Syntax Outranks Phonology: Evidence from Ancient Greek

0 Introduction

What influence do syntax and phonology have on one another? Two types of answer to this question appear in the literature. The consensus view is probably best expressed by Zwicky & Pullum (1986; see also Myers 1987; Vogel & Kenesi 1990), who claim that the relation is one way: although phonological phrasing above the word is affected by syntactic structure, syntax itself is phonology-free. The result is guaranteed in a derivational model of grammar by ordering all syntactic rules before any phonological (prosodic and segmental) rules:

(1) Syntax derivationally precedes phonology (rule-based model)

Zec & Inkelas (1990) have argued that this sort of model is overly restrictive and does not allow for cases where prosodic restrictions partially determine syntactic well-formedness. They propose a constraint-based model in which syntax and phonology interact bidirectionally:¹

(2) Syntax constrains prosody and vice versa (constraint-based model)

I’ll argue in this paper that (1) and (2) are not good ways to model the syntax-phonology interface. (1) does not allow enough prosodic influence on syntax (as argued by Zec & Inkelas) and (2) allows far too much. (1) is able to capture the general case (phonology has little influence on syntax) but not the exceptional cases (it does have some); conversely, grammars conforming to (2) are able to capture the exceptional cases (some influence) but fail to capture the general case (not very much).
I will show that all clear cases of prosodic influence on syntax in the literature involve pairs of syntactic constructions, one of which is ruled out by a prosodic constraint. That is, prosody chooses between structures which are equally well-formed syntactically. In this paper I’ll try and substantiate this generalization and argue that an account in terms of Optimality Theory (Prince & Smolensky 1993) provides a ready solution for it. In particular, I’ll argue that components of the grammar are ranked with respect to one another, much as constraints are ranked with respect to one another in Optimality Theory. In short, I’ll argue that (3) best represents the influence of syntax and phonology on one another:

(3) Syntactic constraints outrank prosodic constraints

To make the issues clearer I begin with illustrative cases from English and German. Consider the pair of NPs in (4):

(4)  a. The [video [of Macbeth]]
    b. The [[Macbeth] video]

(4a) has a postnominal PP modifier whereas (4b) has a prenominal NP modifier; both are fully grammatical. But if we replace Shakespeare’s Macbeth with Joyce’s The Dead, the second construction is ill-formed:

(5)  a. The [video [of The Dead]]
    b. *The [[The Dead] video]

(5b) is syntactically well-formed but prosodically ruled out because of the sequence *the the. The role of the constraint that rules out *the the is to choose between syntactic structures which are equally well-formed: there is a grammatical way to say (5b) and that is (5a).²
Now consider a similar case from German. German syntax readily admits strings of adjacent homophonous function words as in (6):

(6)  die, die die Blumen gekauft haben
    those who the flowers bought have
    ‘those who have bought the flowers’

(7)  daß das das Problem ist
    that this the problem is
    ‘that this is the problem’

(6) and (7) are fully grammatical. What accounts for the difference between the acceptability of (6) and (7) versus the marginality of (5b) is that German offers no simple alternatives to the constructions above. There is no way to postpose the relative clauses or the noun phrases they contain such that the various *die*’s and *das*’s won’t occur next to one another.³ In (6) and (7) the phonology does not have a chance to decide between two well-formed syntactic alternatives and has no effect on grammaticality; in (5) the phonology decides between two equally well-formed syntactic structures and decisively rules in favor of the one that violates the syntactic constraints least. Prosody determines well-formedness if and only if syntax does not.

