses of any specific type. Rather, we can also see from the following that strings of embedding functional elements are also erased in finite clauses, most notably in relative clauses. This is demonstrated below:

- (204) There is no world
except __ felt,

Robert Creeley, 'After'

- (205) A sign of more in __ not mentioned.

TB: 5

- (206) It makes mercy and relaxation and even a strength to spread a table
fuller. There are more places __ not empty.

TB: 5

- (207) at
that always vague edge is
the public so-called condition,
which nobody knows enough
ever, even those __ are supposed to be it.

Robert Creeley, 'Blues'

In each of these examples, some part of the embedding structure of the relative clause seems to have been erased. In (204), the *except*-adverbial is followed only by the adjective *felt*, which seems to modify the head noun *no world*; therefore the adjective is interpreted as the predicate in a copular sentence where the subject is a deictic pronoun that is coindexed with the NP that the *except*-clause modifies. That is, the sentence is interpreted as *there is no world except that (one) which is felt*, where the string *that (one) which is* has been erased. (205) is broadly similar, where the string that has been erased is *that which is*, this time within a PP. In (206) the erased string does not include a deictic pronoun but just the copula and the relative pronoun: the string *not empty* is again

interpreted as a predicate of the NP *more places* in the existential construction, so this material must be in a relative clause modifying the NP, like *more places that are not empty*. In (207) it seems that only the relative pronoun *that* has been erased; this cannot be interpreted as a standard example of deletion of a relative pronoun, as in examples like *I visited the restaurant (that) you told me about*, since this cannot occur with relativized subjects, which is what we get in (207).

Therefore all of the (204)-(207) involve erasure of embedding material by nonstandard means, and in most cases the material that is erased is functional or ‘meaningless’ in the sense described above. Not all of the elements erased in these examples are what we would call meaningless, however: the instances of the deictic pronoun *that* in (204) and (205), for example, are not meaningless, since they are nominals that refer to a narrow range of entities that match with them in terms of their semantic content, which specifies that they be non-human and singular. There are other instances of the erasure of elements that are not technically meaningless but which nevertheless do not cause great problems for interpretation. Below are three examples of erased prepositions:

(208) These decibels
 Are a kind of flagellation, an entity of sound
 Into which being enters, and is apart __ .
John Ashbery, ‘The Skaters’

(209) The fly beckon on the window
 The kids came and we all went __ the briars.
John Ashbery, ‘Night’

(210) It feels __ things
 are muddled again
Robert Creeley, ‘Later’

In (208) the second conjunct *and is apart* seems to be missing the preposition *from*. This is somewhat confused because the two conjuncts are headed by the relativizer *into which*, where the preposition *into* has been pied-piped in relativization from the first conjunct only. In a standard example the pied-

piping would occur ‘across the board,’ in that the second conjunct would also involve a preposition phrase that included *into*; however, the second conjunct does not contain such a conjunct here, since *apart into* is ungrammatical and meaningless. Since the sentence cannot be interpreted as if the preposition *from* has also been relativized, it seems that *from* has been erased by nonstandard means. (209) and (210) are much similar, in that they involve unambiguous erasure of prepositional complements of verbs: (209) is missing *to*, and (210) is missing *like*. Both examples are clear-cut since it is part of the lexical meaning of the verbs that they come with these kinds of complements.

Prepositions are not meaningless in any regular sense, but of all of the main lexical categories, the class of prepositions is the lexical class that generally contributes least to lexical semantic meaning; furthermore some prepositions like *of* are more or less entirely meaningless, showing up only when it is required for Case or some other functional role. Nevertheless what we see in these examples is that the erasure in poetic texts is not solely reserved for purely functional elements; rather, it seems that erasure of words with some degree of lexical semantic content can also occur.

2.2.3 Failed ellipsis

In this subsection I examine cases of erasure of meaningful elements where it seems that the erasure may be described as a case of ‘failed ellipsis’; that is, deletion in a situation where the constraints that normally apply to standard ellipsis, licensing and recoverability, are not satisfied. We will see that violations of these different constraints lead to different kinds of problems for interpretation.

Failed NP-ellipsis

As mentioned above, NP-ellipsis only occurs when there is a determiner or quantifier of a specific kind to license deletion of the head noun; (211a) demonstrates the good cases, and (211b) the bad ones:

- (211) a. A number of protestors gathered in the street.
 {Many/most/some/four/one ___} was/were furious.
 b. A number of protestors gathered in the street.
 {*The/*a/*every ___} was/were furious.

There are no fully worked-out explanations of this pattern in the literature,²¹ but we can take this as the descriptive baseline of what is and is not possible for NP-ellipsis in English.

NP-ellipsis is generally characterized by the presence of other NP/DP-related elements like determiners, so in deviant examples of NP-ellipsis we would expect to see deletion of an NP with a non-licensing element, like the examples in (211b). This is what we see in the following poetic examples:

- (212) The eyesight, seen as inner ___,
 Registers over the impact of itself
 Receiving phenomena,

John Ashbery, 'Tapestry'

- (213) His love boiling up to me
 Forever will I be the only ___
 In sofa I know
 The darkness on his back

John Ashbery, 'Rain'

- (214) A little ___ called anything shows shudders.

TB: 15

In (212) the NP within the parenthetical is missing its NP head, and since the parenthetical modifies the matrix subject *the eyesight* the deleted noun is interpreted as *eyesight*. Yet the adjective *inner* cannot license ellipsis, and hence the deletion is deviant, an instance of erasure.²² The erasures in (213)

²¹Lobeck (1995) is the most substantial attempt to explain, though see Thoms (to appear) for discussion of the theoretical problems and possible ways forward. See also Sleeman (1996) for discussion of NP-ellipsis in Romance.

²²Adjectives can sometimes license ellipsis in English when the circumstances are right; for example, in (i), where both the antecedent and elided NP also contain a definite article:

- i. The outer layer was peeled off quickly. Peeling the inner ___ took a bit longer.

These cases are still not perfect, however, and it is not clear why they are better; it cannot be because the definite article is the licenser, as (211b) shows that definite articles do not generally license NP-ellipsis. Furthermore, we can see that not all adjectival modifiers allow for this kind of ellipsis, as (213) attests: *only* never licenses ellipsis, even when circumstances like those in (i) are set up.

and (214) are more problematic. They both involve deletion without a licenser, as the adjective *only* cannot license ellipsis in (213), nor can *little* in (214), regardless of the presence of the preceding articles *the* and *a*. However, these NP erasures also lack clear antecedents, and because of this it is not clear what the erased nouns would be. One may propose that the missing noun in (213) is *one*, since *one* is the default noun that appears in NPs when ellipsis isn't licensed (the examples in (211b) would typically occur with *one* in the position of the gap). A similar proposal may be made for (214), but the fact that the rest of the sentence indicates that this subject NP is *called anything* indicates that the NP *one* may also be a viable interpretation for the gap.

These examples show that violations of the licensing requirements do not necessarily cause serious problems for interpretation, but violation of the recoverability conditions can cause more significant problems. It seems that when recoverability is violated, the only way to fill in the noun is to interpret a generic and nearly-meaningless proform like *one* in that position, or perhaps a related form depending on the immediate context.²³ In this respect the latter kinds of examples resemble more closely the set of examples discussed in the previous subsections, where functional elements with minimal semantic content are erased. The difference between licensing violations, where full lexical content is erased, and recoverability violations, where only semantically low-content elements can be erased, is an issue to which we will return.

Failed VP-ellipsis

VP-ellipsis in English was demonstrated above in the sample sentence (169c), and this example demonstrated the well-known fact that VP-ellipsis requires an auxiliary licenser in English. VP-ellipsis is not licensed by main verbs, adverbs or NP-subjects, as demonstrated below:

- (215) a. *John put a book on the shelf, and Bill put __ , too.
 b. *John has quickly read the book, and Bill has quickly __ , too.

²³See Schütze (2004) for arguments that *one* is a 'last resort' element for the nominal domain, similar to *do*-support in the verbal domain.

- c. *John read the book, and Bill __ , too.

VP-ellipsis is also restricted in a number of other situations, for example in infinitive complements of different kinds (see Lobeck 1995, Thoms to appear). As with other kinds of ellipsis, VP-ellipsis is also subject to recoverability, and it has been observed that the recoverability condition that applies to VP-ellipsis is more restrictive than for other kinds of ellipsis; related to this, VP-ellipsis sites are also restricted in the kinds of dependencies that they allow, typically not working with wh-movement:²⁴

- (216) *John met one of his friends earlier today, but I don't know which one he did.

Here the wh-phrase is moved from within the elided VP, and this is ungrammatical, even though the full unelided sentence would be fine, as would the sluiced alternative *I don't know which one*. Regardless of the technical analysis of these restrictions, we can see that VP-ellipsis cannot just occur with any VP, and that it is subject to a number of well-formedness constraints in standard use.

As with NP-erasure, there are also examples of VP-erasure without an adequate licenser. This is demonstrated below:

- (217) Where the sweet william grew and a few other cheap plants __
The rhythm became strained,
John Ashbery, 'Haunted Landscape'

- (218) I want the world
I did always __ ,
Robert Creeley, 'For Pen'

- (219) I wonder if I will have any friends there
Whether the future will be kinder to me than the past, for example,
And am all set to be put out, finding it to be not __ .
John Ashbery, 'The Skaters'

In (217) the second conjunct is interpreted as *a few other cheap plants grew*, and this must be due to erasure since a coordination analysis of the structure

²⁴For discussion of the restrictions on VP-ellipsis with respect to recoverability, see Lasnik (1995) and Hartman (2009). For discussion of the restrictions on extraction from VP-ellipsis sites, see Takahashi and Fox (2005), Merchant (2008), Schuyler (2001) and Hartman (to appear).

is not possible: the NP *a few other cheap plants* would need to be adjacent to the conjunction *and* for this to obtain. Nevertheless the coordination of the two sentences with two similar subjects makes it clear that the verb that has been erased must be the lexical verb *grew*, the same as the first conjunct. The sentence is thus an example of ill-formed erasure rather than well-formed ellipsis because there is no licenser for ellipsis in the second clause: as shown by (215c) above, NPs like the subject *a few other cheap plants* cannot license VP-ellipsis. In (218), the VP is erased next to the adverb *always*, and as we saw in (215b) adverbs cannot license ellipsis either; if this example were to be an example of ellipsis, the adverb would need to precede *did*, as it does in the legitimate ellipsis form *I always did*.²⁵ In (219) there is a deleted VP immediately adjacent to the negation in the infinitival clause, and it is interpreted as the adjectival predicate *kinder* (examples like this with the copula are typically analysed as VP-ellipsis, even though they do not involve the deletion of full VPs). While negation does license ellipsis in finite clauses, this kind of negation following the infinitival auxiliary in an infinitive doesn't,²⁶ and as a result the deletion here is ill-formed. The recovery of the meaning of the erased constituent aided by the fact that the subject is a deictic pronoun which is coindexed with the subject of the previous copular sentence *the future will be kinder to me than the past*, and the erased construction itself is also a copular sentence; this parallelism ensures that *it* should be coupled with the predicate *kinder* rather than the other candidate for the antecedent to the erased constituent, *put out*. Here a VP is erased in an unusual manner; parallelism helps the reader to derive the meaning

²⁵This cannot be analysed as an example of a right-adjoined TP adverbial that falls outside of the ellipsis site, as *always* cannot occur in a sentence-final position in full sentences and this is always the case with such adverbials like *already*:

- i. *I liked you always
- ii. I always liked you.
- iii. I have already been to France.
- iv. I have been to France already.
- v. John wants to go to France, but I have been ___ already.

²⁶For discussion of negation and the complexities and quirks of its ellipsis licensing behaviour, see Potsdam (1997), Johnson (2001), van Craenenbroeck (2004) and Thoms (to appear).

of the erased constituent, although it should be noted that such examples are more opaque and difficult to interpret than their standard non-deviant ellipsis equivalents.

In addition to these examples of erasure that violates the licensing conditions, there are also examples where the recoverability conditions for VP-ellipsis are not met, resulting in ill-formed VP-erasure. We have already seen examples of such VP-erasure in the analysis of ‘light verb erasure’ in section 2.2.2 above: in those examples, the minimal VP created by the verb *do* is erased. These examples are comparable with the examples of NP-ellipsis above, where the missing element is a proform with minimal semantic content; the *do*-erasure examples also involve erasure of a proform with minimal lexical content where there is no clear antecedent for the VP.

The following examples illustrate related problems for recoverability:

(220) You come out of love. But are ____ .
John Ashbery, ‘A Box and Its Contents’

(221) It stands. What? Yes. Say it stands. Had to ____ up in the end and stand.
WH: 8

(222) Say yes that the bones may ____ pain till no choice but stand. Somehow ____ up and stand.
WH: 8-9

(223) No knowing how know only no ____ out of. ____ Into only.
WH: 7

(224) Where then but there see now another. Bit by bit an old man and child ____ .
WH: 13

In (220) the erased constituent is interpreted as *out of love*, the complement of the lexical verb *come* in the first conjunct; however this kind of antecedence relation is not possible in English. This is due to the fact that VP-ellipsis requires identity of two VPs, not just of arguments within the VPs; thus, for the

identity relation to hold, the VP following the copula must match the antecedent VP in both its verbal content and in the content of its argument. Since this does not obtain, recoverability is not possible and hence ellipsis is not possible. This mismatch between full VPs and copular clauses containing similar arguments is robust: (220) is plainly ill-formed, and following examples show that this is a general trend rather than a peculiarity of this example:

- (225) a. *I said I would punch a fascist. I know that John is __ , but I won't punch him as he's really big.
 b. *I wanted to watch the Charles river turning green, but John tells me it is __ already.

The underlined constituents are potential antecedents for the elided argument in the copular clauses, but nevertheless we can see that these are not possible ellipses. This shows that the recoverability problem suffered by (220) is a general one.

It is interesting, then, that (220) nevertheless receives the interpretation where *out of love* is the complement of the copula, since the sentence may have had an alternative interpretation, where what is missing is the full VP *coming out of love*; in this case recoverability would be satisfied, since this VP clearly matches the immediate antecedent in the relevant way. This is because of the coordinator *but*, which clearly implicates that the second conjunct should contradict the first one in some way. This kind of interpretation is not available for a version of (220) where it is the full *come*-VP that is filled in for the gap, since there is no possible reading where the second conjunct would contradict the other: *you come out of love, but are coming out of love* is semantically ill-formed in this respect. The interpretation where it is just the PP that is filled in for the gap (that is, the deviant erasure version) is still not entirely well-formed with respect to this need for contradiction, but it seems more plausible: *you come out of love, but are out of love* would be more sensible if it was followed by an adverb like *already*, and it is perhaps for this reason that the latter interpretation is the one that readers are likely to assign to the deviant sentence in (220). We will return to this kind of semantic deviation and the

problems it incurs in section 4.3.3, but for now we can observe that the reading of the sentence where recoverability has been violated is preferred to the reading where no such conditions have been violated, since this erasure reading produces a sentence that is closer to a semantically well-formed utterance.

In (221)-(224), from Beckett's *Worstward Ho*, we see sole verbs erased without clear antecedents. Erasure of the verbs in some of the examples leaves behind their PP arguments, thus making those examples look more like gapping or pseudogapping than standard VP-ellipsis. In most cases, however, it is almost entirely unclear what kind of verb has been erased. Thus (221), the second gap in (222) and both gaps in (223) the presence of the PPs indicates that there must be verbs missing in the positions of the gaps, and that they must be verbs that select for these kinds of PPs; in most of the gaps the verb *stand* would work, yet in both (221) and the second gap in (222) these missing verbs are VP-coordinated with intransitive *stand* in the second conjunct so an interpretation where the first verb is *stand* would be somewhat unusual. The first gap in (222) is preceded by the modal verb *may* and followed by the direct object *pain* and a time adverbial *till no choice*

In (224), the fact that the sentence is modified by the adverbial ensures that it is interpreted as missing some verbal form; the context of the previous sentence (and the preceding discourse) indicates that the narration relates the perspective of some watching figure, and given the meaning of the modifying adverbial, it seems inevitable that this sentence should contain some verb that describes the gradual appearance of the old man and the child on the horizon; a verb like *appear* perhaps. This remains unclear, however, and thus the lexical content of the missing verbal form in (224) remains unspecified. Note also that erasure here occurs in the absence of a licenser (the subject NP cannot license ellipsis, as we saw earlier), so we can see that violations of both conditions need not lead to complete uninterpretability.

In addition to these examples of faulty VP-ellipsis, there are also examples of related phenomena going wrong, namely gapping. Whether or not gapping

is to be subsumed to VP-ellipsis is a subject of continuing debate,²⁷ but for our purposes here I will group them together as a pair of mechanisms for missing out words in standard English. Gapping is a construction in which all auxiliaries, the main verb and all other VP content save argument is missed out, and it is demonstrated below:

(226) John will order the steak, and Mary ~~will order~~ the fish.

Gapping is subject to a number of constraints that are not experienced by ellipsis, such as the fact that it can only occur in coordinate structures and that it cannot occur in strings that precede the antecedent. Thus VP-ellipsis can occur in an *although* clause, but gapping cannot:

(227) a. Although he knows he shouldn't ~~hate Mary~~, John hates Mary.
 b. *Although Bill ~~hates~~ Mary, John hates Dolores.

These facts motivate most of the modern analyses of gapping (discussed in more depth in section 3.1.1 in the next chapter).

Austin (1984) notes that Pope's poetry provides a number of examples of gapping that disobeys these constraints, and he calls these examples 'backwards gapping.' Below are two examples from 'The Rape of the Lock':

(228) When Florio speaks, what virgin could withstand,
 If gentle Damon did not squeeze her hand?
 With varying vanities, from ev'ry part,
 They shift the moving Toyshop of their heart,
 Where wigs __ with wigs, with sword-knots sword-knots strive,
 Beaux banish beaux, and coaches coaches drive. *TR1: 97-102*

(229) While fish __ in streams, or birds delight in air,
 Or in a coach and six the British fair, *TR3: 163-164*

In these examples the gap appears before its antecedent, in contravention of the restriction on gapping: in (228) the gap is interpreted as *strive*, and in (229) the gap is interpreted as *delight*. Both examples within wh-adverbials, and in (228) the coordination requirement is satisfied too, and while these are examples

²⁷See Coppock (2001) and Johnson (2009) for two recent contributions to this debate.

of unsuccessful gapping, they are still interpretable as something like gapping: that is, there is little ambiguity about which verbs should be interpreted in the gaps. This is helped by the fact that both examples also occur in parallelism with other sentences with similar structures.

(228) is particularly interesting since the antecedent for ellipsis, *with sword-knots sword-knots strive*, also displays a topicalization-like displacement, since the VP-argument PP *with sword-knots* appears sentence-initially. This is more like displacement (as identified in the previous section) than regular topicalization, since it occurs within an embedded clause. What is interesting is that the gapped sentence and antecedent differ with respect to word order, since this is generally disallowed in gapping (see Johnson 2004: 101). This goes the other way too: if a first conjunct involves some sort of movement, then the second one must have it too. Thus a topicalized sentence cannot antecede gapping in a non-topicalized sentence, as shown by (230a), but it's OK if both sentences involve topicalization, as shown by (230b):

- (230) a. *The beans, John cooked *t*, and Bill ~~cooked~~ the potatoes.
 b. The beans, John cooked, and the potatoes, Bill ~~cooked~~ *t*.

This resembles a general constraint on ellipsis-like operations, and its explanation probably has the same roots as the ban on extraction from ellipses discussed earlier. Interestingly, there are poetic examples where these constraints on extraction is disregarded:

- (231) With store of pray'rs, for mornings, nights, and noons,
 Her hand is fill'd; her bosom ___ with lampoons. *BR4: 29-30*

- (232) Some place the bliss in Action, some in Ease,
 Those call it Pleasure, and Contentment these ___ ; *ES4: 22-23*

(231) is similar to (230a) in that a PP has been displaced to a sentence-initial position in the first conjunct while there is no parallel topicalization in the second one. (232) is more like (228) since it is the gapped conjunct that displays some sort of displacement, since the gapped NP *Contentment* appears before

the subject. Note, that it is impossible to know what kind of nonstandard operation is in action in these examples: they could either be analysed as good topicalization sentences with erasure, or good gapping examples with displacement. There is no way to determine which analysis is correct, just as with the examples that feature multiple displacements.

All of these examples of bad gapping can be analysed as cases of nonstandard erasure or ‘faulty ellipsis’, alongside the faulty VP-ellipsis cases discussed above, since they involve missing words out in ways that cannot be done in standard English. Whether or not standard gapping does actually involve ellipsis is irrelevant to our purposes here, since all that is required is that we identify types of deviation that occur in poetic texts with some useful descriptive terminology; it is enough to identify these sentences as examples of ‘deviant gapping’ to satisfy this purpose. Nevertheless, in chapter 3 we will see that the specifics of the syntax of gapping become more important, since it becomes obvious that the gapping examples from the poetic texts aren’t just sub-optimal versions of normal gapping sentences, but rather underivable strings that bear no relation to gapping other than one of surface resemblance.

Improper argument ellipsis

This subsection concludes by discussing a few examples of a rarer erasure phenomenon which nevertheless bears a resemblance to some of the examples discussed so far, which I will call ‘improper argument ellipsis.’ This is when an argument that is obligatorily realized in standard English is erased in a poetic example. Argument ellipsis (taken as excluding subject omission) is not attested in English and is generally rare in the related language families, so it is perhaps expected that argument erasure in English poetic texts should be particularly difficult to interpret. The following two examples demonstrate two different kinds of argument erasure from experimental texts:

- (233) months passed,
 things happened in ___ .

Robert Creeley, 'Say Something'

(234) All this and not ordinary, not unordered in not resembling ____ .
TB: 3

In (233) we see that the PP complement of *happened* is missing its obligatory NP argument; in the context, where the sentence follows immediately after *months passed*, it seems that the missing argument should be *them*, which refers to *months*, although this reading is far from clear. In (234) it is the gerundial verb *resembling* that is missing an NP argument, but in this case there is no clear antecedent. In the syntactic context of the negated antecedent, the missing argument is interpreted as the negative polarity indefinite *anything*; this may also be influenced by the pragmatic factors at play, since the the previous sentence *all this and not ordinary* is naturally compatible with a situation where whatever is being discussed, the thing that is not ordinary, does not resemble anything.

In both of these examples, we see arguments missed out in a way that is not ordinarily seen in English, yet it seems that we are compelled to fill these structures in with some sort of generic, fit-for-purpose indefinite or pronoun in the interpretation. This is because the predicates which take these arguments take them obligatorily, and it is impossible to interpret these sentences without the presence of some argument in these positions. This shows that we cannot just alter the meaning of predicates just by missing out their arguments, but rather the lexical meaning of such predicates will remain stable in the face of deviant erasure. Quite why this should be the case is an interesting question in itself, and it is one to which we will return in what comes.

Before concluding this section, it is worth bringing attention to one kind of apparently allowable argument ellipsis in English, namely subject omission. As discussed above, erasure of the subject is common in the poetic texts, in particular the more experimental ones. Omission of subjects is permitted in many of the world's languages, such as Italian, yet in English this is generally

impossible except with imperatives, as demonstrated below:²⁸

- (235) a. (Lui) pensa che (lui) é bravo
 (He) thinks that (he) is great
 He thinks that he is great
 b. *A quale degli amici di Maria tu lo hai presentato?
 to which of-the friends of Maria (you) him have presented-2S
 ‘Which of Maria’s friends have you introduced him to?’
- (236) a. *(He) thinks that *(he) is great.
 b. *Which of Mary’s friends have *(you) introduced to?
 c. (You) run away!

These examples represent the standard view from descriptive and theoretical linguistics; as a result English is known as a ‘non-pro-drop language,’ and Italian is called a ‘pro-drop language.’

However it is know that, in certain styles of written and spoken English, pronoun subjects are sometimes omitted optionally; this is known as ‘diary style’ in written English (Haegeman 1990, 1997, Haegeman and Ihsane 2001), and it is also widely reported in spoken English too (Biber et al 1999: 1048, Thrasher 1977). (237) demonstrates some examples of diary register, and (238) some examples of spoken ‘pro-drop’:

- (237) a. John came in late last night. (He) must have been at his friend’s all night.
 b. (I) finished work. (I) went home directly. (I) thought *(I) better do some reading.
- (238) a. Q: While you’re here, would you be able to do my grammaticality test?
 A: (I’ve) got to go!
 b. Q: Do you have any idea where John is?
 A: (I) dunno.
 c. (I’m) afraid there’s not much we can do. (Haegeman and Ihsane 2001)

²⁸A note on presentation: a * outside the bracket means omitting the element makes the sentences ungrammatical, i.e. **thinks that Giovanni is great*; a * inside the bracket indicates including the element in the bracket makes the sentence ungrammatical. No star with brackets indicates the element in the bracket is optional.

While Haegeman (1990, 1997) argues that diary drop is a root phenomenon – that is, restricted to matrix clauses – Haegeman and Ihsane (2001) provide evidence that shows that some diary registers also allow for omission of subjects in embedded clauses too. Below is one of their examples from *Bridget Jones's Diary* (Fielding 1996):

(239) Think will cross that last bit out as contains mild accusation.
(Fielding 1996: 227, cited in Haegeman and Ihsane 2001)

Analysing diary drop in a small corpus, Haegeman and Ihsane (2001) conclude that it probably involves some kind of argument ellipsis rather than a null proform, and they show that it is subject to a degree of dialectal variation.

Given that subject ellipsis is attested in various registers of English, I will refrain from counting examples of this as evidence that the theory of poetic language should account for. Examples of this kind of diary drop are indeed attested in the poetic texts analysed, but they tend to have a similar character to the diary drop data discussed by Haegeman's work. It may be the case that what we see in poetry and what we see in diary style has the same basic explanation, since the writing styles associated with diary drop are often consciously literary ones. Haegeman and Ihsane (2001: 333 fn.1) note that, while American English speakers tend to find diary style ungrammatical (unlike many British English speakers), this style is to be found in the diaries of the American experimental Allen Ginsberg (Ginsberg 1995), and they describe the style of Ginsberg's writing in his diaries as "generally fragmented," excluding the data from his diary for this reason. However, it is also of relevance that most of the other kinds of erasure described in this section are not found in diary style; if we were to assume that diary drop and erasure have the same basis, it remains a mystery why diary style does not also avail itself of the various different kinds of deletion found in poetic texts. I will therefore put this issue to one side for the remainder of this chapter.

2.2.4 Article erasure

This subsection focuses on the erasure of articles in poetic texts. Omission of articles and determiners is extremely common in poetry, and it is also common in a number of dialects and registers of English; for example ‘headlinese’, the variety of English used in newspaper headlines. Article omission is not as alien a form of deviation as some of the other kinds discussed in this chapter. Nevertheless all of the examples discussed here are clearly deviant, and they are not easily subsumed as simple examples of headlinese or a similar article-light variety,²⁹ so they remain of interest for the purpose of this chapter.

The English articles *the* and *a* distinguish different readings of the same nominals: condensing crudely, *the* creates a definite reading, where *the x* describes the unique entity *x* that can be identified in the context, and *a* creates the indefinite reading, where *a x* describes some unspecified entity *x* that might not be identifiable within the immediate context.³⁰ In situations where articles are missing, a given nominal can become ambiguous between these two readings. In languages like Russian and Chinese, which lack articles altogether, this is a common issue that is resolved with reference to the context of the utterance but in standard English articles are typically used to distinguish one meaning or the other.

Missing out articles in an inconsistent manner may thus cause problems for the interpretation of nominals in poetic texts. This is demonstrated by the following examples:

- (240) Walked to past now dream
of __ previous place [...]

Robert Creeley, ‘In London 3’

- (241) A plate that has a little bobble, all of them, any so.
Please a round it is __ ticket.

²⁹For discussion of the properties of article drop in headlinese, see Weir (2009). I will not go into a detailed explanation of why the article omission we see in these poetic texts is unlike that in headlinese, but I will just note that the forms we get in poetic texts are the kinds of forms that Weir describes as unattested in headlinese.

³⁰For discussion of (in)definiteness that goes beyond this crude summary, see Heim (1982), Reuland and ter Meulen (1987) and the references cited therein.

TB: 17

In (240) singular nominal *previous place* is contained in a PP, and it is missing an article, which it would standardly require. In this particular example it is unclear whether this should be interpreted as a definite or indefinite NP, since the prior context does not give any indication of whether or not a specific previous place may be what is referred to here. In (241) the NP *ticket* is missing an article, and it is similarly ambiguous between the definite and indefinite readings, although its presence as the predicate in a copular construction favours an indefinite interpretation over the definite one. The ambiguity of these erasures contributes to the difficulty of interpreting the texts at the sentence level, and it feeds into the opacity of the texts overall. There are many similar examples in Creeley and Stein's texts, as the omission of articles and functional elements is very common in their poetry, although there are no obvious regularities to this heavy erasure.

Nevertheless, an interesting aspect of examples like this is that they are not more ambiguous than they are, since we may expect missing out an article to introduce other readings that are not actually attested. For example, consider (240): the nominal *place* here cannot be interpreted as a mass noun, like *coffee*, nor can it be interpreted as an abstract noun, like *experience*, even though these readings do become available for *coffee* and *experience* when they appear without articles. The noun *place* could plausibly receive such interpretation, given the basic meaning of the lexical item: *place* is an uncountable and dense much like *coffee*, and likewise *place* can also be thought of as an abstraction broadly similar to *experience*, and it is often used in literary theoretic contexts.³¹ In the context of this particular example, such interpretations would be conceivable in terms of pragmatics, since the first part of the sentence describes someone 'walking to past' and the nominal in question is being dreamt of: this is a situation where an abstract interpretation of *place* would be sensible. Yet this

³¹An example of this comes from the postgraduate curriculum here at Strathclyde University, where there is a literature Masters course called 'Literature, culture and place.'

interpretation is absent, and it seems counter-intuitive to suggest that it would ever have such an interpretation. Deviant examples like (240) are ambiguous, but they are not that ambiguous. *Place* must be interpreted as a nominal that is missing an article, rather than as a different kind of nominal with a mass or abstract interpretation. Examples like this show us that, regardless of deviation or erasure of elements that narrow the semantic interpretation of linguistic expressions, the lexical semantic content of the nominals is always crucial to its interpretation, and the abuses of deviation and erasure in poetic texts cannot alter that.

There are other examples where articles are erased but the lexical meaning of the noun or its syntactic context disambiguates the string in favour of either the definite or indefinite reading:

(242) Out of kindness comes redness and out of rudeness comes rapid ___ same question, out of an eye comes research, out of selection comes painful cattle.

TB: 4

(243) There is ___ hollow belt, a belt is a shawl.

TB: 16

(244) ___ Book was there, it was there. ___ Book was there.

TB: 17

In (242), the nominal *same question* is missing an article,³² but it is interpreted as *the same question* because the modifier *same* does not occur with bare or indefinite forms. In (243), *hollow belt* is necessarily interpreted as *a hollow belt* because it occurs in an existential construction, which cannot host definite noun phrases (except on a ‘presentational *there*’ interpretation) due to the definiteness restriction: *There is the hollow belt* is ungrammatical (Milsark 1974). In (244), *book* appears twice in the text without the article, and the title of the text

³²Here I parse the nominal as *same question* and not *rapid same question*, since this ordering of adjectives would be strange; rather, in this context it seems right to interpret *rapid* as an adverb that modifies the VP. The adverb would normally appear as *rapidly*, but adjectival forms lacking the *-ly* morpheme are frequently used as adverbs in many dialects of English, such as in the sentence *Come here quick!*. This kind of usage is common in Stein’s texts.

is also *Book* without an article. The sentence *book was there* occurs twice in the text, and between these repetitions we have a parallel sentence which has the deictic pronoun *it* in place of *book*; due to this parallelism and the lack of any other obvious antecedent, the most natural interpretation of the text is one where the pronoun is coindexed with *book*, where all three sentences effectively mean the same thing. In this situation *it* must be a deictic pronoun that refers to an entity in the discourse, as it does not appear in the requisite syntactic structure for a bound reading; that is, a reading where it can covary with a quantifying expression like *a book*, as in the sentence *Each person took a book from the shelf and opened it*.³³ Since the pronoun picks out a specific referent, the most natural reading for *book* is a definite one as *the book*, and this reading is preferred for all the occurrences of the bare noun in the sentence.³⁴ Thus the sentence is unambiguous like the other examples, even though the local syntactic context of the nominal does not determine its (in)definiteness entirely. All of these examples show that even when functional elements that make important semantic contributions are erased, in many cases the syntactic context will lead to a single interpretation. Thus in these cases erasure is simply a word order effect, and it has little effect on the interpretation of the examples.

2.2.5 Summary

In this subsection we have seen numerous different kinds of erasure in poetic texts, where words that are interpreted in a given sentence are missed out in the surface string. We have seen that in some cases the erasure of words can cause few problems for interpretation, since the syntactic, semantic and discourse contexts can provide us with enough information to determine what the missing elements should be. We have also seen that the tendency is for erasure of functional elements that are low in functional meaning; in situations where

³³In a case like this, the coordination would need to be asymmetric, allowing the NP *a book* to c-command into the second sentence. For arguments in favour of asymmetric approaches to coordination, see Munn (1993); for discussion of the c-command restriction on bound readings of pronouns, see Reinhart (1983).

³⁴Although see Heim (1982) for discussion of situations where deictic pronouns pick out indefinites.

it seems that elements with lexical content have been erased, the dependency upon utterance context for the meaning of the erased constituent is far greater. Erasure applies to a number of different kinds of elements in various poetic contexts in many different kinds of poetry, and while most of the examples here have been from experimental texts, it is clearly not restricted to texts of this kind, as the work of Austin (1984) and Dillon (1975) attests.

2.3 Overall summary

Throughout the previous subsection I have described the different erasures in terms of their comparison with standard ellipsis phenomena in English, and in numerous places it has become apparent that erasure in poetic texts behave unlike standard ellipsis mechanisms in many ways, targeting non-constituents and erasing elements that are not normally licensed for deletion. Erasure often seems arbitrary, motivated neither by emphasis or by evasion of redundancy or repetition, and it is largely inconsistent, in that sometimes an element will be erased right next to an unerased occurrence of the same element. In this respect, erasure is in fact very similar to the displacement phenomena we saw in section 2.1, since displacement also targets numerous elements that its standard analogue does not target, and it is more proliferate and unconstrained. Both displacement and erasure alter the word order of a given string without making significant changes to its meaning, and both place a heavier dependence upon context and the syntactic information of the residual well-formed parts of a sentence when they deviate further from their standard analogues.

And while displacement and erasure seem similar in the manner of their application, it also seems that they are very different in terms of their distribution. We saw in section 2.1 that displacement almost exclusively affects full phrases or proper parts of phrases, moving them left and right in the sentence to various different positions; we also saw that there was a complete absence of ‘head displacement,’ that is, displacement of individual syntactic heads like verbs or

head nouns. Parallel to this, another descriptive generalization is that displacement always affects constituents with lexical semantic content, as there are no instances of displacement affecting functional elements like auxiliary verbs, low-content prepositions or determiners. What is interesting is that, with both of these generalizations, erasure is in complete contrast with displacement: it mainly targets elements with little or no lexical content, and the general tendency is for it to target syntactic heads like auxiliaries and determiners rather than full phrases.³⁵ The fact that the two phenomena pattern together with respect to the general pattern of their application in poetry on the one hand and their distribution with respect to syntactic elements on the other – modulo some unusual exceptions, such as the fact that they both target non-constituents – is an intriguing fact that needs to be explained by our theory.

This is the main focus of chapter 4. In the next chapter, however, I will show that these wide generalizations, and the smaller empirical generalizations that feed into them, are not dealt with adequately by previous theories of poetic language. I conduct an in-depth critical examination of three major existing theories and from this we learn that certain approaches to the observed patterns will not suffice. This provides the platform for a more radical alternative, pursued in chapter 4.

³⁵On the view that pronouns are determiners, we might say that erasure of pronouns is also erasure of a syntactic head. Recall from section 2.1.5 that pronouns are also targeted by displacement much less than full NPs.

Chapter 3

The theory of poetic language: previous approaches

This chapter critically reviews previous attempts to develop linguistic theories of poetic language. I discuss the issues that these theories aim to tackle, and I then discuss the theoretical and empirical shortcomings that they are unable to address, drawing from the broadened empirical picture sketched in chapter 2. I argue that the theories' shortcomings indicate that an entirely different approach to the problem of poetic language is required, and I set the stage for chapter 4 by reviewing the specific empirical points that a fully adequate theory must deal with.

The chapter is structured as follows. In sections 3.1 I discuss older approaches to poetic language in transformational grammar, and I show that not only are they unable to account for the data in themselves, but also that they are incompatible with developments in modern linguistic theory. In section 3.2 I discuss the account of O'Neil (2001), which proposes that poetic language involves borrowing rules that are available from other languages to generate the relevant

structures. I show that this theory both undergenerates and overgenerates massively, and I also point out some serious theoretical problems that make this proposal untenable. In section 3.3 I discuss Fitzgerald's (2007) OT approach to poetic language (representative of a wider trend in literary linguistics) that generates poetic inversion by allowing syntactic and metrical constraints to interact directly. I show that this theory suffers from internal problems, and I then discuss a number of other issues that compel us to reject Fitzgerald's specific account, and more generally the idea that syntax and metre can interact directly. The final section concludes by diagnosing the fundamental flaws that breed the problems for all of these theories, namely the hypothesis that poetic language sentences are generated just like any other sentences by an extension of the grammar, a 'poetic grammar.' I argue that this assumption must be rejected and summarize the specific empirical and theoretical challenges that the alternative theory needs to meet.

3.1 Earlier approaches: transformational poetic grammar

The most prominent theory of poetic language in generative linguistics over the last few decades can be called the theory of 'poetic grammar.' This theory states that the non-standard forms of poetic texts are generated by the grammar either by a set of extra rules or operations that are not normally available with standard English, or by the omission of a set of constraints that normally apply in English. According to this theory, the grammar used by poets (and, presumably, readers of poetry) is different from the standard grammar, since it includes some extra rules and excludes a set of syntactic constraints; this is the poetic grammar, and the theory of poetic language is concerned with determining the properties of this poetic grammar. Although the other two generative theories reviewed in this chapter are also theories of poetic grammar in some sense, they are associated with altogether different frameworks that bear different relations

to the idea of a poetic grammar (in ways that will become clear in those sections). The work reviewed here was developed in the Extended Standard Theory framework of Chomsky (1965), and although the different accounts exhibit differing degrees of methodological rigor, ultimately they amount to theories that share as their common core the belief that poetic language is generated by the grammar in the same way as standard language. Here I will give an overview of some of these works, and I will identify the flaws they share by examining the differences they claim for themselves.

Dillon (1975) is one of the earliest and most substantial attempts at developing a model of poetic language in the generative era, and it is representative of the other work done in this field at the time (e.g. Thorne 1965, Levin 1967, Banfield 1973). Dillon describes and analyses a number of different kinds of nonstandard inversions and deletions (called displacement and erasure in chapter 2 here), and in his taxonomy he breaks down the kinds of deviant into three basic categories. He argues that these inversions and deletions “can be described by rules which resemble common optional transformations, but differ from them in certain ways” (Dillon 1975: 220), and he describes them as below:

There are rules which

- i. operate as in non-poetic Modern English but
 - a. affect a broader range of elements (e.g. deletions under identity)
 - b. are ordered or written slightly differently (Topicalization, PP-fronting)
 - c. are triggered by a broader range of environments (S/V Inversion)
- ii. differ from any rule of Modern English (Verb-Final)
- iii. affect structures which are normally islands (i.e. immune to chopping)

(Dillon 1975: 220)

Dillon goes on to provide examples of each of these sets of rules, and these examples all overlap to some extent with those discussed in the previous chapter

here. He argues that these rules “characterize a kind of extra syntactic competence required of the reader,” and that they can be “viewed as relaxations of constraints on transformations in Modern English” (Dillon 1975: 221). For Dillon, then, a theory of poetic language requires us to describe the different transformations and rule relaxations that produce the poetic forms; the grammar that encompasses these rules is the poetic grammar. The works of Thorne, Levin and Banfield amount to effectively the same theory, with some non-trivial technical differences that I will not discuss in depth here.¹ The most important point for Dillon is that the ordinary language grammar and the poetic grammar are as similar as possible, and that the linguist should posit as few new rules as possible.

Austin (1984: 25-35) critiques this approach, arguing that such an approach to formulating the poetic grammar leads to a model of the grammar that is unconstrained and unwieldy. He does this by looking at the e.e. cummings poem ‘anyone lived in a pretty how town,’ which is also the subject of analyses by Thorne and Levin. Below is an excerpt from the poem which will suffice for our purposes here:

anyone lived in a pretty how town
 (with up so floating many bells down)
 spring summer autumn winter
 he sang his didn't he danced his did

Women and men (both little and small)
 cared for anyone not at all
 they sowed their isn't they reaped their same
 sun moon stars rain

children guessed (but only a few
 and down they forgot as up they grew
 autumn winter spring summer)
 that noone loved him more by more

¹For example, Levin (1967) argues that the poetic text should be taken not as a variation on the standard grammar, but rather as a sample of a singular dialect that has a different grammar from standard English. Nevertheless, in almost any implementation this would amount to almost the same result as an approach that simply amends the grammar, since one proposes a grammar for the text that is ultimately made up of the rules that generate the sentences that overlap with the standard language plus the ones that generate the non-standard ones. The two grammars are thus descriptively identical. In what follows it will become apparent that the two approaches also share the same theoretical and empirical problems.

when by now and tree by leaf
 she laughed his joy she cried his grief
 bird by snow and stir by still
 anyone's any was all to her

(cumplings 1994: 515)

cummings' poetry is well-known for its syntactic deviation, and the aspect of this particular text that Austin concentrates on is the use of negative polarity items (NPIs) like *anyone* in contexts that do not normally allow for NPIs, that is, outwith the scope of a sufficiently local downward entailing operator (Linebarger 1980). Austin observes that the NPI is interpreted as a 'name' for a male character, and that this parallels the identification of the negative indefinite NP *no one* with a female character. These interpretations are produced by a number of different factors, but primarily by the use of gendered pronouns that seem to be coindexed with the indefinites; for example the first stanza where *he* in line 4 is interpreted as coindexed with the *anyone* in line 1, and this mode of reference is used consistently throughout. To account for this fact, Levin and Thorne would state that the lexical item *anyone* is identified as a name here, or that the rules that typically place restrictions on the distribution of NPIs are relaxed or excluded from the grammar of the poem. Thus all of the uses of NPIs in the text would be judged acceptable and normal within the context of the poem.

However, Austin (1984: 25-34) argues that to do so is to impoverish the text in a way that removes a great deal of the subtlety of the poem's meaning. He tracks an interpretation of the poem which relies crucially on realising that the nonstandard uses of NPIs are indeed nonstandard uses; in this interpretation, the unacceptable uses of NPIs all coincide with situations where the character denoted by *anyone* is experiencing adversity, such as in lines 5-6 where it is revealed that many people around the town *cared for anyone not at all*, and Austin takes the syntactic adversity of the ungrammatical uses of the character's name *anyone* to be a reflection of this predicament. Austin observes that this approach to the interpretation of the poem is unavailable in the accounts

of Levin and Thorne, since they necessarily give the uses of NPIs the same status as all the other well-formed sentences. Austin argues that his analysis should caution us against admitting new rules to the grammar so easily, since to do so without full consideration of the subtleties of interpretation may be to impoverish our analysis.

Austin's criticism is not a theoretical one, however, but a methodological one, in that he is still willing to admit "delimited [adjustments] to some aspect of standard English syntax" (Austin 1984: 50). He subscribes to a theory of poetic grammar that is broadly similar to that of Levin, Thorne and Dillon, but instead promotes a much higher degree of caution in admitting new rules to the grammar, and insists that some sentences should still be recognised as deviant and not as the output of a specifically poetic grammar. Specifically he quotes Dillon's statement that "one must consider the grammar of the poet as a whole" (Austin 1984: 47) and argues that this should be a guiding methodological principles; in effect, this means that the rules that we add to create our poetic grammar should represent regularities in form that are attested throughout the given corpus of texts.² Austin subsequently presents an example of such a trait from Shelley's poetry, which he identifies as "backwards gapping;" this overlaps with some of the data discussed at the end of section 2.2.3 above. Since he finds backwards gapping to be a relatively regular construction in Pope's poetry, he posits that the introduction of a rule similar to that which derives standard gapping would be justified in this case. Thus despite some differences with respect to coverage of the data and methodology, the theories of Austin, Dillon and their contemporaries were united in the assumption that there exists a variation on the ordinary language grammar that is used to generate the unusual sentences of poetry.

²Dillon's quote seems to suggest that we should study the grammar of the poet as an individual, that is, his/her personal idiolect. However it is unlikely that Dillon would have actually meant this, since he cites work by Ronald Emma (1964) (mentioned earlier here) which shows that the S-O-V order that is so common in John Milton's poetry is in fact restricted to his poetry, and is absent in his substantial prose writings. It seems that what is referred to in Dillon's quote is the poet's grammar as a poet, i.e. the grammar of his/her poetic texts.

3.1.1 Empirical problems

The poetic grammar theory suffers from a large number of theoretical and empirical problems, some of which are produced by theory-internal inconsistencies, and some of which have been brought about by the advances of linguistic theory in recent years. I will first outline some of the empirical problems experienced by the different versions of the theory, all the while explaining the theoretical or methodological issues of which they are a symptom.

First, I will review technical and empirical problems with the specific proposals within the poetic grammar approaches; that is, problems for the rules and constraints formulated by the cited authors. One simple but instructive empirical problem for Dillon's paper is that he proposes that some rearranged forms are not attested and hence should be excluded by the poetic grammar, and yet these very forms can actually be found in poetic texts. Dillon discusses the 'Verb final rule,' a rule that produces the S-O-V word order in a given text by inverting the verb and its complement. Below is one of Dillon's examples from Dryden:

- (1) With secret joy indulgent David viewed
 His youthful image in his son renewed
John Dryden, 'Absalom and Achitophel'

Dillon formulates this as a rule that can "apply either to AUX+(ADV)+V or just to V, and to all of the complement to the verb, or just the direct object" (Dillon 1975: 226). This obviously gives the rule a degree of freedom, although it also commits to moving adverbs and all auxiliaries as a single unit.