Neither (1) nor (2) has this as a natural result. But the model in (3) does because the role of dominated constraints in Optimality Theory (OT) is precisely to select between candidates that are not ranked by higher-ranking constraints (Prince & Smolensky 1993). The English data may then be analyzed as follows: the *Macbeth* data violate no constraints, syntactic or phonological, and are both well-formed, as shown in the first tableau in (8).
(8) English

<table>
<thead>
<tr>
<th>SYNTAX</th>
<th>PHONOLOGY</th>
</tr>
</thead>
<tbody>
<tr>
<td>the ‘Macbeth’ video</td>
<td></td>
</tr>
<tr>
<td>the video of ‘Macbeth’</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SYNTAX</th>
<th>PHONOLOGY</th>
</tr>
</thead>
<tbody>
<tr>
<td>the ‘The Dead’ video</td>
<td></td>
</tr>
<tr>
<td>the video of ‘The Dead’</td>
<td></td>
</tr>
</tbody>
</table>

The data for The Dead differ in that the first phrase violates something in the phonology while the second does not. The second wins and is thus the (only) grammatical candidate.

The German data is different because the syntax produces only a single acceptable form:

(9) German

<table>
<thead>
<tr>
<th>SYNTAX</th>
<th>PHONOLOGY</th>
</tr>
</thead>
<tbody>
<tr>
<td>daß das das Problem ist</td>
<td></td>
</tr>
<tr>
<td>daß Problem das ist das</td>
<td></td>
</tr>
</tbody>
</table>

Scrambling the formatives to satisfy the phonology is worse than having the offending phonological string because syntactic constraints are ranked above phonological ones. The opposite ranking would make something like *daß Problem das ist das the grammatical output; the claim of this paper is that such a ranking is impossible in natural language.

The rest of the paper proceeds as follows. Section 1 lays out the prosodic background to the discussion and claims that the prosodic hierarchy (Selkirk 1986; Hayes 1989; Nespor & Vogel 1986) is itself best understood as a result of syntax outranking prosody.
Section 2 introduces a phonological constraint, antihomophony, of Ancient Greek similar to the one behind the English and German data above: the constraint serves only to rule out one of two competing syntactic constructions, as predicted by (3) but not by (1) or (2).

Section 3 looks at how this prosodic constraint interacts with the morphology. Here we find that antihomophony actually forces the appearance of structures that are morphologically ill-formed, suggesting that prosodic constraints are ranked above morphological constraints in Ancient Greek.

Section 4 looks at related phenomena in Japanese, Italian, English and Serbo-Croatian and arrives at the same conclusion: syntax outranks prosody and prosody outranks morphology.

Section 5 considers how the derivational and bi-directional constraint models in (1) and (2) might account for all of this and concludes that (3) gives a more natural explanation for the observed phenomena.

1 The prosodic hierarchy
I assume here the theory of prosody developed in Selkirk (1986), an end-based theory of prosodic domains in which the edges of prosodic domains (Phonological Word, Phonological Phrase) are defined by the edges of syntactic domains ($X^0$, $XP$); but the account offered here is also compatible with the approach begun by Hayes (1989) and furthered by (Nespor & Vogel 1986). We can characterize the major prosodic domains of Ancient Greek as follows, where $X$ ranges over the lexical categories noun, verb and adjective, function words and
their projections being invisible in the mapping of syntax to prosody (Selkirk 1984).

(10) Phonological Word] = X[0]
    Phonological Phrase] = XP]

That is, a Greek utterance has as many phonological words as it has lexical items and the right edge of each phonological word is coterminus with the right edge of a lexical item; likewise, ceteris paribus, for phonological phrases and maximal projections. A phrase like (11) receives the prosodic constituency shown in (12):

(11) apò t-ées arkh-ées]N t-ées metabás-eos]NP
    from the[| beginning the change
    ‘from the beginning of the change’ A.Poet 2.1455b 28

(12) ........................................] ........................................] Ph Word
    .................................................................] Ph Phrase

Prosodically dependent words in Greek are traditionally called appositives. The standard list includes articles, prepositions, relative pronouns, conjunctions, anaphoric pronouns, negatives and enclitics (Fränkel 1960, 142)—ie, Selkirk’s class of function words. Evidence that appositives formed phonological words with following content words comes from four sources: vowel coalescence, nasal assimilation, poetic meter and inscriptions.