We have already seen numerous similar examples in the previous chapter that were analysed as situations where the complement had been displaced, and that the verb had stayed put; although we saw that it was not always possible to know which analysis was correct, the fact that argument displacement could displace one VP-internal argument while leaving another in a postverbal position indicated that it was the argument and not the verb that had been displaced (see section 2.1.2 in particular for discussion). Below is one example:

- (2) Close by those meads, for ev'r crown'd with flow'rs,
 Where Thames with pride surveys his rising tow'rs, *TR3: 1-2*

This is not captured by the formulation of Dillon's rule, since it is a PP *his pride* and not a direct object that is displaced. This may be accounted for by tinkering with Dillon's rule somewhat, but this is missing the point, since what examples like (2) and the numerous others in chapter 2 showed was that the displacement of NPs and PPs from their standard position is largely unconstrained. Numerous other cases cannot be accounted for by this V-final rule, such as the reversal of the order of arguments within the VP (see for example (117) in section 2.1.2), or the reversal of the order of VP-arguments through multiple displacement to a pre-verbal position (see for example (88) in section 2.1.2); to account for these examples, at least within Dillon's account, we would need an additional set of rules or rule relaxations.

Dillon does propose that the application of the reordering rules is constrained, citing the non-occurrence of examples of the form PP-S-O-V as an apparent constraint; he speculates that this may be because the two dislocations proceed in different directions, with the verb moving rightward and the PP moving leftward. However, it turns out that there is no such constraint, as there are examples of this form to be found in Pope's work:

- (3) Favours to none, to all she *smiles* extends;
 Oft she rejects, but never once offends. *TR2: 11-12*
- (4) With hairy sprindges we *the birds* betray,
 Slight lines of hair surprize the finny prey, *TR2: 25-26*

In addition, the following demonstrates an example of O-S-(aux)-PP-V, which is effectively the same only the two different arguments have changed positions:

- (5) This the Beau-monde shall from the Mall survey,
 And hail with Musick its propitious Ray.
This the blest Lover shall *for Venus* take,
 And send up Vows from Rosamonda's Lake. *TR5: 133-136*

This shows that Dillon's 'double dislocation' constraint does not hold for Pope's

poetry, and that formulating constraints based on which direction the displacements proceed is likely to be misguided. Dillon's Verb Final rule may capture some surface orders adequately, but it is just one of a number of ways of accounting for these surface forms. These forms could also be accounted for if we were to propose that the arguments were being displaced leftwards by some movement rule to a position between the verb and the subject, as was implicit in the description of the data in the previous chapter, and given that these rules or similar ones are needed to account for some other forms, it seems that the Verb Final rule would be redundant and undesirable.

There are other reasons to doubt the validity of this particular rule that come from general considerations of how linguistic rules should be formulated. Dillon proposes that the poetic grammar should be as similar to the ordinary grammar as possible, so we would expect that these rules would behave broadly like normal rules, obeying constituent structure for example. Indeed Dillon makes a related observation, noting the fact that the examples he discusses all involve obedience of constituent structure and commenting that this indicates these rules have a great deal in common with normal linguistic rules, indicating that his approach is on the right track (Dillon 1975: 231).³ The problem, then, is that developments in linguistic theory since the Extended Standard Theory have made derivations like the one proposed for Verb Final more or less impossible: rewrite rules were replaced by derivations that involve movement of phrases to specific landing sites, constrained by conditions on locality. Rightward movement derivations do exist in the literature, but these are typically restricted to rightward A-bar movement of argument NPs and PPs to adjunction positions, where the arguments still move up the tree up to right-adjoined positions that are linearized in positions to the right of the TP.⁴ I am aware of no examples of rightward head movement in the descriptive or theoretical literature, so a derivation where the verb is moved rightward by head movement seems implausible.

³Though see the discussion of pronoun displacement in section 2.1.5.

⁴Though see Kayne (1994) for theoretical and empirical arguments against the existence of rightward movement.

A phrasal movement analysis is equally implausible, since this would necessarily involve movement of the direct object and other VP-internal arguments with the verb, making the movement string vacuous.

It is also difficult to see what phrasal constituent could be moved rightward to derive the S-O-aux-(adv)-V order, since the verb and auxiliary do not form a constituent that excludes the VP arguments. A multiple rightward head movement analysis for this order (i.e. aux and V both move by rightward head movement past the object) would be equally implausible, independent of the oddity of rightward movement in itself, since it would surely make predictions that are not fulfilled. We know from the data reviewed in the previous chapter that multiple leftward A-bar movement of argument PPs and NPs can often reverse their order, so if we were to allow for multiple rightward head movement we would surely predict that it would be possible to reorder the verb and its auxiliaries, to produce orders like S-O-V-aux, i.e. *John the ball kicked has* (see sections 2.1.5 and 4.2.5 for discussion). However, I have found no examples of this order in the corpus that I have analysed, and I have found no attestations of this order in any of the other works cited either. To account for this asymmetry we would have to make extra stipulations to account for the differences between rightward head movement and leftward A-bar movement, further weakening the case for the rightward head movement account of the verb final order. It seems that the Verb Final rule is too troubled by technical and empirical problems to remain a plausible rule in the poetic grammar, and that, within the framework of the poetic grammar approach, a reanalysis of the evidence for Verb Final in terms of (multiple) leftward displacement of arguments would be preferable.

It is not just Dillon's apparently "incautious" approach to rule formation that suffers from problems, however. Austin proposes an account of "backwards gapping" that should be admitted to the poetic grammar, and this also suffers a number of problems in the context of modern accounts of gapping. The phenomenon is demonstrated by the example below:

- (6) While fish ___ in streams, or birds delight in air,

Or in a coach and six the British fair,

TR3: 163-164

The gap in the first sentence is interpreted as *delight*, under identity with the second conjunct. The sentence is unusual since gapping normally has the gap in the second conjunct, as in (7) (from Austin 1984: 49):

- (7) a. Pope lived near London and Swift lived in Ireland.
 b. Pope lived near London and Swift ___ in Ireland.
 c. *Pope ___ near London and Swift lived in Ireland.

The poetic example (6) (as well as the other examples cited by Austin and presented in section 2.2.3) represents the ungrammatical version in (7c). Austin proposes that we should “extend the grammar of standard English so as to permit leftward or ‘backward’ gapping” (Austin 1984: 49) in order to account for these examples within the poetic grammar. This was a plausible account in the EST framework within which Austin was working, as gapping was seen as a simple case of a string deletion rules, more specifically as a sub-form of ellipsis.

However, in recent times it has become increasingly apparent that such an approach would not work for gapping. The expanding literature on gapping, most notably the work of Kyle Johnson (2004, 2009), has shown that an ellipsis analysis is largely unworkable, since gapping is provably different from VP-ellipsis in a number of ways; the majority of these comparisons also separate gapping from pseudogapping, where pseudogapping is seen as a sub-type of VP-ellipsis, as argued by Lasnik (1999).⁵ First, gapping can only occur with coordination, whereas VP-ellipsis and pseudogapping can occur in constructions with subordinating conjunctions:⁶

- (8) a. John ordered mussels because Mary did ~~order mussels~~.
 b. John has ordered mussels because Mary has ~~ordered~~ tuna.
 c. *John has ordered mussels because Mary ~~has ordered~~ tuna.

⁵Though see Thoms (to appear) for discussion of an alternative analysis in light of Gengel’s (2007) focus movement analysis and the fact that some languages that lack VP-ellipsis still allow for pseudogapping (i.e. Norwegian).

⁶The data presented here are versions of those presented in Johnson (2009: 293), which also provides a number of references for other discussions of these data points.

Second, pseudogapping and VP-ellipsis are possible in embedded contexts, but gapping is not:

- (9) a. John ordered mussels and Mary claims that others did ~~order mussels~~, too.
 b. John has ordered mussels and Mary claims that others have ~~ordered~~ tuna.
 c. *John has ordered mussels and Mary claims that others ~~has ordered~~ tuna.

Third, the antecedent for pseudogapping and VP-ellipsis can occur in an embedded context, but this is again not possible for gapping:

- (10) a. I hope that John arrives early. I know Mary will ~~arrive early~~.
 b. ?She's said Peter has eaten his peas, and Sally has ~~eaten~~ her green beans, so now we can have dessert.
 c. *She's said Peter has eaten his peas, and Sally ~~has eaten~~ her green beans, so now we can have dessert.

In addition, the availability of gapping as an ellipsis strategy would leave unexplained the fact that NP-subjects do not normally license ellipsis. That is, if gapping is deletion of a verb and additional auxiliaries under identity with an antecedent, why is this only available when there is a VP-internal argument to survive the deletion, and never with intransitives? (11) demonstrates:⁷

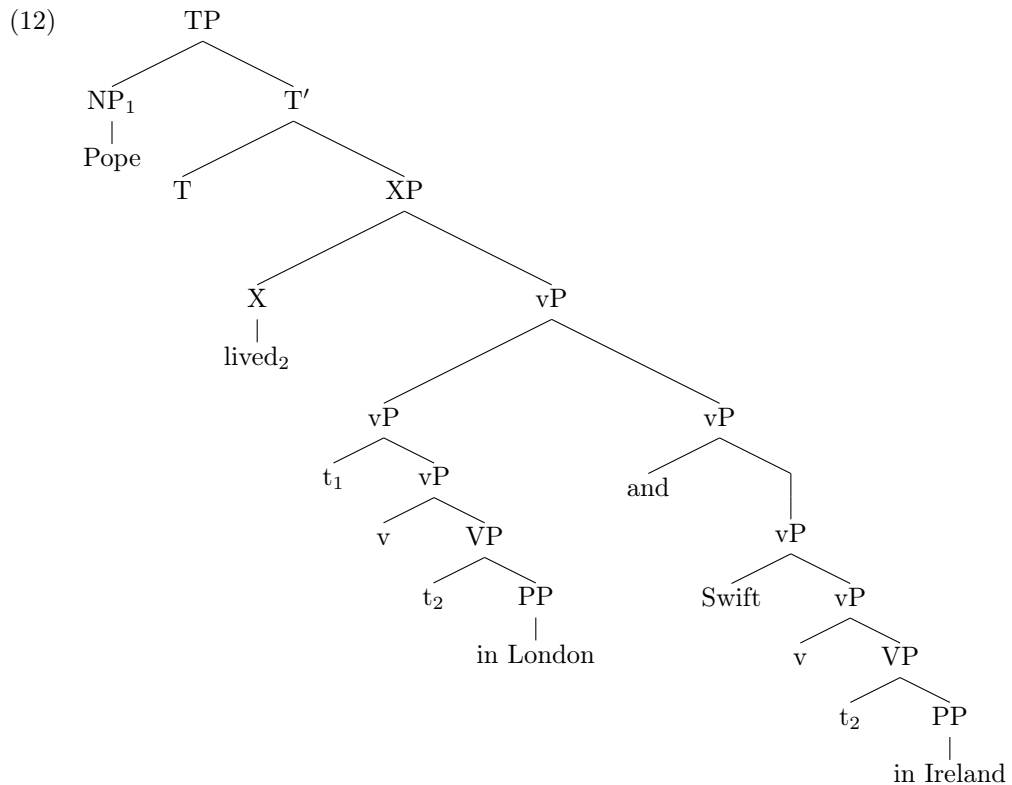
- (11) a. John has arrived, and Mary has ~~arrived~~, too.
 b. *John met Sally, and Mary ~~has arrived~~, too.
 c. John has arrived early and Mary ~~has arrived~~ late.

If the NP subject *Mary* can license ellipsis in the gapping sentence (11c), it remains unclear how we can explain the ungrammaticality of (11b).⁸

⁷Examples like (11b) can be made acceptable by placing heavy focal stress on the subject, and this kind of construction is known as stripping. However, stripping is best explained as something like a fragment answer, where the surviving subject *Sally* moves to a focus position before ellipsis applies to the full sentence; Merchant (2003) discusses these examples and shows that stripping remnants exhibit most of the same form-identity trends as fragment answers. If this is so, stripping can be disregarded with respect to the point being made here.

⁸This additional point on licensing is less significant than the others on antecedence, however, since licensing seems to be at least partially a phonological phenomenon (see Lobeck 1995, Thoms to appear) and hence more subject to crosslinguistic variation. The conditions on antecedence are more general, with no obvious relation to morphosyntactic or lexical variation, so one cannot just deal with these by appealing to exceptional dialectal variation (as one might for licensing).

Facing these problems, Johnson (2004, 2009) has developed an alternative account that aims to explain gapping in terms of movement rather than ellipsis. He proposes that gapping constructions involve coordination at the vP level, with the subjects originating in Spec,vP; then the gapped constituents move to right-adjointed positions within the VPs, and the VP predicates undergoing Across-the-Board (ATB) predicate movement to an XP projection above the vP projection (identified in Johnson 2004 as PredP); finally the subject of the left conjunct undergoes A-movement to Spec,TP, while the subject in the right conjunct stays in situ. Below is a simplified tree structure for (7b) which demonstrates this:⁹



I will not discuss the details of Johnson's account here (they vary between different variations of this analysis), but it is enough to note that each aspect of

⁹Here I factor out some complications for exposition, such as the rightward movement of the PPs within the VPs. This component is crucial to Johnson's account but not crucial to the explanation here, since we are largely concerned with capturing the ATB verb movement, which is important for what follows.

the account is motivated independently and, while there are still some remaining problems (some of which Johnson confronts), this account is the closest there is in the literature to a workable account for gapping, in English at least.

The important thing here is that under this account is that it cannot derive backwards gapping with any plausible reconfigurations or alterations in the poetic grammar: there is no combination of moves that can derive the required order. If we were to propose that predicate movement was suspended in both clauses, we would derive a simple coordination with no gap; if we were to move just the verb from the first conjunct (violating the ATB constraint on coordination extraction), we would derive the same structure; if we were to extract just the verb from the second extract, we would derive a sentence where the verb appears twice in the first conjunct and not at all in the second (*while Fish delight delight in streams or fish in air*). Since no alteration can be made to derive backwards gapping from the same mechanisms as normal gapping, a defender of poetic grammar would have to explain backwards gapping by some other, unrelated mechanism, counter to the guiding principles of the approach, which endeavours to keep the poetic grammar as close to the standard grammar as possible.

This would mean that Austin's backwards gapping rule would fall under category (ia) in Dillon's typology of rules, where backwards gapping involves ellipsis affecting a broader range of environments than normal.¹⁰ Dillon's Verb Final is part of his category (ii) set of rules, which are rules that are unlike others in the language, and I have argued that the empirical and technical problems encountered by this rule indicate that it would be better analysed as a series of leftward NP/PP movements. Maintaining this general approach to poetic language, the plausible analysis for this would be that these leftward movements were a subtype of other kinds of leftwards movements, such as the standard rules for PP- and NP-fronting, i.e. topicalization. This would put these examples under category (ib), along with the other examples of double

¹⁰This is broadly similar to the proposal put forward in chapter 4, though different in its technical details. See section 4.3.2 for extensive discussion.

topicalization discussed by Dillon. Given this, we may speculate as to whether all of our problematic cases can be plausibly reduced to variations on standard rules with some simple alterations: extension of landing sites for movement, extension of the possible number of movements, extension of the kinds of deletable constituents, etc.

For concreteness, let us consider an example of what would previously have been analysed as Verb Final, which would now be analysed as a series of leftward movements in the poetic grammar framework.

- (13) Fair tresses man's imperial race ensnare,
And beauty draws us with a single hair. *TR2: 27-28*

In this example, we would propose that the NP *man's imperial race* has been moved leftward to a position between the subject and the verb. We may assume that this is a sub-type of topicalization, given that this is the only other leftward movement operation available for non-wh-NPs in English; Dillon does something similar in accounting for examples of double leftward displacement, like (14) below:

- (14) *With anxious beating hearts* the dire event they wait,
Anxious, and trembling for the birth of Fate. *TR2: 141-142*

In this example we would have extended use of topicalization, with the NP *the dire event* undergoing topicalization to a position below the PP *with anxious beating hearts*, which has already been moved to Spec,CP. With these two examples in mind, we would propose that topicalization has been doubly extended, to allow for double topicalization and to allow for 'intermediate' topicalization to a pre-verbal position. This would explain the fact that we also get standard topicalization alongside displacement to the pre-verbal position, as well as double topicalization to the pre-verbal position. Extending topicalization would seem cover a significant amount of the data.

However, there are a number of problems for this explanation. The first problem is a simple mismatch between standard topicalization and these examples with respect to their semantic and prosodic properties. If we were to

explain these displacements as instances of topicalization, we would expect that they would all have most of the semantic and prosodic properties normally associated with standard topicalization. However, most of the evidence indicates that this is not necessarily the case. Constable and Aoyama (1999) provide statistical evidence that shows there is no clear correlation between poetic inversion structures and the prosody of the poetic line, arguing against the claim that poetic inversion should be associated with specific phonological properties. It is also clear that poetic examples like (13) and (14) do not require the comma intonation associated with standard topicalization; this would be particularly unnatural for examples like (13). It is also not clear that these examples would necessarily be restricted to the topicalization interpretation of the moved NPs and PPs. Topicalization involves moving the topic of the discourse – the thing that is ‘old’ information – to the front of the sentence. The topic therefore has a specific semantic role in the context, and as a result it can be identified by tests that identify such an element, such as the “what about X” test (Gundel 1974); if one could plausibly interject “what about X” before a sentence containing X, that X element can be identified as a topic. We can see in the following that the test environment is compatible with the topicalized elements, but not the other elements in the sentence:

- (15) A: What about beans?
 B: Beans, I like.
- (16) A: What about the old books?
 B: #John, I gave the old books to.
 B': The old books, I gave to John.

We do not need to go into the full details of the analysis of topicalization (see Gundel 1974, Reinhart 1981 among many others), but it is enough to note that this kind of semantic interpretation is not at all required for the examples from poetic language discussed above. The surrounding discourses of the poetic texts almost always indicate that these interpretations are not intended; for instance, in (13) the displaced argument *man's imperial race* is not salient as a topic in

the surrounding discourse, and as such it would probably fail the “what about X” test. It seems, then, that these examples of displacement have very little in common with the operations on which they are supposed to be parasitic, other than the fact that they involve leftward movement.

At this point we can note that this difference between standard language topicalization and the unusual inversion structures in poetry is not just restricted to these unusual structures. Within the poetic texts, there are numerous examples of NP- and PP-displacement that could be analysed as topicalization since, on the surface, they resemble standard topicalization perfectly in terms of word order. Yet these examples are just as different from standard topicalization as the unusual leftward movement examples are, with respect to the semantic and prosodic properties of topicalization. This indicates that, even when an existing movement operation would derive the word order we find in a given poetic text, it is not necessarily the case that the word order of the poetic text was derived by this operation. The lack of a given semantic interpretation is just as relevant to the analysis as the correlation in word order, so it seems that examples of apparent topicalization may also be analysed as examples of deviant displacement. Given that we found many examples of apparent topicalization in environments that don't normally allow for topicalization, it seems likely that a large number of what we call topicalization is in fact something else entirely.

There is also the question of what kinds of landing sites are made available for movement. In the modern analysis, topicalization is either to the all-purpose Spec,CP position or to a more specific position in the left periphery, such as Spec,FocP or Spec,TopP (see Rizzi 1997 and the references cited therein for more finegrained analyses of the left periphery). If we were moving multiple phrases to this position, we would have to say that several Specs were made available for movement. To account for the movements to the pre-verbal position, we may posit that this is a case of movement to Spec,vP; given that most modern analyses of A-bar movement propose movement to Spec,CP via an intermediate landing site in Spec,vP (Fox 1999, Chomsky 2000, 2001, 2004, den Dikken 2007,

Thoms 2010c), we may propose that this is what we see in these examples; thus in (13) the NP *man's imperial race* would reside in Spec,vP.

All of this would be perfectly plausible and perhaps implementable without trouble, but it would still fail to account for a great deal of the data seen in chapter 2. First, we saw that there are examples of leftward displacement of PPs and NPs to clause-edge positions in infinitives, which are typically taken to lack any CP-layer and hence any potential landing for topicalization (see Chomsky 1981, Bresnan 1994); these examples could not be explained without significant stipulation or alteration to the theory. Second, we saw that displacement to an intermediate position between the verb and the subject was not a unified phenomenon, since there was variation in the position of the displaced phrase with respect to the position of the auxiliary. The examples where the displaced phrase precedes the auxiliaries could not be explained as movement to Spec,vP. This means that the topicalization analysis would still leave a significant body of data unaccounted for. Given that so much of the data is so similar, it would seem that we are missing a key generalization by giving them all separate explanations.

Thus we can see that the analysis of deviant leftward movements as an extension of standard topicalization for the poetic grammar is empirically flawed, just like the more specific proposals made by Austin and Dillon. The point of critiquing these specific rules is not to simply pick holes in Dillon and Austin's analyses, as they were proposed in a different framework with different working assumptions. Rather, what these empirical issues show us is that, within the highly constrained modern framework, it is difficult and sometimes even impossible to alter the grammar in minor ways to generate a varied set of examples attested from poetic texts. To argue that the attested poetic sentences are generated by the grammar by ordinary means, we then have to tolerate a number of other types of bad examples which are not attested in the poetic texts, presenting an overgeneration problem; we can also see that just extending the grammar a little, by altering an existing rule to account for some of the data, is also not

enough, since the data is too diverse for minor alterations to suffice. These problems are not only experienced in the realm of poetic grammar, of course, since all proposals in syntactic theory have a tendency to make incorrect predictions: it is the nature of the generative project to tolerate such problems in the hope that future research shall find a way to repair them. However, the proposed rules for poetic grammar are especially stipulative, since they aim only to describe a small subset of data (specific constructions) attested only in a particular situation, and as a result they tend to be more obviously problematic than those rules proposed to deal with standard data. The two specific construction-based rules discussed here, Verb Final and Backwards Gapping, are a good demonstration of the stipulative nature of the poetic grammar rules, since they are both effectively underivable with current technology; given that the general tendency in generative linguistics (since at least Chomsky 1977) has been to explain movement rules as general rather than construction-specific, these stipulative poetic grammar rules are problematic. The extension of topicalization would be implementable with some stipulation, but it would barely provide the right reward since it only covers a well-behaved subset of the relevant data.

I have critically reviewed just three of the rules proposed to account for forms within poetic texts. Given the massive diversity of the examples cited in the previous chapter, we would need to propose a number of different kinds of new rules which would be similarly stipulative and problematic. Furthermore there are a number of examples of non-constituents being affected by erasure and displacement, and these processes are often as regular as backwards gapping and hence as deserving of a rule-based analysis; yet it is impossible to derive rules which affect non-constituents (cf. Dillon's own comment cited above), and hence impossible to account for these examples in the same way. These examples in particular should be taken as instructive, since they are found across types of deviation (both displacement and erasure), so affecting non-constituents is a characteristic of poetic language that needs to be accounted for by our theory, albeit in a constrained and principled way.

It seems, then, that the poetic grammar theory experiences a number of empirical problems that indicate that it cannot be implemented effectively in the generative framework, at least in the manner suggested by Dillon and Austin. The problems are that the theory is too tightly constrained to allow for this kind of tinkering, and that the data is too diverse for a rule-based explanation to have a real chance of providing an adequate explanation. The only hope for such an approach is to handpick well-behaved data sets that can be explained in this manner for the poetic grammar explanation, and to put the remaining data to one side as some other issue. However this would be highly suspicious from a theoretical and methodological point of view, and ultimately it would seem to be missing an important generalization that can be drawn from the overview: that the data is not well-behaved, and in need of a more radical explanation than a simple alteration of the standard grammar.

3.1.2 Theoretical problems

We can see that many of the specific proposals made by the poetic grammar theory suffer from technical problems in the contemporary framework, and these problems are symptomatic of a wider mismatch between the poetic grammar theory and the more constrained theory of Minimalism. Technical problems can typically be solved by some kind of technical reworking, however: to fix the problems of Verb Final, we can replace it by a series of leftward movement rules for VP-internal arguments; to fix backwards gapping, we can posit that it is simple an instance of deletion, unrelated to gapping as it is derived in the standard grammar. Even some instances of non-constituent-affecting operations can be implemented with the contemporary technology, as this is what we see with pseudogapping, where the ellipsis of a non-constituent is produced by a specific ordering of movement and deletion.¹¹ This kind of approach might be plausi-

¹¹For Jayaseelan (1990) and others, this is rightward movement (HNPS) followed by deletion. For Lasnik (1999), Jayaseelan (2001), Gengel (2007) and Thoms (to appear) this is leftward movement followed by deletion, though the details of the landing sites, kinds of movement and licensors vary between the analyses. The majority of these are remnant movement analyses of a kind, although Lasnik (1999) and Thoms (to appear) derive the necessity

bly extended to some of the cases surveyed here, such as subject-plus-auxiliary deletion (discussed in section 2.2.2), perhaps by leftward movement of the VP followed by ellipsis, although it is not clear how it could deal with some others. However these technical issues are not the only problems for the poetic grammar theory within the contemporary framework, as there are deeper theoretical issues that are not solved by minor alterations to the implementation. The technical constraints of contemporary syntactic theory are the result of increasingly tight constraints on the form of the theory that can be developed, guided by general Minimalist principles. Here I will review some of the theoretical problems faced by the poetic grammar approach to poetic language, showing that such an approach is ultimately incompatible with the contemporary approach.

In Minimalism, the grammar is essentially the product of conditions on the interfaces between the Conceptual-Intentional system the Articulatory-Perceptual system, also known as LF and PF respectively. The syntax itself is virtually non-existent as an independent system; it is composed of the minimal set of conceptually-necessary operations for interfacing these two systems, such as the operation Merge, which combines lexical items, and Agree, which establishes relations between features of lexical items. All other apparently syntactic behaviour is to be explained in terms of the properties of the individual interfaces and their legibility conditions. Morphosyntactic features on lexical items drive the establishment of certain kinds of dependencies, and whether or not these dependencies translate to movement or other such manifestations depends upon the language-specific properties of the interface for that language;¹² for example, the dependency between the C head and *wh*-phrases is established by Agree, and whether or not that translates to overt *wh*-movement is dependent upon whether or not the PF interface of the given language requires pied-piping. Dif-

of deletion from the movement operation.

¹²Note that 'language-specific' here is not a backdoor way of asserting stipulative differences between languages, but rather an admission that different languages have different phonologies and sets of lexical items that lead to different requirements for convergence. These language-particular interface properties are typically trends relating certain aspects of phonology and syntax; for example, the correlation between rich Case agreement and the availability of scrambling, or rich verbal morphology and pro-drop.

ferences between languages and operations are thus explained in terms of variation in morphosyntactic features (variation in the properties of lexical items) and properties of the given interface and not the presence or absence of some set of core syntactic rules. To explain a given linguistic property as the product of the grammar of a given language, one must explain the presence of a given morphosyntactic feature or interface condition that produces this property; independent rules like ‘topicalization’ or ‘move α ’ have no status in this theory.

The rule-based approach has a natural appeal in the view of the poetic grammar, since intuitively to learn a rule is to learn a procedure, and what we often see in the poetic texts seems to be the application of a set of well-defined procedures. This characterization of the application of rules in generative grammar is somewhat misleading, though, since it seems to indicate that rules are consciously applied sets of procedures, contrary to the core assumptions of the generative approach. The change in the theory from the use of explicit rules to feature-checking and interface conditions changes things: to explain a given displacement operation as movement within the poetic grammar, we need to posit the availability of a given morphosyntactic feature on a moving element and at the landing site; to explain erasure as ellipsis, we would need to make the E-feature (the feature responsible for marking a constituent for ellipsis according to Merchant 2001 and Aelbrecht 2009¹³) widely available on a number of lexical items that do not normally bear it.

This reorientation presents two main problems to the poetic grammar theory. The first problem is that we cannot posit the poetic grammar to be the same as the ordinary language plus a few extra rules, but rather we must propose that the poetic grammar is the ordinary grammar plus an extra set of lexical items and morphosyntactic features and a different set of language-specific interface

¹³Though the existence of the E-feature has been challenged both empirically and theoretically in recent work: Thoms (to appear) proposes an alternative account, where ellipsis is derived as a reflex of copy deletion and ellipsis is licensed only by movement. To explain widespread deletion with this technology, one would simply propose more movement-driving features, although this may experience non-trivial technical questions too. See chapter 4 for an alternative perspective where movement-driving features are rendered irrelevant.

conditions. In modern terms, this means that the poetic grammar is another language, and as such it must be acquired by those that use it (both writers and readers). Reading poetic texts thus involves some kind of code-switching, and that this particular language is used only in the poetic situation. As mentioned above, this kind of approach was acknowledged explicitly by Thorne (1965), and, in the proposal that we consider the “grammar of the poet as a whole,” implicitly by Dillon (1975) and Austin (1984).

However, such an approach seems implausible. Given that different poets use different rules in their texts, we would have to propose that readers of literary texts know a number of different, sometimes non-intersecting grammars, and that they use a different one when they read each text. But how could the speakers acquire these languages? The issue of acquisition is hugely important to generative linguistics, as the theoretical dictum of explanatory adequacy is rooted in the question of how children acquire languages so quickly and effortlessly. To acquire a language, one must learn the relevant lexical items through experience as a child, and guided by Universal Grammar (UG), this will converge upon the acquisition of the language. Yet most readers will not experience these texts until they are in adulthood, long after the acquisition period. Clearly, what adults do and what they experience when they get to know the language of a given poet is substantially different from what a child does when he or she learns a language. There is no fix for this problem, as one can't just bolt on extra morphosyntactic features to lexical items that were acquired in adulthood, nor can one just alter the interface conditions to produce some new forms. Simply put, the theory of poetic grammar fails to satisfy explanatory adequacy, if the poetic grammar is taken as a theory of a given language, as it must be.

The second problem is the general form of the theory developed. As mentioned earlier, a number of the proposals required to make the poetic grammar work require a high degree of stipulation, in order to prevent over-generation. For example, we would have to stipulate that the deletion of verbs in back-

wards gapping is restricted only to those contexts where this exceptional verb deletion is possible, since backwards gapping is only attested in finite clauses in immediately adjacent coordinate structures. We would also have to stipulate that features are borne on a number of functional projections that do not normally bear them; for example, in order to explain the full range of data from leftward argument displacement (i.e. PP and NP displacement to pre-verbal positions below the subject), we would have to stipulate that pretty much all of the functional projections in the inflectional layer would bear some feature F that could Agree with the same feature on the lexical items. While it is plausible to propose that lexical items can be learned bearing new features, to propose these additional features on a number of arbitrary landing sites is much more problematic, since most other approaches to movement have sought to derive the landing sites of movement from independent principles;¹⁴ but it would be almost entirely impossible to explain these highly idiosyncratic and varied kinds of displacement in terms of independent principles, so such an explanation would not be available. Numerous other stipulations would be required to restrict the examples of non-constituent movement and erasure, since these examples are relatively limited in their variation and require a number of previously unattested movements to derive their word orders. This level of stipulation is highly undesirable within the Minimalist program, but it is a necessary characteristic of the poetic grammar approach, since the poetic grammar is effectively a set of additional rules that need to be stipulated for in a given environment. When we try to do this within the Minimalist framework, we immediately run up against these overgeneration and stipulation problems.

One potential solution to the overgeneration part of this problem may be to deny that stipulations are required to account for overgeneration, and that the unattested forms are indeed possible but simply not attested due to chance or size of the the sampled texts. In effect this exploits a core methodological

¹⁴See for example the work on Object Shift in Scandinavian that seeks to derive the landing sites in terms of phasehood and linearization or general prosodic conditions: M. Richards (2004), Fox and Pesetsky (2005), Erteschik-Shir (2005), Wallenberg (2009).

problem for the theory of poetic language, that we are limited in our theory construction to the data we find in texts and that we cannot test predictions by producing test sentences. However, this leaves unexplained the question of why the given forms are unattested: if they can all be generated by the same mechanisms, why should one form be common and another entirely absent? This is a guiding principle for other kinds of corpus-based work in generative linguistics (i.e. diachronic and historical linguistics), and as such it should be taken as a guiding principle for theory construction in this framework. If the variation were not rooted entirely in the grammar, this could be explained or at least ignored with some degree of methodological consistency, but in the poetic grammar approach it is fundamental that the variations we aim to explain are explained in terms of grammatical features. As such, unattested patterns are as important to theory construction as attested patterns, and hence the poetic grammar requires a set of stipulations that render its theoretical claims suspect from the view of guiding theoretical principles of parsimony.

3.1.3 Summary

Therefore it seems that the poetic grammar theory and contemporary Minimalism are fundamentally incompatible. The restrictive nature of Minimalist theory prevents us from stipulating additional rules or relaxations of constraints, and this is precisely what the transformational poetic grammar theories propose to do. These theories also encounter non-trivial problems regarding explanatory adequacy, in that there is no way to explain how these grammars are acquired; as such, they should not be considered grammars at all from the generative perspective.

Given the accumulation of theoretical and empirical problems for poetic grammar, it seems that we should abandon this approach and seek out an alternative that departs radically from this way of doing things. Before I go on to develop such a proposal, however, I will review two more generative approaches to poetic grammar that differ from this ‘classical’ transformational approach.

These approaches maintain that the deviant forms in poetic texts are generated by the grammar, but they explain the availability of these grammatical operations not in terms of addition of new rules, but in terms of the borrowing of rules from other languages (section 3.2) and in terms of reranking of constraints on outputs (section 3.3). I will show that these approaches experience similar empirical and theoretical problems, since they share the same fundamental basis as the classical transformational theory, and in doing this I will ultimately reject an approach to poetic language that aims to generate the surface strings from the syntax in a standard manner.

3.2 Exploiting UG: O’Neil (2001)

Here I discuss the account of O’Neil (2001), which is built upon an analysis of the Icelandic poetic tradition of *dróttkvætt*. O’Neil’s proposal can be boiled down to the following: poetic texts written in a language X can sometimes draw upon linguistic operations that are available in another language Y. In this view, poets use a form of language that goes beyond the source language in some way, but these extensions are effectively exploitations of UG, where the poet utilises a set of operations that are available to language in other circumstances. O’Neil’s article is the most extensive defense of this version of the theory of poetic language, but it is representative of a way of thinking about poetic language that has appeared in different forms since Kiparsky (1981). The most important thing about O’Neil’s proposal is that it is not without precedent in syntactic theory: the idea of ‘borrowed rules’ has been raised elsewhere in the literature, albeit in relation to different issues. Here I will outline the basic form of O’Neil’s argument, first describing his account of *dróttkvætt* before going on to suggest how such an account might be extended to explain some of the poetic language phenomena we have seen so far. I will discuss remaining empirical problems for such an account, and then discuss its theoretical motivation and how it ultimately cannot be maintained as a Minimalist theory of

poetic language.

Dróttkvætt (“court-metre”) is an Icelandic poetic tradition found between the 9th and 14th centuries. The form is subject to a number of compositional constraints, and for this reason it has been the subject of a great deal of discussion in the literature on poetic language. Below is a list of the constraints on literary form; the first four are obligatory, and the other three are optional (from O’Neil 2001: 340):

- (17)
- each stanza consists of eight lines made up of syntactically and semantically complete four-line half stanzas (*helmingar*);
 - each line consists of six syllables, with three main stresses, and ends on a trochee;
 - pairs of lines are bound together by alliteration on three stressed syllables (two in odd-numbered lines, one in even-numbered lines);
 - each line contains internal rhymes on two stressed syllables (near rhyme in odd-numbered lines, perfect rhyme in even-numbered lines);
 - the poetry often contains elaborate figures of speech (*kennings*) of up to seven terms, based in Scandinavian and Germanic mythology and elite social convention;
 - word order may be quite free compared with that of everyday language;
 - the several clauses of a stanza may be intercalated, or interlaced.

What is of particular interest here is the free word order component of the texts, which leads to a significant degree of linguistic deviation. O’Neil spends most of his article describing the ways in which the language of the texts deviate from the standard Icelandic of the time, so I will present some of his analyses before reviewing his theoretical conclusions.

O’Neil observes that the word order in *dróttkvætt* can be radically different from that of the Old Icelandic of the time, and that the reordering seen in the poetry can affect full constituents and non-constituents alike. The following example, from *Egils saga 51* (Nordal 1933: 269), demonstrates the extent of the dispersal of clausal constituents. This example also demonstrates the interlacing of different sentences; O’Neil tracks this by marking the parts of the first sentence $A_1, A_2 \dots$, the second sentence $B_1, B_2 \dots$, and so on, and I preserve

that system here:

(18) Þverra nú, þeirs þverrðu,
(A₁) Grow-fewer now, (B₁) those-who decreased

þingbirtingar Ingva,
(A₂) assembly-luminaries Ingvi's

hvar skalk manna mildra,
(C₁) where shall-I men munificent,

mjaðveitar dag, leita ...?
(B₂) mead-carving's day, (C₂) look-for ...?

'(A) Now chieftains [=Ingvi's assembly-luminaries] grow fewer in number, (B) who squandered gold [=mead-carving's day]. (C) Where am I to find munificent men...?

Nordal (1933) provides a reconstructed version of the original word order of this verse:

(19) (A) Þverra nú Ingva þingbirtingar, (B) þeirs
(A) Grow-fewer now Ingvi's assembly-luminaries, (B) those-who
þverrðu mjaðveitar dag; (C) hvar skalk leita mildra
decreased mead-carving's day; (C) where shall-I look-for munificent
manna ...?
men ...?

We can see, then, that the word order of the poetic text differs significantly from that of the standard version. Describing the rearrangement of the text,¹⁵ O'Neil says that the author has

- suspended the main clause after the normal tensed-V, adverbial opening;
- suspended the relative clause after a normal relative subject, tensed-V opening;
- completed the main clause by providing its subject;
- suspended the next main clause (a question) after giving the question word, the modal with a cliticized subject, and the object;
- finally completed the question by providing its main verb – the infinitive *leita* – clause-finally.

(O'Neil 2001: 346)

¹⁵O'Neil does not comment on the inversion of the possessor and possessed NP in *þingbirtingar Ingva*, so I assume that this kind of movement was tolerated in standard Icelandic as well.

O’Neil notes that “basic constituents are not themselves broken up and distributed through the clause” (O’Neil 2001: 346), but he does not give details on how would derive the broken-up linear order of the relative clause, which is identified as clause B and which modifies the NP subject of clause A. The relative clause is distributed across two separate positions: one part, comprised of the relative pronoun subject and verb (*þeirs þverrðou*), appears to the left of the NP that it modifies (*Ingva þingbirtingar*), rather than in its standard position to the right; the other part, comprised of the object of the verb within the relative clause (*mjaðveitar dag*), which occurs to the right of the modified noun, also following a part of the interrupting clause labelled C. It would seem that the only way to derive this order (paying attention only to A and B, which are syntactically related) by syntactic movement would be to move the relative pronoun and the verb of the relative clause to some position above the head noun. This would in fact involve non-constituent movement, since these two elements do not form a constituent to the exclusion of the object; the only way to do this by constituent movement would be for (a) the object of the relative clause to undergo rightward movement to some position in the matrix clause, and then (b) for the relative clause to undergo leftward movement to some position to the left of the noun it modifies. To generate this, we would have to tolerate the violation of a relative clause island by rightward movement in (a), and then posit some sort of relative clause movement to a position below the verb and adverbial (which are presumably in the complementizer field). We would thus require a significant degree of contrivance and a number of strange movements to generate this structure, but it is at least conceivable if we invoke a remnant movement analysis.

However, the contrivances required to give (18) a movement analysis seem wasted when we are then presented with examples of what O’Neil takes to be clear non-constituent movement. Below is a three-word clause from *Grettis Saga 4*, provided by O’Neil as an example of non-constituent movement (the dots represent a break caused by another, unrelated interrupting clause):

- (20) Skáldi sígr . . . þvísa
 Poet slides . . . this
 ‘This poet is going downhill’

O’Neil analyses this as movement of the head noun *skáldi* to Spec,CP, abandoning the determiner *þvísa*,¹⁶ and movement of the verb to C⁰.¹⁷ We might plausibly analyse this as an instance of phrasal NP-movement if we accept the DP-hypothesis (Abney 1987), as is standard in most modern Minimalism,¹⁸ and thus analyse this as violation of a DP-island rather than non-constituent movement. However, O’Neil asserts that non-constituent movement is a regular occurrence in the *dróttkvætt* poetry, providing more radical examples to make this point clear. The following is an example where three parts of an NP (numeral, genitive determiner, head noun) have been separated:

- (21) Tolk höfum gröf hjá gjalfri
 twelve have-we grave by sea

 gunnelds búit runnum.
 war-fire’s prepared trees.

‘I have prepared a grave by the sea for twelve warriors [= war-fire’s trees]’

The NP *tolk gunnelds runnum* has been dispersed into three separate positions in the sentence. O’Neil argues that this string is derived by the following steps: (a) movement of the tensed verb to C⁰; (b) movement of the numeral to Spec,CP; (c) rightward movement of the head noun out of the NP to a right-adjoined position, stranding the genitive possessor in the base position. There seems

¹⁶A potential confound for this analysis would be if the determiner could also be used as a proform, in which case it may not be syntactically related to the other NP.

¹⁷Old Icelandic, like its modern counterpart, was a V2 language, thus requiring movement of the verb to second position. Whether or not this involves movement to the C-field (as O’Neil assumes) is perhaps not a settled matter, since Icelandic is typically identified as an ‘IP-V2’ language, in contrast to ‘CP-V2’ languages like German; this distinction comes from the fact that Icelandic, but not German, exhibits V2 in embedded clauses with an overt complementizer. The specific landing site for movement is not a particularly important matter here, however, assuming the subject moves from a vP-internal position.

¹⁸Though see Bruening 2009 for some fresh challenges to the DP-hypothesis. Throughout this dissertation I have ignored the NP/DP distinction since it does not impact upon the discussion directly for the most part, and ‘NP’ is a simpler and more familiar notion for the non-expert; I will remain agnostic on the issue for now, and will continue to refer to ‘NPs’ rather than ‘DPs’ with no real commitment to the theoretical loadedness of the terminological choice.

to be no prospect for analysing this example in terms of remnant movement or any other set of plausible movements. O’Neil observes that there are also numerous other cases of similarly unusual non-constituent movement, and even examples that split up compound nouns; he asserts that these instances of non-constituent movement are only ever to positions at the edge of the sentence (Spec,CP or right-adjoined), although he observes that there are some apparent counter-examples in the corpus.¹⁹

We can see, then, that the *dróttkvætt* texts display a significant amount of movements that are not found in the contemporary Old Icelandic, including non-constituent movement. How should these movements be analysed? O’Neil’s proposal is that what we see in Skaldic poetry is effectively Japanese scrambling. Discussing (21), he observes that “Universal Grammar allows for similar sorts of movement in languages that exhibit scrambling. . . Skaldic poetry thus allows in part what is possible in these languages though impossible in everyday Icelandic” (O’Neil 2001: 349). In this view, what the composers of *dróttkvætt* do is effectively exploit the potential of UG, tapping into its potential to ‘borrow’ a set of operations that are typical to other grammars like Japanese. This means that the poetic forms are generated by the grammar, and that the ‘poetic grammar’ conceived of here would be the standard Icelandic grammar plus other part of the Universal Grammar that typically remain untapped in the use of Icelandic.

O’Neil argues that borrowing Japanese rules allows us to generate the forms in *dróttkvætt*, so we may wonder if a similar rule borrowing would be able to explain what we have seen in the English language poetry so far. We might argue that these texts also involve borrowing of Japanese scrambling, or perhaps German scrambling. Another potential rule-borrowing could be borrowing Icelandic Object Shift, as this involves movement of objects to Spec,vP; this would cover a number of the cases of objects being displaced to a pre-verbal position. To account for some of the cases of unlicensed argument ellipsis, we

¹⁹Indeed this proposal would be in doubt if we were to analyse (18) as involving movement of a non-constituent, however, as the movement in that example was to a position between the subject NP and an adverbial.

may argue that this involves borrowing the mechanisms made available in other languages too; for missing subjects, we could be borrowing the mechanisms used in Italian and Spanish; for missing objects, we could be borrowing the mechanisms used Korean and Hebrew for argument drop; for missing determiners, we could be switching to a Chinese nominal system; and so on. Universal Grammar makes available a massive number of different operations that may plausibly generate a wide array of strings that differ from those seen in standard English, so allowing a poet to exploit the full potential of UG (as conceived here) would seem to provide a way for even some of the most unusual forms found in poetic texts to be generated by the grammar.

In what follows I will show that this promise is illusive. The theoretical basis for the ‘exploiting UG’ approach is in fact highly problematic and founded on questionable assumptions, and even if this is allowed to pass, UG still cannot generate some of the structures found in poetic text, at least not without admitting significant unwelcome alterations to its internal form.

3.2.1 Empirical problems

The fundamental problem with O’Neil’s account of *dróttkvætt* is that Japanese scrambling could not in fact generate the structures found in the texts. The consensus in the literature is that there are two types of scrambling in Japanese: clause-internal ‘short’ scrambling, and ‘long’ scrambling, which can move an argument across a clause boundary. (22) demonstrates two different permutations of clause-internal scrambling, where the object can scramble over the dative goal argument and over the subject, and (23) demonstrates scrambling of an object out of an embedded clause to the left periphery of the matrix clause:

- (22) a. John-ga Mary-ni piza-o ageta.
 John-NOM Mary-DAT pizza-ACC gave
 ‘John gave Mary pizza’
 b. John-ga piza-o_i Mary-ni t_i ageta.
 John-NOM pizza-ACC Mary-DAT gave
 c. Piza-o_i John-ga Mary-ni t_i ageta.
 pizza-ACC John-NOM Mary-DAT gave

(Miyagawa 1997: 1)

- (23) Sono hon- o_i Hanako-ga Taroo-ga t_i katta to omotteiru
 that book-ACC Hanako-NOM Taroo-NOM bought COMP think
 (koto)
 fact
 ‘Hanako thinks that Taroo bought that book’
 (Saito 1992: 69)

There is a great deal of debate about the status of these different kinds of movement, since they differ in certain respects, but for our purposes here I will put these aside and treat them as part of the same large category of scrambling, which is taken to be a kind of phrasal movement akin to standard A-bar movement.²⁰ We can see, then, that scrambling allows one to reorder the word order of a sentence quite freely. It does not lead to entirely free word order, however. Scrambling is a kind of phrasal movement, and it always involves leftward movement of full phrases; it does not affect non-constituents,²¹ like all other movement rules in syntax, and scrambling never proceeds rightward (Saito 1985, Tanaka 2002).