Vowel coalescence (Haas 1988) merges the features of adjacent vowels within a prosodic word: this occurs within words (13) and across the boundary of a function word and a content word (14) with which it is phrased; but it does not occur across the boundary of two content words.

(13) Vowel Coalescence (within word)
    deelo-omen -> [deelóumen] (o + o = ou)
reveal-1 pl

tima-ete honor-2 pl  ->  [timáate]  (a + e = aa)

(14) Vowel Coalescense (function word + content word)

to ónoma  ->  [toúnoma]  (o + o = ou)

hà egoó  ->  [haagoó]  (a + e = aa)

Nasal Assimilation, in which a dental nasal takes on the place features of a following stop, has the same domain of application: word-internally (15), between a function word and a content word (16), but never between two content words.

(15) Nasal Assimilation (within word)

en-piin-oo in-drink-1 pres  ->  [empiínoo]
en-khe-oO in-pour-1 pres  ->  [eNkhéoo]

(16) Nasal Assimilation (function word + content word)

teen polin in the city (acc)  ->  [teèmpélin]
ton kalon in the good (acc)  ->  [tòNkalón]

Additional evidence for the phonological dependence of Greek function words on following content words comes from their behavior in meter (Fränkel 1960; Bulloch 1970; Devine & Stephens 1978, 1981, 1983). The evidence is based on the observation that word-breaks are preferred at some points in a line (caesurae) and dispreferred at others (bridges). Crucially, the notion ‘word-break’ includes
the break between two content words but not the break between a function word and another word, showing that function words are prosodically dependent.

The final type of evidence for the general dependence of function words comes from epigraphy. Most Ancient Greek inscriptions included no word-breaks at all. But a number of inscriptions put word breaks at the right edges of content words, writing function words as part of the following word: eg, ESGEEN for ες ‘into’ geen ‘country.’ Again, this strongly suggests that these function words were prosodically dependent on what followed.

Not all function words in Greek were proclitic: the language contained a small set of Directional Clitics as well, function words whose direction of cliticization must be lexically listed (Nespor & Vogel 1986). This set included tonally dependent function words (so-called enclitics such as te ‘and’ or ge ‘indeed’) as well as tonally independent function words (so-called postpositives such as dé ‘and’ or gár ‘for’). All evidence—segmental, metrical and epigraphic—points to these words forming the same type of constituent with their hosts that right-leaning function words form with theirs; nor do these words exhibit any special syntactic behavior which might motivate a special syntactic constituency. I won’t discuss Directional Clitics further here (they play no role in what follows) except to note that they must be listed as exceptions to the constraints which govern the formation of prosodic constituents.

Returning now to the general dependence of phonological phrasing on syntax, it should be clear that it is consistent with the claim that syntactic constraints outrank phonological constraints (3). The central observations in this area are (i) that (left or right) syntactic and prosodic edges are aligned and (ii) that syntactic
edges are invariant. That is, alignment forces prosodic edges to align with syntactic edges, never the reverse.

The Greek phrasing algorithms in (10) can be recast as two (ranked) constraints. The first (17) requires that the right edge of a $X^0$ coincide with the right edge of a phonological word; the second requires the same for XPs and phonological phrases:

(17)  \[ \text{ALIGN}(X^0, R, \text{PR\textsc{word}}, R) \]

\[ \text{ALIGN}(XP, R, \text{PR\textsc{phrase}}, R) \]

The dominance of syntax over prosody in (17) is clear once we realize that the equivalence relations are definitions for prosodic words and phrases in terms of $X$ and XP; they are not, however, definitions for $X$ and XP in terms of prosodic words and phrases. A further constraint, \text{PARSE} $\Box$, requires that syllables be prosodically licensed (Itô 1986, 1989; Prince & Somolensky 1993; McCarthy & Prince 1993):

(18)  \[ \text{PARSE} \Box: \text{syllables belong to prosodic words} \]

In what follows I’ll refer to the constraints in (17) jointly as ALIGN and to (18) simply as PARSE. Given the right-branching syntactic structure in (19), we may consider how different candidate parses fare with respect to the ALIGN and PARSE constraints.
The first candidate in (20) is not optimal because it contains a lexical word (arkhées ‘beginning’) whose right edge does not coincide with the right edge of a phonological word. The second candidate is out because it contains a phonological word (apó ‘from’) whose right edge does not align with the right edge of a lexical word. The third candidate fails because it contains unparsed material, leaving the fourth candidate as the optimal one.

2 Antihomophony and syntax

I turn now to the central data of this article, in which a prosodic constraint rules out an otherwise acceptable syntactic construction. The grammar of Ancient Greek prohibits adjacent homophones within a phonological word.
Morphological effects of this constraint will be given in section 3. The syntactic effects are the topic of the present section.

Consider the center-embedded noun phrases in (21) and (22). Each begins with a definite article (téen, tā) that is immediately followed by the definite article for the subordinated NP (tóu, tóon). In (21) the article téen is marked accusative feminine singular (a:f) in agreement with the noun phúsin ‘nature’; tóu is genitive masculine singular (g:m) in agreement with its noun prosoópou ‘face’. Similarly, in (22) tā (accusative neuter plural) agrees with prágmata ‘affairs’ and tóon agrees with póleoon ‘cities’.

\[
(21) \quad \text{[t-éen [t-óu prosoóp-ou] phús-in]}
\begin{align*}
\text{the } & \text{ the } & \text{ nature } & \text{ } \\
\text{f} & \text{ m} & \text{ a:f} & \text{ } \\
\text{of the face} & & & \\
\text{P.Pol.257d} & & & \\
\end{align*}
\]

\[
(22) \quad \text{[t-à [t-óon póle-oon] prágm-ata]}
\begin{align*}
\text{the } & \text{ the } & \text{ affairs } & \text{ } \\
\text{n:p} & \text{ m:p} & \text{ a:n:p} & \text{ } \\
\text{of the cities} & & & \\
\text{P.Pol.291c} & & & \\
\end{align*}
\]

Adjacent articles such as these result from genitive NPs that are center-embedded in other (in this case accusative) NPs. This is shown in (23), the constituent structure for both (21) and (22).

\[
(23) \quad \text{Center-embedding in noun phrases}
\]

\[
\begin{array}{c}
\text{NP} \\
\text{NP} \\
\text{Article} \quad \text{Article} \quad \text{Noun} \quad \text{Noun}
\end{array}
\]

Center-embedding in NPs is quite common and occurs with up to three NPs as shown in (24-25) and schematised in (26).

\[
(24) \quad \text{[t-ò [t-ées [t-óu ksaín-ont-os] tékhn-ees] érg-on]}
\begin{align*}
\text{the } & \text{ the } & \text{ card-er } & \text{ art } & \text{ work } \\
\text{n:n} & \text{ f} & \text{ m} & \text{ f} & \text{ n:p} \\
\text{of the cities} & & & & \\
\end{align*}
\]

\[
\begin{align*}
\text{P.Pol.257d} & & & & \\
\text{P.Pol.291c} & & & & \\
\end{align*}
\]
‘the work of the art of the (wool-)carder’ P.Pol.281a

(25) [t-à [t-ées [t-ón poll-ón] psukh-ées] ómmat-a]
the_n:n:p the_g:f the_g:n:p many_g:m soul_g:f eye_n:n:p
‘the eyes of the soul of the many’ P.Sophist 254a

(26) Multiple center-embeddings within noun phrases

Center-embeddings like this are systematically unattested when two homophonous articles would be brought together. As Smyth (1920:§1162) puts it, “the order bringing together the same forms of the article...is avoided, but two or three articles of different form may stand together” (emphasis in original). To empirically substantiate Smyth’s observation, I conducted a computer search of all of what remains of Ancient Greek literature.