Referring to the reordered text in (21), O’Neil (2001: 346) notes that the non-constituent movement in that example resembles the kind of movement we see in Japanese scrambling. In terms of the effect on surface structure, this is correct, since Japanese allows for scrambling of a numeral modifier to the sentence-initial position:

- (24) San-satu Hanako-ga hon-o katta (koto)
 3-CL Hanako-NOM book-ACC bought (fact)
 ‘Hanako bought three books.’
 (Kawashima 1998: 1)

²⁰There is dispute over whether the short-distance scrambling seen in (22) is actually A-movement; see Miyagawa (2001, 2003, 2005) and Saito (2006) for discussion.

²¹Some have contended that something like non-constituent scrambling is possible in situations where a pair of VP-internal arguments that form a prosodic phrase are fronted as one; Agbayani et al (2009) argue for a form of PF movement that derives this effect, in order to avoid positing non-constituent movement in the narrow syntax. This is broadly similar to the ‘PF movement’ account of scrambling given by Sauerland and Elbourne (2002), and as such it is equally open to the empirical challenges presented by Miyagawa (2005). It seems to me that numerous alternatives can be proposed that avoid non-constituent movement; see Kempson and Kiaer (2010), Takano (2002) and Koizumi (2000) for alternatives.

Here the numeral classifier *san-satu* has been scrambled from its position modifying the NP object *hon-o* to a left-periphery position. Based on the range of different kinds of quantifiers that can scramble, Kawashima (1998) argues that these structures are derived by a kind of remnant movement (similar to the proposal for (18) above), where the NP object is first scrambled out of the containing numeral phrase to a VP-adjunction position, and then the vacated phrase is itself scrambled to the sentence-initial position.

We may expect, then, that a similar set of movements would allow us to move the numeral to the sentence-initial position in (21). I schematize a derivation that would separate the numeral *tolk* and the NP *gunnelds runnum* in (25):

$$(25) \quad [_{CP} [_{NumP} \text{Tolk } t_i]_j \dots [_{VP} [_{NP} \text{gunnelds runnum}]_i [_{VP} t_j \text{V}]]]$$

This would be possible if the full NP was represented in the base position, to the left of the verb, but we can recall that (21) is not that simple, as the nominal argument has three parts and all three have been separated: only the genitive possessor remains in the base position (or thereabouts), and the head NP has been moved to the right edge of the sentence. (26) presents a conceivable derivation: (a) the NP containing both the genitive possessor and the head noun scrambles to the VP-edge; (b) the vacated NumP scrambles to Spec,CP; (c) the possessed NP undergoes rightward scrambling to the right-edge, stranding the genitive possessor NP in the VP-adjunction position:

$$(26) \quad \begin{array}{l} \text{a. } [_{CP} \dots [_{VP} [_{NP} \text{gunnelds } [_{NP} \text{runnum}]_k]_i [_{VP} [_{NumP} \text{Tolk } t_i]_j \text{V}]] \\ \text{b. } [_{CP} [_{NumP} \text{Tolk } t_i]_j \dots [_{VP} [_{NP} \text{gunnelds } [_{NP} \text{runnum}]_k]_i [_{VP} t_j \text{V}]] \\ \text{c. } [[_{CP} [_{NumP} \text{Tolk } t_i]_j \dots [_{VP} [_{NP} \text{gunnelds } t_k]_i [_{VP} t_j \text{V}]]] [_{NP} \\ \text{runnum}]_k] \end{array}$$

This derivation is not possible with Japanese scrambling, however, as the step in (26c) suffers from two problems. First, we have already noted that scrambling only proceeds leftward, and that rightward scrambling is not possible in Japanese: the verb always appears in a sentence-final position. Second, scrambling out of a possessor-possessed NP is not possible in Japanese, as the following pairs demonstrate. (27b) shows that scrambling of a possessed NP is not possible,

and (28b) shows that scrambling of the genitive NP is not possible either. (29) demonstrates that the full constituent can still in principle be scrambled.

- (27) a. Tomoko-ga kosi-o itam-e-ta
Tomoko-NOM back-ACC hurt-tr-Past.
'Tomoko hurt (her) back.'
- b. *Kosi-o Tomoko-ga itam-e-ta
back-ACC Tomoko-NOM hurt-tr-Past.
(Hasegawa 2001: 22)
- (28) a. Mary-ga [_{NP} Amerika-e-no ryokoo-keikaku-o] tate-ta.
Mary-NOM America-to-GEN travel-plan-ACC
'Mary made travel plans to the US.'
- b. *Amerika-e-no_i Mary-ga [_{NP} t_i ryokoo-keikaku-o] tate-ta.
America-to-GEN Mary-NOM travel-plan-ACC
(Kishimoto 2006: 790)
- (29) a. Kyoko_i-ga zibun_i-no neko-o sagasi-te-i-ru.
Kyoko-NOM self-GEN cat-ACC look-for-Prog-Pres
'Kyoko is looking for her cat.'
- b. [zibun_i-no neko-o]_j Kyoko_i-ga t_j sagasi-te-i-ru.
self-GEN cat-ACC Kyoko-NOM look-for-Prog-Pres

Given these problems, it seems that there is no way to derive the structure seen in (21) from the mechanisms made available by Japanese scrambling.

Similar problems befall many of the structures seen in the *dróttkvætt* poems, since they would require kinds of non-constituent movement or rightward movement that are not made available by Japanese scrambling. Therefore a poetic grammar consisting of the standard grammar of Old Icelandic (as it is conceived by Nordal 1933) plus the extra capabilities of Japanese scrambling would undergenerate with respect to the data that is supposed to be covered. There is an additional dimension to this mismatch between Japanese scrambling and *dróttkvætt* that is worth mentioning. O'Neil (2001: 348) notes that the dispersal of "proper parts of a constituent" is always to the left or right edge of a clause, and no further. However it is well-known that Japanese scrambling is not restricted to the boundaries of a clause edge, but rather it can move arguments out of an embedded clause and to the edge of the matrix clause; this was demonstrated by (23) above. If the reordering in *dróttkvætt* is a use of Japanese scrambling, then

why is it restricted to clause-internal scrambling? One might propose that this is because *dróttkvætt* only involves borrowing one of the two kinds of scrambling, namely the clause-internal type that is analyzed as A-scrambling in the recent literature. But this would then reduce significantly the power of the mechanisms, as this kind of scrambling is restricted in certain ways and not capable of deriving some of the structures discussed so far; for example, the scrambling required for moving the numeral to a sentence-initial position must be a type of A-bar scrambling (Kawashima 1998), and hence could not be covered by the A-scrambling option.

We can see, then, that there are significant mismatches between Japanese scrambling and the unusual displacements we see in *dróttkvætt* syntax, and that borrowing of scrambling would not derive the required structures. This is symptomatic of the fact that the rearrangement we see in *dróttkvætt* is nothing like standard natural language movement operations, but rather a more arbitrary and unprincipled kind of procedure. Given the discussion of the previous section, we can see that the same would hold if we were to try to derive the English language cases from scrambling too. The poetic language phenomena are simply too unusual and diverse to be conceived of as the output of syntactic operations, at least as they are conceived of in a restrictive theory of syntax like those developed under the umbrella of Minimalism. It would be a similarly fruitless task to adopt different sets of rules from other grammars; for example, we can conceive of adopting Germanic Object Shift to generate the examples of movement to a pre-verbal position, this would fail to account for examples where the shifted object crossed the auxiliary, since Object Shift is typically to a lower position like Spec,vP. The UG-exploiting approach will always struggle to generate the unusual structures found in the poetic texts since the theory of UG is necessarily a restrictive theory.

3.2.2 Theoretical problems

Here I will show that the ‘UG exploitation’ approach to poetic grammar also suffers from theoretical problems, and that these problems have a similar flavour to those suffered by the transformational poetic grammar described in the previous section. I will first outline the possible reasons for believing that this approach is in fitting with the general mission of Minimalism, before discussing the problems inherent in such an approach.

The UG exploitation theory has a number of different appeals. The theory presumes that the forms produced by poetry involve the poet (and reader) tapping into some ‘underlying’ or ‘potential’ part of our linguistic capability in order to produce and interpret poetic texts. As a description of poetic practice, this is intuitively appealing, since it relates well to folk notions of poetic composition as an especially introspective mode of operation; see for example the quote from Samuel Beckett in the introduction, where he describes a similar ambition. It also has its theoretical appeals, since it allows us to account for the strange data of poetic texts as linguistic data, but not as data of the same kind as ordinary language data. Furthermore, to the syntactician this kind of proposal sounds familiar, since it is often proposed that languages ‘borrow’ rules and operations from different languages; for example, Johnson and Tomioka (1998) argue that English Quantifier Raising “is Dutch scrambling,” and Lasnik (1999) argues that the operation that moves the remnant of pseudogapping in English is the same as Scandinavian Object Shift.²² Given this, it seems reasonable to assume that we can exploit UG in poetic language just like we do in standard English.

As mentioned above, the core motivation for the theory of Universal Grammar is the intention to account for the phenomenon of language acquisition: the poverty of the stimulus indicates that children must be guided as they acquire their language from the data they experience, and this guide is UG. Modeling the content and form of UG is the main mission of theoretical linguistics, and the

²²Though we will see in section 4.3.4 that a more accurate description for the movement in pseudogapping might be better described as “Malayam scrambling,” in line with Jayaseelan (2001) and Thoms (to appear).

current consensus within Minimalism is that UG effectively consists of a set of parameters, an inventory of morphosyntactic features, and a few core syntactic components. This modern view of UG is different from that which was widely held in the Government and Binding era of generative linguistics, where UG was a rich and highly developed system that contained syntactic primitives like the operation $\text{move } \alpha$, the ‘Case filter,’ theta theory, the Binding Theory, and so on (see Chomsky 1981). The move to Minimalism has encouraged linguists to explain away this rich inventory of unlearned mechanisms as natural products of the design of the system of language as an interface between the Conceptual-Intentional and Articulatory-Perceptual systems; in this view, the interfaces can only allow for the growth of grammars with certain properties, and as such the interfaces serve the same purpose as an independent UG system which contains a predetermined set of operations and filters.

As an endowment, then, UG is simply a guide for acquisition, and not an independent endowment of rules or operations. Whether or not a language makes use of the operation known as ‘scrambling’ depends on the setting of a few parameters (see Baker 2001 for an articulated system of parameters), and the setting of these parameters will be determined by the input for acquisition. ‘Scrambling’ as an operation does not exist in UG. As a result, it makes no sense to talk of making use of scrambling by accessing the potential of UG, as is done in the approach advocated by O’Neil: the only potential possessed by UG is a potential to guide acquisition to forming one grammar or another, nothing else. We can see, then, that the same acquisition problem that plagues the transformational poetic grammar theory plagues this version too: it is not plausible to propose that language users acquire the poetic grammars, so it is implausible that the poetic grammar is a grammar in the sense intended in Minimalism. Note that this criticism does not apply to the cases from syntactic work where it is proposed that our language ‘borrows’ some rule from another, since in this case the borrowed rule would be acquired just like any other rule

in our language.²³

I will end this subsection by mentioning briefly another theoretical issue that is more difficult to pin down but which is nevertheless worth mentioning in this context. One of the issues that arises when we say that *dróttkvaett* involves Japanese scrambling is that of the arbitrariness of the choice of Japanese. Why would *dróttkvaett* involve Japanese scrambling, and why does it not exploit any of the numerous other operations available? The same question can be asked in the context of the transformational poetic grammar approach discussed in the previous section: why invent these rules and not others? Why do certain kinds of rules not occur? These are questions that neither of these approaches can answer, since there are no proposed meta-rules for the selection of possible rules or borrowings; rather, Japanese scrambling is chosen simply because it seems to get the right results in some cases. This lack of an explanation does not sit well with the theory, and it leaves a number of questions unanswered and unanswerable. The theory of poetic language should aim to explain why some forms occur but others do not, on the basis of some sort of principle or independent explanation.

3.2.3 Summary

We can see, then, that this alternative way of implementing poetic grammar within a mainstream generative grammar framework is plagued by both empirical and theoretical problems, and that these problems have a similar root to those suffered by the transformational poetic grammar approach. It seems that a generative approach is generally incompatible with poetic grammar, in terms

²³The Johnson-style way of speaking about these cases is merely a descriptive statement about the similarity of pairs of operations across languages. Obviously English QR is not the same operation as Dutch scrambling, as the former is covert and the latter is overt, but since they both target similar constituents and move them in a similar manner (i.e. to a specific landing site, below negation), we can say that they are effectively the same thing. If we were to stake the theoretical claim that they were the same thing, then we would have to say that the differences between the two operations would be derived from some other general principle or parameter setting (for a proposal along these lines in a closely related field, see Bobaljik and Wurmbrand 2008). But in this case the relevant rule would cease to be ‘Dutch scrambling’ and would instead become ‘Germanic scrambling,’ and its presence in English wouldn’t as much be a borrowing but rather a core property of the grammar as defined by its parameter settings.

of both its technical potential and its theoretical content. These two theoretical approaches do not exhaust the possibilities within the broad church of generative linguistics however, as there are different ways to model UG and the grammar and thus different ways to possibly account for poetic language. In the next section I will review one such alternative, and I will show that, although it makes different kinds of predictions from the Minimalist and EST-style approaches, it is equally incapable of accounting for the full range of data with a well-defined and predictive theory.

3.3 An Optimality Theoretic approach to poetic grammar: Fitzgerald (2007)

The previous two sections have shown that poetic grammar and the mainstream approaches to generative syntax are incompatible. However, within the broad church of generative linguistics there are a number of different theories of the architecture of UG, and these different architectures propose different forms for the grammar which in turn propose different approaches to structure-building. One such architecture is Optimality Theory (Prince and Smolensky 1993), a generative theory that proposes that the grammar is composed of a set of ranked constraints, and that the forms that are generated are those that are the optimal candidate for the given set of constraints; the constraints are part of UG, a universal endowment, and the ranking is developed during acquisition, where a different language consists of a different ranking. The fundamental innovation of this account is that the definition of (un)grammaticality is radically different from that of the mainstream approach (which I will take to be Minimalism): whereas in Minimalism grammaticality is defined by the simultaneous satisfaction of all UG constraints, in OT a grammatical sentence can violate a number of constraints, so long as there is no alternative that violates fewer of them.²⁴

²⁴Samek-Lodovici (2006) notes that OT and Minimalism are not in fact incompatible, since Minimalism is typically taken to be a research program guided by a core set of fundamental principles, rather than a single well-defined theory; certain conceptions of UG are compatible

This shift to violable ('soft') constraints changes the status of ungrammatical data in a way that is of some importance in the context of this dissertation, since what we are concerned with explaining is the occurrence of certain forms that are ungrammatical in the standard language.

Most of the work on OT has been in phonology, but in recent years the literature on OT syntax and semantics has increased significantly (see Grimshaw 1997, Legendre et al 1998, Grimshaw and Samek-Lodovici 1998, Ackema and Neeleman 1998, Dekkers et al 2000, Legendre 2001). Parallel to this, a number of linguists have developed approaches to poetic meter in the OT framework (Hanson and Kiparsky 1996, Kiparsky 2006, Hayes and MacEachern 1998, Golston 1998, Golston and Riad 2005); this was a natural development, since early work in generative metrics first introduced constraints as a tool in generative linguistics (Halle and Keyser 1971). OT approaches to meter assume that there is a metrical template and that metrical constraints match the phonological form of the text to the template; for example, the constraint `MATCHSTRESS` induces a violation if the stress in a polysyllabic word does not match a stressed position in the metrical template.

Building upon this OT work in syntax and metrics, and the work of Youmans (1982, 1983, 1996), Fitzgerald (2007) proposes an account of poetic inversion, where word order is changed in certain ways to produce metrically well-formed lines. Fitzgerald proposes that the poetic grammar involves the ranking of metrical constraints over some syntactic constraints; interestingly, she proposes that some syntactic constraints still outrank the metrical constraints, thus explaining the fact that some inversions occur for metrical purposes, but not others. Specifically, she adopts the syntactic proposals of Grimshaw (1997), who proposes the

with the core tenets of Minimalism, and Samek-Lodovici argues that Minimalism would be better off to develop in the OT direction. The reader should bear in mind, then, that when I refer to Minimalism in this section and henceforth, I take this to denote a standard theory of sorts that is representative of the work done in this research program; that is, a theory that favours a 'strongly derivational' model of the grammar, one composed of a set of inviolable UG constraints and lacking the soft constraints of the OT type. A significant amount of work in Minimalism implicitly or explicitly rejects the soft constraints and global calculations of economy promoted in OT (for example the 'crucially derivational' work of N. Richards 2001 and Landau 1999 or the local economy calculations in Fox 2000 and Thoms 2010c), and indeed Chomsky (1995b) rejects such an approach explicitly.

following constraints on the alignment of specifiers, heads and complements in phrases:

- (30) a. HEADLEFT (HDLFT): Every X^0 is at the left edge of an X^{max} .
 b. SPECLEFT: Every specifier is at the left edge of an X^{max} .

To these, Fitzgerald (2007: 210) adds the following constraint to account for the structure of NP modification in English:

- (31) ADJ-N: Adjectives precede the nouns that they modify.

Thus the ranking for a language like English that tends to display Spec-head-comp alignment, the ranking of the constraints in (30) is SPECLEFT >> HDLFT, as is demonstrated by the tableau below:

(32)

/input/	SPECLEFT	HDLFT
a. Spec-H-Comp		*
b. Spec-Comp-H		**!
c. H-Comp-Spec	*!*	
d. Comp-H-Spec	*!*	*

Satisfaction of SPECLEFT necessarily involves violation of HDLFT, but since SPECLEFT outranks HDLFT, this is the preferred option and the Spec-H-Comp order is produced.

Fitzgerald proposes that the metrical constraint MATCHSTRESS “interweaves” with these syntactic constraints, outranking some but not others, and that this explains the occurrence of poetic inversions in a set of examples from Shakespeare’s verse (from *Hamlet* and the *Sonnets*). First, she proposes that MATCHSTRESS outranks ADJ-N, on the basis of examples like (33) where a noun and its adjectival modifier invert to provide a metrically well-formed line; the (a) example provides the attested metrical line with inversion, and the (b) example the unmetrical line without inversion:²⁵

²⁵A brief note on the conventions for scanning metrical lines (which are preserved from Fitzgerald’s presentation): the top line is the representation of stressed and unstressed syllables in the line, with ‘X’ indicating high stress and ‘.’ low stress; the third line represents the metrical template (here iambic pentameter), with ‘W’ representing weak positions and ‘S’ strong positions.

- (33) a. . X . X X
 When rócks imprégnable are not so stóut
 W S W S W S W S W S
 (Sonnet 75, line 7)
- b. . . X . . X X
 When imprégnable rócks are not so stóut
 W S W S W S W S W S

In (33a) the noun *rocks* and the modifying adjective *impregnable* undergo inversion to avoid the main stress in the polysyllabic adjective falling into the weak third position in the metrical template; thus the deviant syntactic form is tolerated in favour of preserving the line's iambic pentameter metre. A simplified version of the tableau is given below:

(34)

/input/	MATCHSTRESS	ADJ-N
a. rocks impregnable		*
b. impregnable rocks	*!	

This account of adjective-noun inversion in poetic lines avoids committing to some inversion rule that is instantiated in every noun-adjective pair; for example, in an NP like *red car* neither of the words contains a polysyllable that is controlled by MATCHSTRESS, so a candidate with the uninverted form would lack a violation of ADJ-N and would thus prevail over the inverted form *car red*, which is predicted to be unattested.

Fitzgerald proposes that MATCHSTRESS also outranks HDLFT, and she provides examples where the verb and its complement are inverted to turn an unmetrical form into a metrical one; these examples are similar to many we have seen already, so I will not replicate them here. What is important is that Fitzgerald proposes that constraints like SPECLEFT are not outranked by metrical constraints, as she says that this allows us to account for the fact that many inversions that involve Spec and Head do not occur. She gives one example modeled on (33), which I present below:

- (35) . X . . X X
 Imprégnable rócks are when not so stóut
 W S W S W S W S W S
 (Fitzgerald 2007: 213)

Fitzgerald indicates that the inversion of *when* and the string *impregnable rocks are* would alleviate the MATCHSTRESS violation, so the fact that lines like this do not occur is evidence for the proposed ranking of SPECLEFT above MATCHSTRESS. I am not entirely sure how this structure would be derived, since the assumed syntactic framework dictates that the structures generated are of basic X-bar-theoretic form (this is a stated assumption in Grimshaw 1997); given this, it is unclear how the specifier could be linearized to appear in a position between the elements in its complement. Nevertheless the point would remain if a more plausible structure were proposed, such as (36), where *when* occurs in a right-specifier, violating SPECLEFT while ameliorating the MATCHSTRESS violation:

- (36) . X . . X . X . X X
 Imprégnable rócks are not so stóut when
 W S WS W S W S W S

Examples like this are equally unusual and are not attested, to my knowledge. I will return to the significance of such examples below.

In this view, then, the poetic grammar is a grammar in which the metrical constraints interweave with syntactic constraints in the ranking. It is implicit in Fitzgerald's presentation of this theory that the only way in which the poetic grammar differs from the standard grammar is in the interweaving of the metrical constraints with the syntactic ones; presumably the relative ranking of the syntactic constraints remains constant, and the metrical constraints are part of the universal endowment of UG but just not ranked above any relevant syntactic constraints in the standard grammar, in English at least. Although the empirical coverage within her article is only presentational, Fitzgerald's proposed system makes a strong set of predictions and claims to account for the occurrence of some forms and non-occurrence of others.

3.3.1 Empirical problems

Although it is developed in a different framework from the previous two theories of poetic grammar, the OT theory still shares fundamental assumptions about the status of data in poetic texts: most significantly, the assumption that the deviant forms of poetic texts are generated by ordinary means in the grammar. Therefore the OT approach still suffers from many of the same empirical and theoretical problems suffered by the other generative approaches to poetic grammar.

The first issue is identified by Fitzgerald (2007: 209), concerning subject-auxiliary/subject-verb inversion. We saw many examples of this so far, and below is an example where inversion alleviated a violation of the Monosyllabic Word Constraint, another metrical constraint proposed by Fitzgerald (modeled after Kiparsky 1977):

- (37) a. . X . X . . X X . X
 But die thy thóughts when thy first lord is déad.
 W S W S W S W S W S

(*Hamlet* 3.2.214)

- b. . . X X . . X X . X
 But thy thóughts die when thy first lord is déad.
 W S W S W S W S W S

In (37a) the subject *thy thoughts* and the verb *die* have been inverted to avoid the stress on *thoughts* falling on the unstressed position, in accordance with the Monosyllabic Word Constraint (see Fitzgerald 2007: 206-210 for discussion). Since subjects are typically taken to occur in Spec,TP, and the verb in this construction would be in T' (given that Shakespeare's English still retained verb movement to T), this would constitute a violation of SPECLEFT; in Fitzgerald's system we would need to rank the metrical constraint that controls for the Monosyllabic Word constraint above SPECLEFT, contra Fitzgerald's own proposals.

Fitzgerald acknowledges this apparent problem, but she dismisses it without

much discussion and plays down its significance. First, she comments that subject-verb verb examples “occur infrequently” (2006: 209), implying that the empirical weight of this challenge may not be significant if the frequency of data tokens is taken into account (more on this later). Then, she observes that

in Grimshaw’s [1997] model of OT syntax, question-driven [inversions] of the subject and verb do not result in violations of SPECLEFT. The movement of the head may result in a violation of HEADLEFT, depending on the particular configuration. (Fitzgerald 2007: 214)

With this observation in place, she then proceeds to example cases like (37a) as violations of HDLFT rather than SPECLEFT. Yet in her conclusion, she observes that “subject-verb inversion is the only type of specifier-head inversion” and comments that “this is unsurprising given that there has been considerable research into subject movement in syntactic theory” (2006: 215). In the end it is unclear whether Fitzgerald has committed to one analysis or another of subject-verb inversion; although it seems that the desire is to eliminate this possibility, it remains as a potential problem with many details not worked out.

It seems that Fitzgerald’s analysis of subject-inversion as a violation of HDLFT follows from a misreading of Grimshaw’s account.²⁶ For concreteness, let us consider the simplified pair below:²⁷

- (38) a. $[_{VP} [\text{thy thoughts}]_{DP} \text{die}_V]$
 b. $*[_{CP} \text{die}_V [_{VP} [\text{thy thoughts}]_{DP} t_V]]$

²⁶The only place in Grimshaw’s paper where SPECLEFT features is p.407, where she discusses the difference between types of embedded interrogatives (with and without C heads), as in the following contrast:

- i. I wonder when I will see such a sight again.
 ii. *I wonder when that I will see such a sight again.

Grimshaw accounts for this contrast in terms of an extra violation of HDLFT in (ii), due to the inclusion of the complementizer head with the *wh*-specifier. However, although it is represented in the relevant tableaux, as outranking HDLFT, SPECLEFT is not at all implicated in the contrast (it is not violated in either input), nor is it represented in the overall ranking proposed at the beginning of the paper (p.375).

²⁷An alternative way of presenting (38b) would be to simply represent it as a (base-generated) Hd-Spec-Comp order, with no movement. This would constitute an outright violation of SPECLEFT. Fitzgerald assumes that this does not occur, but my point here is that, however we derive this word order, we will end up with a violation of this constraint if we get subject-verb inversion, given the assumptions of Grimshaw’s framework.

In Grimshaw's framework, a sentence without an auxiliary projects as a minimal VP containing the subject and any internal arguments, and in a situation where the verb precedes the subject (subject-verb inversion), this is to be derived by movement of V to CP, since SPECLEFT outranks STAY. Since (38b) incurs a violation of STAY, (38a) is the preferred candidate in the derivation, and hence it (38a) is grammatical and (38b) is not. But what about HDLFT? (38a) clearly contains a violation of this, because of the subject in the specifier, but (38b) does not, since the verb is the head of the CP projection after movement; that a moved head is the head of its target projection is a necessary result in Grimshaw's framework (see p.376 for Grimshaw's discussion of extended projections), so the result is that the V head is the leftmost element in the TP layer and thus no violation of HDLFT is incurred.²⁸ But (38b) does involve a violation of SPECLEFT, since the specifier associated with the VP is no longer to the left of all the projections of its head (which now include the CP layer). This is all represented in the tableau below:

(39)

/input/	SPECLFT	HDLFT	STAY
a. $\text{[VP [thy thoughts]_{DP} die]}$		*	
b. $\text{[CP die}_i \text{ [VP [thy thoughts]_{DP} t_i]]}$	*!		*

Grimshaw's tableau representing the contrast between examples like those in (38) does not represent SPECLEFT and HDLFT, as she focuses upon the extra violation of STAY involved in the inverted example, but what is demonstrated in (39) follows necessarily from her core assumptions and the rankings she proposes.

It is important to clarify that the constraints on the ordering of Spec, Head and Comp do not just apply at the level of base-generation, but they also count for their derived positions, at least in part; if this were not the case, then the moved verb in (39) would not count for a violation of SPECLEFT. The fact that (38a) is grammatical and (38b) is not is evidence for this fact itself, since there

²⁸Grimshaw (1997: 408) notes that the traces of heads do not count for the calculation of HDLFT.

is no other way to distinguish the two derivations: Grimshaw's discussion only addresses the difference between the two derivations with respect to STAY, but since STAY is outranked by HDLFT, there must be some additional violation in (38b) that accounts for the optimality of (38a), and the tableau shows that this is SPECLEFT that is violated due to the movement of the verb. Grimshaw does not address the issue of whether moved heads count for violations of SPECLEFT. She does argue that moved heads do not figure in the calculation of HDLFT (p.408), but it is clear that moved heads (what Grimshaw calls 'imperfect' heads) should count for the calculation of other constraints, given their contribution to the satisfaction of OBHD, a constraint that requires projections to have heads. (39) shows that it is necessary to include moved heads in the calculation of SPECLEFT. Thus in Grimshaw's (1997) framework, it is impossible to avoid analysing subject-verb inversion as a case of a SPEC-LFT violation.

If this is the case, where does this leave the theory? The immediate result is that subject-verb inversions like that in (37a) involve a violation of SPECLEFT, and hence this constraint must be outranked by the metrical constraints. Fitzgerald notes that ranking metrical constraints above SPECLEFT means that we overgenerate strings like (35) which are not attested; above I noted that such strings are not in fact generable given Grimshaw's assumptions, but that the same point can be made with superficially similar examples like (36). Nevertheless, we could still rule out this particular example, provided that SPECLEFT outranks ADJ-N: the reordering found in the original poetic example avoids the MATCHSTRESS violation as well, so it would be the optimal candidate in this case. Below I present a simplified and revised tableau to make this point:

(40)

/input/	MATCHSTRESS	SPECLEFT	ADJ-N
a. ♣ When rocks impregnable are not so stout (33a)			*
b. When impregnable rocks are not so stout (33b)	*!		
c. Impregnable rocks are not so stout when (36)		*!	

Thus the inclusion of SPECLFT does not actually introduce as many problems as Fitzgerald suggests, at least with that basic data.

However we can find examples where a violation of SPECLFT would actually obtain a more metrical line, thus becoming the optimal candidate, while the grammatical form itself is highly unusual and seemingly unattested. (41) is one of Fitzgerald's examples (from p.214, taken from *Sonnet 12*):

- (41) . X . X . X X . . X
 When lofty trees I see barren of leaves
 W SW S W S WS W S

Fitzgerald shows that this line incurs an extra violation of HDLFT, because of movement of the complement *lofty trees*. It also violates the metrical constraint *CLASH, which rules against sequences of adjacent stresses (modeled on Nespor and Vogel 1979); this is due to the adjacency of *see* and *barren*. MATCHSTRESS is also violated once, due to the placement of the stressed syllable in the polysyllabic word *lofty*.

Nevertheless (41) seems to be the optimal candidate, since the non-inverted alternative, presented below, incurs an extra violation of *CLASH (adjacent *see* and *lófty* is one, adjacent *trees* and *bárren* another), in addition to a violation of MATCHSTRESS (stress on *barren* matched to a weak position). (43) summarises this in a tableau:

- (42) . . X X . X X . . X
 When I see lofty trees barren of leaves
 W W S SW S WS W S

(43)

/input/	MATCHSTRESS	*CLASH	HDLFT	ADJ-N
a. ☞ When lofty trees I see barren of leaves (41)	*	*	*	
b. When I see lofty trees barren of leaves (42)	*	**!		

Since the metrical constraint *CLASH outranks the syntactic constraint HDLFT, (41) is the optimal candidate between this pair.

Now consider the following altered version of the Shakespearean line:

- (44) . X . X . X . . X .
 I see trees lofty barren of leaves when
 W S W SW S W S W S

This version incurs a violation of two syntactic constraints: SPECLEFT, since the wh-phrase *when* occurs in a right-specifier; ADJ-N, due to the inversion of *trees* and the adjective *lofty*. Yet this line no longer incurs a violation of MATCHSTRESS, the highest-ranked metrical constraint, and no violations of *CLASH either; a violation might be incurred if *trees* is stressed rather than *see*, but this would still only present one violation, fewer than (41). Since SPECLEFT must be outranked by the metrical constraints (in accordance with our conclusions above), we would expect that this more metrical line would be the optimal candidate. This is demonstrated in the updated tableau, where I present only the relevant constraints:

(45)

/input/	MATCHSTRESS	*CLASH	SPECLEFT	HDLEFT	ADJ-N
a. (41)	*!	*		*	
b. (42)	*!	**			
c. (44)			*		*

The problem is that this optimal candidate is not the attested form, and it looks very unlike any sort of form that occurs in the poetry. As a version of “when I see lofty trees barren of leaves,” it is a veritable word salad, and our theory should not rule in such examples because they are not attested. This shows that accepting SPECLEFT as a constraint that is outranked by the metrical constraints causes a lot of empirical problems for the theory, not just ruling in bad examples but also ruling out good ones like the attested (41).

One possible way to reduce this problem is to play down the significance of subject-verb inversion as a construction in poetic language, and thus to deny that SPECLEFT is outranked by the metrical constraints. As mentioned above, Fitzgerald seems to do this somewhat tentatively, asserting that subject-verb inversion is uncommon. Yet we have already seen in section 2.1.4 that, while

subject displacement is less common than other kinds of displacement, it still occurs across different kinds of texts. As it happens, there are other cases of putative *SPECLFT* violation within Fitzgerald's own paper, in the cases of adjective-noun inversion: it is often argued on the basis of syntactic and semantic evidence that (some kinds of) adjectives are NP-specifiers (Cinque 1993, Alexiadou 2001 among others), and if this is the case, adjective-noun inversion like that in (33) and (44) would be on a par with subject-verb inversion as a case of a violation of *SPEC-LFT*; indeed the independent constraint *ADJ-N* would seem to be unrequired in this view. Thus it seems that there would be a strong case for admitting *SPECLFT* as another syntactic constraint that is outranked by the metrical constraints. Yet if we were to amend the poetic grammar in this fashion, we would then make a large number of predictions regarding possible orders that do not occur, such as those mentioned by Fitzgerald, which seem to stand as genuine data points as far as I know. The poetic grammar would thus require a much more refined readjustment in order to make the correct distinctions, although it is not clear what the alterations would be.

However, even if we are to ignore *SPECLFT* and dismiss the analysis of adjectives as specifiers, we still find that the constraint rankings proposed by Fitzgerald overgenerate massively. Fitzgerald proposes that the metrical constraints outrank *HDLFT*, and that this accounts for displacement of the verbal complement(s) that alleviates a violation of the metrical constraints. An additional benefit is that it also accounts for a number of the other kinds of deviant displacement of NPs and PPs seen in the previous chapter; unlike the rule-based system discussed in section 3.1, this explanation does not need to stipulate that the displacement we see is a sub-type of topicalization or some other standard movement rule, since movement is not constrained in the same way in the OT system. The problem, however, is that we predict that a number of entirely unfamiliar structures should be just as available as these well-attested structures. For example, we would seem to predict that determiners and their NP complements could be inverted, at least if we assume the DP hypothesis (as Grimshaw

does). We can find a potential example of this in Fitzgerald's data: (46a) is the poetic example from *Hamlet*, which contains a violation of HDLFT induced by inversion of the verb and its nominal argument; the uninverted form (46b) incurs a violation of MATCHSTRESS with the stress in *dalliance* falling in a weak position. In (46c) I provide an example where the D head of the DP and the NP complement have been inverted, perhaps by movement of NP to Spec,DP; this example, like (46a), avoids the violation of MATCHSTRESS by virtue of the inversion²⁹, as demonstrated by the tableau in (47):³⁰

- (46) a. . X . X . X . X . X
 Himself the primrose path of dalliance treads
 W S W S W S W S W S
- b. . X X . X . X . X .
 Himself treads the primrose path of dalliance
 W S W S W S W S W S
- c. . X . X . X . X . .
 Himself treads primrose path of dalliance the
 W S W S W S W S W S

(47)

/input/	MATCHSTRESS	HDLFT
a. ☞ (46a)		*
b. (46b)	*!	
c. ☞ (46c)		*

This derives the unwelcome situation where there is no optimal candidate, at least according to the present set of rankings, and it is especially unwelcome since this kind of inversion is not found in the poetry of Shakespeare. Yet given the set of rankings, this form should be as common as the attested form in (46a).

Fitzgerald considers the general impossibility of inverting determiners and their nominal complements briefly, and she states that such inversions may

²⁹The verb *treads* can be destressed in this position too, thus avoiding a violation of *CLASH; if this wasn't possible, the optimality of (46a) would follow.

³⁰All of these examples also involve extra violations of STAY, but I factor this out of the tableau since it does not impact upon the discussion here (since it is low-ranked).

be avoided by formulating a constraint that ensures that functional categories occur to the left of lexical categories. Properly formulated, such a constraint would deal with the data above. However, there are other situations where this kind of explanation is not possible. In particular, we struggle to control kinds of movement that are sometimes licensed in restricted circumstances by virtue of the interaction of low-ranked constraints like *STAY* and *HDLFT*. One example comes from auxiliary movement. The normal ranking of constraints for English is designed to allow for some kinds of movement of auxiliary verbs, for example in *wh*-questions. I will briefly introduce Grimshaw's (1997) account of subject-auxiliary inversion in *wh*-questions, before then showing how adding higher-ranked constraints like the metrical ones undercuts this analysis and hence introduces a number of overgeneration problems for the OT theory of poetic inversion.

Grimshaw's system makes use of three core constraints in deriving the syntax of English matrix *wh*-questions: *STAY*, *OBHD*, and *OPSPEC*; the latter dictates that an operator (like a *wh*-phrase) occupy a specifier position, in which it can form an operator-variable relation. The ranking is *OPSPEC*»*OBHD*»*STAY*. The basic paradigm is represented in (48):

- (48) a. What will she say?
 b. *What she will say?
 c. *She will say what? (as full *wh*-question)

(48a) is the standard case with *wh*-movement and subject-auxiliary inversion; (48b) lacks aux-raising and (48c) lacks both aux-raising and *wh*-movement. Failure to move the *wh*-phrase incurs a violation of *OPSPEC*, in (48c), and failure to provide the CP layer with a head;³¹ the lack of head movement incurs a violation of *OBHD*. The tableau below demonstrates this:

³¹Grimshaw proposes there is no C head in the CP in the absence of an overt complementizer.

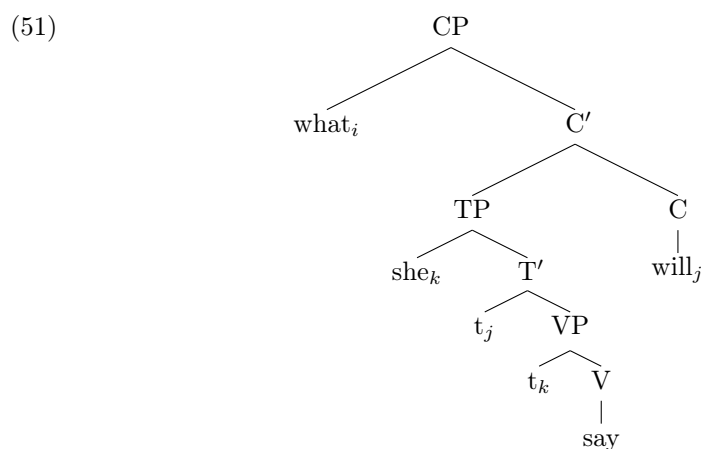
(49)

/input/	OPSPEC	OBHD	STAY
a. What What will she say?			**
b. What she will say?		*!	*
c. She will say what?	*!	*	

What is important here is that the constraint that makes the difference between (48a) and (48b) is OBHD. In the ranking proposed by Grimshaw's (1997) system, OBHD is low in the ranking, outranked by constraints like HDLFT; in the present context, this means it would be outranked by the metrical constraints.

This makes the unwelcome prediction that head movement to CP is relatively unconstrained, as long as it leads to satisfaction of a metrical constraint. For example, we cannot rule out a situation where the moved auxiliary occurs in a right-headed position. An example of such a structure in a wh-question is given in (50), and its full structure is provided in (51):

(50) *What she say will?



Recall that Grimshaw's system controls spec-head-comp order via the alignment constraints SPECLFT and HDLFT (and the corresponding right-headed versions), rather than by the (inviolable) alignment parameters used in Minimalism. Recall also that movement of the head to C satisfies OBHD, which is outranked by HDLFT; thus we may expect that the OT system will rule out (51) by proposing that the moved head violates HDLFT in its derived position.³²

³² In fact, Grimshaw's system is actually unable to prevent this structure as it is formulated

Now consider the following example from Pope, discussed in chapter 2:

- (52) But still this world, so fitted for the knave,
 Contents us not. A better shall we have? *ES4: 131-132*

The yes-no question in (52) contains a deviant example of NP-displacement, where *a better* has been moved to a sentence-initial position. Since English typically does not allow for topicalization to occur in a yes-no question, and topicalization itself does not trigger subject-verb inversion, this is an unlicensed movement. This example could be explained easily by Fitzgerald's system, since it is a perfect example of poetic inversion. Below I give the basic metrical analysis of the inverted and uninverted forms, and the tableau for the comparison:

- (53) a. . X . X . X . X . X
 Contents us not. A better shall we have?
 W S W S W S W S W S
- b. . X . X X . X . X .
 Contents us not. Shall we have a better?
 W S W S W S W S W S

(54)

/input/	MATCHSTRESS	HDLFT	STAY
a. Contents us not. A better shall we have?		*	**
b. Contents us not. Shall we have a better?	*!		*

Note that one of the violations of STAY in both examples is caused by the movement of the verb to C, to satisfy the need for a head in that projection. I assume here that the NP *a better* moves to a higher projection above CP, since Spec,CP is filled by the null operator that is typically assumed to occupy this

in Grimshaw (1997), since later (p.409) she states that HDLFT applies only to 'perfect' heads, that is, heads that match their projections in all their features; V or T are not a perfect heads in C. I am unaware of any solution to this problem in the literature; Grimshaw (2001) takes the opposite view, arguing that moved heads must count for the alignment constraints, but such an alteration completely undermines the empirical basis of the analysis in Grimshaw (1997) and Grimshaw's discussion there makes no reference to the empirical picture. We might recover this data by adding a constraint that regulates for unwelcome right-heads (and formulate it so it applies to moved heads), but this would be stipulating to cover for what seems to be a systematic flaw that stems from making the alignment of X-bar projections contingent upon constraint satisfaction.

position in yes-no questions (as is the case in Grimshaw's framework).³³ The crucial point is that MATCHSTRESS outranks HDLFT, and since the movement of the object NP leads to satisfaction of MATCHSTRESS, the example with the inversion is the optimal candidate. Thus Fitzgerald's theory correctly predicts that the poetic inversion should occur in this case.

Now consider the following alternative version of (53a). In this example, there is no fronting of the NP, but the auxiliary verb has undergone movement to a right-headed position in C, as opposed to the left-headed position that is standard. This incurs an extra violation of HDLFT, but it leads to satisfaction of MATCHSTRESS in the same way as deviant NP movement does in (53a). The tableau compares the uninverted line and the attested one with this alternative version, and we see that the alternative is actually the optimal structure since it incurs one less violation of STAY:

- (55) . X . X . X . X . X
 Contents us not. We have a better shall?
 W S W S W S W S W S

(56)

/input/	MATCHSTRESS	HDLFT	STAY
a. Contents us not. A better shall we have?		*	**!
b. Contents us not. Shall we have a better?	*!		*
c. Contents us not. We have a better shall?		*	*

This is the wrong result in this particular case, and it indicates that in other cases the theory would predict forms like this to occur in poetry. But right-headed forms like this never occur, either in the most contrived metrical poetry or in the most experimental avant garde texts.³⁴ This is a clear case of an over-generation that we should strive to avoid, since such lines are almost impossible to interpret and thus outwith the remit of the theory of poetic language (as it is set out in the introduction).

³³I ignore the question of whether or not the movement of the verb induces extra violations of HDLFT, since this has no bearing on the comparison being made here.

³⁴See sections 2.1.5 and 4.2.5 for discussion and explanation.

This overgeneration is indicative of a larger flaw in the OT-driven poetic grammar. A core property of OT syntax is that its candidate generator GEN can generate a large array of linguistic forms for comparison in the EVAL component; the forms that are generated are only restricted by the limitations of basic X-bar-theoretic phrase structure and the lexicon. This generative power is kept in check by the constraint ranking that describes a particular grammar, and this constraint ranking is delicately balanced to describe the language in question: the interaction of low-level constraints can make the difference between forms that are perfect and those that are ungrammatical. However, once we allow metrical constraints to outrank syntactic constraints, these distinctions are lost and we henceforth predict a massive range of ungrammatical forms to be attested and ‘grammatical’ in the context of metrical realignments. Importantly, there is no way to distinguish between the good predictions and the bad ones, since the acceptability of the output of an OT grammar is blind to the kinds of constraints that are violated in a comparison of non-optimal candidates. Here I have reviewed just a few overgenerations, but the generative power of GEN would force us to generate and expect attested versions of forms that are barely interpretable as sentences of English.

We can see, then, that the specifics of Fitzgerald’s proposals are flawed, and that the theory will necessarily over-generate or under-generate depending on how we intend to deal with the issue of subject-inversion (and the related issue of Adj-N inversion). We have reviewed just a few internal empirical problems for the theory, and shown that they cause serious problems. But the main empirical problems for Fitzgerald’s theory of poetic grammar come from the outside: there is a massive amount of poetic language data that cannot be dealt with by the proposed mechanisms. Keeping to the realm of traditional poetry, there is a very strong tendency for inversions for the sake of rhyme; we saw plenty of this in the previous chapter, and it is one of the most well-known motivations for poetic inversion in most English traditions. Yet rhyme is not conditioned by any sort of metrical constraint, nor by any linguistic rule, since it is a paralinguistic

trait that holds between linguistically unrelated non-linguistic objects – that is, separate lines.³⁵ It is thus impossible to construct a plausible set of linguistic constraints that will regulate poetic inversion in service of rhyme in a ranked constraints system.