Greek has four inflectional cases: dative, accusative, nominative and genitive. An NP in any of these cases can take a center-embedded possessor as long as the resulting string does not include adjacent homophonous articles.

The ubiquity of different adjacent articles in possessed NPs may be seen in (27), (29), (31) and (33). Rows in (27) contain the number of occurrences of pairs of articles in which the first is a dative singular (either feminine or masculine/neuter—masculine and neuter are indistinct in the dative singular) and the second is a genitive of some sort; columns indicate the complete corpora of four prose authors. Thus, the first row says that the ordered pair of articles téei
tées (dative feminine singular, genitive feminine singular) occurs 14 times in the collected works of Aristotle, 4 times in Isocrates, 33 times in Plato and 4 times in Xenophon, for a total of 55 occurrences.

(27) \[ \text{NP dative [NP genitive]} \]

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>I</th>
<th>P</th>
<th>X</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>d:f g:f</td>
<td>14</td>
<td>4</td>
<td>33</td>
<td>4</td>
<td>55</td>
</tr>
<tr>
<td>d:f g:mn</td>
<td>64</td>
<td>4</td>
<td>56</td>
<td>10</td>
<td>134</td>
</tr>
<tr>
<td>d:f gp</td>
<td>2</td>
<td>1</td>
<td>47</td>
<td>16</td>
<td>66</td>
</tr>
<tr>
<td>d:mn g:f</td>
<td>17</td>
<td>2</td>
<td>21</td>
<td>3</td>
<td>43</td>
</tr>
<tr>
<td>d:mn g:mn</td>
<td>39</td>
<td>0</td>
<td>45</td>
<td>3</td>
<td>87</td>
</tr>
<tr>
<td>d:mn gp</td>
<td>31</td>
<td>4</td>
<td>29</td>
<td>7</td>
<td>71</td>
</tr>
</tbody>
</table>

An example of an NP of this sort is given below.

(28) \[ \text{[t-éei [t-ées huphántikees] dunámei]} \]

\[
\text{thed:f theg:f weaving} \quad \text{powerd:f}
\]

‘with the power of weaving’ \hspace{1cm} P.Pol.281b

Notice that (nearly) all of the cells in (27) are filled. Only one cell here is empty (underlined); this is just the result of the small corpus size for Isocrates, however, as the corresponding cells for the other authors are all filled. The final column in (27) shows that each of the combinations is well attested across different authors.

(29) shows the number of occurrences of pairs of articles in which the first is an accusative singular (masculine and neuter are distinct here) and the second is a genitive of some sort. The pattern here mirrors that of (27): all the cells are filled and the total column is well-populated.

(29) \[ \text{NP accusative [NP genitive]} \]

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>I</th>
<th>P</th>
<th>X</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>a:f g:f</td>
<td>119</td>
<td>16</td>
<td>84</td>
<td>11</td>
<td>230</td>
</tr>
<tr>
<td>a:f g:mn</td>
<td>270</td>
<td>19</td>
<td>218</td>
<td>29</td>
<td>536</td>
</tr>
<tr>
<td>a:f gp</td>
<td>239</td>
<td>66</td>
<td>304</td>
<td>47</td>
<td>656</td>
</tr>
<tr>
<td>a:m g:f</td>
<td>39</td>
<td>4</td>
<td>21</td>
<td>11</td>
<td>75</td>
</tr>
<tr>
<td>a:m g:mn</td>
<td>2</td>
<td>2</td>
<td>48</td>
<td>9</td>
<td>61</td>
</tr>
<tr>
<td>a:m gp</td>
<td>41</td>
<td>11</td>
<td>51</td>
<td>11</td>
<td>114</td>
</tr>
</tbody>
</table>
An example of such an NP is given in (30).

(30)  [t-eèn  [t-óon  stroomátoon]  sún-thesen]
    the_a:f  the_g:f:p  bed_g:f:p  together-putting_d:f
    ‘the construction of beds’  P.Pol.280b

Sequences of articles in which the first is nominative and the second genitive

(31)  again mirror the patterns shown earlier for datives and accusatives:

(31)  [NP nominative [NP genitive]]

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>I</th>
<th>P</th>
<th>X</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>n:f g:f</td>
<td>106</td>
<td>0</td>
<td>30</td>
<td>4</td>
<td>140</td>
</tr>
<tr>
<td>n:f g:mn</td>
<td>287</td>
<td>0</td>
<td>79</td>
<td>6</td>
<td>372</td>
</tr>
<tr>
<td>n:f g:p</td>
<td>207</td>
<td>4</td>
<td>82</td>
<td>17</td>
<td>310</td>
</tr>
<tr>
<td>n:m g:f</td>
<td>55</td>
<td>1</td>
<td>12</td>
<td>2</td>
<td>70</td>
</tr>
<tr>
<td>n:m g:mn</td>
<td>75</td>
<td>1</td>
<td>34</td>
<td>10</td>
<td>120</td>
</tr>
<tr>
<td>n:m g:p</td>
<td>41</td>
<td>3</td>
<td>20</td>
<td>12</td>
<td>76</td>
</tr>
<tr>
<td>n:n g:f</td>
<td>137</td>
<td>7</td>
<td>119</td>
<td>3</td>
<td>266</td>
</tr>
<tr>
<td>n:n g:mn</td>
<td>196</td>
<td>1</td>
<td>114</td>
<td>7</td>
<td>318</td>
</tr>
<tr>
<td>n:n g:p</td>
<td>290</td>
<td>11</td>
<td>112</td>
<td>34</td>
<td>447</td>
</tr>
</tbody>
</table>

(Again, the zeros in Isocrates are due to the small size of his collected works.) An example of this sort of NP is given in (32).

(32)  [h-ee  [t-óu  pleéth-ous]  arkh-eé]
    the_n:f  the_g:n  crowd_g:n  rule_d:f
    ‘the rule of the crowd (democracy)’  P.Pol.291d

But compare (27), (29) and (31) to (33), which shows the number of occurrences of pairs of articles in which both are genitive.

(33)  [NP genitive [NP genitive]]

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>I</th>
<th>P</th>
<th>X</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>n:f</td>
<td>137</td>
<td>7</td>
<td>119</td>
<td>3</td>
<td>266</td>
</tr>
<tr>
<td>n:mn</td>
<td>196</td>
<td>1</td>
<td>114</td>
<td>7</td>
<td>318</td>
</tr>
<tr>
<td>n:p</td>
<td>290</td>
<td>11</td>
<td>112</td>
<td>34</td>
<td>447</td>
</tr>
</tbody>
</table>
An example of this kind of NP is given in (34).

(34)  [t-ées [t-ón himatí-oon] ergasí-as]
      the clothing production
      ‘of the production of clothing’ P.Pol.280e

In (27), (29) and (31) none of the pairs of articles is homophonous and (almost) every cell is filled. In (33), on the other hand, a number of cells (those in the rectangular boxes) contain homophonous pairs of articles and exactly these cells are empty.9

Expanding the search to cover all of attested Greek literature from Homer (mid-8th century BCE) to Chares (3rd to 2nd century BCE) yields not a single pair of adjacent homophonous articles:

(35)  Greek literature to the 3rd century BCE
      all authors 0 0 0
      toú toú tées tées tóon tóon

This search spans 500 years and 67 authors and includes genres from comic and tragic plays to biography, philosophy, history, poetry, and oration in a number of Greek dialects including Aeolic, Attic, Doric, and Ionic. This confirms that the generalization is not merely a trend, nor is it limited by dialect, genre or period.10

So far I have shown that there is an asymmetry in the attested types of center-embedded noun phrases in Greek such that NPs with adjacent homophonous
articles are systematically unattested. This is not to say, however, that non-adjacent homophonous articles are not allowed. Identical forms of articles often appear within a possessed NP, but never \textit{adjacent}. The internal syntax of noun phrases in Greek allows an alternative to center-embedding of a possessive NP: one commonly finds genitival NPs postposed (36) rather than center-embedded (37).