This is not the only gap in the theory's coverage, however, as there is also a great deal of inversion that has no effect on metrical structure (Fitzgerald acknowledges the existence of such data), and indeed there a great deal of inversion and deviation in poetic lines that have no obvious metrical form (as mentioned in the previous chapter): we may see this in the more experimental texts, or in poetic prose like that of Samuel Beckett. It would seem strange to propose a theory of poetic language that is able to explain only one corner of the entire empirical picture, and to do so would be clearly missing important generalizations that can be extracted from considering the data as a whole, so I consider this to be a significant problem for the poetic grammar theory formulated in this way. Fitzgerald notes that the formal inversions that are best explained by her model are those that occur most frequently and, as mentioned earlier, she plays down the significance of a counter-example by noting its (purported) infrequency. One might surmise that this desire to favour the explanation of forms that are more commonly attested in the texts in favour of wider coverage of the less common and more diverse data is derived from the tendency in OT to favouring of probabilistic models of the grammar, such as that proposed by Bresnan and Hay (2007);³⁶ thus the role of the poetic grammar would be to explain broad tendencies in poetic language. This seems to miss the point of theorizing poetic language, however, as it is the diversity of the kinds of deviation that is its most striking aspect.

Thus the OT approach suffers problems both internally in its details and externally in its relation to the wider data picture.

³⁵For discussion see Fabb (2002, 2009).

³⁶Though see Bruening (2010) for a recent rebuttal of the structural analysis proposed by Bresnan in this work and others.

3.3.2 Theoretical problems

As we may expect, the OT approach suffers from some of the same theoretical problems as the theories discussed previously, as well as some others that are unique to this specific approach. I will overview these problems briefly, concentrating on the source of the problems unique to this approach and what they teach us about the nature of poetic language.

The main shared problem is the one discussed previously, that of acquisition and explanatory adequacy. In OT, grammars are rankings of universal constraints, and the rankings are set during acquisition. The problem, then, is familiar: how can we acquire the relevant ranking of metrical constraints against the syntactic constraints when exposure to the relevant data (in very small amounts) only occurs long after the critical period? It seems implausible that the ways in which we become familiar with poetic inversions are the same as the ways in which we learn syntax and phonology. This is no less of a problem for OT than it is for the other generative theories, and as such it is a serious problem for Fitzgerald's approach.

The OT approach also shares with the 'UG exploitation' approach a certain degree of arbitrariness in its empirical coverage, in that the ranking of metrical constraints above some syntactic constraints and below others is effectively accidental. Why, for example, should *MATCHSTRESS* outrank *HDLFT* but not *SPECLFT*? No explanation is given; rather, this is just the result of the ranking of constraints. A similar problem was exposed in the previous subsection, where we saw that some forms that violate syntactic constraints in the name of satisfying metrical constraints are not attested, even though they are predicted to be just as common and available as the attested forms. The present dissertation has set the goal of trying to explain these correlations, rather than just to describe them, so as such this seems like a further weakness.

Perhaps the most interesting theoretical problem for the OT approach is its use of verse lines in the computation of syntactic conditions. The fundamental assumption of Fitzgerald's theory is that metrical conditions can influence the

operations of the syntactic engine, thus conditioning whether or not a given syntactic form is well-formed (and hence likely to be produced/generated). The metrical conditions are tied to the theory-internal notion of a poetic line, specifically the metrical template. We must assume that the set of metrical templates has the status of a set of linguistic primitives, like lexical items, since otherwise they could not be involved in the computation of syntactic well-formedness in the grammar. I will put to one side here the exact ontological status of the metrical template, which is never made clear in Fitzgerald's work or elsewhere in the literature. What is clear is that it needs to be part of the set of linguistic primitives, and one may assume that a given metrical template is stored in memory like a lexical item (or perhaps a functional projection), and to use one particular template is to access that macrostructural object for the linguistic computation. This is the very least we need to assume to derive a theory where the syntax and the meter interact.³⁷

A fundamental assumption of all work in generative linguistics is that the maximal object of linguistic description is the sentence: the form of one sentence is never influenced at the level of grammatical computation (phonological, syntactic or semantic) by the form of preceding or following sentences. We thus assume that the grammar generates just sentences, and that the grammatical engine thus 'forgets' the previous form once it moves on to generate another. With this basic set of assumptions, we are thus committed to assuming that metrical templates and grammatical objects are effectively isomorphic: that is, a line coincides with a single discrete syntactic object. It is impossible for a grammatical primitive like a metrical to straddle separate grammatical objects, since the first object is cleared from the workspace of computation once it is completed.

³⁷Note that this is a heavier ontological burden than that which is required by alternative theories like Fabb and Halle (2008). Fabb and Halle propose a rule-based system for meter, where a metrical grid is generated by a set of specific metrical rules that are not part of the phonology proper. As such, this system does not require the notion of a template, since the grid is the object that is generated by the rules. In contrast, the OT system needs to assume that there are both metrical constraints operating in the grammar and metrical templates for comparison.

That metrical templates and grammatical objects are isomorphic is clearly wrong. Look again at (53a) from above, in its full couplet form:

(57) . X . X . X . X . X
But still this world, so fitted for the knave,

W S W S W S W S W S

. X . X . X . X . X
Contents us not. A better shall we have?

W S W S W S W S W S

This couplet displays an example of enjambement, where the subject of the first sentence, *this world*, is severed from the rest of the sentence by the line boundary. This shows a simple non-isomorphism between the line and the sentence. The importance of this non-isomorphism becomes clearer when we look at the second line, which is composed of the second half of one sentence and a full second one containing a poetic inversion. Whether or not the inversion is allowed depends on the relative ranking of the syntactic and metrical constraints and the violations incurred on these constraints; for the calculation of this, we must access the metrical template and compare it with the prosodic structure of the sentence. Crucially, this involves accessing the prosodic structure of a part of the previous sentence: the structure of the inverted sentence is dependent upon the grammatical form of the previous section. This kind of interdependence of separate grammatical forms is impossible under the most basic assumptions of generative linguistics, since we assume that sentences are wholly encapsulated with respect to grammatical properties and computations. To allow for this interdependence, we would need to give up these basic assumptions and move to a framework that allows for such non-modular computations.

What this indicates is that lines and syntactic structures are entirely unrelated objects, and that we should not propose dependencies between them. There is plenty of evidence for this. Another aspect of the theory's use of metrical templates in determining syntax is that we predict that poetic inversions

should always be line-internal; that is, that it should not be possible to invert phrases that span a pair of lines by enjambement. It is certainly true that there are fewer examples of cross-line inversions, but nevertheless these inversions are attested. Below is one such example from chapter 2:

- (58) Triumphant Umbriel on a Sconce's Height
 Clapp'd his glad Wings, and sate to view the Fight, *TR5: 53-54*

Here displacement of the PP straddles the line break but it provides metrical well-formedness for both lines. This inversion is broadly similar to all the others reviewed by Fitzgerald, but it is impossible to model it within her system.

Another interesting mismatch between lines and syntax is the fact that non-constituent displacement occurs often when a constituent straddles a line break. This is what we see in the following examples from chapter 2: line-internal inversions separate the parts of a syntactic constituent for the sake of creating metrical well-formedness:

- (59) For this, ere Phoebus rose, he had implor'd
 Propitious heav'n, and ev'ry pow'r ador'd,
 But chiefly Love – to Love an Altar built,
Of twelve vast French Romances, neatly guilt *TR2: 35-38*
- (60) Here Britain's statesmen oft the fall foredoom
Of foreign tyrants, and of nymphs at home; *TR3: 5-6*
- (61) One self-approving hour whole years outweighs
Of stupid starers and loud huzzas: *ES4: 255-256*

The majority of examples of non-constituent movement gathered in section 2.1.4 involve inversions of this kind. Here syntax and lineation mismatch in two different ways: the basic alignment of phrases in the uninverted forms does not match a sentence or constituent to a line. This seems to be a clear demonstration that line-internal inversion does not involve syntactic computations, but non-syntactic reorderings of words on the page, and that syntactic and metrical computations are entirely divorced. There is no direct one-to-one relationship between verse lines and syntactic structures, and the computations involving

syntax and meter are effectively unable to ‘see’ one another. Given this, the theory should not attempt to yoke syntax and versification together into a single set of mechanisms, since they are clearly unrelated.

3.3.3 Summary

We have seen that the OT approach proposed by Fitzgerald (2007) suffers from a number of empirical and theoretical problems that make it unviable as a theory of metrically motivated poetic inversion and hence as a general theory of poetic language. Once more, most of the problems stem from the assumption that the forms found in poetry are generated by a poetic grammar in a manner that is largely the same as with the ordinary grammatical operations. In formulating grammatical constraints to allow for a given set of poetic examples, we necessarily predict a massive range of unattested forms to be equally available, all the while failing to account for the majority of basic data. There are different ways to interpret these problems. One reaction is see them as strong predictions of an underdeveloped theory, and that this is part of the process of theory development in generative linguistics; we learn from these false predictions and redesign to cover the given cases and improve our understanding of the general mechanisms.

However, we have seen that the empirical problems identified above are wide-ranging and often unfixable, and that they are coupled with deep theoretical problems that cannot be resolved by tinkering. Furthermore, in preserving the poetic grammar theory by continuing complications it would seem that we would be missing an important generalization about poetic language that may influence how we orient our theory at the very beginning: that it behaves unlike normal language in many different ways and thus seems to be something else altogether. Within the framework of OT, it may well be possible to redesign the theory to account for a reasonable amount of the empirical problems discussed above, since the machinery of OT is extremely powerful, but I believe that the diversity of problems discussed in this section, and in the others above, indicates

that this would be missing the point.

3.4 Overall summary

In this chapter I have critically reviewed three theories of poetic language and shown that they all suffer from theoretical and empirical problems. All of these theories are unified in the assumption that the deviant forms we see in poetic language are generated by the grammar, specifically a special variation known as the ‘poetic grammar,’ and this has effectively been the standard assumption in linguistic analyses of poetic language at least since the inception of generative linguistics. Yet we have seen that the three theories reviewed here, which constitute the clearest statements of poetic language theories in the literature, suffer from a number of internal problems, in terms of their specific proposals and the predictions made, and has become apparent that it is very difficult to implement a poetic grammar theory within the modern framework. Minimalist syntax proved too restrictive to accommodate the kinds of amendments required of the poetic grammar, and Optimality Theory syntax proved too powerful, overgenerating wildly after minor poetic amendments. None of the generative approaches are able to generate the examples where non-constituents are affected by erasure or displacement, at least without significant and suspicious contrivance.

What this review has shown us is that the diversity of phenomena in poetic language is simply not amenable to a fully linguistic explanation. Although the orthodoxy is to assume that such an explanation should be desirable, it must be stressed that there are no compelling reasons to assume *a priori* that such a solution should be correct. Indeed, there are many reasons to believe that the language of poetry should actually be different from natural language, that it should be produced and interpreted in a different way and theorized in a different way. We have seen in the discussion of the metrically driven OT theory in particular that the reordering of words in poetic texts is often driven by considerations of meter and other aspects of literary form, yet we have

also seen that it is impossible to model a direct connection between reordering *as syntactic reordering* and metrical form. Instead, what we seem to see is simple rearrangement of words and phrases as sequences of words in a line, by a conscious and deliberate procedure. Indeed the contrivance of literary form, with its control of syllable counting, sound patterning, structural parallelism and lineation, is all extraneous to the core computations of human language, unsurprisingly given that they are only employed in a very specific and contrived context. Therefore with such a high degree of contrivance in the arrangement of parts of sound, it is perhaps to be expected that the arrangement of words and sequences of words associated with phrases is also contrived and outside of the domain of core grammatical procedures.

The most important function of the detailed critiques in this chapter is that it identifies a number of specific empirical issues that need to be dealt with by the theory of poetic language. This sets the stage for the next chapter, where I propose a theory that breaks with the approaches reviewed here to provide an alternative theory that is better equipped to deal with the empirical challenges and more readily compatible with the modern view of linguistic theory.

Chapter 4

A new theory

This chapter presents a new theory of poetic language that deals with the empirical challenges identified in chapters 2 and 3. I propose that we dispense with the poetic grammar theory and reject the idea that poetic language is generated in the same way as ordinary language. Instead I propose that the forms of poetic language are produced by what I call ‘PF-uneconomical derivations.’ Displacement is derived by concatenation of well-formed chunks of syntactic structure (‘Pieces’), employing independently needed basic operations for the sake of deriving meanings for pieces of syntactic structure that cannot compose into larger syntactic units. Erasure is derived as over-application of the operation Delete in the context of an economy-based theory of ellipsis and copy deletion, with some derivations overlapping with those involved in displacement. This theory is situated in the context of the interface-centred view of the grammar, and it follows entirely from Minimalist assumptions without any additional rules or stipulations. The theory explains the fact that economy doesn’t block these structures by appealing to an enriched ‘output expectation’ at the interfaces, and I conclude by speculating on how this might allow us to account for apparent interactions between syntax and external modes of formal organization.

The chapter is structured as follows. Section 4.1 reviews the problems with

the poetic grammar approach and advocates an alternative hypothesis about how poetic language is generated, which I call the ‘Non-uniformity Hypothesis.’ Section 4.2 introduces the initial proposal, ‘Bypass Theory,’ which deals with the problems posed by displacement; I provide full details on the justification for the different parts of the proposal and indicate how it makes the right predictions with respect to the distribution of displacement, as it was identified in chapter 2. Section 4.3 then provides details of the relevant derivations through some examples. It also discusses the adjustments required to account for erasure and presents a generalized form of the theory in section 4.3.4. Section 4.4 sums up.

4.1 The Non-Uniformity Hypothesis

We have seen that poetic grammar is not a viable theory of poetic language. Modern generative frameworks are technically and theoretically incompatible with poetic grammar; it is simply not possible to configure a Minimalist syntax to generate the various different kinds of forms we see in poetic language. We may wonder where this leaves the generative theory of poetic language, since it has been assumed implicitly and explicitly throughout the literature for decades that the only way in which generative linguistics may be able to explain the language of poetry is to propose that it is generated by the grammar in the same way as ordinary language. In this chapter I will show that this is not the case, and that we can develop a theory of poetic language without presuming that poetic language is generated in the same way as ordinary language.

Before developing such a proposal, we first need to understand what it means in theoretical terms to claim that poetic language is generated in the same way as ordinary language. In this section I will review the fundamental roots of this proposal, what Fabb (2010) calls ‘the Development Hypothesis,’ and I will then propose that the evidence presented so far provides strong motivation for abandoning this hypothesis. I will then develop an alternative proposal, which I call ‘the Non-Uniformity Hypothesis,’ which will provide the background for

the theory to be developed.

4.1.1 The Development Hypothesis and its problems

The core idea of the poetic grammar theories is that the language of literature is a ‘development’ of ordinary language. Fabb (2010) calls this ‘the Development Hypothesis,’ and he states this explicitly as the following:

- (1) The Development Hypothesis: Literary language is governed only by rules and constraints which are available to ordinary language, and which refer only to representations which are present (at some stage in a derivation) in ordinary language.
(Fabb 2010: 1220)

The works reviewed in chapter 3 all assume this working hypothesis, in ‘strong’ and ‘weak’ forms. According to Fabb (2010: 1220), the strong form assumes that “a literary language is a development of its source language,” and this is the version that is assumed by the works reviewed in section 3.1. The weak form of the Development Hypothesis “allows a literary language be a development of the universal possibilities underlying all languages,” and this is demonstrated by O’Neil (2001), discussed in depth in section 3.2. The proposals of Fitzgerald (2007), discussed in section 3.3, may be identified as a weak form of the Development Hypothesis, since in that system the source language is effectively a set of ranked universal constraints, and the literary variation on the source language involves ranking constraints that are universally available in a different way. Given that the relative ranking of the source language remains stable, the possibilities that are made available by the additional ranking are technically a variation on the source language.

Fabb (2010: 1220) observes that the Development Hypothesis depends on distinguishing between two different kinds of literary language, ‘natural literary language’ and ‘artificial literary language;’ it is assumed that only ‘natural literary language’ conforms to the Hypothesis, and that artificial literary language is something else that is not to be explained by the linguistic theory of literary language. Artificial literary language is exemplified by the unusual language

of avant garde writers, and Fabb notes that Sapir's (1921) analysis of Walt Whitman's poetry is an early discussion of this division between artificial and natural literary language. Fabb notes that there is precedent for this separation of artistic work into two categories, one that fits with cognition, one that does not: he cites Lerdahl's (1988) discussion of the cognition of music, in which he argues that the avant garde composition of Pierre Boulez's work is "cognitively opaque" and thus qualitatively different from the work of Beethoven and other such composers. Lerdahl's work with the linguist Ray Jackendoff (Lerdahl and Jackendoff 1983, Jackendoff and Lerdahl 2006) develops a generative theory of musical structure, so this partitioning of the avant garde on one side and conventional music on the other is effectively a way of maintaining a similar division to that which is proposed within the Development Hypothesis.¹ In the study of poetic language, then, the examples of unusual erasure found in the avant garde texts would not be considered in the formation of a theory. This work would be classified as artificial literary language, and it would be put to one side. We may presume that the theory would be concerned largely with metrically motivated inversion and other such 'well-behaved' deviations.

One of the most important points about my discussion of the theory of poetic language so far is that I have assumed that the theory should give a uniform explanation for the various different kinds of deviation we see across different kinds of poetic texts. In doing so, I have rejected the assumption that there are different categories of deviation in poetic language, and specifically I have rejected the idea that some kinds of deviation are artificial and others are natural. I have assumed that the non-constituent-affecting processes of displacement and erasure are to be explained by the same mechanisms that explain the metrical inversions discussed by Fitzgerald (2007), rhyme-driven inversions and the various other kinds of more regular and predictable forms of deviation. The reasons

¹Note that there are important differences between Lerdahl and Jackendoff's work on the one hand and the work on poetic language on the other, since Lerdahl and Jackendoff do not set out to study music as a special variant of some other basic component of cognition. Effectively their object of study is the standard language or prose equivalent in musical cognition, rather than some deviant form like literary language.

for adopting this assumption were discussed in chapter 1, where I argued that arguments from both linguistic and literary theory militate against separating the traditional and the avant garde when it comes to defining the empirical basis of a theory of poetic language. The analysis in the preceding chapters has added weight to that argument, as we have seen that many kinds of deviation are found in both traditional and avant garde texts. Indeed one might even argue that the language of the traditional poetry is stranger and more “artificial” than that of the avant garde texts, since the forms of deviation that are most difficult to account for with standard mechanisms, like non-constituent displacement, are found primarily in the traditional texts. It would be highly suspicious to separate off the non-constituent-affecting set of deviations within the corpus of traditional poetry as “artificial” and concentrate on just the obedient cases, not least since they often share some crucial similarities. The fact that most manipulation of linguistic form in traditional language can be described in terms of metrical or rhyme-driven “motivation” is irrelevant, since we also saw in chapter 3 that it was impossible to model a causal formal relation between metre and syntax, since the two systems access unrelated objects and cannot ‘see’ the inner operations of one another.

The problem is that the categorization of some sets of examples as “artificial” and others as “natural” is almost entirely arbitrary and based on extralinguistic factors. Such a partitioning is both theoretically and methodologically unsound, and I believe it should be avoided.² If the partitioning of poetic language into

²It should be stressed that the artificial-natural distinction may still have some part to play in defining the boundary conditions for the theory of poetic language. Recall that in section 1.1 I set out the scope of the theory of poetic language as the set of deviant sentences in poetic texts; that is, the set of sequences of words that do not correspond to normal sentences of English but which nevertheless map onto sentential meanings. In setting this definition, we thus exclude a whole class of examples in literary texts from the theory, namely those sets of words that do not map onto sentential meanings. This class is found commonly in avant garde literary practise, in particular in the cut-up works of William Burroughs, John Cage, Jackson Mac Low and their contemporaries. Such data cannot be described as linguistic data, since the correspondence from sound (or orthographic representation of sounds) to meaning (describable in terms of truth-conditions as propositions) is lacking; these sequences can be described as “artificial” since the only thing that they share in common with natural language data is surface appearance, as a sequence of words. However, this partitioning of natural and artificial works with a clear definition of *natural* language, and as such it is very different from the distinction at the heart of the Development Hypothesis, which distinguishes between literary cultures rather than formal properties of texts.

natural and artificial along these lines is rejected, the Development Hypothesis is made to look increasingly suspicious. Given that the poetic grammar theories also suffer from a number of theory-internal problems, I propose that we reject the Development Hypothesis altogether and thus reject the idea that poetic language sentences are generated as well-formed outputs of the grammar just like ordinary language sentences. This means that we assume that poetic language is fundamentally different from ordinary language, and that the theory of standard language does not provide a uniform explanation for standard language and poetic language both. I formulate this below as the ‘Non-Uniformity Hypothesis’ below:

- (2) The Non-Uniformity Hypothesis: poetic language and ordinary language do not receive a uniform explanation in terms of linguistic theory.

An important aspect of this Hypothesis is that it does not exclude the possibility that some sort of linguistic explanation should be provided for poetic language. What it does do, however, is ensure that the explanation for poetic language is not the same as that for ordinary language, as is necessarily assumed by the Development Hypothesis and theories based on it. Thus the Non-uniformity Hypothesis rejects the idea that poetic language is generated by the same means as ordinary language.

As far as I am aware, the only work which has proposed to reject the Development Hypothesis is Fabb (2009).³ Considering what he calls the “crazy syntax” of poetry and its relation to the other modes of formal organization imposed upon texts’ structure, in particular lineation, Fabb rejects the idea that poetry is a rearranged version of some well-formed prose equivalent. Instead he proposes that

[verse] lines are composed directly as independent units. . . the rules of syntax are not part of the composition of a line of verse: instead words and phrases and other pieces of language are concatenated together, to produce a flat chain of elements which together constitute the line.

³This is prefigured somewhat in the theoretical discussion in Fabb (2004).

(Fabb 2009: 54)

This is a radical departure from the standard account of how poetic language is generated, and successfully explains the “crazy syntax” of poetry and the existence of lines as formal objects, but it introduces a number of problems in addition. In particular, it raises the question of how we are able to understand lines of poetry and attribute to them truth-conditional meanings as sentences of English.

Fabb acknowledges this issue and proposes that syntax “plays a supplementary role in keeping the lines regular and giving them a meaning:” parallel to the production of lines, we generate a separate piece of ordinary language text, which the generated lines are checked against, roughly in the same manner that they are checked against other constraints like metre and rhyme. Specifically Fabb says that

Depending on the poetic tradition, there will be a requirement that the sequence of words in the line matches the sequence of words in the independently-generated text, but with certain mismatches allowed. This is what makes the line syntactically regular, and also gives the line an interpretation, which it borrows from the separately generated match.

(Fabb 2009: 55)

Fabb’s proposal is capable of capturing the relevant facts about the syntax of verse lines, since it divorces the word order of the line from that of the sentence that corresponds to its meaning, that is, the match sentence. We do not need to rely upon a specific set of linguistic rules for deriving one from the other, but we do need to propose a set of “matching rules” that will ensure that the matching sentence and the line match up in the right way. Fabb does not provide a full explanation of how this would work, but one may imagine that the matching rule would involve some sort of identity relation that ensured that the two did not differ in basic lexical content.

There are a few problems with this account that do not have an immediately obvious solution. The main source of these problems is the match sentence. It

seems that there is a redundancy in generating an entirely separate match sentence along side the fragments of the line. Consider the situation where a reader is interpreting a line of poetic language; in this situation, the role of the match is much less clear than it is in the production of lines by a poet, which is the main focus of Fabb's proposal. The reader is presented with a set of concatenated words, and she must derive the match sentence from this line by some means, since it cannot come from nowhere.⁴ But once this match sentence is generated, why must the reader then go back and compare the match to the line, if the meaning is taken from the match? Moreover, what purpose does the match sentence serve when it is compared to the line in this situation? There are no obvious answers for these questions, and it seems that there is a significant degree of redundancy in this model of interpretation. Besides this, there are a number of other issues for Fabb's theory that remain outstanding, for example the status of the matching rules and how they may be constrained to provide the right empirical coverage; that is, that they only find a match with a well-formed sentence in a limited set of cases. Otherwise we would expect any jumble of words to find a match in any context, and this is an unwelcome prediction.

Despite these problems, I believe that Fabb's proposal presents a viable alternative to the poetic grammar theories of poetic language, and that the problems can be avoided if we dispense with the match sentence and propose an alternative method for generating a well-formed LF representation from the concatenated pieces of language. In the next section (4.2) I develop a proposal for doing this. This allows us to keep the gains of Fabb's proposal – in particular, an explanation for lineation, which is lacking in alternative proposals⁵ – while also providing a theory that explains the “crazy syntax” of poetic language, at

⁴See below for arguments that the meaning of the sentence is not just contextually derived by pragmatics.

⁵Here I refer to the poetic grammar theories reviewed in chapter 3. None of these theories attempts to explain the existence of poetic lines or derive it from any other basic operations. In the other theories, lineation is effectively an accidental property of poetry, and the co-occurrence of lineation with crazy syntax is a coincidence. This even applies to Fitzgerald's (2007) theory, which relates syntax to metre; in this system, lineation is stipulated by the use of a metrical template, but there is no explanation for where the metrical grid and hence the line comes from; also, we have already seen that the relation between syntax and the line in this model is deeply flawed.

least with respect to displacement. I will then generalize this account in section 4.3 to account for the wider range of data, in particular the data from erasure.

Before moving on, I will consider one more alternative theoretical variation that needs to be addressed before moving on. This is an extreme interpretation of the Non-Uniformity Hypothesis that differs from that of Fabb (2009) in its exclusion of the semantic component from the theory altogether. A review of the theoretical and empirical weaknesses of this alternative motivates a minor reformulation of the Non-Uniformity Hypothesis.

4.1.2 The Pragmatics Hypothesis and its problems

In the preceding discussion I have argued that poetic language is not generated in the same way as ordinary language. I formulated this as the Non-Uniformity Hypothesis in (2), and the main purpose of this Hypothesis is to provide the basis for a theory of poetic language that rejects the Development Hypothesis and the theories of poetic grammar that it entails. However the Non-Uniformity Hypothesis has been formulated in an open-ended way that allows for a number of interpretations. The interpretation in Fabb (2009) involves deriving the meaning of deviant sentences from well-formed match sentences that are generated independently in parallel, and as such this approach maintains that the interpretation of deviant sentences involves the semantic component of the linguistic module, Logical Form (LF).⁶

Other interpretations of the Hypothesis may reject this assumption and propose that the interpretation of deviant sentences does not involve LF in this way, and instead propose that the processes involved are almost entirely in the pragmatics.⁷ We may formulate this as a narrower alternative hypothesis:

- (3) The Pragmatics Hypothesis: the propositional meanings of pieces of poetic language is produced largely in the pragmatics.

⁶The discussion here is somewhat untechnical, so I do not commit to a position on the question of whether or not LF is an independent ‘level of representation’ in the sense of Chomsky (1986) or other formulations (i.e. Jackendoff 1997) as it does not impact upon the matters discussed here.

⁷It should be noted here that ‘pragmatics’ is used here in the broad sense of domain-general cognition, as it is defined in Sperber and Wilson (1986), for example.

Consider the following implementation. We may assume that the ‘pieces’ of poetic language that are relevant are well-formed phrases that can be parsed and assigned a full and independent syntactic and semantic structures, broadly similar to the concatenated chunks in Fabb’s (2009) proposal. These linguistic objects would map onto individual conceptual interpretations in the pragmatic component; for example, the word “John” would map onto a conceptual entry that corresponds to some male individual whose name is John, and the word “smokes” would map into a conceptual entry that corresponds to an act of smoking. In the pragmatics, inferencing on the basis of these pieces of evidence and the communicative context – where we assume that the speaker/writer is trying to communicate an encapsulated proposition – would produce a proposition where the entity identified as John partakes in the act identified as smoking. This proposition would not be mediated by a full LF that identifies *John* as an argument of the predicate *smokes*; rather, the proposition would be ‘figured out’ by general inferencing, just as we may figure out the same proposition upon being presented with at a picture of John holding a lit cigarette. This kind of symbolic communication occurs frequently in literature: for example, Peter Handke’s novel *The Goalkeeper’s Fear Of The Penalty* ends with a series of symbols communicating the protagonists’ heightened emotional state, and Kathy Acker’s novels interweave pictures and symbols with text regularly, using these symbols to communicate abstract meanings that are open to interpretation. These sections involve the kind of pragmatic interpretation just described.⁸ The question is whether this is adequate for explaining the interpretation of poetic language.

The formulation presented in the previous paragraph is very crude, but it is a plausible proposal, insofar as it is capable of producing the right general effect, that is, to produce propositional content from the sequences of words we find in poetic texts. Theories derived from the Pragmatics Hypothesis in (3) would be

⁸It is questionable whether the meanings that we produce from such sequences should be called “propositions” in the technical sense, but nevertheless one may still ascribe to them meanings that perform roughly the same communicative function as a linguistic proposition.

variations on this basic outline. However, such theories suffers from a number of theoretical and empirical problems. Perhaps the biggest theoretical problem is that the theory is almost entirely unconstrained and, as a result, largely uninteresting from the perspective adopted in the introduction to this thesis. The introduction set out the goal to develop a theory of poetic language that can explain the kinds of poetic language we find and the kinds that we do not find, and as such the theory must be strongly predictive and constrained. Yet it is difficult to envisage constraints on the very general operations of pragmatics that will adequately constrain the theory, and in the absence of such constraints we would be unable to control the power of the pragmatic component.

This is not to say that pragmatic inferencing is unconstrained, since much work on pragmatic theory documents constraints on inferencing, most notably the work in the Relevance Theory framework (Sperber and Wilson 1986; Carston 2002). However, the constraints on interpretation discussed by Relevance Theory are not well-suited to describing constraints on the reordering of words in poetry, since this work tends to focus on how pragmatic inferencing works at the utterance level; as such, it does not hold much promise for distinguishing different kinds of linguistic deviations, and as a result we would predict a vast array of rearrangements that do not occur. For example, we would not be able to distinguish rightward displacement of an auxiliary verb and leftward displacement of an object, since both are simple displacements that can plausibly be ‘put back together’ by pragmatic reasoning; yet the former kind of displacement is not attested, while the latter is extremely common. Similarly we would expect pragmatic enrichment to be able to fill in a much wider range of contextually plausible meanings for missing strings of words,⁹ thus predicting much more

⁹The fact that the pragmatics only has a minor role to play in (well-formed) ellipsis resolution was one of the central insights of Hankamer and Sag’s (1976) seminal study. They observed that VP-ellipsis anaphora differs from NP-anaphora in the fact that it almost always requires a linguistic antecedent, even if the context makes the action that is referred to absolutely clear:

- i. [Hankamer attempts to stuff a 9-inch ball through a 6-inch hoop]
Sag: #It’s not clear that you’ll be able to. (Hankamer and Sag 1976: 392)

I will return to the issue of antecedents for erasure later in the chapter.

widespread erasure than what is attested. One might tackle this problem by saying that the fact that so many examples do not occur is down to accident or some other unknown factors. This is much more plausible with a pragmatic explanation than a syntactic one, since the syntactic theory controls for ‘other factors’ in a way that the pragmatic theory cannot; that is, the different forms generated by a poetic grammar are born equivalent to one another,¹⁰ whereas the forms put together by pragmatic inferencing may differ in many ways. However, this defense does not get us anywhere in producing a focused and predictive theory.

To demonstrate the empirical problems that this theory suffers, I will briefly discuss a specific case that the pragmatic theory would struggle to deal with, drawing from data in section 2.2.4. Recall that there were many cases of the erasure of articles in the more experimental texts, in particular the work of Robert Creeley. This sometimes caused problems for the interpretation of nominals that would normally take articles in English. However, recall also that missing articles did not always create the ambiguities that we might expect if the meanings of the deviant sentences were recovered solely by pragmatic means. Consider the following:

- (4) Walked to past now dream
of __ previous place was about to
get all the confusions at last
resolved when he then woke up.

Robert Creeley, ‘In London’

The NP *previous place* is ill-formed in this context, because it is singular and missing an article/determiner, and English singular nouns of this kind occur with an article like indefinite *a*. In a wholly neutral context, then, this example could have two sources of deviation: either it is a singular NP which has had its article deleted, or it is a plural NP that is missing plural inflection. Only the

¹⁰This is not strictly true, as section 4.2.7 discusses the potential issue of competition between different kinds of competition. Nevertheless the point remains that explaining poetic language as the output of the grammar is much more black-and-white, in that these theories make clear predictions about what is possible and what is not, whereas pragmatic theories are much less clear.

former interpretation is judged to be available in this case, however. Moreover, a third possible interpretation, where *place* is interpreted as a mass abstract noun like *experience*, is also missing, even though *dream of previous experience* would be a perfectly well-formed usage, especially in the context of the predicate *dream*. The theory should aim to explain this fact, but it seems unlikely that a pragmatic theory would offer an explanation or indeed the correct prediction. In this case, inferencing based on the bare NP in its full context would make the plural/mass interpretation slightly more likely, since the conceptual content of the NP easily maps onto a plural meaning and the sentence that follows refers to “all the confusions,” which may seem like a reference to the “previous places” denoted by the plural-interpreted NP.

What we see in this case is the linguistic material of the deviant sentence constraining the interpretation of the sentence much more than it would do if the meaning of the sentence was derived by simple inferencing. There are many more such cases to be found, since the mechanisms of pragmatic inferencing are so powerful. In addition to these cases of overgeneration, we can also find evidence for deviant sentences exhibiting behaviour typical of ‘LF representations’ or sentences that correspond to such representations, involving scopal interactions, binding relations and so on. Below is one such example, which was discussed as an example of a ‘metrical inversion’ in the previous chapter:

- (5) But still this world, so fitted for the knave,
 Contents us not.—A better shall we have?

In the deviant question *A better shall we have?* there are two scope-taking elements, the modal *shall* and the displaced object indefinite *a better*. These elements have been reordered by displacement of the indefinite, but nevertheless the indefinite takes scope under the quasi-epistemic modal verb;¹¹ the interpretation corresponding to “shall it be the case that we will have a better (one)” (narrow scope for the indefinite) is available, while the interpretation corre-

¹¹I say “quasi-epistemic” since it seems that the interpretation here should be an epistemic one, whereas some authors (such as Copley 2004) have denied that *should* and *shall* can have epistemic interpretations. See Kratzer (1981) for criticism of the hard-and-fast categorization of modals.

sponding to “is there a particular better (one) that we shall have” (wide scope for the indefinite). I will not go into the details of why this should be the case (see von Stechow and Iatridou 2003 for discussion of the curiosities of epistemic modals), but it is clear that such restrictions on scopal interpretations are caused by conditions that apply to LF representations. Such a restriction would not necessarily apply, however, if the meaning of this sentence were pieced together solely by pragmatic reasoning, since this bypasses LF and its curiosities and cuts straight to a set of relatively unconstrained possible propositions. The lack of ambiguity with deviant examples like (5) indicates that the meaning of the deviant sentence is mediated by an LF representation.

There is more to be said about the possible merits and problems of a pragmatic theory of poetic language, but for now I will put such a theory to one side. In doing so, I reject the extreme interpretation of the Non-Uniformity Hypothesis that is represented by the Pragmatics Hypothesis. I thus revise the Non-Uniformity Hypothesis to exclude such extreme interpretations and implicate the necessity of the system generating *sentences*, which necessarily correspond to LF representations:¹²

- (6) The (revised) Non-Uniformity Hypothesis: the sentences of poetic language sentences are not derived in the same way as the sentences of ordinary language.

With this in place, we can proceed to formulate a theory based on this Hypothesis that is capable of handling the challenges identified.

4.2 The initial proposal: Bypass Theory

If we adopt the (revised) Non-Uniformity Hypothesis in (6), we must then propose a way to produce poetic language sentences and clarify a number of issues.

In particular we must clarify how this differs from the way in which ordinary

¹²This may seem to indicate that the *LF representations* of poetic language sentences are produced by non-standard means. I discuss this in section 4.2.5 for displacement; however it will become apparent in sections 4.3.2-4.3.3 that not all poetic language sentences are like this.

language sentences are produced and how it is possible to generate sentences by different means. This section addresses these issues, introducing the core proposal and discussing direct empirical predictions made by the different sub-parts of the proposal in turn. Note that the proposal here is designed to deal with the cases of displacement; in section 4.3 the proposal will be amended to deal with the other cases as well.

4.2.1 Theoretical context

To put the proposal in context, let us first consider the ways in which poetic language differs from ordinary language. The data surveyed in chapter 2 was separated into two main categories, displacement and erasure phenomena. Displacement involves words and phrases appearing in positions that they do not normally occur in. It is a word order phenomenon, and its description will vary from language to language, since different languages often have different word orders. Erasure involves words being interpreted but not pronounced in certain positions; we observed that most examples of erasure could be described as ‘failed ellipsis,’ since it resembles the natural language process for omitting words, failing only due to the non-satisfaction of ‘licensing’ conditions.¹³

What unites these phenomena? In the contemporary framework assumed here, they may all be plausibly analysed as broadly phonological phenomena: that is, matters that are assessed and derived in the mapping to PF. This is perhaps most obvious in the case of erasure, where the only difference between the well-formed sentence and the deviant poetic sentence is the non-pronunciation of a given lexical item. As mentioned above, the problematic cases of erasure seem to be like cases of ellipsis that fail to satisfy the so-called ‘licensing conditions’ on ellipsis, and these conditions are typically taken to be phonologically determined to a large degree (i.e. Lobeck 1995, Thoms to appear, van Craenenbroeck 2010). Characterizing displacement as a largely phonological

¹³Some of the examples in section 2.2.3 were not described as licensing failures, but rather as failures of some more general antecedence relation. This will be clarified and corrected in section 4.3.3.

phenomenon may be less obvious than with erasure, but nevertheless there is theoretical impetus for such a characterization in Minimalism. First, there is a growing literature on instances of movement which occur ‘at PF,’ that is, after spellout on the way to PF; this kind of analysis has been used to describe head movement (Boeckx and Stjepanović 2001, Chomsky 2000, Brody 2000, Adger 2006), ‘lowering’ operations (Embick and Noyer 2001), and even some instances of phrasal movement (Aoun and Benmamoun 1998, Sauerland and Elbourne 2003, Ertshik-Shir 2005, Agbayani et al 2009). Second, a number of theories have argued that movement is determined in large part by conditions on linearization of chains; most such theories have been in the broad framework of Kayne (1994) (i.e. Nunes 2004, Fox and Pesetsky 2005, Wallenberg 2009), or in direct reaction to Kayne’s proposals (M. Richards 2004), and all of this work assumes that the well-formedness of movement dependencies and word order are largely conditioned by a linearization algorithm that applies in the mapping to PF. Third, studies of ellipsis have shown that some kinds of deviant movement are “salvaged by PF deletion,” indicating that the well-formedness conditions on these operations are PF conditions: see Merchant (2001) on “PF islands;” Lasnik (1999) on “overt object shift” in English pseudogapping; Merchant (2001), N. Richards (2001), Lasnik (2007), Park and Kang (2007), Toosarvandani (2008), Thoms (to appear) on variations in *wh*-movement (multiple- in single-*wh* movement and single in *wh*-in-situ movement); Merchant (2001), van Craenenbroeck and den Dikken (2006) on the Extended Projection Principle (though see Lasnik and Park 2003, Lasnik 2010); All of this work suggests that characterizing the conditions on overt movement (and hence the ungrammaticality of what we called displacement in chapter 2) as ‘PF phenomena’ may be correct, at least in part.

4.2.2 Proposal

We can assume, then, that the ungrammaticality phenomena we find in poetic texts may be described as “PF ungrammaticality.” This notion of PF ungram-

maticity is explained in more detail later in the chapter, but for now I will continue without further elaboration. Given this, the theory of poetic language needs to be constructed so that the need to generate well-formed and unified PFs is somehow ‘bypassed,’ while still ensuring that the system can generate well-formed LFs from the poetic language text. Therefore the proposal is as follows. To create a poetic language sentence, we begin with a Numeration, which contains the basic lexical items to be used in the derivation. Some of the lexical items are merged into phrases in the workspace of narrow syntax,¹⁴ by standard means; this means that the phrases formed are all well-formed in terms of internal structure. The phrases and words are then spelled out as individual items, each mapping onto individual LF and PF representations; we may call the separate LF-PF pairs ‘Pieces’ (I capitalise this term throughout for clarity). The PF representations undergo a process of ‘PF Concatenation,’ which involves random combination of the elements into linear objects, which may correspond to poetic lines; this process is the same as the concatenation process proposed in Fabb (2009). The LF representations undergo a process of ‘LF Combination,’ where the separate parts are combined randomly to form larger LF representations; if a given combination produces a single well-formed LF representation that corresponds to a sentence, this combination provides the meaning of the poetic language sentence.

Under this proposal, the separate Pieces are not combined together in the narrow syntax, where Merge normally combines lexical items and phrases into a single syntactic structure which is linearized as a single unit at PF. Rather, the Pieces are linearized independently, and thus their order with respect to one another is determined only by the random order of PF Concatenation, which is not the same as the linearization algorithm that determines the PF representation of fully formed syntactic structures. The LFs of the independent Pieces can nevertheless be put together in the LF component by combinatory procedures, since this component makes use of the required technology in standard use (this

¹⁴From hereon I use the term ‘narrow syntax’ to refer to the stage of a syntactic derivation prior to spellout.

will be justified and demonstrated below). As a result, the production of the poetic language sentence bypasses narrow syntax, since it does not rely upon narrow syntactic operations to produce an LF. I will therefore call this proposal for generating poetic language ‘Bypass Theory,’ since it is this bypassing which separates the generation of poetic language sentences from the generation of ordinary language sentences. This is stated below for clarity:

- (7) Bypass Theory: poetic language is produced by PF Concatenation and LF Combination of independently-formed Pieces of syntactic structure, bypassing the generation of a full syntactic structure.

This theory exploits the interface-driven model of the grammar that has been developed in recent Minimalism, in particular the assumption that the narrow syntax plays a minimal role in the production of PF and LF objects. I will demonstrate how this is done by discussing a simple example, before going on to probe the theoretical roots of this proposal in more detail.

To illustrate, consider a simple example of displacement from Pope, where an object NP is not in its standard position:

- (8) Her lively looks a sprightly mind disclose, *TR2: 9-10*

First, we should clarify why this surface structure is not derived by standard operations, since this is important for what follows. In the Minimalist framework, there is no general operation ‘move α ’ that can move phrases freely and produce a variety of forms to be sorted by filters, as in GB. Instead, movement is purely derivational, occurring only when it is required for feature-checking in higher positions in a structure; the fact that movement only occurs when it needs to is derived from basic considerations of derivational economy, a guiding principle.¹⁵ English lacks a form of movement that would generate the word order in (8), because there is no feature in the VP-edge position that could attract the object and thus derive the correct word order; this is a fact about English and the VP-edge projections. If there is no such well-formed derivation for English,

¹⁵I return to this issue in section 4.2.6.

then we cannot produce an LF corresponding to the sentential meaning of (8) from this string by normal computations of the grammar. Therefore we assume that this has not been generated by standard means.

However, this string could be created by an alternative derivation, such as one of the kind outlined by Bypass Theory above. According to this theory, the sentence would be generated in the following way. The Numeration contains the lexical items $\langle her, lively, looks, disclose, a, sprightly, mind \rangle$ and the functional elements that need to be part of the Numeration for the generation of a well-formed, such as T, C and whatever other operators and morphemes are required: the ‘little letter projections’ n and v, if such a system is presupposed, event operators if they are assumed to be part of the narrow syntax, and so on. The well-formed nominals *her lively looks* and *a sprightly mind* are generated by merger of all of the relevant projections. The verb *disclose* is then merged with its v projection, without an object in the object position. The subject NP *her lively looks* is merged in Spec,vP, the standard subject position, and then the rest of the derivation for this constituent proceeds as normal to form a CP. The only problem with this structure is that the VP is missing an object argument. I assume that this is a failure to satisfy semantic conditions on selection, and that this semantic condition applies at LF; therefore the derived CP will converge at PF when it is spelled out, but further action will be required in the LF component to ensure that the given derivation creates a legitimate LF object.

This is where LF combination comes in. In the LF component, the two separate parts, the CP and the unattached NP object, are combined in random configurations by LF-merger, which is simply Merge in the LF component (the technical details of this procedure are spelled out in 4.2.5). LF-merger will slot the NP into various positions in the CP structure, but the only convergent derivation will be one where the NP is interpretable, i.e. when the NP is merged into the base position of objects in VP or some higher derived position.¹⁶ In

¹⁶If the NP is merged into a higher position like AgrOP, upon interpreting the LF the semantic component will fill empty argument positions with variables of the relevant type (e

the case of (8), this will lead to the same interpretation that would be derived from a well-formed sentence like *her lively looks disclose a sprightly mind*.

The sketched derivation skips over a number of very important details, and I will address these in turn in the following subsections. In doing so I will outline the specific predictions that the theory makes due to the proposed derivation, and I will indicate how this allows us to derive some important empirical generalizations.

4.2.3 Pieces

In the sample derivation we see that there are two Pieces: the dislocated NP and the otherwise well-formed CP. The assumption behind this is that phrases that show up in non-standard positions are always independent Pieces, and that the other constituents are generated as standard with gaps (empty complement or specifier positions) where these elements would normally occur. Having gaps in these positions may incur violations of selectional restrictions; I assume, with Pesetsky (1982) and others (e.g. Hale and Keyser 1993, 2002; see also Chomsky 2004: 111, and Heim and Kratzer 1998 on the Theta Criterion in particular), that these selectional restrictions are semantic and are assessed at LF, and that apparently syntactic selectional restriction effects can be attributed to independent principles of the grammar, like Case. This means that the constituents with missing arguments are syntactically well-formed, in a manner of speaking; they are similar to constituents with gaps in them that correspond to the gaps of movement, the only difference being that the movement cases will have traces or marked-for-deletion copies in the empty positions. The unattached element, on the other hand, may be ill-formed; for example, dislocated objects will bear uninterpretable Case features (*uCase*) that need to be checked or deleted.¹⁷ Since

in the case of the object in (8)), just as it does for uninterpreted copies in standard movement chains (i.e. Heim and Kratzer 1998), and the variables are assigned a random indices within the assignment; the only interpretable derivation will involve giving the variable the same index as the object.

¹⁷For discussion of the interpretable/uninterpretable feature split and the notation relating to *uCase*, see Pesetsky and Torrego (2001, 2004). For an articulated model of such a feature system, see Adger (2003).

these objects have not occupied a Case position at any point in the derivation so far, the uCase features on these constituents will remain unchecked, and they will need to be deleted at LF to prevent a crash. I return to this issue in section 4.2.6.

This way of conceiving of Pieces has a few consequences. One of the consequences is that we should expect chunks to be maximal pieces of syntactic structure (modulo missing arguments). This may follow from a fully articulated model of the parser, which will automatically parse the largest full structure that it can; however, I will not commit to a particular model of the parser to yield this result, and instead in section 4.2.6 I will suggest an alternative explanation for this tendency based on derivational economy. From the perspective of composition, we may say that the Piece of structure that is to be displaced is ‘put to one side’ in the derivation, perhaps by being spelled out very early, and the rest of the derivation proceeds as if it is blind to the presence of this syntactic material; therefore it creates a maximally well-formed Piece from this material, thus creating the well-formed maximal unit.

There is a technical question about how Pieces can be put to one side within the derivation. It is assumed in generative approaches that the narrow syntax is ‘blind,’ generating structure without regard to external factors beyond the interfaces. This means that, given a Numeration, we would expect the syntactic engine to proceed until it had generated the full structure; thus putting a given Piece to one side would be beyond its capabilities, under standard assumptions. To implement the Bypass Theory Piecing procedure, then, we need to explain the technology involved in allowing the separation of Pieces in the syntax; otherwise we must give up the assumption that the poetic sentences are built from individual Numerations, just like ordinary language sentences, and this would lead to a large number of problems that I will not discuss here.¹⁸ The explanation I invoke here is borrowed and adapted from McGinnis (2004) and is discussed in more detail in Thoms (2010a). McGinnis introduces the notion

¹⁸See Chomsky (2000: 100-104) for discussion of the importance of the Numeration (or ‘lexical array’) in syntax.

of the “syntactic address” to the theory in order to explain Chain Condition violations and related issues in a way that is compatible with the Copy Theory of Movement (i.e. Rizzi 1986) and Bare Phrase Structure (i.e. Chomsky 1995a). The address of a syntactic element is an index that is used to distinguish different copies in a given chain, and McGinnis (2004: 67) argues that the address of relevance is the sister of the element in the structure; thus an element in the specifier of XP will have the address [XP], a head will have as its address its complement [YP], and two elements occupying multiple specifiers in a projection will have the same address. The address is indexed onto the copy, and it is used along with the numerational index – an index that indicates identity between different occurrences (copies) of the same element from the numeration – to identify and connect different copies in a chain. Thus copies in a standard chain will have the same numerational index but different addresses.

The address locates a syntactic element within the syntactic structure, so it is sensible to assume that all elements enter the derivation without a specified address, and that Merge is an address-setting process, in part.¹⁹ For a derivation to be complete, all of the elements in the Numeration need to be assigned an address, as otherwise they are not ‘in’ the derivation in any meaningful sense; effectively this would mean that they had not been selected, and Chomsky (1995b: 225) argues that the definition of a derivation requires that all of the elements in the derivation are selected. I assume that providing the Piece with an initial address is a “costless” operation, in Chomsky’s terms. Chomsky (1995b: 226) states that operations are costless if they are required for the definition of a derivation; all other operations that only ensure convergence are “costly,” and they count when it comes to the application of economy metrics. I return to the issues of economy and syntactic addresses in section 4.2.5 below, but for now the above characterization of the role of addresses will suffice.

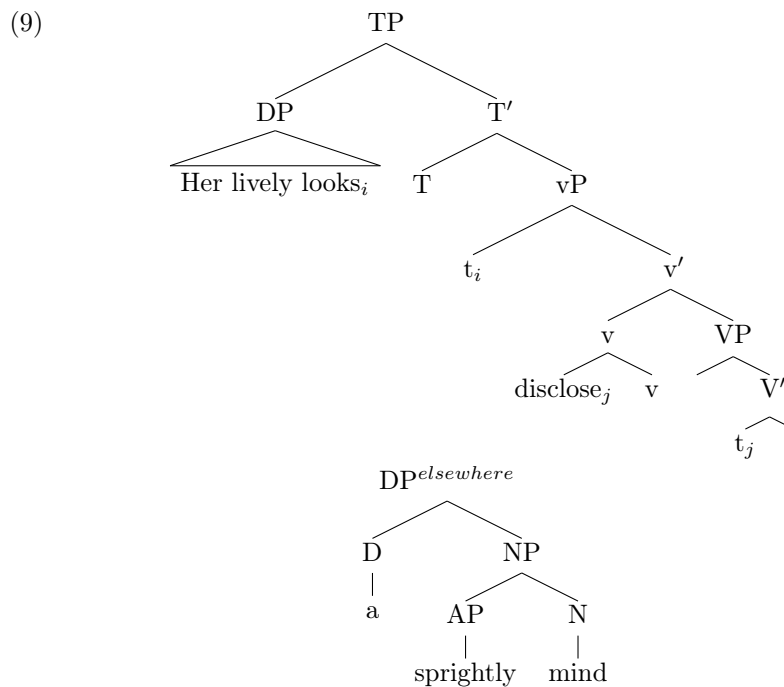
¹⁹Hornstein (2009) decomposes Merge to two sub-operations, Concatenate and Label, where Combine simply combines the two units into a sisterhood relation, and Label establishes which of the units projects to ‘label’ the unit (deciding which argument projects). In this view, Concatenate would be the address-setting component of Merge. See Thoms (2010a) and section 4.2.5 for discussion.

The fundamental characteristic of Pieces that are set to one side is that they are pre-destined to occur in some other location other than in the structure built in overt syntax. I propose that setting to one side involves assigning an [ELSEWHERE] address to the maximal projection of the Piece once it is completed in the narrow syntax workspace. This ensures that the phrase is not available for first-time Merge in the rest of the derivation, and hence that it is only available for Move – that is, ReMerge – later in the derivation; however, since the phrase does not check any other features further in the structure (this is why it is not eligible to move in the first place), derivational economy dictates that the phrase is not integrated into the structure in the narrow syntax. As far as the blind derivation is concerned, the Piece has been assigned an address, so it need not be merged into the full structure, and hence the derivation can freely proceed to spellout, where the two units are passed on as separate Pieces. I assume that the [ELSEWHERE] address can be assigned freely in the derivation; whether or not such an addressing will lead to a convergent derivation is not important, as convergence is decided at the interfaces. Given this, the generation of Pieces is a wholly regular narrow syntactic operation, rather than anything specific to the theory poetic language. This is a welcome result, since Minimalist principles dictate that the narrow syntax should be maximally simple, and that variation should be due to factors that are realized at the interfaces.

An additional technical detail of this analysis needs to be clarified before we can move on. In generative syntax, derivations are often represented in terms of tree structures. These tree structures diagram the information that is present in a derivation. In the model described above, the relevant information is the syntactic address, projection preferences (labelling), the numerical index and the lexical item; all of this is represented in a standard tree. Given the conventions of how syntactic relations are typically diagrammed, we may think of a single derivation as being exhausted by a single syntactic tree: if a given item is part of the derivation, it must be in the tree somewhere.²⁰ This

²⁰This seems to be a real theoretical motivation for the approaches to parentheticals that advocate integrating them in the syntactic structure; see section 4.2.4 below for discussion.

makes the role of Pieces in a syntactic derivation somewhat mysterious, since they are not technically integrated into the single tree structure that has been constructed. Pieces are in the derivation, but they are not assigned an address that corresponds to a place in the unified tree structure; thus an approximate tree structure representation for (8) at spell out could look like this:



The fact that this does not resemble a standard unified syntactic tree is not a serious issue, however, as trees are just one way of diagramming the relations that syntactic operations construct between syntactic elements. Chomsky (1995b) argues that Merge combines two elements to form an unordered set; this unordered set is typically diagrammed as a tree, but it is not necessary. Address-setting is a sub-operation of Merge, so we may assume that it always adds elements to the set that represents the derivation; in effect, this would perform the function of the operation Select, which Chomsky argues is necessary for taking an element from the Numeration to the derivation. Whether or not the elements in the unordered set correspond to a single tree depends on

Some other works in theoretical syntax make explicit reference to the notion of the ‘root of the tree’ in derivations, and indeed in Epstein et al (2009) it is proposed that derivations with no single root are “halted” (though it is not explained why this should be so).

the information that is read from the address indices and labels; in the case of Pieces that are [ELSEWHERE], that information does not determine a place that corresponds to a tree structure position. I conclude that the tree structure issue is just a notational issue and that it does not cause problems for the theory proposed here.

We can see that the notion of the address allows us to explain the existence of Pieces in a derivational model of syntax without unwelcome stipulation. That the operation Address should be a sub-part of Merge is argued for by McGinnis (2004) and Thoms (2010a), where it is shown that this allows us to explain some significant differences between overt and covert movement. I return to these issues in section 4.2.5, where the address is exploited further to explain the properties of LF Combination.

4.2.4 PF Concatenation

Here I will describe the theoretical context and technical details of PF Concatenation as it is used in the present theory. This notion will be refined and updated in section 4.3, where I will discuss how the theory deals with the cases of erasure discussed in chapter 2, but this update will not affect any of the technical details in the present subsection.

The idea of PF Concatenation is borrowed from Fabb (2009). It involves taking the separate Pieces that have been sent to PF by the narrow syntax and combining them in a random order into a single PF unit that can be pronounced as a single sound stream like an ordinary sentence. In the case of verse, the reordering that is chosen will be one that also produces a metrically well-formed or rhyming line, but this is not always necessarily the case, as there is no direct causal relation between PF Concatenation and other modes of organization. This is what is suggested by Fabb (2009), who takes the extra conditions of literary form to be a further set of ‘filters’ that act upon the poetic language forms generated by concatenation. I return to this issue below.

PF Concatenation is taken to be a combinatory operation that occurs in the

PF branch of the derivation. As mentioned in section (4.2.1), there is a great deal of precedent for combinatory operations that occur in the PF branch (see the references cited in section 4.2.1). Most of the cases dealt with are cases of movement operations that do not affect interpretation and which therefore seem to be only phonological; this includes verb movement, which doesn't have a obvious effect on interpretation (though see Lechner 2006, Hartman to appear for contentions), and kinds of phrasal movement that obligatorily reconstruct for scope and binding. If we assume that at least some of these proposals for PF movement are valid, it follows that UG makes available some set of operations that (re)merge linguistic units in the PF branch. This gives some justification to the proposing the operation of PF Concatenation, which we may take to be an example of such an operation. The important thing is that the one major difference between poetic language sentences and their ordinary language equivalents is phonological form. Given this, we might assume that PF Concatenation is the only major difference between ordinary language and poetic language. I will adopt this assumption for now, and in section 4.2.6 I will detail how adopting this allows us to explain the fact that poetic language sentences are judged as ungrammatical in most standard situations. In section 4.3 this will be refined and generalized.

Despite the name, PF Concatenation need not line up units one after another, but rather it can place one Piece within another in its sequence; this is what we see in (8) above, where the Piece *a sprightly mind* is placed within the larger Piece *her lively looks disclose*. This is to be expected, since similar interspersing of separate 'Pieces' occurs in standard language use too, such as with parenthetical elements, demonstrated below in (10)-(11):

- (10) Her lively looks, an endowment she thanks her mother for, disclose a sprightly mind.
- (11) Her lively looks – this is what we're supposed to believe! – disclose a sprightly mind.

It is assumed that both the parenthetical and the interjection are not combined

with their host sentences in the narrow syntax, as they fail a number of tests for syntactic dependencies between merged elements,²¹ so their appearance within the pronounced forms must involve an insertion of sorts similar to that seen in PF Concatenation. The strongest assumption would be that PF Concatenation and the phenomena demonstrated by (10)-(11) (I will put these under the umbrella term of parenthesis in what follows) rely upon the same general-purpose mechanisms for combining independently-formed linguistic objects at PF; that is, parentheticals are also integrated into the single PF representation as Pieces that are PF Concatenated. I adopt this assumption, and in doing so I capture the fact that they behave as if they have not been integrated in the narrow syntax by standard Merge operations; as such this account resembles to a great degree the proposal in Ackema and Neeleman (2004), which proposes a ‘PF insertion’ analysis of parentheticals.²² More important for the present discussion, I also arrive at an account for the fact that PF Concatenation does not always

²¹De Vries (2007) lists ten different diagnostics that indicate that parentheticals are invisible to c-command relations and thus not combined with their hosts by standard merge. Nevertheless he rejects the idea the “radical orphanage” approaches of Haegeman (1991), Fabb (1990), Burton-Roberts (1999) and others, proposing instead that parentheticals are combined with their hosts by a special kind of merge (“b-merge”) that creates a syntactic dependency that blocks c-command. See footnote 22.

²² This requires the additional assumption that the parenthetical would be LF-Merged in an appropriate position in the structure, perhaps as an adjunct to the matrix CP; we may postulate that this is the only point in the structure where the independent unit is semantically interpretable, since it has independent force and mood (de Vries 2007). This would capture the fact that the element is not within the c-command domain of any element in the structure.

Note that this account is effectively a hybrid of the “radical orphanage” approaches of Haegeman (1991), Fabb (1990), Burton-Roberts (1999) on the one hand and the “syntactic integration” accounts of de Vries (2007), McCawley (1982), Emonds (1979) and others: the parentheticals are syntactic orphans, since they are not merged into the structure in the narrow syntax, but they are integrated into the syntactic representation in some stage in the derivation. I briefly discuss their relationship to economy and interpretation in footnote 36 below.

A benefit of this account is that we not need to invoke alternative merge operations like de Vries’ “b-merge” to account for parentheticals, avoiding potential theoretical complications raised by such an operation. One such complication is restricting the operation’s application: if b-merge is available freely like standard merge, it should be available to combine the adjunct with the matrix sentence as an “invisible constituent” in (i), thus avoiding a Condition C violation:

- i. *He_i liked to play tennis [when John_i was younger].

The fact that such examples are ungrammatical indicates the application of b-merge is not available freely. However, given that the syntax is typically assumed to be blind, it is difficult to see how one could formulate a narrow syntactic operation that only applies in certain situations.

lead to entirely random orderings of separate Pieces.

The mechanisms for inserting parentheticals are constrained, as it is clear that some kinds of parenthetical interruptions are far more acceptable than others. For example, they tend to be far more acceptable if they occur at phrase boundaries. Consider the cases in (12), where interruptions occur at points in the phonological string that do not correspond to phrase boundaries; all are unusual and awkward (indicated by the diacritic #).

- (12)
- a. #Her – this is what we’re supposed to believe! – lively looks disclose a sprightly mind.
 - b. #Her lively – this is what we’re supposed to believe! – looks disclose a sprightly mind.
 - c. #Her lively looks disclose a – this is what we’re supposed to believe! – sprightly mind.
 - d. #Her lively looks disclose a sprightly – this is what we’re supposed to believe! – mind.

In contrast, compare these with the following examples, where the interruptions occur at phrase boundaries.

- (13)
- a. Interruptions – this is what we’re supposed to believe! – tend to be far more acceptable if they occur at phrase boundaries.
 - b. Interruptions tend – this is what we’re supposed to believe! – to be far more acceptable if they occur at phrase boundaries.
 - c. Interruptions tend to be far more acceptable – this is what we’re supposed to believe! – if they occur at phrase boundaries.
 - d. Interruptions tend to be far more acceptable if they occur – this is what we’re supposed to believe! – at phrase boundaries.

All of these are acceptable, and the contrast quite strongly with those in (12).

The difference between (12) and (13) is to be expected if we assume that the operation that combines parentheticals and their hosts occurs at PF. PF operations necessarily manipulate PF representations, and these representations are interpretations of the syntactic structure produced at spellout; importantly, there are mismatches of sorts between the constituency of PF representations and syntactic representations (see Selkirk 1984, 1995). The former is sorted into phonological phrases, which correspond to syntactic phrase boundaries,

and then the phonological phrases are sorted into intonational phrases and subsequently utterances; however the grouping into intonational phrases does not necessarily correspond to the recursive syntactic phrase structure, hence the mismatches between the different structures. These PF structures are subject to well-formedness conditions, and many of the proposed PF operations (in particular those involving clitic clusters) involve movements which lead to the satisfaction of those well-formedness conditions. I assume, then, that the PF operation that inserts parentheticals is subject to a well-formedness condition like the following:

- (14) Parentheticals must be inserted at a phonological phrase boundary.

This condition is perfectly natural and is assumed implicitly in much work on parentheticals; Dehé (2006, 2009) shows that parentheticals do not necessarily fall upon intonational phrase boundaries, as has been assumed in much of the literature, but all of the examples she presents conform to the condition in (14) and she does not address the issue of why this might hold.²³

With (14), and the assumption that parenthetical insertion and PF Concatenation are effectively the same operation, we are in a position to explain some facts that are important for the theory at hand. In poetic language, we do not find the ‘dislocated’ Pieces appearing at non-phrase boundaries within other Pieces, just as we do not find interruptions occurring at non-phrase boundaries. None of the examples discussed in chapter 2 involve dislocated Pieces occurring in non-phrase boundary positions, and I have come across no examples of this in the data I have surveyed. This is a major distributional trend that needs to be accounted for by the theory of poetic language, and we can do this here within the present theory by saying that PF Concatenation is subject to a version of

²³Note that this well-formedness condition needs to be stipulated for the narrow syntactic accounts of parenthesis as well as others. Consider, de Vries (2007) which is perhaps the most thoroughly worked-out of the narrow syntactic explanations of parenthesis. De Vries notes the relative freedom of parenthesis insertion and proposes that they are adjoined to any projection by a special merger operation (b-Merge) in the narrow syntax. The theory provides no restriction to prevent some of the cases in (12), since these interruptions could be generated by adjunction, so it must explain this with a condition like (14) which applies at the PF-interface.

(14). I restate this in (15) for clarity:

- (15) The PF Insertion Constraint: if one PF object is embedded within another (by parenthetical insertion of PF Concatenation), the embedded object must occur at a phonological phrase boundary.

This constraint bears a lot of the empirical burden of the present theory, since it stops PF Concatenation – a potentially very powerful operation – from overgenerating wildly. Speaking informally, we may say that PF Concatenation buys the theory a great degree of randomness, accounting for a lot of the diversity of poetic language data, but that the constraint in (15) helps us to control that randomness. The only innovation that is required, however, is for us to posit the operation of PF Concatenation, one that follows from the model of syntax adopted here and which can be motivated independently in the analysis of parentheticals.

4.2.5 LF Combination

While PF Concatenation ensures that the pieces of structure that are generated separately in the derivation are eventually put together to form a single phonological (\approx graphological) unit, LF Combination ensures that the units combine together to produce a well-formed LF representation that maps onto a sentential meaning. As with PF Concatenation, I assume that the operations that allow for LF Combination are general and do not require additional technology. The specific explanation given here depends again upon the notion of the ‘syntactic address’ introduced above (and is heavily indebted to the proposals in Thoms 2010a), although alternative formulations could conceivably achieve similar results.

Recall from section 4.2.3 that some Pieces are constructed in the narrow syntax with missing arguments. This requires the assumption (mentioned earlier) that selectional restrictions are semantic, rather than in the narrow syntax, and that semantic conditions like the Theta Criterion are either interface conditions or collapsible to the more general operations of the semantic component (see the

references cited at the beginning of section 4.2.3); otherwise the derived Pieces would be semantically ill-formed and the theory would require a set of unusual and perhaps unwelcome ‘repair operations’ in order to compose the separate Pieces into well-formed LFs.

How, then, can we produce a well-formed LF from separate Pieces if they have not been integrated in narrow syntax? I propose that the separate Pieces are integrated in the covert cycle (‘at LF’) by the application of LF-Merge.²⁴ In the derivation of an interpretable poetic language sentence, a given Piece will be merged into a position where it would normally receive its interpretation; for example, an object Piece will be merged into the object position (i.e. into the specifier or complement position that is projected for the argument). This requires us to reject the Strict Cycle Condition of Chomsky (1995b), and its updated form, the Extension Condition (Chomsky 2000, 2001, 2004), since it involves merging pieces of structure that do not extend the root; precedent for rejecting this comes from N. Richards (2001) and Nissenbaum (2000) among others, which show that “tucking in” derivations in *wh*-movement provide direct empirical arguments against such a model of the syntax.²⁵ If we reject the Extension Condition, nothing prevents us from (Re)Merging constituents into argument positions in this manner in the LF cycle. In effect, LF-Merging a constituent into an argument position in this manner is equivalent to ReMerging a *wh*-phrase into a specifier position (phase edge) by using standard covert movement operations, or indeed ReMerging a quantifier into an adjunction position

²⁴The covert cycle here is the component of the syntactic derivation that follows spellout, affecting only LF structure. This requires us to assume an architecture of the grammar broadly similar to the ‘Y-model’ of Chomsky and Lasnik (1977), and to reject modern alternatives to that model, such as Bobaljik’s (2002) ‘Single Output Syntax’ model (and indeed models like it, such as that proposed by Chomsky 2008). In Thoms (2010a) I present a number of empirical arguments for adopting a modern version of the Y-model in favour of Bobaljik’s system, and I show that Bobaljik’s arguments for Single Output Syntax (specifically his identification of ‘lookahead’ problems for the Y-model) are not convincing. I therefore adopt the Y-model for the theory developed here.

²⁵There are ways to model the present system without rejecting the Extension Condition; for example, we may propose that the relevant arguments are merged into specifier positions in a phase-by-phase derivation; the missing argument positions would then be filled by variables coindexed with the relevant argument in the semantic component, much in the same manner that is required for uninterpreted base copies in certain chains (i.e. quantifiers in positions where they encounter a type mismatch).

by Quantifier Raising (cf. May 1985). Many details of this proposal need to be explained, and I will proceed to do that in what follows.

First, what is LF-Merge, and is it any different from standard Merge? Answering this requires an excursus into the definition of Merge, as the theory proposed here depends on a specific view of the basic operations. I assume that the definition of Merge in Chomsky (1995b: 243-246): informally, Merge combines two syntactic elements to form a single syntactic element, in which one of the elements projects to label the new larger element. Merge is typically taken to be a basic operation in narrow syntax. However, I follow Hornstein (2009) in assuming that Merge can be decomposed into two separate basic operations. Hornstein proposes that the operation typically called Merge can be decomposed into the operations Concatenate and Label; Concatenate takes two elements and combines them into one (pairing them as sisters), and Label decides which of the two Concatenated parts projects as the label of the new element. Hornstein proposes that some apparent instances of Merge actually involve only one of the sub-operations, and in doing so he produces explanations for a number of syntactic phenomena, based on the full definitions of the sub-operations.

I follow Hornstein's lead in decomposing Merge, although the terms used here are slightly different; in particular, I exploit the notion of the syntactic address introduced above, and argue that it should be given a crucial role in defining the basic operations of a syntactic derivation.²⁶ Recall from section 4.2.3 that McGinnis (2004) takes the address to be an index that is added to copies in movement chains in order to distinguish the different copies in the chain: different copies have different addresses. The address is taken to be the sister: so, for an element X in the specifier of a projection of Y, [Y] is the address of X. Above I extended this definition somewhat, arguing that *all* elements in a derivation receive an address index, not just elements in non-trivial chains. Specifically, I proposed that Merge is an address-setting operation in part: when we Merge two items X and Y together from the Numeration, what we are doing

²⁶The discussion that follows here overlaps significantly with Thoms (2010a).

is saying that X has the address [Y] and Y has the address [X], since Merge of X and Y makes them sisters of one another. We may say, then, that there is a sub-operation of Merge which sets the address of the combined elements, and I will call this sub-operation *Address*. This corresponds to Concatenate in Hornstein’s decomposition of Merge; here I will continue to use the term Address to distinguish Hornstein’s theory from mine, although I believe they are largely compatible. I assume, following Hornstein, that Merge requires another sub-operation, corresponding to his operation Label, and I will use the same term in what follows. In the standard case, Merge involves both Address and Label.

I follow Chomsky (1995b) in assuming that Merge is a “costless” operation in the definition of economy metrics. Chomsky (1995b: 225-226) argues that different operations in the syntax can be distinguished as “costless” or “costly” for the calculation of derivational economy depending on whether or not they are required for the definition of a derivation: operations that are required to form a derivation are costless, while operations that are only required to ensure convergence are costly. I assume that the sub-operation Address is required for the definition of a derivation, since producing a derivation requires selecting and combining elements from the Numeration, and when we combine elements, we assign Addresses; without any address, an element cannot be said to have entered into the derivation. I assume that Label is also required for a derivation, although I will not dwell on this interesting matter. We can see, then, that Merge is a costless operation.

With this in place, consider the case of ReMerge, one of the names given to Move in recent Minimalist theories. As the name ReMerge involves taking a sub-part of the derivation and Merging it again. ReMerge is never required for the definition of a derivation, since all of the sub-parts have already been assigned addresses; rather, ReMerge is only ever required to ensure convergence (by checking features, creating an interpretable configuration, etc). Nevertheless, recall that McGinnis’ initial motivation for introducing the notion of the address was to distinguish different copies of a given syntactic element in a

chain. Altering McGinnis' proposal slightly, I propose that ReMerge involves 'updating' the address of a given element, adding the new address to the copy alongside its old one, in the manner of a passport stamp; thus if an element X is Merged with Y and then ReMerged later with Z , X will have the address $[Z, Y]$. This involves another instance of the operation Address. However, whether this instance of ReMerge is costless depends on its contribution to the derivation. It is not required for a derivation, since all of the elements are already part of the derivation; rather, it is required in order to distinguish the different copies of the element in the chain. Why must the copies be distinguished? I propose that this is to ensure that the given derivation can be linearized at PF: Kayne (1994), Nunes (2004) and many others assume that, in order for a given derivation to be linearized, different copies of the same element need to be distinguishable from one another, as linearization depends on precedence and copies of the same element cannot precede one another. The address allows us to distinguish between copies: different copies have different addresses. I propose, then, that this is the purpose of address-setting in ReMerge: Address allows the given chain to be linearizable at PF. I propose (*pace* Nunes 2004: 166) that linearizability is not a condition for the definition of a derivation, but rather a condition for convergence, and as a result the extra instance of Address that is involved in ReMerge is costly, rather than costless.²⁷ This derives the familiar Move Over Merge constraint (Chomsky 1995b, Hornstein et al 2005) as a reflex of derivational economy: Merge will always be preferred to ReMerge since ReMerge is

²⁷ In his discussion of the definitional properties of a derivation, Chomsky (1995b: 226) assumes that the only truly costless operations are Select (the operation that selects an element from the Numeration) and Merge. Nunes (2004: 166) gestures towards Chomsky's suggestion that a computation must reach the pair (π, λ) to be a derivation, where π is a set of instructions to the articulatory-performance system (PF) and λ is a set of instructions to the conceptual-intentional system (LF) (Chomsky 1995b: 219); Nunes takes this statement to be an indication that an unlinearizable PF representation does not meet this basic requirement. I do not agree with this, however, since an unlinearizable PF is still a "set of instructions" to the articulatory-performance system; the fact that this is a contradictory set of instructions is besides the point. Indeed Chomsky acknowledges this possibility, saying that the "legitimate objects" that are interpreted at the interfaces may be interpreted "perhaps as gibberish" (Chomsky 1995b: 219). Contradictory pronunciation instructions would be a clear example of this gibberish. I therefore reject Nunes' claim that linearizability is a condition for the definition of a derivation; rather, it is a condition for convergence at the interface. See also section 4.3.3 and Thoms (to appear) for relevant discussion.

costly while Merge is not. More generally, we derive the fact that ReMerge is costly.

This is not the full picture. ReMerge is costly because of the extra instance of the sub-operation Address (I assume that there are no other differences between Merge and ReMerge) and this extra instance of Address is required for convergence at PF. However, there are instances of ReMerge that do not care about PF convergence: ReMerge in the LF branch of the derivation. ReMerge in this component does not need to update the address of the constituent, since the address-updating is only required for PF-convergence; this is not an issue for LF-Merge, since it is post-spellout. Given derivational economy, we may assume that since the additional instance of Address is not required for LF ReMerge, it is therefore not possible with this kind of movement due to economy. We can therefore conclude that LF-Merge is costless, unlike ReMerge in the overt syntax.²⁸

With this background, we can now return to the issue of LF-Merge in the derivation of poetic language sentences. Recall that Pieces are ‘put to one side’ in the narrow syntax by setting the address of the given Piece to [ELSEWHERE]; this ensures that the Piece is part of the derivation (to adopt Chomsky’s terms, they have been Selected), but it also ensures that it is not Merged into the standard position where it is selected,²⁹ since this would require an instance of ReMerge; specifically, we can say that we would require an additional instance of Address to add the Piece to its standard position in the narrow syntax, and

²⁸As noted by Thoms (2010a), this solves the ‘trigger problem’ for LF movement operations like QR, since QR no longer needs to be motivated by morphosyntactic feature-checking. I also argue there that this model can explain two significant asymmetries between overt and covert movement: (i) the fact that overt movement can bleed binding relations while covert movement cannot, and (ii) the fact that overt movement is subject to anti-locality (Abels 2003) while covert movement is not. As mentioned above, these empirical arguments militate against the Single Output Syntax model and in favour of the classical Y-model, and they also provide support for the decomposition of merge into Address and Label.

²⁹Here and onwards I talk in terms of selected arguments, but a similar logic can be extended to adjuncts. VP-adjuncts that show up in non-VP-internal positions do not cause problems for semantic selection, but there are obviously other conditions that control the fact that VP-adjuncts are indeed restricted to appearing in the VP; for example, they may contain event variables that need to be bound by a certain event operator in the syntax. The result would be that they need to be in the VP at LF, requiring some kind of VP-adjunction via LF-Merger.

since this would be more costly than just leaving it to one side, we opt to leave it to one side. However, at LF, ReMerge of this Piece into its standard position is not costly, since LF-Merge does not require address-updating. In effect, this involves the same kind of operation as LF-Merge operations like covert wh-movement or QR, since a copy of the Piece is Merged into the ‘main structure’ to produce a single LF representation. Merging the Pieces into the correct argument positions will produce a well-formed LF and the derivation will converge at the LF interface with the same truth-conditional interpretation that we would get with the ordinary sentence.

This way of describing LF Combination has a number of consequences. One is that there is no constraint on the number of Pieces that need to be LF-Merged. If LF-Merge was a costly operation, we would predict that derivations that recombine one Piece would outcompete those that recombine two Pieces. Derivational economy would count one instance of non-feature-driven ReMerge in the former but two in the latter, and therefore it would make single displacement preferable to double displacement, at least on the assumption that the two versions would compete in these terms.³⁰ However, LF-Merge is costless in the present theory, so this predicts no such competition. This is a welcome result, as we saw in chapter 2 (especially sections 2.1.2 and 2.1.4) that double displacement was almost as common as single displacement in poetry. We saw examples of NPs being displaced alongside PPs, double PP displacement, NPs being displaced alongside APs, and so on. Most of the possible combinations of multiple phrasal displacement were found in the analysed corpus. I will return to some possible constraints on certain kinds of double displacement below, but for now it is enough to note that there is no categorical constraint on double displacement, as would be predicted if LF-Merge was a costly operation.

Another consequence of this way of formulating LF-Merger and Piecing is that the theory rules out certain kinds of ‘head displacement,’ that is, forms

³⁰Chomsky (1995b) and many others working in Minimalist syntax assume that such economy-based reference-set computations are a core feature of the grammar, so I adopt that assumption here. I return to the issue of derivational economy in section 4.2.6.

of deviant head movement where certain syntactic heads are displaced from their normal positions in the same manner that the object NP *a sprightly mind* is in (8) above. Specifically, the theory rules out LF-Merging heads into ‘in between’ positions in the syntactic structure. By an ‘in between position’ I mean a position in the syntax where a head X would normally project and both select a complement and be selected by a higher projection. To demonstrate this, consider the following potential example of a poetic language sentence:

- (16) #John heard that record has. (“John has heard that record”)

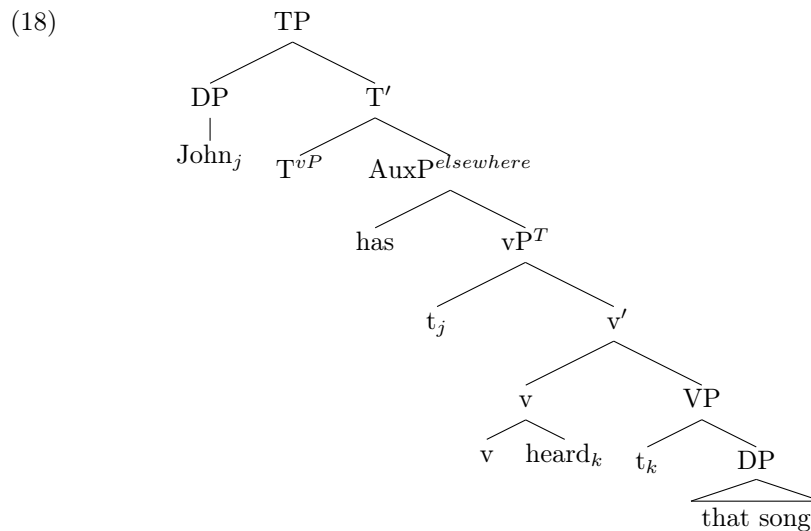
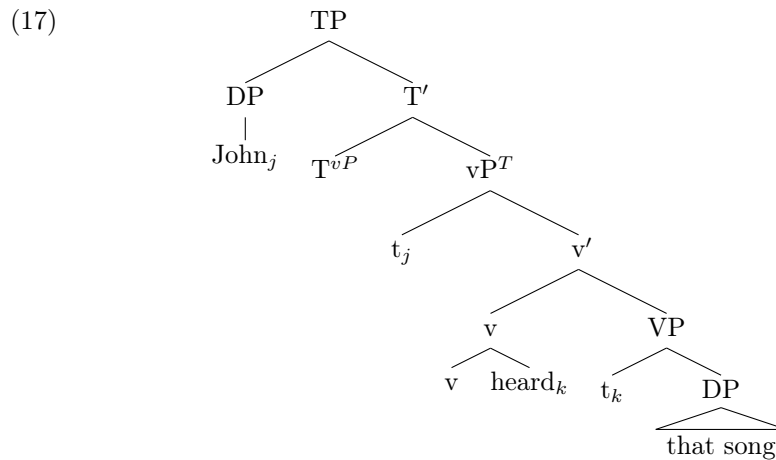
This sentence is intended to have the same meaning as ‘John has heard that record,’ and it is marked with the diacritic # because it is not an attested form in the poetic examples. As a poetic language example, it would be analysed as a case of ‘head displacement,’ where the syntactic head *has* has been set aside as a Piece and undergone PF Concatenation to occur at the end. All of this is allowed by the theory, so we should expect that the LF Combination process will help us to rule out such examples.

This is what happens. Above I proposed that LF-Merge puts the unintegrated Pieces into their standard positions in the covert cycle of the derivation; without doing this the derivation would produce a structure that fails at the semantic interface, since the selectional requirements on some part of the structure would not be met. This is simple enough with phrasal arguments, since they are simply merged into the unfilled specifier or complement position in the structure projected by the not-fully-satisfied phrasal projection (i.e. the VP). However, this is more complicated for heads, especially for heads that appear in the spine of the syntactic structure. Consider the case of the auxiliary *has* in (16). This auxiliary normally selects for a verbal complement like vP, and it is typically selected itself by the functional projections that make up the inflectional layer, like TP. In the situation in (16), this AuxP is ‘set to one side’ as a separate Piece in the narrow syntactic derivation, and its selectional restrictions remain to be satisfied; parallel to this, we generate the maximal structure for

the other Piece, which will be a full sentence: the V *heard* Merges with the DP object *that record*; this Merges with v, and then the subject *John* Merges into Spec,vP; then T Merges with vP and the subject raises to form the well-formed Piece corresponding to the sentence *John heard that record*. Note that this sentence is fully well-formed in itself, since there is no problem for selection: vP is selected by T directly in many cases. Note that the address assigned to T in the narrow syntactic derivation is [vP], and the address assigned to vP is [T], since they are sisters.

The problem occurs when we try to find a place for the unintegrated Aux head. This head only selects for the vP, so we would have to try to Merge this into an ‘inbetween position’ between the TP and vP projections; furthermore, this AuxP would have to project in order for the selectional requirements to be met. Say we allowed this to happen. We know that LF-Merge does not invoke the Address operation as standard, so it would not change any of these details in the syntactic derivation. The result of projection of the AuxP would be that the T head ‘thinks’ its address is [vP] and that it is selecting a vP sister, yet its sister is actually an AuxP projection. I diagram this in the two tree structures below; (17) demonstrates the tree prior to insertion of the AuxP, and (18) the tree afterwards; the superscripts indicate the addresses that have been added to each element in the derivation (added where appropriate).³¹

³¹Here I use the standard X-bar notation for phrase structure, but this could be redrawn in line with a Bare Phrase Structure approach without any difficulty.



I propose that this kind of inconsistency in the syntactic information encoded in the addresses of the elements will produce an uninterpretable LF representation, since the conceptual-intentional system is provided with conflicting information about the semantic content of the different nodes.³²

The result is that LF-Merging of heads into these kinds of positions is predicted by the theory to be impossible because the resulting derivations crash at the LF interface.³³ This shows that the theory is not entirely unconstrained

³²Something like this explanation may be similar to that which lies behind what Chomsky (2008) calls the 'No Tampering Condition.' Chomsky (2008) also mentions that errors in labeling and node information are generable in the syntax, and he says that such derivations fail at the interfaces, similar to the explanation here.

³³Note that this formulation does not rule out all LF-Merger of heads in all cases. It might be possible for head movement to occur at LF, so long as the head's projection is already part

in what it allows to be ‘set aside’ in the derivation, since only certain kinds of constituents can be LF-Merged without causing serious complications. The problem in the case discussed above was that the relevant element necessarily projected in the structure, and that the derivation of the rest of the structure could proceed to completion without this projection; as a result, it needed to be LF-Merged into an inbetween position. We can state this clearly as a theoretical claim about LF-Merger in (19), with the ancillary definition of an inbetween position in (20), and then we can propose the corollary claim for the proposed theory of poetic language in (21):

- (19) The LF-Merge Condition: a syntactic element X cannot be LF-Merged into an Inbetween Position between YP and ZP if Z^0 also selects YP.
- (20) A syntactic element X is in an Inbetween Position if it selects a YP and X is selected by an immediately dominating head Z^0 .
- (21) The LF-Merge Condition Corollary: syntactic elements can only be set to one side in a derivation if they do not need to be LF-Merged into an Inbetween Position.

This is another welcome result, as this kind of head displacement is also not attested in poetic language, as mentioned in section 3.3.1. We do not need to stipulate a specific ban on head displacement of the kind seen in (16), but rather it follows from the way in which the theory is formulated.

We can see that the formulation of LF-Merger has some important consequences for how the theory is constrained and what kinds of predictions it makes when it comes to the forms of poetic language that can be generated by the theory. Perhaps the most important fact about LF-Merger in this theory is that it is costless, and the outcome of this is that any problems that are caused by setting aside a certain syntactic element can be remedied for the LF side of the derivation by subsequent applications of LF-Merge after spellout. However, this then raises the question of how derivations that set aside Pieces can be distinguished from standard derivations that proceed to produce ordinary language of the ‘spine’ of the tree; in that case, its selectional requirements will have been satisfied.

sentences; after all, the processes that apply at LF do not distinguish the two different derivations, yet they clearly need to be distinguished, since the poetic language examples are ultimately ungrammatical while the ordinary language ones are not. I turn to this question in the following subsection.

4.2.6 Blocking poetic language and blocking ordinary language

Above I have described how the proposed theory allows for Pieces to be set aside in the narrow syntactic derivation, and how they can then be integrated at LF to create a unitary semantic representation and PF-Concatenated at PF to create a single object corresponding to a poetic language sentence. I argued that all of the relevant steps make use of mechanisms that are available for ordinary language: setting aside a Piece by giving it an [ELSEWHERE] address is naturally available and can be used to generate parentheticals; LF-Combination relies upon mechanisms that are readily available in ordinary derivations and which receive independent support; PF Concatenation involves the same mechanisms that puts together a parenthetical and its host. This raises an important question: if all these mechanisms are normally available, why do these sentences only show up in poetry? Or to put it another way, why are these sentences deviant? It is important that the theory make this distinction, since otherwise the theory cannot be called a theory of poetic language, but rather an unwelcome extension of the ordinary language grammar. In effect, we must ‘block’ poetic language from occurring all the time.

I propose to do this by appealing to derivational economy. Before explaining the specific proposal, it is useful to define derivational economy as it is used here. Chomsky’s Minimalist Program (Chomsky 1993 *et seq*) provides the broad framework for the generative theory of poetic language developed here, and it is driven by the proposal that the faculty of language is an optimal solution to the design problem of mapping sound to meaning. From this core proposal, there follows a number of natural assumptions about how the computational system

of language should work, and one of these is that the computational is efficient and that its operations should be minimal. Chomsky (1995b: 220) states this clearly in the following passage:

It seems that a linguistic expression of L cannot be defined just as a pair (π, λ) [“ π is a PF representation and λ an LF representation”] formed by a convergent derivation. Rather its derivation must be *optimal*, satisfying certain natural economy conditions: locality of movement, no “superfluous steps” in derivations, and so on. Less economical derivations are blocked even if they converge.

The language L thus generates three relevant sets of computations: the set D of derivations, a subset D_C of convergent derivations, and a subset D_A of admissible derivations. FI [“Full Interpretation,” the condition that all elements in the (π, λ) pairs receive an interpretation at their respective interfaces] determines D_C , and the economy conditions select D_A . . . economy conditions only hold among convergent derivations; if a derivation crashes, it does not block others. Thus, D_A is a subset of D_C .

The theoretical and empirical basis of economy conditions has been a major topic of discussion in generative linguistics in recent years (see e.g. Brody 1995, Collins 1997, Bošković 1997, Espstein et al 1998 Fox 2000, Nunes 2004; in OT, see Prince and Smolensky 1993, Grimshaw 1997, Legendre 2001). This work requires working assumptions about the relative “cost” of operations; for example, according to Chomsky (1995b) the operation Move is taken to be more costly than Merge, and based on this, Fox (1995, 2000) proposes that the movement operation Quantifier Raising is only allowed if it derives an interpretation that was not available prior to movement.³⁴ Thus Fox’s work has shown that derivations compete relative to a set intended truth-conditional interpretation (see also Adger 1994, Grimshaw 1997), as well as a specific selection of lexical items or Numeration (as assumed by Chomsky, Grimshaw and others).

Now consider the case of a poetic language sentence like (8), which I repeat here for reference:

(8) Her lively looks a sprightly mind disclose,

³⁴In Thoms (2010a,c) I propose that Fox’s empirical results can be reframed as constraints on reconstruction (the operation Reconstruct) rather than QR. Crucially, Reconstruction is not an LF-Merge operation, but rather a kind of LF-deletion operation, and as such it is not costless like LF-Merge. Fox’s insights are thus preserved.

Above I argued that this sentence was not generated by standard means; specifically, I proposed that the S-O-V order cannot be derived by standard means because there is no operation that creates such a word order in English. Why not? Movement, as a costly operation, must always be driven by feature-checking. If an element moves to check a feature, this derivation converges and the derivation belongs to the set D_C ; if there is no other way of producing the same effect with less cost – for example by directly merging an element to check the feature, moving a closer element to check the feature, etc – the derivation will also belong to the set D_A , i.e. it will be grammatical. However, if an element moves to a position where it does not check a feature, as in (8), this derivation may well belong to D_C , but it will not belong to D_A , since a derivation that missed out the movement step would map onto exactly the same truth-conditional meaning from the same Numeration, but it would be shorter and thus it would be more economical than one with the extra movement. In effect, the well-formed sentence *her lively looks disclose a sprightly mind* blocks (8), hence the difference in grammaticality.

Given this, we might propose that poetic language is simply described as the set of uneconomical but convergent derivations, which we might call D_U ; using Chomsky's terms, this would be the complement or set difference of D_C and D_A . This would be empirically inadequate, however, in light of the data survey we saw in chapter 2: some of the kinds of displacement, such as non-constituent displacement, could not be generated by this means, as the fact that non-constituents cannot be moved is a fact about how movements are defined, rather than an issue of economy. Furthermore we would run into difficulties with multiple displacement examples, since these would necessarily be less economical than single displacement examples; we would thus predict competition between these different derivations, and we know from the discussion in the previous section (and the data in chapter 2) that such competition is unlikely. Given that the goal of this dissertation is to develop a unified theory of poetic language, these empirical issues indicate that describing poetic language as uneconomical

but convergent ordinary language is not the right way to go. Rather, we seem to require the ‘putting to one side’ derivations proposed by Bypass Theory, since these derivations are perfectly capable of dealing with all of the relevant data (as will be shown in the remainder of the chapter). I will therefore reject the possibility of derivations like the one just described, where economy constraints are violated in the narrow syntactic derivation.