\begin{align*}
\text{(36)} & \quad \text{[h-ee tólm-a] [t-óon leg-ónt-oon]} \\
& \quad \text{the}_{n:f} \text{ courage}_{n:f} \text{ the}_{g:m:p} \text{ speak-ing}_{g:m:p} \\
& \quad \text{‘the courage of those speaking’} \\
& \quad \text{L.12.41}
\end{align*}

\begin{align*}
\text{(37)} & \quad \text{[h-ee [t-óon leg-ónt-oon] tólμ-α]} \\
& \quad \text{the}_{n:f} \text{ the}_{g:m:p} \text{ speak-ing}_{g:m:p} \text{ courage}_{n:f} \\
& \quad \text{‘the courage of those speaking’} \\
& \quad \text{[construct]}
\end{align*}

Such postposed genitives are always available for cases in which center-embedding would result in adjacent homophones (38).

\begin{align*}
\text{(38)} & \quad \text{[[t-óon oikeí-oon] tin-ās] [t-óon ekeín-oon]} \\
& \quad \text{the}_{g:f:p} \text{ slave}_{g:f:p} \text{ some}_{f:a:p} \text{ the}_{g:m:p} \text{ those}_{g:m:p} \\
& \quad \text{‘some of the slaves of those [people]’} \\
& \quad \text{P.A.33d}
\end{align*}

Were (38) to be center-embedded, it would contain adjacent homophones, as the ungrammatical construct in (39) shows.

\begin{align*}
\text{(39)} & \quad \ast \text{[[t-óon [t-óon ekeín-oon] oikeí-oon] tin-ās]} \\
& \quad \text{the}_{g:f:p} \text{ the}_{g:m:p} \text{ those}_{g:f:p} \text{ slaves}_{g:m:p} \text{ some}_{f:a:p} \\
& \quad \text{‘some of the slaves of those [people]’} \\
& \quad \text{[construct]}
\end{align*}

Cases like (38) abound while cases like (39) are unattested. For instance, cases of two \textit{non}-adjacent feminine genitive singular articles are common enough (40, 42), although adjacent instances of homophonous articles do not occur (41, 43) as has been seen.

\begin{align*}
\text{(40)} & \quad \text{[[t-ées arkh-ées] [t-ées pól-eoos]]} \\
& \quad \text{the}_{g:f} \text{ dominion}_{g:f} \text{ the}_{g:f} \text{ city}_{g:f} \\
& \quad \text{‘of the dominion of the city’} \\
& \quad \text{P.Pol 275a 8}
\end{align*}
(41) *[t-ées [t-ées pól-eoos] arkh-ées]]
    the\textsubscript{g:f} the\textsubscript{g:f} city\textsubscript{g:f} dominion\textsubscript{g:f}
    ‘of the dominion of the city’ [construct]

(42) [apò [[t-ées arkh-ées] [t-ées metabás-eoos]]]
    from the\textsubscript{g:f} beginning\textsubscript{g:f} the\textsubscript{g:f} change\textsubscript{g:f}
    ‘from the beginning of the change’ A.Poet 2.1455b 28

(43) *[apò [t-ées [t-ées metabás-eoos] arkh-ées]]]
    from the\textsubscript{g:f} the\textsubscript{g:f} change\textsubscript{g:f} beginning\textsubscript{g:f}
    ‘from the beginning of the change’ [construct]

Cases like these show that only adjacent homophones are ruled out.

Why should adjacent definite articles be prohibited just in case they are homophonous? I would like to propose that this is motivated by the Obligatory Contour Principle (OCP). The OCP prohibits consecutive identical autosegments (Leben 1973, 1978; Goldsmith 1976, 1984; Pulleyblank 1986), segments (Steriade 1982; Schein & Steriade 1986; Hayes 1986; Prince 1984) or syllables (Yip 1993). It is generally taken to constrain the application of rules as well as the shapes of underlying morphological representations (McCarthy 1988; Clements 1988; Davis 1991).