Nevertheless, we can still bring economy to bear on the question of how to block poetic language from occurring in all situations. Consider again the case of (8), and the comparable well-formed sentence *her lively looks disclose a sprightly mind*. Both versions start out with the same Numeration, and both map onto the same semantic interpretation; none of the displacement operations alter the meaning of the ‘original’ sentence. The narrow syntactic operations that would have built the full structure (i.e. the standard derivation) would not involve any more or any less costly syntactic operations than the narrow syntactic part of the poetic derivation, since the only difference between the two is an extra instance of the costless operation Merge (Merging the object with the verb). Both versions of the sentence also involve very similar LF-Merge operations, since the object NP that has been displaced in the poetic example is quantificational, therefore it would undergo QR to a vP-adjunction position in order to be interpretable (cf. May 1985, Heim and Kratzer 1998, Fox 2000); this is not important, however, since even if there were a disparity in the number of LF-Merge operations, this would not be relevant for the calculation of economy metrics, as LF-Merge is costless. The difference lies in the PF branch: the poetic language sentence requires the operation of PF Concatenation to integrate the separate pieces into one unit, while the standard language sentence does not. This follows from the fact that PF Concatenation is a costly operation; this is the case since it is required for convergence in a Bypassing derivation, but not for the definition of a derivation in Chomsky’s sense.³⁵ This means that the poetic

³⁵See also footnote 27 in this chapter. I propose that an unintegrated PF representation is problematic in the same way as an unlinearizable derivation, since there is no way of knowing which element precedes which when it comes to the performance of the articulatory-perceptual system (i.e. actual pronunciation). PF Concatenation fixes this by integrating them into one

language derivation is eventually a more costly derivation, and as a result it is blocked by the standard derivation, which does not require this extra operation to be linearized. This is perhaps to be expected, since PF Concatenation is the only operation that is involved in poetic language derivations that isn't normally involved in other standard derivations; thus it has proven that this extra operation is the one that makes the difference between ordinary language and poetic language.³⁶

We can see that Bypassing Theory follows exactly from Minimalist assumptions. The core assumption of the Minimalist approach is that the syntax is the optimal system for matching sound to meaning. The syntax does this by combining a full set of lexical items by way of its core set of operations, i.e. Merge, ReMerge, Select. It follows, then, that producing a match between sound and meaning by any other means – for example by the poetic language derivations sketched above, where not all the lexical items are integrated in the syntax – should result in sub-optimal and hence dispreferred derivations. Most Minimalist theories of syntax would not rule out such derivations (this would require unwelcome stipulation), but they would predict that these derivations should be ill-formed and beholden to external factors not related to narrow syntax. If the faculty of language was not the optimal solution to the sound-meaning

linearizable unit, and as such it is an operation that creates convergence; as such, it is costly, in Chomsky's terms.

³⁶ Recall that in section 4.2.4 I argued that PF Concatenation may also be involved in inserting parentheticals into sentences. We may wonder, then, whether this runs into a problem with respect to economy; that is, whether derivations that produce parenthetical-containing sentences might sometimes be blocked by derivations that integrate the given element into the host sentence. Two relevant cases would be with restrictive/non-restrictive relatives pairs like (i)-(ii), and with parenthetical adjectives like (iii)-(iv):

- i. The sprinter, who is Jamaican, will win the race.
- ii. The sprinter who is Jamaican will win the race.
- iii. The dog, black, jumped back into the pond.
- iv. The black dog jumped back into the pond.

It is well-known, however, that these pairs do not share the same meanings, as indicated by the different names for restrictive and non-restrictive relative clauses (see Fabb 1990, Haegeman 1988; see also the discussion in section 2.1.3 relating to iii-iv). Since they do not have the same meanings, they would not compete in terms of economy, and as a result we would not expect the non-parenthetical versions to block the parenthetical versions.

problem, we might expect that the poetic language derivations may be on a par with, or perhaps even better than those produced by the regular operations of the faculty of language.

There remains a question of how these derivations are allowed to surface at all; that is, why don't we always just produce the well-formed sentences instead? The pre-theoretical description of how this is done is simple: it is the speaker's choice to produce a poetic language sentence, rather than a well-formed ordinary sentence. We cannot impose the volitional notion of choice upon the computational system, but we can add it as an interface condition of sorts. Recall that whether or not a given derivation is allowed ahead of others (ie. is in D_A) is relative to at least two conditions: we start with the same Numeration (see Chomsky 2000, 2004, 2008) and we end up with the same expected result, i.e. a given truth-conditional meaning. We may call the first of these the Numeration Condition, and the second the Expectation Condition. These are not interface conditions in the technical sense but global conditions that must hold in order for economy metrics to be calculated. It is obvious that the Numeration Condition should hold in the case of poetic language derivations, so I propose that the difference between ordinary language and poetic language has its root in whether or not the Expectation Condition is satisfied. Specifically, I propose that, when a speaker generates a poetic language sentence, she starts with the 'output expectation' that the form of the utterance be one that has been distorted by PF Concatenation, in addition to the output expectation of a given truth-conditional sentence. This seems intuitively correct, since producing a poetic language sentence seems to involve putting the pieces of the sentence together in an unusual fashion (an intuition captured by Fabb's (2009) theory); we thus generate the sentence with the expectation that it should be disordered in this fashion. The present proposal also explains why the poetic language sentences are clearly bad as ordinary language but fine as poetry: in one context we begin with the standard output expectation, and in the other we have an additional output expectation of some kind of disorder.

This expectation of disorder is not stated in structural terms, as it is external to the grammar's computations; rather, it may be regarded as a meta-linguistic condition on the output of the grammar, selecting a particular formal object based on resemblance to other objects identified as "deviant." The notion of "form" here is thus a non-linguistic one, but rather a conceptual one.³⁷

A bolder theory might propose that the additional output expectation that is set for poetic language isn't just for 'disorder' or PF Concatenation as stated above, but rather for a particular kind of phonological form, such as one that will adequately map onto a particular metrical structure; that is, the output expectation could be a metrical grid corresponding to iambic pentameter. This would provide a way of modeling an indirect relation between metre and syntax without falling prey to some of the problems experienced by other theories discussed in the previous chapter, and it may also be able to explain the indirect relation between syntax and other kinds of formal organization, like rhyme and alliteration. In effect, this would be broadly similar to the approach suggested by Fabb (2009), which proposes that metre, rhyme and other such poetic form conditions apply as 'filters' on the output of Concatenation. I will not commit to this bolder theory now, as the intention here is to construct a general theory of poetic language, rather than one exclusive to metrical verse. However, I will demonstrate below that this may have some promise in explaining some interesting lacunae in the distribution of types of displacement in the poetic texts surveyed in chapter 2. For now, the important thing is that the implementation would involve articulating and expanding upon the Expectation Condition as it is formulated here.

The result of this way of thinking is that the well-formed sentence and the poetic version do not compete in terms of derivational economy, since they have different output expectations; they cannot compete for economy any more than

³⁷For discussion of form as a kind of content, see Fabb (2002, 2004), which argues that certain aspects of form in poetic texts, such as "iambic pentameter" or "line," are understood as conceptual content by "interpretive resemblance" (as in Sperber and Wilson 1986). Thus a line of iambic pentameter verse communicates the contentful predicate "in iambic pentameter" not by virtue of some fundamental link between the formal properties of the metre and content, but by the fact that the given line resembles other objects identified as "in iambic pentameter."

a pair of sentences with different truth-conditional meanings can compete with each other. Thus a generated poetic language sentence is technically well-formed as a poetic language sentence, even though it would be ill-formed as an ordinary language sentence. For clarity, I state this as the proposal in (22) below:

- (22) Poetic language is required to satisfy an additional output expectation which dictates that the form of the sentence should be distorted.

This proposal is somewhat crude, and the notion of ‘output expectation’ used here is not fully defined formally, but it captures the core intuitions and allows us to explain how poetic language is ever allowed to surface. I adopt it tentatively, as I believe that it is not an essential part of the theory; after all, one might simply ignore the apparent problem of ordinary language blocking poetic language by assuming that blocking has no effect on the kinds of forms that can be generated by a poet for use in poetry. The most important thing about this proposal is that the factors that determine whether or not a given poetic language sentence should be allowed to surface are external to the grammar: that is, it is not the result of core computations in the narrow syntax, but general economy considerations that apply at the output of the PF branch of the grammar.

With (22) in place, I will now turn to a set of sub-predictions that follow from this version of the theory. Since (22) is adopted tentatively, the following subsection is speculative, and the empirical basis for testing these predictions remains somewhat underdeveloped. Nevertheless this is an important exercise in testing how far the linguistic theory of poetic language can get us in explaining the distribution of kinds of deviation

4.2.7 Poetic language blocking poetic language?

Although it is intended to close off one set of questions, the proposal in (22) raises a number of new questions in addition. In particular, we may wonder if different poetic language derivations may compete between one another: that is, does a poetic language derivation A block a derivation B if A is more econom-

ical, relative to the same basic output expectation? To examine this issue, we must first establish what the basis would be for comparison between competing derivations. It is obvious that the additional operation of PF Concatenation would not separate A and B, since they would both necessarily involve this operation, both being poetic language sentences. However, a given derivation may require some other extra operations to ensure convergence.

Here I will consider two particular syntactic issues that may be relevant to the calculation of economy in poetic language derivations. The first issue is the subject condition known as the Extended Projection Principle (Chomsky 1981), which is implicated in derivations that affect subjects. Recent work has established that the EPP should be divorced from Case-licensing and instead be given independent status as an additional condition; for example in Chomsky (1995b) it is an “uninterpretable D-feature on T,” which means that Spec,TP must host a DP (NP) which will check this feature.³⁸ Now consider the case of a poetic example of subject displacement, like the following:

- (23) Already see you a degraded Toast,
And all your Honour in a Whisper lost!

Here the subject *you* has been displaced from its standard position in front of the verb, and to derive this structure we would propose that the subject is set to one side in the narrow syntactic derivation and then PF Concatenated to produce the word order in (23). A natural consequence of this is that the D-feature on T goes unchecked in the narrow syntactic derivation, since the subject

³⁸ More recently, Landau (2007) has argued that the EPP is in fact a p-selection condition at T: this is a phonological selection condition which requires Spec,TP to be filled by a phrase with an overt head. While Landau’s theory is attractive, it suffers from a number of serious empirical problems. For example, in order to account for the fact that Germanic languages like English allow for bare noun subjects (*Men like beans*), he has to propose that the ‘anchoring’ head for selection is parametrized, where the head that counts for p-selection is N, not D (as in Romance). However, this then begs the question of how Germanic also allows for pronoun subjects, if pronouns are determiners (Postal 1969, Abney 1987, Wiltschko 1998), and there are also unanswered questions about how we could account for subjects with NP-ellipsis (*Many people were convinced but some weren’t*). I will therefore put Landau’s theory to one side for now, noting that its promise may well be realized by advances in future work; in particular, I believe that the EPP may well be reduced to a phonological condition, given the evidence in Merchant (2001) and van Craenenbroeck and den Dikken (2006) for subject movement being suspended by ellipsis (though see the references in footnote 44 for an alternative view).

is not merged in T. If this feature were to remain unchecked at the interfaces, the derivation would crash. Note that this dilemma is not encountered by PPs and objects, since they are not subject to the EPP.

This may lead to one of two different results: the derivation corresponding to the poetic language crashes, or the uninterpretable D-feature is deleted by an operation at PF in order to ensure convergence. Nunes (2004) proposes that an operation of formal feature deletion (FF-Deletion) is required to explain certain facts about linearization and the Copy Theory of Movement, so I will assume that such operations are in principle available.³⁹ Thus subject displacement sentences involve an extra instance of FF-Deletion, a costly operation, and as a result they are more costly in terms of derivational economy than derivations that involve object or PP displacement.⁴⁰

From the perspective of poetic language, the relevant empirical fact is that subject displacement is less common in poetry than other kinds of displacement like object and PP displacement, as noted in chapters 2 and 3 and by Fitzgerald (2007). We might propose that this is because of economy: a derivation which displaces the subject is less economical than a derivation that displaces the object because the subject displacement requires an instance of the extra operation FF-delete, and as a result the non-subject displacement derivations block the subject displacement derivations. This would extend the empirical coverage of the theory, but it would then open up the question of why we still find examples like (23) and the others cited in section 2.1.4 and more generally why there is just a strong preference for non-subject displacement over subject

³⁹Note that Nunes' theory of linearization and economy is not wholly compatible with the assumptions adopted here. However, a similar kind of operation, called "feature suppression," is motivated in far more compatible terms in Truswell (2009) (see also Abels 2003).

⁴⁰This may be contended if we assume that the EPP is in fact driven by the Case needs of the subject, and that objects also need to move to an AgrO projection for Case reasons, as originally proposed by Chomsky (1993) and assumed by others (e.g. Lasnik 1999, 2003, Hornstein 1995). However, this issue would be moot in the context of the present theory, since these proposals for abstract Case assume that Case can be licensed at LF. Given that I have proposed that LF-Merge does not incur derivational cost, the extra step required to ensure an object gets Case would be unimportant for calculation of economy or any other matters. Furthermore, the arguments for stating the EPP in terms of Case-licensing have been seriously undermined in recent years: see Marantz (1991), Harley (1995), Wurmbrand (2006), Bobaljik (2008) among others.

displacement, rather than an absolute restriction.

There are a few different ways to approach this issue. One way (entertained by Fitzgerald) would be to dismiss the data in section 2.1.4. There may be grounds for doing this in part, since some of the examples discussed might be plausibly dismissed, as acknowledged in section 2.1.4. To illustrate, let us consider again the examples:

(24) But errs not Nature from this gracious end,
From burning suns when livid deaths descend, *ES1: 141-142*

(25) Why has not man a microscopic eye?
For this plain reason, man is not a fly. *ES1:193-194*

(26) And out over the park where crawled roadsters
The apricot and purple clouds were
And our blood flowed down the grating [...]
John Ashbery, 'Two Sonnets'

(27) still, the rods, could not they take long
More anthems until Dust
flocks disguised machine. The stone
the valentine couldn't save... Hooks
John Ashbery, 'Europe'

Note that the issue here is not whether or not subject-verb inversion can be generated at all, but whether or not subject displacement competes with non-subject inversion. Therefore we can exclude (26) from the present discussion: the verb in the sentence is intransitive and there are no other constituents which could be displaced ahead of the subject, so there is no competing derivation which could block the inverted form we see here.

The examples in (25) and (27) are both questions, so the auxiliaries in these examples would be head-adjoined to C^0 by standard head movement, thus explaining the fact that they precede the subject. The reason that these seem like examples of subject displacement is that the subject also appears to the right of *not*; given that verb movement to C usually only pied-pipes contracted negation (*don't you understand?*), this indicates that the subject is not in its standard

Spec,TP position. However, this might not necessarily be so. First, the fact that negation shows up as full *not* rather than *n't* may simply be a reflection of the 'high style' of the text, where convention overrides normal usage to exclude representing negation as the spoken form *n't*; thus we could say that negation is parsed as *n't* even though it is written as the full-form *not*⁴¹ Second, the restriction against pied-piping *not* is a modern English phenomenon, as imperatives in EME would often appear with full negation to the left of a pronounced subject, as in the following example from *Much Ado About Nothing*:

- (28) *Lucio*: My lord, here comes the rascal I spoke of ; here with the provost.
Escalus: In very good time: — speak not you to him,
 (*Much Ado* 5.1.282-285)

Most importantly, we can also see this occurring in questions, exactly like what we see in (25) and (27):

- (29) *Leonato*: Sweet prince, why speak not you?
 (*Much Ado* 4.1.57)

Given this, we may surmise that (25) and (27) involve imitation of the EME style of pied-piping full negation in T-to-C head movement, and thus we could deny that these are examples of the kind of deviation that is of interest for the theory developed here.⁴²

If we were to accept this explanation, (23) and (24) would still require some alternative explanation. To this we can add Fitzgerald's example from *Hamlet*, discussed in 3.3.1. It is repeated below:

- (30) But die thy thoughts when thy first lord is dead.
 (*Hamlet* 3.2.214)

Recall that Fitzgerald (2007) explains (30) as a metrically-motivated inversion,

⁴¹We may assume that this occurs at the level of graphological representation, where convention dictates that cliticized negation is written out as full negation, even though interpretation and pronunciation would dictate otherwise.

⁴²The notion of imitating a given dialect of English is itself problematic, as it is not entirely clear what kind of knowledge would be used in producing such sentences. If the author does not have the grammar that would generate such representations, they must be derived at least in part from some other aspect of his knowledge. For discussion of the relation of written languages, standardization and their relation to the grammar, see Weik(2007) and Hope (2000).

where inverting the subject and verb produces a metrical line from a previously unmetrical one. In section 3.3 I argued that the OT implementation proposed by Fitzgerald was untenable for empirical and theoretical reasons, and I argued that it is not possible to produce a model of the grammar which allows metre to influence whether or not a given syntactic structure is well-formed.

However, in section 4.2.6 above I suggested a way in which one may reinstate Fitzgerald's proposal in a way that is compatible with the present theory: allowing a metricality condition to be part of a more articulated output expectation. Below I present alternative derivations for (23), (24) and (30) where a non-subject element is displaced; these are the derivations that we would expect to be outcompeting the attested versions in terms of economy, since they would not involve the extra step of FF-Deletion (note also that (33) is a regularly well-formed fronting structure):

- (31) Already you a degraded Toast see,
And all your Honour in a Whisper lost!
- (32) But Nature from this gracious end errs not
From burning suns when livid deaths descend,
- (33) But when thy first lord is dead thy thoughts die

All of these examples are less well-formed than their counterparts in terms of metrical form. In (31) the main stress in the polysyllabic word *degraded* falls on an odd position in the iambic pentameter template. (33) is on a par with the well-formed uninverted sentence in terms of metricality, since both fall foul of the Monosyllabic Word Constraint (see section 3.3.1), unlike the inverted form in (30). In (32) the main stress that would normally fall on the main verb *errs* is aligned to an odd-numbered position; the line may be well-formed if *not* is stressed instead of the verb, but this is not the preferred prosodic structure for the text, and it may also change the meaning of the sentence by putting stress on negation (in which case it would not compete for the calculation of economy). We can see, then, that the subject displacement derivations for these

examples would be more metrical than their counterparts. If a well-formed metrical template was part of the output expectation for the sentences, the subject displacement derivations would not compete with the alternatives in terms of economy, and we would thus expect these forms to appear.

This may seem to be a way of sneaking Fitzgerald's results into the present theory through the back door, as it allows metre to determine the distribution of 'well-formed' poetic language sentences by narrowing the conditions for economy competition. While this may be true to an extent, the proposal is formulated in a way that avoids some of the crucial problems for Fitzgerald's proposal, and as such it is significantly different from an OT-theoretic approach to poetic language. First, this metrical output expectation is an add-on for Bypass Theory, rather than an alternative, so it allows us to retain the explanation of non-metrically-motivated deviations. Second, the output expectation does not interact with the computations of the narrow syntax, and as a result it does not suffer from the overgeneration problem suffered by Fitzgerald's theory; rather, output expectations have the effect of a filter, choosing between generable derivations.⁴³ Third, the output expectations are not part of the narrow linguistic computation as such, so they do not need to be restricted to computations over linguistic primitives; this means that non-linguistic notions such as rhyme and alliteration may plausibly be included as part of the output expectations.

This last point has particularly broad implications, as it is well-known that poetic texts can be composed to meet a wide range of arbitrary non-linguistic constraints, not just metre. For example, Fabb (1997) argues that many poetic traditions constrain composition with arbitrary notions of what "correct grammar" should be (in the "school grammar" sense or others derived from guides for good practice); we also find many poetic texts in the 'Oulipo' literary school which are written in a way that obeys constraints on the occurrence of words

⁴³The theory would also not be subject to the internal problems suffered by the OT theory, such as the difficulty of ruling out head movement to right-headed positions in a left-headed language (see section 3.3.1, footnote 32).

containing certain graphological letters (such as Georges Perec’s ‘lipogram’ *La Disparition*, written without the letter *e*), on the occurrence of word-initial graphological letters (such as Walter Abish’s *Alphabetical Africa*) and various other arbitrary aspects of composition (see Mathews and Brotchie 1998 for an overview of the Oulipo and its predecessors). There remain a number of real issues that may not admit to easy solutions in terms of output expectations, and the terms of the potential explanation given here have not been clearly defined. Furthermore, the empirical basis for these added conditions require further empirical testing, and they are contingent upon a specific analysis of the EPP, an issue which is far from settled.⁴⁴ For these reasons, I will not commit to a final decision here on whether a more articulated output expectation should be adopted as a sub-part of the theory, noting its promise for expanding the theory’s empirical coverage. I return to this issue later in the chapter.

The second potential case of economy competition between different poetic language derivations involves Case. In section 4.2.3 I mentioned that argument NPs are often assumed to bear an uninterpretable Case feature *uCase* which needs to be checked off before it reaches the interfaces, in accordance with Full Interpretation. It is also typically assumed that *uCase* can only be checked either by entering into a local Spec-head relation with an appropriate Case-assigning head (Chomsky 1995b), or by entering into an Agree relation with the relevant Case-assigner (Chomsky 2001). If we accept that this is so,⁴⁵ we could assume that an argument NP that is put to one side in a poetic language derivation will need to have its *uCase* feature deleted later in the derivation, in the same manner

⁴⁴ For example, there are a number of alternative analyses of the EPP that would not require us to make an economy distinction between subject displacement and non-subject displacement. Chomsky (2001, 2004), Lasnik (2001) and Lasnik and Park (2003) argue that the EPP is not a feature-driven form of movement but rather a “configurational requirement” for certain functional heads to have a specifier; as such, it applies at LF and could be fulfilled by (costless) LF-Merger under the present proposals. Alternatively, Alexiadou and Anagnostopoulou (1998) argue that the EPP is a different kind of condition that is satisfied by some element being evacuated from the VP; under a certain interpretation, the analysis proposed here may be taken to be directly compatible with this. See also footnote 38 for references to another set of intriguing alternative analyses.

⁴⁵ This is far from an innocent assumption. In recent years abstract Case has been criticized in a number of ways, and many linguists have suggested that it should be dispensed with as a theoretical notion altogether. See the references mentioned in footnote 40. For an alternative approach to Case effects, see Pesetsky and Torrego (2001, 2004).

as the D-feature on T in the explanation of the EPP derivations given above (see also Nunes 2004). Specifically, this would take place in the PF branch, since it would not be required in the LF branch if LF-Merger comes for free, as we could simply LF-Merge the NP into the correct position to check Case. Only FF-Deletion in the PF branch would incur derivational cost.⁴⁶ These poetic language sentences would still converge, but they would be less economical than derivations in which an element without Case was set to one side.

This would lead to a few specific predictions about the co-occurrence of NP displacement and other kinds of phrasal displacement in poetic language derivations. One prediction is that, in a situation where just one phrase is being set to one side, displacement of a non-Case-related phrase would outcompete displacement of an NP when it is possible to displace either of them. So, given a sentence with a basic S-V-O-PP word order, if only one of the VP-internal phrases were to be set to one side, it would be the PP, and as a result we would expect something like S-PP-V-O. Furthermore, we would expect this S-PP-V-O derivation to block the S-O-V-PP derivation, since the former would be more economical. This prediction is confirmed by one of the surprising gaps in the data gathered in chapter 2, discussed briefly in section 2.1.5. We saw in section 2.1.1 that S-PP-V-O was a common attested structure; below is one of the five examples gathered:

- (34) Close by those meads, for ev'r crown'd with flow'rs,
Where Thames with pride surveys his rising tow'rs,

However, the corpus of texts examined contained no examples of the corresponding S-O-V-PP, in which an object has been displaced in favour of a PP. In fact, this is the only rearrangement of the S-V-O-PP template that is not attested (keeping the S-V order stable), as we find examples of S-O-PP-V (35), S-PP-O-V (36), PP-S-O-V (37), O-S-PP-V (38), O-PP-S-V (39) and PP-O-S-V (40):

⁴⁶Note that this analysis departs from the earlier model in Chomsky (1993, 1995b) by assuming that unchecked Case features are relevant at both PF and LF; on Chomsky's analysis, Case could be checked by covert movement, indicating that Case features were not relevant at PF.

- (35) Remembrance and reflection how allied!
What thin partitions *Sense* from Thought divide! *ES1: 225-226*
- (36) Love *in these labyrinths* his slaves detains,
And mighty hearts are held in slender chains. *TR2: 23-24*
- (37) Favours to none, *to all* she smiles extends;
Oft she rejects, but never once offends. *TR2: 11-12*
- (38) *Thy voice* I seem in ev'ry hymn to hear,
With ev'ry bead I drop too soft a tear. *EL: 269-270*
- (39) *The drops to thee*, Brilliante, we consign;
And Momentilla, let the watch be thine; *TR2: 113-114*
- (40) A heav'nly Image in the glass appears,
To that she bends, *to that* her eyes she rears; *TR1: 125-126*

This makes the S-O-V-PP gap in the empirical picture particularly intriguing, and the fact that it follows straightforwardly as a prediction of the theory proposed here (in collusion with certain assumptions about Case) is a strong indication that poetic language derivations may indeed compete with each other in terms of economy.

A potential problem here is that examples (35), (36), and (37) all display displacement of objects to a pre-verbal position; we may expect, then, that they would be outcompeted by the corresponding derivations that leave the object in the VP. However, we can observe that these examples are all unmetrical without displacement of the object, similar in some respect to the subject displacement examples discussed above. Below are the relevant alternative versions:

- (41) Remembrance and reflection how allied!
What thin partitions from Thought divide *Sense* !
- (42) Love *in these labyrinths* detains his slaves,
And mighty hearts are held in slender chains.
- (43) Favours to none, *to all* she extends smiles;
Oft she rejects, but never once offends.

In (41) the stress in *divide* falls in an unstressed position in the template, and the same happens with *detains* in (42) and *extends* in (43). This would seem to provide further evidence for allowing metrical structure to decide which poetic language derivations compete in terms of economy (although rhyme is also strongly implicated here).

As with the EPP analysis given above, I will refrain from claiming this as a clear empirical result for the theory proposed in this chapter, for similar reasons. First, the corpus from which these examples have been taken is relatively small, so we may well find more examples of the apparently unattested form if we were to examine more texts. These kinds of claims about distribution only become convincing in the face of much wider confirmation. It is noteworthy that there is no significant difference between the unattested S-O-V-PP form and the common S-PP-V-O form in terms of intuitive judgments of acceptability; for example, a version of (2) in which the object has been displaced instead of the PP is not obviously worse than the original (it is also metrical):

(44) Where Thames his rising tow'rs surveys with pride,

Second, as acknowledged above, the required assumptions about Case are not at all innocent, and they are tied to a specific view of formal feature checking that may or may not be correct. This characterization of Case facts is dependent upon the dichotomy between interpretable and uninterpretable features, where uninterpretable ones like *uCase* need to be checked to ensure a derivation converges. However, Preminger (2010) has challenged the existence of these sorts of dichotomies between interpretable and uninterpretable features, working within the domain of phi-agreement in Hebrew; he shows that there are agreement patterns that cannot be dealt with by such dichotomies, but which instead implicate “obligatory operations” as part of the syntax. Given that a great deal of recent work on Case has posited that Case agreement is intrinsically related to phi-agreement (see e.g. Chomsky 2000, 2001, 2004, Boeckx 2003, Řezáč 2004, 2008), such a reanalysis of phi-agreement may have significant implications for

the analysis of Case-like effects proposed here.

Despite the above disclaimers, the important result of this section is that we can see that the present theory of poetic language can make clear and testable predictions about the kinds of poetic language forms that may be possible, in collaboration with stated assumptions about relevant theoretical points. This is done by assessing the relative cost of suitably similar poetic language derivations and proposing that more economical derivations outcompete others. As such, the theory assumes that there is a set of possible (convergent) poetic language derivations, PD_P , which is defined by the form of the grammar and generated by procedures like those described by Bypass Theory. All things being equal, we can define a subset of well-formed poetic language derivations in PD_P , which we may call PD_W ; these are the most economical among the derivations in PD_P , and the complement of PD_P and PD_W is the set of less economical poetic language derivations, PD_L ; all the members of PD_L are outcompeted by members of PD_W .

We also saw that a more articulated model of poetry-specific output expectations could allow for externally-defined notions of form – metricality, rhyme, etc – to narrow the scope of comparison of members of PD_L and PD_W ; in effect, the output expectations may filter out members of PD_W and thus allow for members of PD_L to appear. The definition of output expectations is the realm of the poet’s choice, and it is guided by non-linguistic notions of form that are amenable to conscious manipulation; for example, the poet would choose “in iambic pentameter” or “rhymed,” rather than some more technical notion like “displace the object to the edge of vP” or “apply object shift.” The most important thing about this fact, however, is that these output expectations allow for the choice between members of PD_P , rather than for the creation of impossible forms; as such, poets may manipulate external conditions to affect which members of PD_P can occur, but they cannot manipulate the computations of the grammar to produce derivations that fall out of PD_P . This is a very significant theoretical point that I will return to in chapter 5.

4.2.8 Summary

In this section I have introduced a theory of poetic language that accounts for why certain kinds of poetic language sentences occur and others do not. The proposed system generates the linguistic forms of poetic language sentences by setting Pieces of structure to one side in the narrow syntactic derivation and then recombining the Pieces into one full structure by independent processes in the LF and PF branches of the derivation. I proposed to derive a number of empirical results directly from this proposal, such as the ban on ‘head displacement’ and the fact that displaced constituents occur at phrase boundaries. I then argued that poetic language differs from ordinary language in terms of derivational economy, since poetic language derivations involve an extra costly derivational step in the PF branch. I proposed to deal with the fact that ordinary language does not block poetic language in poetry by appealing to the notion of an output expectation (required for any use of an economy metric), which was taken to be a condition which demanded that the output form be distorted in some way.

4.3 Deriving the data

The discussion above has explored the details of the proposal by looking at a few simple examples of displacement, in particular, examples of NP and PP displacement. In this section, I will show how a generalized version of this theory can derive the other attested forms identified in chapter 2. First I explain how the theory deals with the other cases of displacement, including the important category of non-constituent displacement, which cannot be explained by other approaches but which follow straightforwardly from the present theory. Then I discuss in detail the important category of erasure, explaining the theoretical background of phonological deletion in more detail and generalizing the present theory to account for this data. This provides us with a unified account of the phenomena discussed in chapter 2.

4.3.1 Different forms of displacement

Here I will round off the discussion of displacement by showing how the theory as presented in the previous section can be used to derive all of the remaining core cases of displacement discussed in section 2.1. First, consider the remaining cases of constituent displacement. Below is an example of TP displacement:

- (45) No bandit fierce, no tyrant mad with pride,
 No cavern'd hermit, rests self-satisfied;
 Who most to shun or hate mankind pretend,
 Seek an admirer, or would fix a friend. *ES4: 41-44*

In this example, the phrase *to shun or hate mankind* is a full TP (control infinitive) that has been displaced from its regular position as a selected complement of the verb *pretend*. This would be generated much in the same way as an example of object displacement. First, the full TP would be constructed by normal means and then set to one side by being assigned an [ELSEWHERE] address upon completion of the full structure; we may assume that this structure also encompasses the PRO subject of the infinitive.⁴⁷ Then, the rest of the matrix clause would be constructed by normal means, with a gap in the position normally filled by the infinitive. The two would then undergo PF Concatenation at PF to derive the word order, and at LF the phrase would be merged into the standard complement position, as attested by the fact that the disjunction falls within the scope of the adverb *most(ly)*.

The examples of AP displacement would be generated in a similar manner. Below are examples that exemplify the two main kinds of AP displacement, AP complement displacement (46) and AP specifier displacement (47):

- (46) Oh had I rather unadmir'd remain'd
 In some lone Isle, or distant Northern land; *TR4: 153-154*
- (47) Of systems possible, if 'tis confest
 That wisdom infinite must form the best, *ES1:145-146*

⁴⁷It is perhaps significant that all of the cases of TP displacement discussed in section 2.1.4 involve control infinitives rather than raising infinitives, since the latter clearly have A-chain dependencies between the TP and the matrix clause, while the former may not (depending on one's analysis of control); there may be difficulty in deriving the correct A-chain in a derivation where the base of the chain is located in a phrase that is set to one side.

The structure in (46) would be generated much in the same way as (45) and other such examples of complement displacement: the AP would be constructed separately (including its adverb modifier *rather*) as a Piece and then recombined in the different branches of the derivation by PF Concatenation to the relevant phrase boundary and LF-Merger into the complement position. The example of AP-specifier displacement would differ only in the fact that the AP is LF-Merged into a Spec,NP position in the NP, rather than a complement position.

The examples of adverbs would be given a similar explanation to AP-specifiers. Below is an example:

- (48) Some secret truths, from Learned Pride conceal'd,
 To Maids alone and Children are reveal'd: *TR1: 37-38*

Here the adverb *alone* takes scope over the whole of the conjunct, so its occurrence in the surface position to the right of *Maids* suggests it has been displaced from its standard position, which would be right-adjoined to the larger NP (*Maids and Children alone*). This would be generated by setting the adverb to one side and then LF-Merging it into the correct adjunction or specifier position⁴⁸ and PF-Concatenating to put it in the position to the right of *Maids*. The conjunction *and* would be grouped with the second conjunct *Children* in the phonological phrasing of the constituent (Taglicht 1998), so the appearance of the adverb in this position is predicted by the theory of PF Concatenation outlined above. Finally, it should be stressed that this kind of reordering cannot be done by standard movement rules (as mentioned in chapter 2), so it is the kind of arbitrary reordering that is well served by the present theory.

Finally, we can show that the theory also deals easily with the problem case of non-constituent displacement. This is exemplified by the following example:

- (49) For this, ere Phoebus rose, he had implor'd

⁴⁸It should be borne in mind that the constraint on head displacement from ‘in between positions’ (discussed in section 4.2.5) thus makes the present analysis of adverb displacement incompatible with the cartographic approach to adverb order which holds that adverbs occupy specifiers of dedicated functional projections (Cinque 1999), unless the set of projections is assumed to be a part of the structure that is always projected. See Nilson (2003) for an alternative approach to Cinque’s hierarchy, Bobaljik (1999) for problems and Abels (2010) for discussion of some general problems for cartography that are relevant here.

Propitious heav'n, and ev'ry pow'r ador'd,
 But chiefly Love – to Love an Altar built,
 Of twelve vast French Romances, neatly guilt

TR2: 35-38

Here the NP *an Altar* is separated from its argument PP complement *of twelve vast French romances* (the separate PP complement *to Love* is also displaced from its VP-internal position but we can ignore that here). This is generated by setting both the NP and the PP complement to the side as separate Pieces, constructing them as usual, with a gap in the NP's argument position. At LF, the NP is LF-Merged into the standard VP complement position and the PP is LF-Merged into its position as an argument of that NP. At PF, they are Concatenated to produce the word order in (59). As a result, all selectional restrictions are satisfied and the sentence is interpreted like normal. This requires no stipulations or extra and problematic 'relaxations' of standard rules, the likes of which would be required by all other theories of poetic language.

Thus the theory proposed in section 4.2 can generate all the classes of displacement discussed in chapter 2 without any stipulation. The rest of the section concentrates on extending the theory to deal with the data from erasure.

4.3.2 Uniting erasure and displacement?

The erasure phenomena discussed in chapter 2 do not follow in an obvious way from the theory proposed in section 4.2. The theory proposes that Pieces are set aside and later LF-Merged into their standard positions and PF Concatenated to some other position, where the effect of PF Concatenation is to (randomly) set a position for the pronunciation of the Piece within the larger string formed by the rest of the derivation. Because the crucial difference between erasure and displacement lies in what happens in the PF branch, we may presume that PF Concatenation is the part of the proposal that needs to be amended.

One might respond to this challenge by simply extending the remit of the operation called 'PF Concatenation,' by allowing it to delete Pieces as well as insert them into the structure. The core property of PF Concatenation as it

was defined in section 4.2.4 is that it settles the linear order of two given pieces, saving a linearization crash: without this operation, the articulatory-perceptual system would not have instructions for fully ordering the output of the derivation, and as a result the instructions to the system would be “gibberish.” We might propose that simply deleting a Piece is a way of achieving the same effect as inserting it, since it technically provides unambiguous instructions to the articulatory-perceptual system. If we accepted this, ‘PF Concatenation’ may be reconfigured as a more general ‘PF Ordering’ operation which is able to delete as well as insert Pieces.

This is attractive from the theoretical perspective, since it provides a simple unified account of poetic language. However, it is problematized by the fact that erasure and displacement generally target very different syntactic elements. We saw in chapter 2 that the majority of displacement involves NPs and PPs, often more than one at once: displacement seldom (if ever) affects syntactic heads, a fact explained in section 4.2.5 as a result that follows from problems that occur from trying to recombine these Pieces by LF-Merge. In contrast, the overview of erasure in chapter 2 showed that erasure mainly affects head elements, like dummy *do*, infinitival *to*, the articles *the* and *a*, and auxiliaries like *have* and *be*. In sum, it seems that the kinds of elements affected by erasure and displacement are almost in complementary distribution. This is a clear disincentive against providing the two phenomena with a completely uniform explanation.

The explanation for the non-existence of head displacement hinges on the impossibility of LF-Merging head Pieces, so we can surmise that erasure does not involve generating separate Pieces. Nevertheless, we can still develop a unified explanation of the erasure and displacement if we propose that the difference between an erasure sentence and its ordinary language equivalent is that the erasure derivation is less economical, in particular due to economy violations incurred in the PF branch. This is what I will suggest in what follows, and in the end the unified explanation of poetic language is that poetic language derivations are ‘PF uneconomical’: displacement involves extra and unjustified

instances of Concatenation, and erasure involves extra and unjustified instances of the PF operation Delete. An important part of this proposal is that deletion as a phonological operation should be costly, and showing this requires us to provide an introduction to the theory of ellipsis licensing.

4.3.3 Erasure, ellipsis and economy

As mentioned in section 2.2.1, well-formed ellipsis is subject to a syntactic requirement known as ‘licensing,’ which requires that the ellipsis site occur adjacent to certain kinds of syntactic elements. For example, it is well-known that, in English, ellipsis sites can occur adjacent to finite auxiliaries (50) and wh-phrases (51), but not next to prepositions (52), finite verbs (53), (unfocused) NPs (54):

- (50) Mary will arrive on time, and John will ~~arrive on time~~, too.
- (51) They kicked someone out, but I don’t know who ~~they kicked out~~.
- (52) *When I saw Michael, I thought he was cycling to his mum’s house, but it turns out that he was actually coming home from ~~his mum’s house~~.
- (53) *John seemed to like cake, and Mary seemed ~~to like cake~~, too.
- (54) *John liked cake, and Mary ~~liked cake~~, too.

In the literature, the term ‘licensor’ is used to describe the element that allows for ellipsis at its edge, and paradigms like this are typically taken to be evidence that finite auxiliaries and wh-complementizers⁴⁹ are licensors, but prepositions, finite verbs and NPs are not. More specifically, the licensor is taken to be a specific syntactic head, known as the licensing head.

⁴⁹It is typically assumed that it is the wh-complementizer that is the licensor rather than the wh-phrase because wh-phrases themselves do not license ellipsis when they do not occur in Spec,CP. This is largely due to theoretical considerations (Lobeck 1995), but there is evidence for this, such as the fact that wh-in-situ not license ellipsis:

- i. *John said a friend of his believes some film to be a message from the dead, but I don’t know which friend of his believes which film ~~to be a message from the dead~~.

Lobeck (1995) explains the category of ‘licensing head’ by appealing to notions of government, arguing that the empty categories represented by ellipsis sites are null proforms that are subject to the Empty Category Principle (ECP), much like other empty categories like traces. However, the move to Minimalism abandoned notions like government and the ECP, and as a result researchers were forced to rethink the licensing conditions in more modern terms. Merchant (2001) proposed that ellipsis was produced by a feature, known as the E-feature, which is optionally added to parts of the structure and checked in an adjacency relation with the lexical items in the class of ellipsis licensors. In his account, checking of the E-feature means the whole constituent is marked for deletion at PF,⁵⁰ and the fact that this feature is only checked by licensors in a local relation ensures that ellipsis only occurs adjacent to the licensor. Aelbrecht (2009) updates Merchant’s theory, arguing that ellipsis is licensed by an Agree relation between the E-feature, which is borne on certain projections, and the licensing head; this allows for deletion when the licensing head and E-feature are in a non-local relation, and it also provides an explanation for some unusual extraction patterns in Dutch Modal Complement Ellipsis (MCE).

In Thoms (to appear) I argue that the E-feature account of ellipsis licensing is both theoretically and empirically flawed, and that the category of licensing head is little more than a stipulation and thus not in keeping with the general spirit of Minimalist inquiry. On the basis of dialectal variation in English, I propose instead that ellipsis is not licensed by an E-feature, but rather by overt movement. I propose that ellipsis licensing can in fact be derived from the Copy Theory of Movement and the assumption that a chain cannot be linearized if one pronounced copy c-commands the other. I thus argue that ellipsis is a ‘repair strategy’ of sorts that is required to save a linearization failure. Specifically, ellipsis occurs at the edge of a moved element when the base element in the movement chain is not deleted locally, that is, at the point when it is moved. When the element is moved into its new position, deletion of

⁵⁰See Merchant (2001) for extensive arguments for the PF deletion account of ellipsis licensing in favour of the proform account of Lobeck (1995).

the entire complement of the landing site is required to ensure that the structure can be linearized, since otherwise the higher copy will c-command the undeleted lower copy causing a linearization failure.⁵¹

Importantly, this means that only moved elements will ever be able to license ellipsis, and the process of deletion that results in what we know as ellipsis is essentially the same as the deletion operation that deletes a lower copy in a chain. This reduces the optionality of ellipsis to the option of applying the operation Delete locally or non-locally: the difference is implemented by different orderings of the sub-operations of movement. Thus if move is taken to involve two sub-operations, Copy and Merge, local deletion would involve the order *Copy* > *Delete* > *Merge*, with Delete targeting the site targeted by Copy whereas non-local deletion would be *Copy* > *Merge* > *Delete*, with Delete targeting the site targeted by Merge, that is, the new sister of the Merged element (see Thoms to appear for details). This does not mean, however, that all movement licenses ellipsis. A given movement process needs to be able to avoid local deletion – that is, it needs to be able to leave an undeleted lower copy – in order to later produce ellipsis by deletion of the complement of the destination position. In Thoms (to appear) I point out that this is tied to the ability to pronounce lower copies, and argue that this explains the fact that A-bar movement and verb movement can license ellipsis, but A-movement cannot: A-bar and verb movement often permits non-deletion of base copies, as in covert movement and various kinds of doubling phenomena, whereas A-movement never allows for base copies to be pronounced (see Thoms to appear and the references cited therein). The theory thus derives the facts in (50)-(54) and other related facts from the presence or absence of certain kinds of movement: (50) and (51) involve verb movement and A-bar movement, whereas the elements at the edge of the ellipsis site in (52)-(54) have either undergone A-movement or have not moved at all.⁵²

⁵¹This is in accordance with the general thinking of works on linearization like Kayne (1994), Uriagereka (1999), Fox and Pesetsky (2005), M. Richards (2004), Nunes (2004) and others.

⁵²Examples like (54) are acceptable if the subject of the second sentence is given special focus; these instances are known as ‘stripping,’ and they are often analysed as instances of focus movement-driven ellipsis similar to sluicing. See Thoms (to appear), Merchant (2003) and references cited therein.

Recall that in section 4.2.5 it was argued that linearization is an issue of convergence, and that operations that ensured convergence were taken to be costly operations. Given this, Delete is a costly operation that should only occur when it is required for convergence, so optional application of this operation should thus be blocked by economy considerations in the PF branch. In the case of ordinary ellipsis, we say that Delete is required to ensure convergence, as an alternative mode of copy deletion. It follows, then, that ellipsis in the absence of a licenser⁵³ constitutes a violation of derivational economy, since the given derivation involves an extra and unwarranted instance of the operation Delete in the PF branch.

With this established, we can explain how erasure can be dealt with in the context of the present theory of poetic language. Examples of erasure involve deletion in the absence of a licenser, so we say that the derivation of these poetic language sentences involves Delete applying without the necessary ‘motivation.’ That is, Delete is not applied to save a linearization failure, but rather it is simply applied to erase a given set of words, and for this reason the derivation is less economical than the standard sentence in which Delete does not apply; thus the elements that precede erasure sites in poetic language need not have moved to their surface position. Since both lead to convergent derivations that map onto the same meaning, they compete in terms of economy, and therefore the erasure sentence is blocked by the non-erased alternative because of the extra instance of the costly operation Delete. This explains the fact that erasure sentences are ungrammatical in ordinary language contexts. Just like with displacement, we are required to propose an enriched output expectation, where the expectation of distortion (perhaps defined differently for this case) allows this blocking to be lifted. Given this, displacement and erasure are united as derivations that can be described as ‘PF-uneconomical,’ in that they involve the application of extra operations at PF that would not be required in their ordinary language

⁵³Note that the term ‘licenser’ no longer refers to a specific class of elements; instead, the elements that are termed ‘licensors’ are simply elements that have moved and which allow for non-local copy deletion. Nevertheless I will continue to use the term ‘licenser’ for clarity of exposition.

equivalents. We can state the unified theory of poetic language in the following generalization:

- (55) Poetic language derivations are PF-uneconomical derivations that are allowed to surface by externally imposed output expectations.

While displacement and deletion involve different derivations, (55) proposes that they are of the same essential character. As such, (55) constitutes a unified theory of poetic language. I will unpack some of the implications of (55) and provide an explicit reformulation of the theory in section 4.3.4, but for the remainder of this section I will concentrate on demonstrating how the proposed derivation for erasure fits with the data.

Simple cases of deletion of a single element follow straightforwardly from this theory as cases of deletion targeting the single head or phrase. Consider the following examples from chapter 2, which exemplify a number of different cases of erasure:

- (56) ... rather, it is in the disrepair
Of these lives that we ___ not find despair

LI1: 53

- (57) What can the rain that fell
All day on the grounds
And on the bingo tables ___ ?

John Ashbery, 'Album Leaf'

- (58) And after ___
Taken out behind the stairs and ___ stood them
In the kitchen... the flowers blowing in the wind
Felt funny just the same...

John Ashbery, 'Night'

- (59) No words for what when words ___ gone.

WH: 28

- (60) How ___ be
young and yet to be loved?

Robert Creeley, 'Ice Cream'

- (61) The kids came and we all went ___ the briars.
John Ashbery, 'Night'
- (62) His love boiling up to me
Forever will I be the only ___
In sofa I know
The darkness on his back
John Ashbery, 'Rain'
- (63) While fish ___ in streams, or birds delight in air,
Or in a coach and six the British fair,
TR3: 163-164
- (64) Where the sweet william grew and a few other cheap plants ___
The rhythm became strained,
John Ashbery, 'Haunted Landscape'
- (65) months passed,
things happened in ___ .
Robert Creeley, 'Say Something'
- (66) A plate that has a little bobble, all of them, any so.
Please a round it is ___ ticket.
TB: 17

All of these examples can be treated as instances of unlicensed (i.e. unnecessary) application of Delete to syntactic heads; in this respect, the Deletion we see is similar to that which would occur at the foot of a movement chain when one of these items underwent head movement. In (56) and (57) the head *do* is targeted by Delete locally. In (58)-(60) it is auxiliary heads that are targeted by Delete, presumably after they have moved to the positions where they bear the relevant inflectional heads. In (61) the preposition *to* is targeted by Delete, while the complement of the head is left intact. We may presume that (62) is also an instance of head erasure, where the head noun *one* is targeted by Delete; if this were an instance of phrasal ellipsis, we may expect the adjective *only* in Spec,NP to be Deleted too. (63) is an example of what was called 'faulty gapping' in chapter 2, but this can simply be analysed as a situation where Delete targets the verb head *delight*, much like (64). Neither case can

be analysed as VP-ellipsis, because they lack an appropriate licenser and the gapping examples involve the survival of a VP-internal constituent. (65) may be an example of phrasal or head deletion, as the missing complement of the preposition is interpreted as a pronoun that is anaphoric with *months*, such as *them*. (66) involves the determiner head *a* being targeted by Delete to the exclusion of the rest of the NP.

The cases in (56)-(66) are representative of the majority of cases of erasure, and they are explained easily as cases of Delete targeting individual heads.⁵⁴ However, there are still a significant number of cases that do not follow so easily from the present theory. For example, we saw that there were many examples of erasure of non-constituent strings. For example, the string *there is/are* was erased in a number of examples, such as (67), as was the similar *it is*, such as in (68); (69) demonstrates erasure of the string *to be* and in (70) the string *that which is* is erased:

(67) this voice is truly changeable of which ___ so little left in me
HI: 15

(68) The time to show a message is when ___ too late and later there is no hanging in a blight.
TB: 4

(69) An occasion for a plate, an occasional resource is in buying and how soon does washing enable a selection of the same thing ___ neater.
TB: 7

(70) There is no world
 except ___ felt,
Robert Creeley, 'After'

There are a number of different options for explaining these examples, which I will consider in turn.

⁵⁴Note that the fact that standard ellipsis doesn't normally target single heads is irrelevant, since we do know that Delete can target heads, as in the case of head movement where the lower copy is Deleted. The fact that ellipsis only targets phrases is due to the fact that movement is always to a position that c-commands a full phrase; therefore, ellipsis always precipitates Deletion of a full phrase.

First, we could propose that they involve Deletion of each individual element; thus in (67) there are two superfluous instances of Delete, one targeting *there* and another targeting *is*. This is problematic, however, as we would expect that this derivation would be blocked by the shorter one that involves just one instance of Delete, targeting just one of the elements. The previous examples show that both of the elements *there* and *is* can be targeted by Delete independently equally freely, so one might argue that, in the absence of a factor deciding between which of these elements should be Deleted, the default action is to Delete both when one has decided to Delete at least one. However, this explanation is unsatisfactory, because it is ad hoc, it appeals to vague notions of ‘default actions’ and it also seems to predict that we would find multiple Deletion of adjacent elements in numerous other cases where we do not find it; for instance, we might predict that (60) would involve Deletion of both the infinitival auxiliary *to* and the verb *be* in the first clause, contrary to fact.

A second option is to propose that Delete targets phonological strings, rather than phrases and constituents; this would allow us to explain all of the cases in (67)-(70) as simple instances of Delete. However, this is highly unlikely since syntactic operations generally target constituents. For example, this would not sit well with what we know about most other instances of Delete, such as that which applies in VP-ellipsis: VP-ellipsis always targets constituents, and not simply phonological strings. This is attested by the fact that the ellipsis site can exclude non-c-commanded adjuncts, as in (71a); these examples cannot be explained as leftward movement followed by full string deletion, since these adjuncts do not normally undergo leftward movement to such a position, (71b):

- (71) a. John will celebrate wildly if a European team wins, but Mary only will ~~celebrate wildly~~ if Holland win.
 b. *John will if Spain win celebrate wildly.

While Lasnik’s (1999) account of pseudogapping proposes something similar, in that case the leftward movement is an attested case of movement to AgroP

(object shift),⁵⁵ whereas the ‘adjunct shift’ seen in (71b) is unfamiliar and thus unattractive as a proposal, so we can reject the idea that ellipsis targets non-constituent strings as empirically flawed. Therefore it seems unlikely that erasure should involve non-constituent string Deletion when standard ellipsis always involves the Deletion of constituents.

A third alternative is to propose that the examples of non-constituent erasure receive a similar explanation to the account of pseudogapping mentioned above, where the surviving constituent moves before application of Delete to the relevant full constituent. In the case of (67), *so little left in me* would move to a position outside of the TP (either by leftward movement to Spec,CP or rightward movement to an adjunction position), followed by Deletion of the whole TP. These derivations would be similar to those involved in fragment answers to questions. However, this would bring the theory closer to those critiqued in chapter 3, since it would require us to posit a number of movement operations that are not attested outside poetry; for example, movement to a focus position in an embedded Spec,CP is not possible in English (although it is possible in others, like Dutch; see Temmerman 2009). Therefore this is also an unwelcome approach, given the conclusions of the previous chapter.

Finally, a fourth option is to propose that the elements that follow the ellipsis – *so little left in me* in (67) – have undergone displacement (as it is defined in section 4.2) as a single constituent, and that the erased string is Deleted as a constituent after PF Concatenation of the phrase to a position at the end of the sentence (outwith the Deletion site defined by the TP constituent). This is a variation on the third proposal above, but it differs from it in that it uses the poetic language theory-related mechanisms of displacement, rather than the ordinary language mechanisms of movement, to explain the non-constituent-targeting aspect of Deletion. This is preferable from a conceptual point of view, since we may well expect that it will be the poetic language mechanisms that derive these unusual cases. There is one potentially serious problem for this

⁵⁵See the discussion of pseudogapping in section 4.3.4 for further relevant discussion of pseudogapping.

account: we may expect that such a derivation would be blocked by economy consideration like the others, since it involves two additional costly PF operations that are not required for convergence, namely PF-Concatenate, which puts the constituent *so little left in me* to the right of the rest of the sentence, and Delete, which ensures that the string *there is* is erased.

However, this is not the case. There are three possible alternative derivations: (i) the well-formed one, (ii) one with just displacement, and (iii) one with just erasure of the TP. As we may expect, option (i) can be excluded from competition, because it does not produce a form that matches the output expectation of distortion. However, option (ii) can be excluded for the same reason, since just displacing the constituent to the sentence-final position where it occurs in (67) will not produce a distorted word order; rather, this word order will be identical to the well-formed one, and as a result, it is not a contender as a poetic language derivation. Importantly, option (iii) can also be excluded, because this would produce a surface form like this:

(72) this voice is truly changeable of which ~~there is so little left in me~~

Since the constituent *so little left in me* has neither been moved or displaced in (72), it remains in the TP and is thus erased. However, this causes severe problems for recoverability (an issue I will return to shortly), since there is no immediate antecedent for the lexical material in the Deletion site, and as a result the sentence would be semantically ill-formed; more specifically, it would fail to satisfy the LF output expectation for a clause corresponding to a meaning like *of which there is so little left in me*. As a result, this is also out as a competitor and hence there is no blocking derivation that prevents the displacement-plus-erasure derivation for (67). We can thus derive the above examples of non-constituent erasure in this manner.

It is worth considering some important characteristics of the derivation for non-constituent erasure that I have just sketched. First, the erasure site must be left-adjacent to a syntactic element that can be readily displaced; in all of

the cases above this is the case, as the adjacent elements are argument APs, which we know to be readily displaceable. Second, the non-constituent that is erased needs to correspond to a continuous string, rather than a set of elements separated by non-erased elements, as otherwise we would experience the same problems as the multiple deletion derivation sketched above. Third, the non-constituent string needs to be contained within a single constituent, since otherwise Delete will not be able to catch all of the elements in the string in one go. This set of conditions contrives to make the most likely non-constituent derivations those that correspond to erasure of clause-initial elements in TP (i.e. a subject and auxiliary, or a pair of auxiliaries), and this is what is reflected in the data discussed above and in chapter 2. We may be tempted to extend this explanation to the cases of ‘subject ellipsis’ discussed at the end of section 2.2.3, since they are broadly similar to the cases just reviewed here; however, I will not go into this question, since it opens up a much larger set of questions that cannot be addressed here.⁵⁶

To conclude this subsection, I will discuss two important issues for the present analysis of erasure. The first issue has been alluded to in the discussion of non-constituent erasure above, namely instances of double erasure. Below is an example from Beckett’s *Worstward Ho*:

(73) So skull ___ not go. What left of skull ___ not go.

WH: 46

The gaps indicate missing words which were proposed to be the functional element *does*. While each of these gaps are in separate sentence, each of the sentences contain other kinds of erasure: the noun *skull* in both sentences is missing an article, and the second clause also seems to be missing a tensed verb in the free relative subject, as it is interpreted as *what is left of the skull*.

However, both of these are regular features of the text, as the entire novella

⁵⁶For example, we would wonder if the ‘diary style’ described by Haegeman (1990, 1997) and Haegeman and Ihsane (2001) is using the same kinds of mechanisms of displacement, setting Pieces to one side, and so on, in conjunction with an output expectation for some kind of reduced style. In answering this question, we would be testing the boundaries of the present theory.

doesn't contain a single occurrence of *the* or *a*, or a single use of tensed *be* (see Thoms 2008: 73). The fact that these are strong regularities in the text may allow us to include them as part of the articulate output expectation for poetic language sentences in this particular context, introducing a condition like "do not pronounce articles" to the output condition; in doing so, we remove the blocking problem for such instances of multiple erasure, as sentences that lack this feature will not enter into competition. The given double erasure example would then win out in competition because its extra instance of non-regular erasure qualifies it as 'distorted,' and thus it satisfies this additional output expectation. Thus we still assume that the theory should be able to explain how we can interpret the regular cases of erasure, but their behaviour with respect to blocking is given an alternative explanation.

This way of explaining examples like (73) appeals to further enrichment of the output expectation, where an artificial condition like "do not pronounce articles" or "do not pronounce tensed auxiliary *be*" is represented alongside the basic condition for distortion introduced earlier. This may seem like an unusual way of accounting for regularities of this sort, as one might propose that these regularities are in fact dialectal features of the text in question; indeed this is the kind of explanation proposed by Levin (1967), discussed in section 3.1.⁵⁷ However, such dialect-based solutions are difficult to maintain in the face of the criticisms outlined in chapter 3, and even though some regularities might seem to admit such an analysis, it is clear that many other artificial conditions on the form of the language could not be described as dialectal features. As mentioned above, by making this "do not pronounce articles" condition a condition in the articulated output expectation we capture the fact that these are regular deviations (regular uses of something that is deviant in the standard language) that occur alongside a number of irregular deviations; they are also similar to the non-regular deviations (examples of erasure) and invite a unified analysis

⁵⁷Note that this issue does not bear upon the present analysis of blocking: if we were to accept the dialect analysis for the present case, it would actually remove the problem of blocking, just like the account proposed here.

of the kind that could not be given by a dialectal analysis.⁵⁸ Finally, there is also motivation from the literary side of things, since it was a stated goal of Beckett's to render in *Worstward Ho* his most pared-down style of writing, and the necessity of finding such a minimal mode of expression is one of the core themes of the text itself. Therefore it seems clear that this regularized erasure is part of a consciously-crafted style of writing rather than a dialect.

The second issue is the recoverability of erased elements. As mentioned in section 2.2.1 and above, ellipsis is subject to an antecedence condition known as 'recoverability,' which requires that a given ellipsis site has an appropriate linguistic antecedent. This was one of the core insights of Hankamer and Sag's (1976) seminal study, which showed that the anaphoric dependency between ellipsis and antecedent is different from other anaphoric dependencies, like that between a pronoun and its referent. This is demonstrated by the following pair (partly repeated from footnote 9 above):

- (74) a. [Hankamer attempts to stuff a 9-inch ball through a 6-inch hoop]
Sag: It's not clear that you'll be able to do it.
b. [Hankamer attempts to stuff a 9-inch ball through a 6-inch hoop]
Sag: #It's not clear that you'll be able to.
(Hankamer and Sag 1976: 392)

In (74a) the pronoun *it* can easily refer to the ongoing action in the context, just like *he* can refer to a male individual in a given context, but this is not as readily available with the ellipsis site, which seems to need to establish an anaphoric dependency with an antecedent, as in (75):

- (75) Hankamer: I'm going to stuff this ball through this hoop.
Sag: It's not clear that you'll be able to.
(Hankamer and Sag 1976: 392)

⁵⁸ This is attested by the fact that the omitted elements do not form a unified syntactic category: for example, while *Worstward Ho* has no examples of finite auxiliary *be*, it does contain examples of the finite auxiliary *have*; likewise with determiners, since it contains examples of the indefinite determiner *some*, which has a great deal of syntactic and semantic similarities with *a*, which is systematically omitted. If the omission of finite auxiliaries and indefinite determiners was a dialectal feature of Beckett's language in the text, we would expect these elements to be omitted uniformly.

Thus Hankamer and Sag (1976) and many since have assumed that ellipsis is only possible if the elided phrase can be in an identity relation with a given linguistic antecedent. This is the condition of recoverability; when there is no antecedent, the meaning of the ellipsis cannot be recovered. In recent years, a lot of research has converged upon the view that this identity relation is a semantic one, although this remains an issue of some debate.⁵⁹ From hereon I will assume that recoverability is a semantic condition on ellipsis, and therefore that ellipses that do not satisfy the recoverability condition are semantically ill-formed in some way.

We may expect that many of the examples above of erasure should be semantically ill-formed, since they are cases of ellipsis-like Deletion (i.e. ellipsis without a licenser) and they occur without clear antecedents.⁶⁰ However, this need not be the case: the semantic identity relation that needs to hold between ellipsis and antecedent is not sensitive to the presence or absence of certain kinds of syntactic elements. Recall the following examples from section 2.2.2:

- (76) a. Fixing a car is easy if you know how ~~to fix a car~~ (Merchant 2001: 22)
 b. I know that John DID arrive late, but I don't know why ~~HE~~ arrived late
 c. There is someone at the door, but I don't know who ~~is at the door~~

The underlined elements in the antecedents are not present in the ellipses (and vice versa). This indicates that the recoverability condition is not sensitive to discrepancies between antecedent and ellipsis with respect to the presence of auxiliaries like *be*, *do* and *to*, pronouns like *he* and expletive nominals like *there*.

There is similar evidence from fragment sentences, which Merchant (2004) and others explain as instances of focus movement to Spec,FocP (in the CP layer) followed by ellipsis. Importantly, some fragments can occur without linguistic antecedents, so long as what remains in the ellipsis site is low in lexical

⁵⁹See Merchant (2001, 2008), Chung et al (1995), Chung (2006), Hartman (2009) and many others for discussion of the identity relation.

⁶⁰I do not provide the full linguistic context for all of the examples for practical reasons, but the reader can assume that such a context is provided by the preceding discourse of each example that has been discussed.

semantic content. So the fragment responses in (77) are assumed to have the underlying syntactic structures represented in (78) (all from Merchant 2004: 661-662), even though only the first example has a linguistic antecedent:

- (77) a. Abby and Ben are at a party. Abby asks Ben about who their mutual friend Beth is bringing as a date by uttering: “Who is Beth bringing?” Ben answers:
“Alex.”
- b. Abby and Ben are at a party. Abby sees an unfamiliar man with Beth, a mutual friend of theirs, and turns to Ben with a puzzled look on her face. Ben says:
“Some guy she met at the park.”
- c. Abby and Ben are arguing about the origin of products in a new store on their block, with Ben maintaining that the store carries only German products. To settle their debate, they walk into the store together. Ben picks up a lamp at random, upends it, examines the label (which reads *Lampenwelt GmbH, Stuttgart*), holds the lamp out towards Abby, and proudly proclaims to her:
“From Germany! See, I told you!”
- (78) a. Alex, ~~Beth is bringing~~ *t*.
- b. Some guy she met at the park, ~~he is~~ *t*.
- c. From Germany, ~~it is~~ *t*.

Examples like these conflict with Hankamer and Sag’s assertion that ellipsis needs an antecedent.

However, Merchant (2004: 716-732) argues convincingly that it is the strict interpretation of Hankamer and Sag’s proposal that is wrong, rather than the ellipsis account of fragments. He argues that examples like (74b) are subject to variation in judgments, at least with respect to the interpretation where there is a “minimal ellipsis” of a constituent like *do it*; while Hankamer and Sag seem to judge such an interpretation infelicitous, Merchant finds it fine and argues that such an interpretation is available for many similar ellipses without linguistic antecedent. He does not reject Hankamer and Sag’s point outright though:

[I believe] there is a real truth lurking behind Hankamer’s intuition. . . That truth essentially is that only the VP *do it* can be made manifest enough to antecede an ellipsis; other linguistic descriptions of pragmatically salient eventualities, with particular lexical items and other structure-specific properties, cannot. This may be due to

the general pragmatic fact that any given situation will support a large number of mutually compatible specific linguistic descriptions, and deciding which among these might be intended by a user of ellipsis is simply impossible. The general action description *do it*, however, subsumes enough of the possible descriptions (all of them, in fact, except statives), that it is appropriate in any of the contexts.

(Merchant 2004: 722)

Merchant concludes that the implicit antecedent *do it* can be made manifest in a number of contexts, while richer antecedents containing lexical content cannot, at least not in the same fashion. Thus in effect he argues that Hankamer and Sag's example (74b) can be interpreted as (79a), but not as (79b):

- (79) a. It's not clear that you'll be able to ~~do it~~.
 b. It's not clear that you'll be able to ~~stuff the ball through the hole~~.

He then argues that a set of 'fossilized expressions' similar to *do it* can also be made manifest as antecedents for ellipsis, such as *he is* and *it is* in the cases of (78b) and (78c) respectively. A similar account could be given for cases of nominal ellipsis, where the missing nominal is interpreted as a proform like *one*, as in the nominal subject in (80), which is interpreted as (81a) but not (81b) or (81c):⁶¹

- (80) John is in a car showroom, looking appreciatively at a line of shiny new Porsches of different colours. Noting that John seems particularly interested in the silver model salesman comes over to him and utters:
 "The silver is our best-seller."

- (81) a. The silver ~~one~~ is our best-seller.
 b. The silver ~~car~~ is our best-seller.
 c. The silver ~~Porsche~~ is our best-seller.

We may assume, then, that such items are always available as antecedents in sufficiently rich discourses. These items are the same kinds of items that can vary between ellipsis and antecedent, and they have minimal lexical semantic content.

⁶¹Note that this preference for the proform cannot be put down to a simple condition of avoiding redundant use of lexical items, as may be proposed for use of pronouns in subsequent sentences (*John arrived. He was late*). This is obviously because no one has used the words *car* or *Porsche* yet, at the time of utterance.

Ultimately we can conclude that recoverability is significantly less strict when it comes to the omission of functional items and lexical items that have minimal semantic content. This explains the fact that our examples of erasure do not fall foul of recoverability, as they all involve these kinds of elements being Deleted just as they are in similar situations in ordinary ellipsis cases. Given this, we are now in a position explain an important fact about erasure observed in the description in chapter 2: it typically targets linguistic elements with little or no lexical semantic content. In the few cases where there is no obvious antecedent (see the discussion of (213) and (214) in section 2.2.3), the low-content meanings are assigned to the erased constituents, just like in ordinary use of language; in these cases we do not as much have violations of recoverability as default cases. We can thus assume that the semantic condition that constrains ordinary use of ellipsis also constrains erasure.

Some cases, like erasure of prepositions, are not obviously subsumed by this explanation. Consider (61) from above:

(61) The kids came and we all went ___ the briars.

However these are easily understood since the prepositions in question are selected by the verbs; thus the presence of the lexical item *to* is encoded by the verb's meaning just like the presence of a nominal complement is encoded by a verb like *hit*; the difference is the specific lexical head of the complement is encoded by *go* in (61), whereas *hit* only selects for the category of NP.⁶² Therefore these do not involve violations of recoverability as such, since the meanings of the ellipses are clearly recoverable from the linguistic context; the fact that this linguistic context is the immediate context, rather than a separate antecedent, is not important.

⁶²Note that recoverability would be problematized if the verb that selects the preposition can select numerous other prepositional complements, as in the following example:

- i. It was Christmas and I got a free meal ___ Max.

The verb *get* can select various prepositions that would be appropriate with the complement *Max*: *for*, *from*, *with* and, less obviously, *through*. This means that the example is ambiguous, and that the linguistic context isn't enough to help.

The fact that erasure is subject to recoverability just like standard ellipsis is not a trivial point, as this might not necessarily have been the case. Indeed we have seen that poetic language derivations are allowed to avoid satisfying certain standards at the PF interface (namely, economy); given this, we might have expected that poetic language is also capable of violating semantic conditions like recoverability. This would necessitate a reanalysis of the theory proposed here, since this assumes that the LFs that are produced by poetic language derivations are the same as those in ordinary derivations. The fact that this is not required is a confirmation of this aspect of the theory, and an indication of the character of poetic language more generally: it allows for deviation in phonological form, but not for deviation in semantic form. This is a significant theoretical point that I will discuss in more depth in the final chapter.

4.3.4 The revised theory

In the previous subsection, we saw that, in order to account for the cases of erasure in poetic language, we would have to generalize the theory of poetic language beyond the specific model proposed as ‘Bypass Theory’ in (7); I repeat that formulation here for the reader’s convenience:

- (7) Bypass Theory: poetic language is produced by PF Concatenation and LF Combination of independently-formed Pieces of syntactic structure, bypassing the generation of a full syntactic structure.

This formulation is only appropriate for deriving cases of displacement, as we saw above; erasure examples may involve steps corresponding to this procedure (as in the cases of non-constituent erasure), but it is not sufficient to explain those examples fully. Erasure examples also involve instances of unlicensed application of the operation Delete, and this is not mentioned in (7). We also saw that the erasure derivations and the ones for displacement proposed in (7) do still share a common core, in that they are both kinds of derivations which violate derivational economy in the PF branch. Indeed this is the only factor that

distinguishes the poetic language derivations from ordinary language derivations, since they do not involve any other operations or derivational steps that are not independently required. We can restate the theory of poetic language to capture the two different kinds of derivations in the formulation in (82):

- (82) The PF Theory of Poetic Language: poetic language sentences are produced by convergent but PF-uneconomical derivations.

This covers both displacement and erasure, thus providing us with a unified theory. (7) supplements this as a way of spelling out the precise details of a displacement derivation, but it is superseded by the theoretical proposal in (82).

(82) is intended to sum up the discussions of this chapter so far, but as a broad theoretical proposal it necessarily opens up a whole new set of questions. In particular, on the basis of this proposal we might predict that poetic language will also allow for other standard PF operations to apply excessively in poetic language derivations. For example, if we are to accept that verb movement in English is a PF phenomenon, we may predict that we will find poetic language derivations where this operation of PF verb movement applies without convergence/feature-checking motivation. If PF movement exists, such derivations are in principle possible but filtered out by derivational economy. In Chomsky's set-theoretic terms, these derivations would correspond to the complement of a subset of set of D_C (convergent derivations) and D_A (admissible derivations), where the relevant subset is those economy violations that occur due to unmotivated PF operations;⁶³ this is the same set as those corresponding to poetic language derivations.

The question of whether this is a problematic prediction, then, depends on what kinds of PF operations are possible in English and whether their unmotivated use would generate structures that are not attested in the poetry. We can consider this by looking at some of the proposals in the literature one by one; here I will restrict my attention to the two most important cases for the present study, phrasal movement and head movement. First, consider Sauerland and

⁶³This isn't strictly true, as there may be competition between these derivations, just like there was with poetic language.

Elbourne's (2003) proposal for phrasal movement. Sauerland and Elbourne argue that some kinds of phrasal movement, namely A-movement that undergoes 'total reconstruction,' can be analysed as cases of PF phrasal movement, and they argue that Japanese scrambling can be given a similar analysis. The empirical basis of Sauerland and Elbourne's proposal is far from solid – see Thoms (2009) for a critique of the evidence from reconstruction in British English and Miyagawa (2005) for evidence against the account of Japanese scrambling – but this is besides the point, as the analysis does not pose problems for the present theory: derivations corresponding to displacement would produce phonological forms homophonous with those produced by this operation. Recall however that we cannot simply replace displacement with PF movement or some other similar operation, since if we did so we would still lack an account of the cases of non-constituent displacement. Therefore this kind of PF movement would not introduce any empirical problems for the theory.

Now consider the more problematic of verb movement. Following the lead of Chomsky (2000), Boeckx and Stjepanović (2001) propose that head movement operations should not be analysed as cases of narrow syntactic movement (as in Pollock 1989, Chomsky 1993, 1995b and many others) but instead as an instance of PF movement. Their primary evidence for this proposal is an analysis of Lasnik's (1999) account of pseudogapping in English, which proposes that verbs do not raise in pseudogapping constructions, but otherwise the evidence for head movement as PF movement is negative: since it has no obvious evidence of head movement affecting interpretation, there is no need to presume that it occurs in the narrow syntax; therefore it can be shifted to PF to remove certain theoretical problems, such as the fact that head movement seems to violate the Extension Condition.⁶⁴ Unlike with the phrasal movement case above, adopting PF head movement would cause much more problems for the present theory, as such forms of unlicensed movement are not attested in the poetic language data; indeed it was taken to be a virtue of the displacement theory that it

⁶⁴Recall from section 4.2.5 that I do not assume the Extension Condition here, given the arguments from N. Richards (2001) for rejecting it.

prevented head displacement. Therefore the arguments to be provided here must be arguments against analysing PF movement as a suitable analysis for head movement, at least in English.

As it happens, there are many such arguments. First, the theoretical arguments against head movement in narrow syntax have been addressed by numerous authors already: for example, Roberts (2005, 2010) and Matushansky (2006) provide accounts of head movement that does not violate the Extension Condition and show that PF movement accounts of head movement are just as problematic in theoretical terms as the narrow syntactic accounts. Second, Matushansky (2006: 99-101) notes that head movement is typically very unlike a PF operation, since it is not conditioned by phonological properties. Note that Matushansky's critiques do not apply to all analyses of head movement as PF operations, but rather the hypothesis that *all* head movement is in the PF branch. There are several analyses in the literature which derive the effects of apparent head movement phenomena from phonological or morphological processes that move or rebracket the output of the syntax: see Marantz (1988), Embick and Noyer (2001), Adger (2006) among others. The important thing for the present analysis is that the PF operations in these accounts all involve extremely local operations over morphological structure in languages with much richer morphology than English; as such, we may expect that the kinds of processes that are analysed in these cases are likely to be largely irrelevant to the analysis of English in the present situation.

Third, the negative argument against narrow syntactic head movement is not compelling, as there are a number of proposals in the literature for semantic effects of head movement: for example, Lechner (2006) proposes that head movement is crucially implicated in the availability of split readings of negative universal quantifiers (*not everyone can be above average*), Truckenbrodt (2006) argues that head movement to C in German is motivated by semantic properties of force, and Hartman (to appear) diagnoses semantic effects of T-to-C movement of auxiliary verbs using the test of ellipsis parallelism. There is also

simple evidence for semantic effects of head movement from negative polarity item (NPI) licensing in English, as attested by the following minimal pair (David Pesetsky, personal communication):

- (83) a. Which book didn't anyone read?
 b. ?*Which book did anyone not read?

In (83a) movement of the verb bearing negation allows for licensing of the NPI *anyone*, whereas in (83b) negation is not pied-piped to C and as a result the NPI is not licensed. Given that NPI licensing is a semantic effect, this is unexpected if head movement is a PF phenomenon, as there would be no difference between the two cases at LF.

Fourth, the evidence from Boeckx and Stjepanović (2001) from pseudogapping is wholly dependent upon two specific proposals that have been challenged in the literature: Lasnik's (1999) analysis of pseudogapping as movement to AgrOP followed by VP ellipsis, and a VP structure for English in which the object undergoes overt movement and the verb moves over it. Gengel (2007) has shown that the movement in pseudogapping is better analysed as a case of leftward movement to a focus projection in the IP field below the subject position, taking precedent from Jayaseelan's (2001) analysis (which finds precedent for such focus projections in the analysis of Malayam scrambling). Gengel argues that this is movement is then followed by standard VP-ellipsis, but Thoms (to appear) shows that the VP-ellipsis analysis of pseudogapping is inadequate, since many languages other than English have pseudogapping even though they lack VP-ellipsis. The following pair is from Norwegian:

- (84) *Mary vil gi mange penger til Susan og Paul vil til Jane.*
 Mary will give much money to Susan and Paul will to Jane.
 'Mary will give much money to Susan, and Paul will to Jane'
 (Gengel 2007:)

- (85) **Mary vil gi mange penger til Susan og Paul vil, også*
 Mary will give much money to Susan and Paul will, too

'Mary will give much money to Susan, and Paul will, too'
 (Terje Lohndal, personal communication)

(85) contrasts minimally with (84) and it shows that the deletion process in pseudogapping is not the same one that is involved in VPE, since it is generally not available in Norwegian. Thoms (to appear) proposes instead that what we see in pseudogapping is better analysed as a case of “sub-slucing,” where it is the movement of the pseudogapping remnant to the focus projection that licenses ellipsis, just as movement of the *wh*-phrase in sluicing licenses ellipsis in that case (as in Thoms to appear’s analysis). The important point for the present study is that this analysis of pseudogapping renders Boeckx and Stjepanović’s (2001) empirical argument irrelevant, since the updated analysis of pseudogapping does not implicate verb movement as a PF phenomenon.

Considering the balance of these arguments, I believe it is justified to assume that head movement cannot be reduced to PF movement, but rather many of the head movement operations we see in English syntax are derived in the narrow syntax. Given this, we do not need to worry about overgeneration of PF head movement derivations as a problem for the theory proposed in (82). Since the other case of phrasal movement does not present problems either, I conclude that (82) does not incur any obvious overgeneration problems by allowing PF movement operations to apply unmotivated in poetic language derivations.

4.4 Summary

In this chapter I have developed a theory of poetic language in which poetic language derivations are PF-uneconomical but convergent derivations that are generated by the grammar. First I argued that the theory of poetic language should be guided by the ‘Non-Uniformity Hypothesis,’ which proposes that poetic language sentences and ordinary language sentences should not be generated in the same way. I then argued that displacement is derived by setting Pieces of syntactic structure to one side in the narrow syntactic derivation and then LF-Merging these Pieces into their appropriate positions and PF-Concatenating the respective Pieces to produce a unitary, linearizable phonological form. I ar-

gued that erasure is derived as unlicensed application of the operation Delete, where ‘unlicensed’ means the application of the operation is not required for convergence.

I showed that the theory as stated is capable of deriving the data presented in chapter 2, and I also showed that the form of the theory also allows us to explain the non-occurrence of other possible kinds of derivations, such as those involving head displacement. I proposed that the fact that poetic language sentences are allowed to occur is the result of an extra output expectation at the PF interface, which dictates that the phonological form of the poetic language sentence be distorted in some way. I suggested that a more articulated model of the output expectation may allow us to explain the interaction of poetic syntax and other aspects of formal organization in poetic language without falling foul of the problems that troubled the narrow syntactic accounts reviewed in chapter 3.

The theory does not require us to propose any stipulative forms of movement or alterations of the grammar. Rather, the theory follows entirely from core Minimalist assumptions about the nature of derivations: specifically, that the narrow syntactic derivation is optimal way to produce sound-meaning pairs, and by taking alternative routes we necessarily produce sub-optimal derivations that correspond to deviant sentences. The theory proposed here can be described as a “phonological theory of poetic language,” since it proposes that the variations we see in poetic language syntax are ultimately the result of different processes applying in the PF branch of derivations.

Chapter 5

Conclusion

Chapter 4 presented a new theory that aims to explain the kinds of linguistic deviation we find in poetic language. In this chapter I conclude by summarising the thesis' findings, putting the theory in the wider context and discussing some of its implications.

5.1 The formal character of poetic language

A significant part of this dissertation is devoted to describing the formal characteristics of poetic language. Chapter 2 provided a comprehensive overview of the different kinds of deviation we find in poetic language: I described two large sub-types of deviation, displacement and erasure, and I discussed their nature and distribution. Here I will summarise this description and its implications for the theory of poetic language, as discussed in chapter 3.

Displacement is when a syntactic element is displaced from its standard position to some other position in the sentence by nonstandard means. Displacement is perhaps the best-known kind of deviation in poetic language, and the survey in chapter 2 showed that it affects all different kinds of constituents – NPs, PPs, APs, TPs, adverbs – and often more than one at the same time. Perhaps the most important finding of chapter 2 was that the way in which displace-

ment affects these elements is very different from the way standard operations like Move affect them. Displacement of NPs and PPs often resembles standard movement operations like topicalization, but we saw that displacement affects phrases in environments where we don't normally get topicalization, such as in relative clauses or *if*-clauses. We also saw that displacement could affect more than one element at once, sometimes allowing them to reverse in order, and that displacement was seldom to a fixed position with respect to other elements in the structure (such as adverbs and auxiliaries). Crucially, we also saw that displacement affected non-constituent elements, and it also shifted constituents to positions in the sentence that would not be accessible by ordinary movement operations.

We concluded that the application of displacement was largely unconstrained, although we saw that there were significant gaps in its distribution; that is, we saw that it tended not to affect syntactic heads, despite the fact that standard movement operations often affect heads. From a pre-theoretical point of view this was unexpected, since it seemed that displacement was a largely random process of reordering; the fact that it was nevertheless limited to affecting non-heads was a surprise, one that had not been discussed previously in the literature on poetic language. Importantly, in chapter 3 we saw that it was impossible to describe these characteristics of poetic language in terms of movement operations already available in ordinary language, and that previous attempts to do this, approaches which I characterized as “poetic grammar” theories of poetic language, were unfeasible for both empirical and theoretical reasons. In chapter 4 I proceeded to develop an alternative explanation that captured these facts.

The other sub-type of deviation described in chapter 2 is erasure, which involves the deletion of elements in a sentence by nonstandard means; that is, descriptively it is the nonstandard equivalent of ellipsis, just as displacement is (descriptively) the nonstandard equivalent of standard movement. We saw that there are many different kinds of erasure in poetic texts: some kinds resembled “failed” versions of standard ellipsis, such as the cases of backwards gapping,

and others were partly systematic in given texts (in particular in experimental texts). We saw that erasure typically targeted functional elements such as auxiliaries and articles, rather than elements with lexical semantic content; this was attributed to the fact that erasure is largely restricted by the ellipsis condition of Recoverability. We also saw that erasure would sometimes target non-constituents, although this kind of deletion was similar to others found in ordinary language.

An interesting characteristic of erasure is that its distribution is almost complementary with displacement, since it frequently affects heads and seldom affects phrasal constituents. The fact that the two major types of deviation are significantly different in terms of their distribution was taken to be a significant characteristic that the theory should account for; that is, it argued against a fully unified theory of poetic deviation. In the end, the theory proposed that erasure is much closer to its ordinary language equivalent than displacement is: whereas displacement involves a very different set of derivational steps, erasure is in effect an “overapplication” of the standard operations involved in ellipsis. We noted, however, that erasure is not entirely unconstrained, as the semantic condition of Recoverability still applies;¹ rather, only constraints related to the phonological side of its application are lifted (in a sense that is defined in section 4.3.3). Importantly, erasure and displacement still received a unified explanation in the theory developed in chapter 4, and it was argued that the two kinds of derivations may even overlap in the situations involving the erasure of non-constituents.

To conclude, deviation in poetic language is much more diverse than has been previously assumed in the literature, and the discussion in chapters 2 and 3 showed that this diversity cannot be explained entirely in terms of ordinary language operations. In many ways this is to be expected, since we know that,

¹There were examples where recoverability did seem to be violated, but we may recall that the relevant examples were more difficult to interpret. Furthermore, in most cases the interpretations given to the examples were in preference to semantically deviant alternatives that might have been available in different linguistic contexts; see the discussion of (220) in section 2.2.3 in particular. The status of these examples is not wholly clear, but it does seem that they are different in some way from the other cases discussed in chapter 2.

in experimenting with the forms of language, the poet is often deliberately distancing herself from the constraints imposed by ordinary language in an attempt to find a new and more vital mode of creative expression; that this would involve doing things that aren't done in ordinary language situations is perhaps not surprising. Nevertheless, it is not certain that stepping beyond the rules of ordinary language actually does allow one to find a new way of creating meaning with language. In the next section I discuss this question in more detail, considering how it relates to the theory proposed in chapter 4.

5.2 PF Economy, possible derivations and the medium of language

The theory developed in chapter four was described as a “PF Economy” theory of poetic language, because it identified the sole difference between poetic language and ordinary language is that the former involves derivations which are uneconomical in the PF branch of the derivation. The theory concludes that the differences between poetic language and ordinary language are only phonological: in effect, the poetic language derivation does “something extra” to the phonology of the sentence, on top of what would normally be done in a standard derivation. We also saw that the theory correctly predicted a number of kinds of derivations to be impossible, such as those involving head displacement or erasure of material containing lexical semantic content without a proper antecedent. These two cases had different explanations: head displacement was ruled out because it would involve a derivation which failed to converge at the LF interface, and erasure of lexical content was ruled out because the deletion process was subject to the inviolable semantic condition of Recoverability. These facts about the theory have a number of interesting implications.

This point is particularly relevant for the literary theoretic view, suggested by the quote from Samuel Beckett in chapter 1. Beckett's quote represents a wider concern that is present the discussion of poetic language by theorists

and practitioners alike: the question of whether one can extend the expressive potential of language by experimenting with form or breaking its rules; or, the question of whether we may be able to do something new or special with language by stepping out of the confines defined by its rules. This ideology was present in Beckett's quote, in the rhetoric of the Futurists in Italy and Russia in the 1910s, in the linguistic gnosticism of the Dadaist Hugo Ball and his contemporaries, and in the practice and theory of many others who have followed them since. It was also present in the linguistic skepticism of Jacques Derrida's deconstructionist approach to literary theory, as elucidated by Attridge's (1988) Derridean look at the history of approaches to poetic language, discussed briefly at the outset of chapter 1.

The theory developed in this dissertation in effect proposes that one is not able to alter the expressive potential of language by experimenting with its form, on the assumption that altering this "expressive potential" would involve creating meanings or forms of communication that are not ordinarily available in the language. This is because the theory proposes that experimentation is not capable of altering semantic structure in any ways that are not available in ordinary language, and it is represented technically in the theory by the fact that the poetic language sentences are defined by a set of possible derivations, PD_P (the members of which then compete amongst one another based upon PF-economy factors). All of the members of PD_P are convergent and generated by the standard grammar, and as such they are all interpreted as standard semantic structures at LF. In effect, the poet is limited to using the sentences that are defined as possible by the grammar; she cannot create new rules, nor can she create new forms that are not usually possible, but rather all she can do is "a little extra" in terms of the phonological form of the sentence. Thus creative use of language is fundamentally constrained by the grammar, even in the extreme case of experimentation with form in poetry.

There is a question of whether this aspect of the theory is in fact a discovery about the character of poetic language, or just an artifact of the scope of

inquiry. Recall in chapter 1 that I argued for a number of restrictions on the scope of the theory: in particular, we ruled out trying to account for deviation in conceptual semantic structure (i.e. metaphor), and we ruled out analysing examples that could not be interpreted as sentences of English. It may be that, by ruling out these cases, we necessarily restrict ourselves to data that will not diagnose any kind of semantic deviation. Nevertheless, these two cases were ruled out on principled grounds because they are essentially unanalysable in terms of linguistic theory. Both kinds of deviation pose methodological problems, since their deviation cannot be adequately described and analysed using the tools of linguistic theory. We saw that the diversity of use and interpretation with metaphor took us well beyond the remit of a theory of poetic language; indeed it became apparent that the relationship between conceptual semantic deviation and metaphorical interpretation may only be indirect at best, and thus unanalysable in terms of the discrete computations of the grammar. We also saw that data that cannot be analysed as sentences of English is in effect unknowable as “linguistic data,” since some sort of truth-conditional semantic interpretation is a pre-requisite for calling a given datum “linguistic.” Without such a restriction, the theory threatens to become meaningless.

The importance of this point about the theory can be understood more clearly if we outline what we may have found in the poetic language data if it were in fact possible to create new meanings. Consider the following prediction: if poetic language derivations could create new meanings, we may predict that this would allow us to separate more or less economical derivations at the LF-interface and thus generate the unattested forms like head displacement. It is possible that the LF syntax could make available a set of operations similar to the PF operation of formal feature deletion (FF-deletion, as in Nunes 2004; see section 4.2.7) that would adjust or correct a semantic structure to make it interpretable; for example, the problem caused by LF-Merging a head in an “in between position” (discussed in section 4.2.5) might be solved by an operation of resetting the address on the problematic syntactic constituents. Assume that

such an operation exists.² This operation, which we can call ‘LF Correction’ for now, would apply in the LF branch of the derivation and we would thus expect it to allow head displacement to occur, but only in situations where it wasn’t blocked by economy: that is, when the derivation D_1 that produces the head movement structure with an extra application of LF Correction is not blocked by another derivation D_2 that does not involve an instance of this (i.e. the ordinary language equivalent).

However, for this to happen there would have to be some difference between D_1 and D_2 with respect to the output expectation at the LF interface; that is, they would have to be semantically non-equivalent in some way. Recall that with the cases discussed in sections 4.2.6-4.2.7, the less economical derivation would be allowed because it satisfied an additional output expectation at the PF branch for some kind of deviation or regulation of form (i.e. metricality). Differences in the phonological form would be irrelevant for the calculation of economy differences between D_1 and D_2 , however, as these calculations would occur in the LF branch where these differences would be irrelevant; therefore, D_1 and D_2 would have to lead to different meanings, as this is the only way the outputs at LF can be teased apart. This would require the poetic language derivations to be creating meanings that are not possible with the ordinary language derivations. The fact that the sentence corresponding to D_1 does not occur implies that the poetic language derivation D_1 does not lead to a meaning that is not expressed by the ordinary language derivation D_2 . Therefore, the poetic language derivations cannot create new meanings in these situations.

There may be many other kinds of forms that we would expect to find if

²There are two different kinds of precedent for this. One is reconstruction as conceived in Thoms (2010a): this is an operation that deletes the address of a given element, which allows general principles to prefer and interpret a lower copy. Reconstruction occurs at LF in that system, and as an address-altering operation it is broadly similar to the mooted operation. Another set of more distantly related precedents are the “type-shifting” semantic operations proposed by Partee and Rooth (1983) and many others since (Groenendijk and Stokhof 1989, Chierchia 1998, Winter 2001); these are “last resort” operations that are used to avoid undesirable meanings as derived from the standard operations of the grammar, changing the semantic type of certain elements to higher-level types. Both types of operations are costly and only used to ensure that a derivation converges upon a particular interpretation.

poetic language derivations could create meanings that are not possible with ordinary language, but the preceding discussion is enough to make the point clearly: *any* kind of change of meaning (among those that can be formulated in terms of LF structure) would be relevant to the case just discussed, as it would separate the problematic derivation D_1 from D_2 , yet there are no such differences, or at least no such differences that can be apprehended and formulated in an output expectation. This case doesn't diagnose the possibility of creating new meanings by describing a set of attested interpretations that are not normally available, but instead it shows that, were there such a set of interpretations, they would interact with the computations of the grammar in a certain way. This allows us to test for poetic language-only meanings without having to tackle a number of difficult data issues, namely specifying exactly how the interpretations given to poetic language sentences differ from those given to ordinary language sentences.

With this point established, we might then go on to speculate about *why* poetic language only allows for PF-related manipulation and not manipulation of semantic structure. I will not go into this issue in great detail here, but instead I will mention just one point that is relevant to this speculation. Consider the role of the output expectation. Recall that in the derivation of poetic language sentences involving displacement, the output expectation may be characterized crudely as "be distorted" or "be in iambic pentameter." The important thing is that these output expectations can be apprehended consciously by the language user, as they do not appeal to abstract notions of grammatical structure; rather they are meta-linguistic descriptions of the surface form of the sentences, and they are easy to manipulate and to use to compare sets of derivations. Now consider the case with an altered semantic structure: how does one describe a poetry-specific semantic structure for the purposes of the output expectation? One can only use sentences of ordinary language to describe the output expectation, and as a result it will be impossible to state an output expectation that will adequately separate the ordinary language meaning from the "exceptional"

poetic language meaning. PF-uneconomical derivations are possible because ordinary language can be used to describe non-ordinariness in phonological form; on the other hand, LF-uneconomical derivations are not possible because one cannot use ordinary language to describe the non-ordinariness of “special” meanings. Or, in other words, one is trapped into using language when it comes to defining an output expectation that will define whether or not a given derivation is licit, and this necessarily restricts these output expectations to second-order statements about the surface form of the sentences; the meta-linguistic statements about the form of the output expectation are themselves bound by the possibilities of language. To paraphrase the quote from Derek Attridge in the introduction, there is nothing in the poet’s armory but language, and as a result her experiments with language are fundamentally constrained.

What the preceding discussion shows is that the linguistic theory of poetic language is able to “feed back” into literary theory in a meaningful way. This is a real result, since the relationship between linguistic theory and literary theory has always been a problematic one, in particular with the case of stylistics. Work in stylistics takes empirical generalizations about linguistic usage and uses them as the basis for claims about the interpretation of given texts. However, this kind of work has been criticized heavily by those working within literary theory, with the most notable critiques coming from Stanley Fish (e.g. Fish 1980). Fish argues that the logic of stylistics is fundamentally flawed, since generalizations about the relationship between style and meaning necessarily begin with unfounded assumptions about the meanings associated with stylistic traits; thus, the claim is that stylistics theories are self-fulfilling prophecies of a kind which do not discover anything about literature other than the prejudices of the scholar (see Stockwell 2002 for a recent defense of stylistics against such criticisms). As a result there are many points of tension between linguistic approaches to literature and the mainstream theoretical approaches, and many works of stylistics set themselves in opposition to postmodern literary theory.

The theory presented in this dissertation is not implicated in this tension,

however, since it is not a work of stylistics. The theory proposed here is entirely formal, eschewing the functionalism of stylistics work, and it has not made any claims about the relationship between the form of language and literary interpretation (beyond standard assumptions about semantic compositionality). This means that generalizations drawn from this theory do not necessarily clash with the fundamental assumptions of literary theory, and therefore that literary theory may be informed by the proposals developed here.

5.3 Poetic language and linguistic theory

The previous section sketches the ways in which the proposals developed in this dissertation may feed back into literary theory, and this section does the same with the other adjacent discipline, namely linguistic theory. The theory presented in this dissertation has been developed within the broad framework of Minimalism, and it has extended the empirical coverage of such theories significantly, since there have been no works in this framework that have been designed to deal with this empirical domain. Successful explanation of this empirical domain must therefore be taken to be evidence for the particular syntactic proposals developed within, just as it might be in the explanation of a previously unstudied set of constructions or family of languages. I will discuss some of the specific syntactic proposals in turn, before discussing the wider empirical and theoretical implications.

The explanation for displacement is built upon a set of syntactic assumptions that received independent support, so the success of that component of the theory in describing the range of possible displacements and their limitations should provide evidence for this set of assumptions. First, it was assumed that Pieces of structure can be “set to one side” in a derivation by setting of the syntactic address to [ELSEWHERE]; this was required to ensure that the derivation can proceed without fully integrating all of the lexical items in the narrow syntax, since other Each of the parts of this theory receives independent

support: the notion of the syntactic address was proposed by McGinnis (2004) to deal with interactions between binding and movement in what are known as “Chain Condition” violations (Rizzi 1986); Thoms (2010a) extends the use of the address and in doing so accounts for differences between overt and covert movement with respect to binding, while also explaining Abels’ (2003) anti-locality constraint in a principled manner and explaining why it doesn’t apply to covert movement; in the discussion in section 4.2.3 I also propose that this use of the address could be extended to deal with the problem of integrating parentheticals into their host sentences. The address-setting theory was also used in the explanation of LF Combination, since the theory required LF (Re)Merge to be costless. This was what was argued by Thoms (2010a). Given the fact that the address-setting theory allowed us to capture a number of facts about the distribution of displacement, we can say that the explanation of displacement provides further support for the notion of the address used here, and related issues like the decomposition of Merge.³

The explanation for erasure is built upon the economy-based theory of ellipsis licensing proposed in Thoms (to appear), where ellipsis is proposed to be a form of copy deletion a la the Copy Theory of Movement. Specifically, Delete is proposed to be a costly operation and ellipsis is taken to be a way of deleting a lower copy in a movement chain in a non-local configuration. Thoms (to appear) presents a number of empirical and theoretical arguments for this proposal, showing that previous approaches to ellipsis licensing are both stipulative and descriptively inadequate. In section 4.3.3 I argued that erasure is simply application of Delete in situations where it is not required for convergence, and in doing this I provided a unified account of poetic language in terms of PF economy, where the possibilities of poetic language are specified as the set of

³The theory also required the operation of PF Concatenation, which was taken to be the same operation as the one that inserts parentheticals, and to be analogous to the other “PF Remerger”-type proposals in the literature. This allowed us to describe constraints on the distribution of displacement, which resembles the distribution of parentheticals in only ever occurring between phonological phrases. This correlation was not explained, however, but only captured by a descriptive generalization. I assume that the correlation may be explained by an adequate theory of the mapping from syntax to prosodic phonology, such as that proposed by Selkirk (1984, 1995).

derivations that are uneconomical in the PF branch. If that unification can be seen to be successful, we may thus argue that the coverage of the theory thus lends support to the economy-based description of ellipsis licensing in Thoms (to appear).⁴

Looking beyond the specifics of these aspects of the analysis, the proposed theory provides evidence for the interface-based approach to syntax proposed in recent Minimalism, where the narrow syntax is an optimal solution to the design problem of matching sound and meaning. I have argued that there are other ways to compose sequences of sounds into meaningful elements than by standard derivations, but what is crucial is that doing so by some route other than the optimal one will necessarily be more costly in terms of calculations of economy. I have also shown that this makes a number of precise predictions regarding the interaction between the grammar and non-grammatical notions of form, modeled with respect to the notion of the output expectation, which was independently motivated (though not explicitly formulated) in the work of Fox (1995, 2000), Adger (1994), Reinhart (2006) and others. This showed that one can model the ways in which syntax interfaces with external pressures on well-formedness without jeopardising standard assumptions about the modularity of the grammar or the blindness of syntax: external factors do not impose upon the operations on the grammar, but they may influence the ways in which convergent derivations are ‘filtered’ with respect to their competing analogues.⁵ Ultimately, then, the theory of poetic language proposed here has shown that economy principles play a crucial part in defining well-formedness of the output of the grammar, thus adding empirical weight to the conjectures of Chomsky’s (1995b) Minimalism.

Finally, we may conclude that the discussion in this dissertation has shown

⁴It should be noted that the specific proposal in section 4.3.3 does not make any predictions about the distribution of erasure, or at least any predictions that would not be made by most other theories of deletion and ellipsis. Thus the fact that erasure tends to target heads is not taken to be evidence for or against the theory of ellipsis proposed in Thoms (to appear).

⁵This set of predictions was left somewhat underdeveloped in its specifics, but what the discussion showed was that an articulated model of the output expectation could sharpen the predictive force of the theory significantly.

is that the formal study of poetic language may make a significant contribution to linguistic theory. Linguistics is concerned with describing and explaining the human capacity for language, and it does this by analysing the diversity of linguistic form within and across languages and asking what general properties unite these diverse phenomena. We have seen here that poetic language represents an extreme case of diversity in linguistic behaviour, but we have also seen that under this diversity lies significant generality, and that this generality is of the kind that we find in many other realms of linguistic inquiry. As such, it is a fruitful area for linguistic inquiry, one that may teach us more about language more generally.

Bibliography

- Abels, K. (2003). *Successive cyclicity, anti-locality and adposition stranding*. Ph. D. thesis, University of Connecticut at Storrs.
- Abels, K. (2010). The italian left periphery: a view from locality. Manuscript, University College London, lingBuzz/001018.
- Abney, S. (1987). *The English noun phrase in its sentential aspect*. Ph. D. thesis, Massachusetts Institute of Technology.
- Ackema, P. and A. Neeleman (1998). Optimal questions. *Natural Language and Linguistic Theory* 16(3), 443–490.
- Ackema, P. and A. Neeleman (2004). *Beyond morphology: interface conditions on word formation*. Cambridge: Cambridge University Press.
- Adger, D. (1994). Meaning, movement and economy. In R. Aranovich and et al (Eds.), *Proceedings of WCCFL*, Volume 13, pp. 451–466. Citeseer.
- Adger, D. (2003). *Core Syntax*. Oxford: Oxford University Press.
- Adger, D. (2006). Post-syntactic movement and the Old Irish verb. *Natural Language and Linguistic Theory* 24(3), 605–654.
- Adger, D., C. de Cat, B. Plunkett, and G. Tsoulas (Eds.) (2007). *Peripheral positions: syntactic edges and their effects*. Dordrecht: Kluwer Publications.
- Aelbrecht, L. (2009). *You Have the Right to Remain Silent: The syntactic licensing of ellipsis*. Ph. D. thesis, HU Brussels.

- Agbayani, B., C. Golston, and T. Ishii (2009). Prosodic scrambling in Japanese. In D. Potter and D. R. Storoshenko (Eds.), *Proceedings of the 2nd International Conference on East Asian Linguistics*, Volume 2 of *Simon Fraser University Working Papers in Linguistics*.
- Alexiadou, A. (2001). Adjective syntax and noun raising: word order asymmetries in the DP as the result of adjective distribution. *Studia Linguistica* 55(3), 217–248.
- Alexiadou, A. and E. Anagnostopoulou (1998). Parametrizing Agr: word order, V-movement and EPP-checking. *Natural Language and Linguistic Theory* 16(3), 491–539.
- Aoun, J. and E. Benmamoun (1998). Minimality, reconstruction, and PF movement. *Linguistic Inquiry* 29(4), 569–597.
- Ashbery, J. (1979). *As We Know*. Manchester: Carcanet Press.
- Ashbery, J. (1997). *The Mooring of Starting Out*. New York, New York: Ecco.
- Ashbery, J. (2008). *Collected Poems, 1956-1987*. New York, NY: Library of America.
- Attridge, D. (1988). *Peculiar language: literature as difference from the Renaissance to James Joyce*. Ithaca, NY: Cornell University Press.
- Attridge, D. (2004). *The singularity of literature*. London: Routledge.
- Austin, T. (1984). *Language crafted: a linguistic theory of poetic syntax*. Bloomington, Indiana: Indiana University Press.
- Baker, M. C. (2001). *The atoms of language: the mind's hidden rules of grammar*. New York, New York: Perseus.
- Banfield, A. M. (1973). *Stylistic transformations: a study based on the syntax of 'Paradise Lost'*. Ph. D. thesis, University of Wisconsin.
- Beckett, S. (1979). *Company*. London: Calder.

- Beckett, S. (1994 [1965]). *How It Is*. New York, NY: Grove Press.
- Beckett, S. (1999 [1983]). *Worstward Ho*. London: John Calder.
- Beckett, S. (2001). *Disjecta: Miscellaneous Writings and a Dramatic Fragment*. London: Calder.
- Belletti, A. and L. Rizzi (1988). Psych verbs and theta-theory. *Natural Language and Linguistic Theory* 6(3), 291–352.
- Biber, D., E. Finnegan, S. Johansson, S. Conrad, and G. Leech (1999). *Longman grammar of Spoken and Written English*. London: Longman.
- Bobaljik, J. (2002). A-chains at the PF-interface: copies and ‘covert’ movement. *Natural Language and Linguistic Theory* 20(2), 197–267.
- Bobaljik, J. D. (1999). Adverbs: the hierarchy paradox. *GLOT International* 4(9), 27–28.
- Bobaljik, J. D. (2008). Where’s Phi? Agreement as a postsyntactic operation. In D. Harbour, D. Adger, and S. Béjar (Eds.), *Phi Theory*, pp. 295–328. Oxford: Oxford University Press.
- Bobaljik, J. D. and S. Wurmbrand (2008). Word order and scope: transparent interfaces and the 3/4 signature. Manuscript, University of Connecticut.
- Boeckx, C. (2001). Head-ing toward PF. *Linguistic Inquiry* 32(2), 345–355.
- Boeckx, C. (2003). *Islands and Chains: Resumption as Stranding*. John Benjamins Publishing Company.
- Bošković, Ž. (1997). *The Syntax of Nonfinite Complementation: An Economy Approach*. Cambridge, Massachusetts: MIT Press.
- Bošković, Ž. (2002). On multiple wh-fronting. *Linguistic Inquiry* 33(3), 351–383.
- Bresnan, J. (1994). Locative inversion and the architecture of universal grammar. *Language* 70(1), 72–131.

- Bresnan, J. and J. Hay (2008). Gradient Grammar: An effect of animacy on the syntax of *give* in New Zealand and American English. *Lingua* 118(2), 245–259.
- Brody, M. (1995). *Lexico-Logical Form: A Radically Minimalist Theory*. Cambridge, Massachusetts: Massachusetts Institute of Technology.
- Brody, M. (2000). Mirror theory: syntactic representation in perfect syntax. *Linguistic Inquiry* 31(1), 29–56.
- Bruening, B. (2009). Selectional Asymmetries between CP and DP Suggest that the DP Hypothesis is Wrong. In L. MacKenzie (Ed.), *U. Penn Working Papers in Linguistics 15.1: Proceedings of the 32nd Annual Penn Linguistics Colloquium*, pp. 26–35.
- Burton-Roberts, N. (1999). Language, linear precedence and parentheticals. In P. Collins and D. A. Lee (Eds.), *The clause in English: in honour of Rodney Huddleston*, pp. 33–52. John Benjamins Publishing Company.
- Carston, R. (2002). *Thoughts and Utterances: the pragmatics of explicit communication*. Oxford: Blackwell.
- Chierchia, G. (1998). Reference to kinds across languages. *Natural Language Semantics* 6(4), 339–405.
- Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, Massachusetts: M.I.T. Press.
- Chomsky, N. (1977). On wh-movement. In P. Culicover, T. Wasow, and A. Akmajian (Eds.), *Formal syntax*. New York, New York: Academic Press.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, The Netherlands: Foris Publications.
- Chomsky, N. (1986). *Barriers*. Cambridge, Massachusetts: MIT Press.

- Chomsky, N. (1993). A minimalist program for linguistic theory. In K. Hale and J. Keyser (Eds.), *The view from Building 20*, pp. 1–52. Cambridge, Massachusetts: MIT Press.
- Chomsky, N. (1995a). Bare phrase structure. In G. Webelhuth (Ed.), *Government binding theory and the minimalist program*, pp. 383–439. Oxford: Oxford University Press.
- Chomsky, N. (1995b). *The minimalist program*. Cambridge, Massachusetts: MIT Press.
- Chomsky, N. (2000). Minimalist inquiries: The framework. In R. Martin, D. Michaels, and J. Uriagereka (Eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, pp. 89–156. MIT Press.
- Chomsky, N. (2001). Derivation by phase. In M. Kenstowicz (Ed.), *Ken Hale: A Life in Linguistics*, pp. 1–52. Cambridge, Massachusetts: MIT Press.
- Chomsky, N. (2004). Beyond explanatory adequacy. In A. Belletti (Ed.), *Structures and Beyond: The Cartography of Syntactic Structures*, Volume 3, pp. 104–131. Oxford: Oxford University Press.
- Chomsky, N. (2008). On phases. In R. Freidin, C. Otero, and M. L. Zubizarreta (Eds.), *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, pp. 133–166. Cambridge, Massachusetts: MIT Press.
- Chomsky, N. and H. Lasnik (1977). Filters and control. *Linguistic Inquiry* 8(3), 425–504.
- Chomsky, N. and G. A. Miller (1963). Introduction to the formal analysis of natural language. In R. D. Luce, R. R. Bush, and E. Galanter (Eds.), *Handbook of Mathematical Psychology*, Volume 2, pp. 269–321. New York, NY: Wiley.
- Chung, S. (2006). Sluicing and the lexicon: the point of no return. In R. T. Cover

- and Y. Kim (Eds.), *BLS 31: General Session and Parasession on Prosodic Variation and Change*, pp. 73–91.
- Chung, S., W. A. Ladusaw, and J. McCloskey (1995). Sluicing and Logical Form. *Natural Language Semantics* 3(3), 239–282.
- Cinque, G. (1993). On the evidence for partial N movement in the Romance DP. *University of Venice Working Papers in Linguistics* 3(2), 21–40.
- Cinque, G. (1999). *Adverbs and functional heads: a cross-linguistic perspective*. New York: Oxford University Press.
- Collins, C. (1997). *Local economy*. Cambridge, Massachusetts: MIT Press.
- Collins, C. (2005a). A smuggling approach to raising in English. *Linguistic Inquiry* 36(2), 289–298.
- Collins, C. (2005b). A smuggling approach to the passive in English. *Syntax* 8(2), 81–120.
- Constable, J. and H. Aoyama (1999). Word length frequency and distribution in English: part ii. an empirical and mathematical examination of the character and consequences of isometric lineation. *Literary and linguistic computing* 14(4), 507–536.
- Copley, B. (2004). So-called epistemic ‘should’. *Snippets* 9, 7–8.
- Coppock, E. (2001). Gapping: in defense of deletion. In M. Andronis, C. Ball, H. Elston, and S. Neuvel (Eds.), *Proceedings of the Chicago Linguistics Society* 37, University of Chicago, pp. 133–147.
- Corver, N. and H. van Riemsdijk (Eds.) (2001). *Semi-lexical categories: on the content of function words and the function of content words*. Berlin: Mouton de Gruyter.
- Craenenbroeck, J. v. (2004). *Ellipsis in Dutch Dialects*. Ph. D. thesis, Utrecht Institute of Linguistics OTS, Utrecht.

- Craenenbroeck, J. v. (2010). Piecing together the ellipsis puzzle: towards identifying and unifying ellipsis licensors. Talk given at UMass, Amherst.
- Craenenbroeck, J. v. and M. d. Dikken (2006). Ellipsis and EPP repair. *Linguistic Inquiry* 37(4), 653–664.
- Creeley, R. (2006). *The Collected Poems of Robert Creeley 1975-2006*. London: University of California Press.
- Culicover, P. W. (2009). *Natural language syntax*. Oxford: Oxford University Press.
- Culicover, P. W. and R. Jackendoff (2005). *Simpler Syntax*. Oxford University Press.
- Culler, J. D. (1975). *Structuralist poetics: Structuralism, linguistics and the study of literature*. London: Routledge.
- Culler, J. D. (1981). Literary competence. *Essays in modern stylistics*, 24–41.
- cummings, e. e. (1994). *Complete Poems 1904-1962*. New York, NY: W.W. Norton & Co.
- Dekkers, J., F. van der Leeuw, and J. van der Weijer (Eds.) (2000). *Optimality theory: phonology, syntax, and acquisition*. Oxford University Press.
- Dikken, M. d. (2007). Phase Extension: contours of a theory of the role of head movement in phrasal extraction. *Theoretical Linguistics* 33, 1–41.
- Dillon, G. L. (1975). Inversions and deletions in English poetry. *Language and Style* 8(3), 220–237.
- Embick, D. and R. Noyer (2001). Movement operations after syntax. *Linguistic Inquiry* 32(4), 555–595.
- Emma, R. (1964). *Milton's syntax*. The Hague, The Netherlands: Mouton.
- Emonds, J. E. (1979). Appositive relatives have no properties. *Linguistic Inquiry* 10(2), 211–243.

- Epstein, S. D., T. D. Seely, R. Kawashima, and H. Kitahara (1998). *A derivational approach to syntactic relations*. Oxford University Press.
- Epstein, S. D., T. D. Seely, and H. Kitahara (2009). The application of Transfer is Case-sensitive and deducible. Paper presented at Minimalist Approaches to Syntactic Locality 2009, Budapest, Hungary.
- Erteschik-Shir, N. (2005). Sound patterns of syntax: Object Shift. *Theoretical Linguistics* 31, 47–93.
- Fabb, N. (1990). The difference between restrictive and non-restrictive relatives. *Journal of Linguistics* 26(1), 57–77.
- Fabb, N. (1997). *Linguistics and literature: language in the verbal arts of the world*. Blackwell.
- Fabb, N. (2002). *Language and literary structure*. Cambridge: Cambridge University Press.
- Fabb, N. (2004). Form as fiction. *Belgian Journal of Language and Literatures* 2, 63–73.
- Fabb, N. (2009). Why is verse poetry? *PN Review* 189, 52–57.
- Fabb, N. (2010). Is literary language a development of ordinary language? *Lingua* 120, 1219–1232.
- Fabb, N. and M. Halle (2008). *Meter in poetry: a new theory*. Cambridge: Cambridge University Press.
- Fielding, H. (1996). *Bridget Jones's Diary*. London: Picador.
- Fintel, K. v. and S. Iatridou (2003). Epistemic containment. *Linguistic Inquiry* 34(2), 173–198.
- Fish, S. (1980). *Is there a text in this classroom? The authority of interpretive communities*. Cambridge: Cambridge University Press.

- Fitzgerald, C. M. (2006). An optimality treatment of syntactic inversions in English verse. *Language sciences* 29, 203–217.
- Fox, D. (1995). Economy and scope. *Natural Language Semantics* 3(3), 283–341.
- Fox, D. (1999). Reconstruction, binding theory, and the interpretation of chains. *Linguistic Inquiry* 30(2), 157–196.
- Fox, D. (2000). *Economy and semantic interpretation*. Cambridge, Massachusetts: MIT Press.
- Fox, D. and D. Pesetsky (2005). Cyclic linearization of syntactic structure. *Theoretical Linguistics* 31(1-2), 1–46.
- Frazier, L., J. Merchant, J. Clifton, Charles, and T. Weskott (2009). Fragment answers to questions: a case of inaudible syntax. Manuscript, University of Chicago.
- Gengel, K. (2007). *Focus and ellipsis: a generative analysis of pseudogapping and other elliptical structures*. Ph. D. thesis, University of Stuttgart.
- Gibbs, R. W. (Ed.) (2008). *The Cambridge handbook of metaphor and thought*. Cambridge University Press.
- Ginsberg, A. (1995). *Journals Mid-Fifties 1954-1958*. London: HarperCollins.
- Golston, C. (1998). Constraint-based metrics. *Natural Language and Linguistic Theory* 16(4), 719–770.
- Golston, C. and T. Riad (2005). The phonology of Greek lyric meter. *Journal of Linguistics* 41(1), 77–115.
- Grimshaw, J. (1997). Projection, heads, and optimality. *Linguistic Inquiry* 28(3), 373–422.
- Grimshaw, J. (2001). Economy of structure in ot. Manuscript, Rutgers University, ROA 434.

- Grimshaw, J. and V. Samek-Lodovici (1998). Optimal subjects and subject universals. In P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis, and D. Pesetsky (Eds.), *Is the Best Good Enough?*, pp. 193–220. Cambridge, Massachusetts: MIT Press.
- Groenendijk, J. A. G. and M. B. J. Stokhof (1989). Type shifting rules and the semantics of interrogatives. In G. Chierchia, B. H. Partee, and R. Turner (Eds.), *Properties, Types, and Meanings v. II*. Dordrecht: Kluwer Academic Press.
- Gundel, J. K. (1974). *The role of topic and comment in linguistic theory*. Ph. D. thesis, University of Texas.
- Haegeman, L. M. V. (1988). Parenthetical adverbials: the radical orphanage approach. In S. Chiba (Ed.), *Aspects of Modern English Linguistics*. Tokyo, Japan: Kaitakushi.
- Haegeman, L. M. V. (1990). Understood subjects in English diaries. *Multilingua* 9(2), 157–199.
- Haegeman, L. M. V. (1997). Register variation, truncation and subject omission in English and French. *English language and linguistics* 1(2), 233–270.
- Haegeman, L. M. V. and T. Ihsane (2001). Adult null subjects in the non-pro-drop languages: two diary dialects. *Language acquisition* 9(4), 329–346.
- Hale, K. and S. J. Keyser (1993). On argument structure and the lexical expression of syntactic relations. In K. Hale and S. J. Keyser (Eds.), *The view from building 20*, pp. 53–110. Cambridge, Massachusetts: MIT Press.
- Hale, K. and S. J. Keyser (2002). *Prolegomenon to a theory of argument structure*. Cambridge, Massachusetts: MIT Press.
- Halle, M. and S. J. Keyser (1971). *English stress: its form, its growth and its role in verse*. Harper and Row Publishers.

- Halle, M. and A. Marantz (1993). Distributed morphology and the pieces of inflection. In K. Hale and S. J. Keyser (Eds.), *The View from Building 20*, pp. 111–176. Cambridge, Massachusetts: MIT Press.
- Halliday, M. and R. Hasan (1976). *Cohesion in English*. London: Longman.
- Halliday, M. and C. Matthiessen (2004). *An Introduction to Functional Grammar* (3rd ed.). London: Arnold.
- Hankamer, J. and I. Sag (1976). Deep and surface anaphora. *Linguistic Inquiry* 7(3), 391–428.
- Hanson, K. and P. Kiparsky (1996). A parametric theory of poetic meter. *Language* 72(2), 287–335.
- Harley, H. (1995). Abstracting away from abstract Case. In J. Beckman (Ed.), *Proceedings of the North East Linguistic Society*, Volume 25, University of Pennsylvania, pp. 207–222. Graduate Linguistic Student Association.
- Hartman, J. (2009). When e-GIVENness overpredicts identity. Paper presented at BCGL4.
- Hartman, J. (to appear). The semantic uniformity of traces: evidence from ellipsis parallelism. *Linguistic Inquiry*.
- Hasegawa, N. (2001). Causatives and the role of v: Agent, causer, and experiencer. In K. Inoue and N. Hasegawa (Eds.), *Linguistics and Interdisciplinary Research*. Kanda University of International Studies.
- Hayes, B. and M. MacEachern (1998). Quatrain form in English folk verse. *Language* 74(3), 473–507.
- Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases*. Ph. D. thesis, University of Massachusetts, Amherst.
- Heim, I. and A. Kratzer (1998). *Semantics in generative grammar*. Malden, MA: Blackwell.

- Hicks, G. (2009). *The derivation of anaphoric relations*. Amsterdam: John Benjamins Publishing Company.
- Hope, J. (2000). Rats, bats, sparrows and dogs: biology, linguistics and the nature of standard English. In L. Wright (Ed.), *The Development of Standard English 1300–1800*, pp. 49–56. Cambridge: Cambridge University Press.
- Hornstein, N. (1995). *Logical Form: from GB to Minimalism*. Cambridge, Massachusetts: Basil Blackwell.
- Hornstein, N. (2009). *A theory of syntax*. Cambridge: Cambridge University Press.
- Hornstein, N., J. Nunes, and K. K. Grohmann (2005). *Understanding Minimalism*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.
- Jackendoff, R. S. (1997). *The architecture of the language faculty*. Cambridge, Massachusetts: MIT Press.
- Jayaseelan, K. A. (1990). Incomplete VP deletion and Gapping. *Linguistic Analysis* 20(1-2), 64–81.
- Jayaseelan, K. A. (2001). IP-internal topic and focus phrases. *Studia Linguistica* 55(1), 39–75.
- Johnson (2004). In search of the English middle field. Manuscript, University of Massachusetts.
- Johnson, K. (2001). What VP ellipsis can do, and what it can't, but not why. In M. Baltin and C. Collins (Eds.), *The handbook of contemporary syntactic theory*, pp. 439–479. Oxford: Blackwell Publishers.
- Johnson, K. (2009). Gapping is not (VP-)deletion. *Linguistic Inquiry* 40(2), 289–328.

- Johnson, K. and S. Tomioka (1998). Lowering and mid-size clauses. In G. Katz, S.-S. Kim, and W. Haike (Eds.), *Proceedings of the 1997 Tübingen workshop on reconstruction*, Tübingen, Germany, pp. 185–206. Sprachtheoretische Grundlagen für die Computer Linguistik.
- Kayne, R. S. (1994). *The Antisymmetry of Syntax*. Cambridge, Massachusetts: MIT Press.
- Kempson, R. and J. Kiaer (2010). Multiple long-distance scrambling: syntax as a reflection of parsing. *Journal of Linguistics* 46, 127–192.
- Kiparsky, P. (1977). The rhythmic structure of English verse. *Linguistic Inquiry* 8, 189–207.
- Kiparsky, P. (1981). The role of linguistics in a theory of poetry. In D. Freeman (Ed.), *Essays in modern stylistics*, pp. 9–23. London: Meuthen.
- Kiparsky, P. (2006). A modular metrics for folk verse. In B. E. Drescher and N. Friedberg (Eds.), *Formal approaches to poetry*, pp. 7–49. Mouton de Gruyter.
- Kishimoto, H. (2006). Japanese syntactic nominalization and VP-internal syntax. *Lingua* 116(6), 771–810.
- Koizumi, M. (2000). String vacuous overt verb raising. *Journal of East Asian Linguistics* 9, 227–285.
- Kolaiti, P. (2005). Pragmatics and the phantasm called Text: a relevance-theoretic approach to Cohesion. *UCL Working Papers in Linguistics* 17, 335–352.
- Kratzer, A. (1981). The notional category of modality. In H. Eikmeyer and H. Reiser (Eds.), *Words, worlds and contexts: New approaches in word semantics*. Mouton de Gruyter.
- Lakoff, G. and M. Johnson (1980). *Metaphors we live by*. Chicago, IL: University of Chicago Press.

- Landau, I. (1999). *Elements of control*. Ph. D. thesis, Massachusetts Institute of Technology.
- Landau, I. (2007). EPP extensions. *Linguistic Inquiry* 38(3), 485–523.
- Lasnik, H. (1999). Pseudogapping puzzles. In S. Lappin and E. Benmamoun (Eds.), *Fragments: Studies in Ellipsis and Gapping*, pp. 141–174. Oxford: Oxford University Press.
- Lasnik, H. (2003). *Minimalist Investigations in Linguistic Theory*. New York, New York: Routledge.
- Lasnik, H. (2007). On ellipsis: The PF approach to missing constituents. Manuscript, University of Maryland.
- Lasnik, H. (2010). EPP, repair by ellipsis, and the organization of the grammar. Handout of a talk given at SyntaxFest, Indiana University.
- Lasnik, H. and M.-K. Park (2003). The EPP and the Subject Condition under Sluicing. *Linguistic Inquiry* 34(4), 649–660.
- Lasnik, Howard with Depiante, M. A. and A. Stepanov (2000). *Syntactic Structures revisited: Contemporary lectures on classic transformational theory*. Cambridge, Mass.: MIT Press.
- Lechner, W. (2006). An interpretive effect of Head Movement. In M. Frascarelli (Ed.), *Phases of interpretation*, pp. 45–71. Berlin: Mouton de Gruyter.
- Legendre, G. (2001). *Optimality Theoretic Syntax*. Cambridge, MA: MIT Press.
- Legendre, G., P. Smolensky, and C. Wilson (1998). When is less more? faithfulness and minimal links in wh-chains. In P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis, and D. Pesetsky (Eds.), *Is the Best Good Enough?*, pp. 249–290. Cambridge, Massachusetts: MIT Press.
- Lerdahl, F. (1988). Cognitive constraints on compositional systems. In J. A. Sloboda (Ed.), *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition*. Oxford: Clarendon Press.

- Lerdahl, F. and R. S. Jackendoff (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.
- Lerdahl, F. and R. S. Jackendoff (2006). The capacity for music: What is it, and what's special about it? *Cognition* 100(1), 33–72.
- Levin, S. R. (1967). *Poetry and grammaticalness*. Boston, MA: Houghton Mifflin.
- Linebarger, M. (1980). *The grammar of negative polarity*. Ph. D. thesis, Massachusetts Institute of Technology.
- Lobeck, A. (1995). *Ellipsis: Functional Heads, Licensing and Identification*. New York, NY: Oxford University Press.
- Mac Low, J. (2008). *Thing of Beauty: New And Selected Works*. London: California University Press.
- Marantz, A. (1988). Clitics, morphological merger, and the mapping to phonological structure. In M. Hammond and M. Noonan (Eds.), *Theoretical Morphology: Approaches in Modern Linguistics*, pp. 253–270. San Diego, California: Academic Press.
- Marantz, A. (1991). Case and licensing. In G. Westphal, B. Ao, and H.-R. Chae (Eds.), *Eastern States Conference on Linguistics*, Cornell University, Ithaca, NY, pp. 234–253. Cornell Linguistics Club.
- Mathews, H., A. Brotchie, and I. Monk (Eds.) (1998). *The Oulipo Compendium*. London: Atlas.
- May, R. (1985). *Logical Form: Its Structure and Derivation*. Cambridge, Massachusetts: MIT Press.
- McCawley, J. D. (1982). Parentheticals and discontinuous constituent structure. *Linguistic Inquiry* 13(1), 91–106.

- McCloskey, J. (1996). Subjecthood and subject positions. In L. Haegeman (Ed.), *Elements of grammar*, pp. 197–235. Dordrecht: Kluwer Academic Publishers.
- McGinnis, M. (2004). Lethal ambiguity. *Linguistic Inquiry* 35(1), 47–95.
- Merchant, J. (2001). *The syntax of silence: sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.
- Merchant, J. (2003). Remarks on stripping. Ms., University of Chicago.
- Merchant, J. (2004). Fragments and ellipsis. *Linguistics and philosophy* 27(6), 661–738.
- Merchant, J. (2008). Variable island repair under ellipsis. In K. Johnson (Ed.), *Topics in Ellipsis*, pp. 132–153. Cambridge, United Kingdom: Cambridge University Press.
- Milsark, G. (1974). *Existential Sentences in English*. Ph. D. thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts.
- Miyagawa, S. (1997). Against optional scrambling. *Linguistic Inquiry* 28(1), 1–25.
- Miyagawa, S. (2001). EPP, scrambling and wh-in-situ. In M. Kenstowicz (Ed.), *Ken Hale: A life in language*, pp. 293–338. Cambridge, MA: MIT Press.
- Miyagawa, S. (2003). A-movement, scrambling and options without optionality. In S. Karimi (Ed.), *Word order and scrambling*, pp. 177–200. Oxford: Blackwell.
- Miyagawa, S. (2005). EPP and semantically vacuous scrambling. In J. Sabel and M. Saito (Eds.), *The free word order phenomenon*, pp. 181–220. Berlin: Mouton de Gruyter.
- Munn, A. B. (1993). *Topics in the syntax and semantics of coordinate structures*. Ph. D. thesis, The University of Maryland.

- Nariyama, S. (2004). Subject ellipsis in English. *Journal of Pragmatics* 36(2), 237–264.
- Nespor, M. and I. Vogel (1979). Clash avoidance in Italian. *Linguistic Inquiry* 10(3), 467–482.
- Newmeyer, F. J. (1983). *Grammatical theory: its limits and possibilities*. Chicago, IL: University of California Press.
- Newmeyer, F. J. (2003). Grammar is grammar and usage is usage. *Language* 79(4), 682–707.
- Newmeyer, F. J. (2005). *Possible and Probable Languages: A Generative Perspective on Linguistic Typology*. Oxford: Oxford University Press.
- Nilsen, O. (2003). *Eliminating positions: syntax and semantics of sentence modification*. Ph. D. thesis, Utrecht Institute of Linguistics OTS.
- Nissenbaum, J. (2000). *Investigations of covert phrase movement*. Ph. D. thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts.
- Nordal, S. (1933). *Egils saga Skallagrímssonar*. Reykjavik: Hið Íslenska Fornritafélag.
- Nunes, J. (2004). *Linearization of chains and sideward movement*. Cambridge, Mass.: MIT Press.
- O’Neil, W. (2001). Grammar games: unscrambling Skaldic syntax. In M. Kenstowicz (Ed.), *Ken Hale: A life in language*, Number 339-354. Cambridge, Massachusetts: MIT Press.
- Ortony, A. (Ed.) (1993). *Metaphor and thought* (Second ed.). Cambridge University Press.
- Park, M.-K. and J.-M. Kang (2007). Multiple sluicing in English. In *Proceedings of the 21st Pacific Asia Conference on Language, Information and Computation*, pp. 394–404. The Korean Society for Language and Information.

- Partee, B. H. and M. Rooth (1983). Generalized conjunction and type ambiguity. In R. Bäuerle, C. Schwarze, and A. von Stechow (Eds.), *Meaning, Use and Interpretation of Language*, pp. 362–383. de Gruyter.
- Perloff, M. (1982). Between verse and prose: Beckett and the New Poetry. *Critical Inquiry* 9(2), 415–434.
- Pesetsky, D. (1982). *Paths and categories*. Ph. D. thesis, Massachusetts Institute of Technology.
- Pesetsky, D. (1995). *Zero syntax*. Cambridge, Massachusetts: MIT Press.
- Pesetsky, D. and E. Torrego (2001). T-to-C movement: Causes and consequences. In M. Kenstowicz (Ed.), *Ken Hale: A life in language*, pp. 355–426. Cambridge, MA: MIT Press.
- Pesetsky, D. and E. Torrego (2004). Tense, case, and the nature of syntactic categories. In J. Guéron and J. Lecarme (Eds.), *The syntax of time*, pp. 495–537. Cambridge, MA: MIT Press.
- Pilkington, A. (2000). *Poetic effects: a relevance theory perspective*. Amsterdam: John Benjamins Publishing Company.
- Pinker, S. (1995). *The Language Instinct*. London: Penguin.
- Pinker, S. (2002). *The Blank Slate: the moder denial of human nature*. New York, NY: Viking.
- Pollock, J.-Y. (1989). Verb movement, UG and the structure of IP. *Linguistic Inquiry* 20, 365–424.
- Pope, A. (1966). *Pope: Poetical works*. London: Oxford University Press.
- Postal, P. M. (1969). On so-called ‘pronouns’ in English. In D. Reibel and S. Schane (Eds.), *Modern Studies in English*, pp. 201–244. Englewood Cliffs, New Jersey: Prentice-Hall.

- Potsdam, E. (1997). NegP and subjunctive complements in English. *Linguistic Inquiry* 28(3), 533–541.
- Preminger, O. (2010). Against “derivational time-bombs” as the (only) basis for obligatory operations in syntax. Manuscript, MIT.
- Prince, A. and P. Smolensky (1993). Optimality theory. Manuscript, Rutgers University and University of Colorado, Boulder.
- Recanati, F. (2004). *Literal meaning*. Cambridge: Cambridge University Press.
- Reinhart, T. (1981). Pragmatics and linguistics: an analysis of sentence topics. *Philosophica* 27(1), 53–94.
- Reinhart, T. (1983). *Anaphora and Semantic Interpretation*. Chicago, Illinois: University of Chicago Press.
- Reinhart, T. (2006). *Interface Strategies*. Cambridge, MA: MIT Press.
- Reuland, E. J. and A. G. B. t. Meulen (Eds.) (1987). *The Representation of (In)definiteness*. Cambridge, MA: MIT Press.
- Rezác, M. (2004). *Elements of cyclic syntax: Agree and Merge*. Ph. D. thesis, University of Toronto.
- Rezác, M. (2008). Phi-Agree and Theta-related Case. In D. Harbour, D. Adger, and S. Béjar (Eds.), *Phi Theory*, pp. 83–129. Oxford: Oxford University Press.
- Richards, M. D. (2004). *Object shift and scrambling in North and West Germanic: a study in symmetrical syntax*. Ph. D. thesis, University of Cambridge.
- Richards, N. (2001). *Movement in language: interactions and architectures*. Oxford: Oxford University Press.
- Rizzi, L. (1986). On chain formation. In H. Borer (Ed.), *The Grammar of Pronominal Clitics*(19). New York, NY: Academic Press.
- Rizzi, L. (1997). The fine structure of the left periphery. In L. Haegeman (Ed.), *Elements of grammar*, pp. 281–337. Dordrecht: Kluwer Academic Publishers.

- Roberts, I. (1993). *Verbs and Diachronic Syntax*, Volume 28. Dordrecht: Kluwer Publications.
- Roberts, I. (2010). *Agreement and head movement*. Cambridge, MA: MIT Press.
- Roberts, I. G. (2005). *Principles and Parameters in a VSO Language: A Case Study in Welsh*. Oxford Studies in Comparative Syntax. Oxford: Oxford University Press.
- Rochemont, M. and P. W. Culicover (1990). *English focus constructions and the theory of grammar*. Cambridge: Cambridge University Press.
- Ross, J. (1967). *Constraints on variables in syntax*. Ph. D. thesis, Massachusetts Institute of Technology.
- Saito, M. (1985). *Some asymmetries in Japanese and their theoretical implications*. Ph. D. thesis, Massachusetts Institute of Technology.
- Saito, M. (1992). Long distance scrambling in Japanese. *Journal of East Asian Linguistics* 1, 69–118.
- Saito, M. (2006). Optional A-scrambling. Paper presented at the 16th Japanese/Korean linguistics conference.
- Samek-Lodovici, V. (2006). Optimality Theory and the Minimalist Program. In R. Vogel and H. Broekhuis (Eds.), *Optimality Theory and Minimalism: A Possible Convergence?*, Volume 25, pp. 77–97. Linguistics in Potsdam.
- Sapir, E. (1921). *Language: an introduction to the study of speech*. New York, NY: Harcourt, Brace and Co.
- Sauerland, U. and P. Elbourne (2002). Total reconstruction, PF-movement, and the derivational order. *Linguistic Inquiry* 33(2), 283–319.
- Schütze, C. (2004). Synchronic and diachronic microvariation in English do. *Lingua* 14(4), 389–542.

- Schuyler, T. (2001). Wh-movement out of the site of VP ellipsis. In S. M. Bhloscaidh (Ed.), *Syntax at Santa Cruz*, pp. 1–20. UC, Santa Cruz: Linguistics Department.
- Selkirk, E. (1995). Sentence prosody: intonation, stress and phrasing. In J. A. Goldsmith (Ed.), *The handbook of phonological theory*. London: Blackwell Publishers.
- Selkirk, E. O. (1984). *Phonology and syntax : the relation between sound and structure*. Cambridge, Mass.: MIT Press.
- Sleeman, P. (1996). *Licensing Empty Nouns in French*. Holland Institute of Generative Linguistics.
- Sperber, D. and D. Wilson (1986). *Relevance: communication and cognition*. Blackwell.
- Stein, G. (1997 [1914]). *Tender Buttons*. New York, NY: Dover Publications.
- Stern, J. (1983). Metaphor and grammatical deviance. *Nous* 17(4), 577–599.
- Stern, J. (2000). *Metaphor in context*. Cambridge, MA: MIT Press.
- Stern, J. (2006). Metaphor, literal, literalism. *Mind and Language* 21(3), 243–279.
- Stern, J. (2009). Metaphor and minimalism. *Philosophical Studies* DOI: 10.1007/s11098-009-9486-3.
- Stockwell, P. (2002). *Cognitive poetics: an introduction*. London: Routledge.
- Taglicht, J. (1998). Constraints on intonational phrasing in english. *Journal of Linguistics* 34(2), 181–211.
- Takahashi, S. (2004). Pseudogapping and cyclic linearization. In K. Moulton and M. Wolf (Eds.), *Proceedings of the Thirty-Fourth meeting of the North East Linguistic Society*, University of Massachusetts at Amherst, pp. 571–585. Graduate Linguistic Student Association.

- Takahashi, S. and D. Fox (2005). MaxElide and the re-binding problem. In E. Georgala and J. Howell (Eds.), *Proceedings of SALT XV*, Cornell University, Ithaca, NY, pp. 223–240. CLC Publications.
- Takano, Y. (2002). Surprising constituents. *Journal of East Asian Linguistics* 11(3), 243–301.
- Tanaka, H. (2002). Rightward dislocation as scrambling. *Journal of Linguistics* 37(3), 551–579.
- Temmerman, T. (2009). More than meets the ear: ellipsis in Dutch and English fragment answers. Paper presented at BCGL4, Brussels.
- Thoms, G. (2008). The syntax of experimental literature: a literary linguistic investigation. Master's thesis, University of Strathclyde, Glasgow.
- Thoms, G. (2009). British 'team' DPs and theories of scope. To appear in *Proceedings of the fourth Language at the University of Essex Postgraduate Conference*. Essex, UK.
- Thoms, G. (2010a). Addressing issues for the covert cycle: asymmetries with binding and anti-locality. Manuscript, Strathclyde University.
- Thoms, G. (2010b). Reducing Last Resort insertion to locality: English *do*-support. Under review for *Proceeding of Minimalist Accounts of Syntactic Locality 2009*. Cambridge: Cambridge University Press.
- Thoms, G. (2010c). Syntactic reconstruction and Scope Economy. Paper presented at GLOW33, Wroclaw, Poland.
- Thoms, G. (to appear). Verb floating and VPE: towards a movement account of ellipsis licensing. *Linguistic Variation Yearbook*.
- Thorne, J. (1965). Stylistics and generative grammars. *Journal of Linguistics* 1, 49–59.

- Thrasher, R. (1977). *One way to say more by saying less: a study of so-called subjectless sentences*. Tokyo: The Eihosha.
- Toosarvandani, M. (2008). Wh-movement and the syntax of sluicing. *Journal of Linguistics* 44(3), 677–722.
- Travis, L. (1984). *Parameters and Effects of Word Order Variation*. Ph. D. thesis, Massachusetts Institute of Technology.
- Truckenbrodt, H. (2006). On the semantic motivation of head movement to c in german. *Theoretical Linguistics* 32(3), 257–306.
- Truswell, R. (2009). reposition-stranding, passivisation, and extraction from adjuncts in germanic. *Linguistic Variation Yearbook* 9, 131–177.
- Turner, F. (1985). *Natural classicism: essays on literature and science*. New York, NY: Paragon House.
- Uriagereka, J. (1999). Multiple spell-out. In S. Epstein and N. Hornstein (Eds.), *Working minimalism*, pp. 251–282. Cambridge, Massachusetts: MIT Press.
- Vlachos, C. (2010). Merchant says that MaxElide works for instances of wh-movement followed by VP deletion but it's not clear how or how it does. *Snippets* 21, 17–18.
- Vries, M. d. (2007). Invisible constituents? parentheticals as b-merged adverbial phrases. In N. Dehé and Y. Kavalova (Eds.), *Parentheticals*. Amsterdam: John Benjamins Publishing Company.
- Wallenberg, J. C. (2009). *Antisymmetry and the conservation of c-command: scrambling and phrase structure in synchronic and diachronic perspective*. Ph. D. thesis, University of Pennsylvania.
- Weir, A. (2009). Article drop in english headlines. Master's thesis, University College London, London.

- Weiß, H. (2007). A question of relevance: some remarks on standard languages. In M. Penke and A. Rosenbach (Eds.), *What counts as evidence in linguistics. The case of innateness*, pp. 181–208. Amsterdam: John Benjamins Publishing Company.
- Wexler, K. (2002). Lenneberg's Dream: Learning Normal Language Development and Specific Language Impairment. In J. Schaeffer and Y. Levy (Eds.), *Language competence accross populations: Towards a definition of Specific Language Impairment*. Mahwah, NJ: Erlbaum.
- Wilson, D. and R. Carston (2007). A unitary approach to lexical pragmatics: relevance, inference and ad hoc concepts. In N. Burton-Roberts (Ed.), *Pragmatics*, pp. 230–259. Basingstoke: Palgrave.
- Wiltschko, M. (1998). On the syntax and semantics of (relative) pronouns and determiners. *The Journal of Comparative Germanic Linguistics* 2(2), 143–181.
- Winter, Y. (2001). *Flexibility principles in Boolean Semantics*. Cambridge, MA: MIT Press.
- Wurmbrand, S. (2006). Licensing case. *Journal of Germanic Linguistics* 18(3), 175–236.
- Youmans, G. (1982). Hamlet's testimony on Kiparsky's theory of meter. *Neophilologus* 66(490-503).
- Youmans, G. (1983). Generative tests for generative meters. *Language* 59, 67–92.
- Youmans, G. (1996). Reconsidering Chaucer's prosody. In C. B. McCully and J. Anderson (Eds.), *English historical metrics*, pp. 185–209. Cambridge: Cambridge University Press.

- Zaenen, A., J. Maling, and H. Thráinsson (1985). Case and grammatical functions: the Icelandic passive. *Natural Language and Linguistic Theory* 3(4), 441–483.