McCarthy (1986) coined the term ‘antigemination’ for a constraint which prohibits syncope rules from creating clusters of identical consonants. McCarthy attributes the blocking affect to the OCP: a derivation that yields a violation of the OCP is blocked. The phenomenon is widespread (Borowsky 1987; Archangeli 1986; Clark 1986; Itô & Mester 1986; Myers 1987; Yip 1988). Returning now to the Greek data, we may introduce the term antihomophony for the blocking of a syntactic construction that would contain a violation of the OCP at the phonological word level of representation: antihomophony bans adjacent
homophonous morphemes within a phonological word. Using \( [\square] \) to indicate some stretch of sound, A and B to indicate some stretch of meaning, and \( \square \) to indicate a phonological word, we may formulate antihomophony as follows:

\[
(44) \text{Antihomophony} \\
* [\ldots \square \square \ldots] \square \\
\quad | \quad | \\
\quad A \quad B
\]

Note that antihomophony is not a restriction against adjacent identical syllables, which may occur across word-boundaries (45) and within word-boundaries as a result of affixation (46) or reduplication (47):

\[
(45) \quad \text{autò \ tò \ kalón} \\
\quad \text{itself\:n\:n \ the\:n\:n \ beauty\:n\:n} \\
\quad \text{‘beauty itself’}
\]

\[
(46) \quad \text{toút-ou} \quad \text{[toú.tou]} \\
\quad \text{loos-en-s-aás-aa} \quad \text{[luu.saá.saa]} \\
\quad \text{‘this\:aor \ f\:dual\:part’}
\]

\[
(47) \quad \text{ke-kéleu-ka} \quad \text{[ke.ké.leu.ka]} \\
\quad \text{dè-de-ka} \quad \text{[dé.de.ka]} \\
\quad \text{‘redupcommand\:1s\:pf’}
\]

Rather, it is a constraint against adjacent homophones. This is clearly brought out in the contrast between [toútou [tóu érgou]] ‘of this work’ (48) and [tóu [tóu érgou]]... (49).

\[
(48) \quad \text{toútou \ tóu \ érgou} \\
\quad \text{this\:g\:n \ the\:g\:n \ work\:g\:n}
\]

\[
(49) \quad \ast \quad \text{tóu \ tóu \ érgou} \\
\quad \text{the\:g\:n \ the\:g\:n \ job\:g\:n}
\]

(48) has no adjacent homophones and thus does not violate the OCP; (49) does have adjacent homophones (tou and tou) and is ruled out. Whereas the structure in (48) occurs quite commonly in Greek, the one in (49) is unattested:

\[
(50) \quad \text{toútou tóu} \\
\quad 52 \quad 2 \quad 70 \quad 23 \quad \text{Total} \quad 147
\]
tóu tóu 0 0 0 0 0 0

Of course, not any two homophonous words are subject to antihomophony. Paul Kiparsky (p.c.) points out that the restriction against homophones does not apply to sequences of content words. Consider the adjacent homophonous content words in (51) and (52), both from Plato’s Symposium.

(51) ho éroo-s éroo-s estin ouden-òs eè tin-òs; the\textsubscript{m} Desire\textsubscript{m} desire\textsubscript{m} is nothing\textsubscript{g:n} or some-thing\textsubscript{g:n} ‘Is Desire desire of nothing or of something?’ P.Symp.199e6

(52) hôti éstin hué-os gè thugatr-òs hopateèr pateéer that is son\textsubscript{g:m} PRT daughter\textsubscript{g:f} the father father ‘that a father is the father of a son or a daughter, right?’ ibid.199d6

What allows (51) and (52) to slip through the ban on adjacent homophones is that the homophony occurs across rather than within phonological words.

One might well suppose that it is not the phonological identity of adjacent definite articles that is prohibited but the identity of the morpho-syntactic features (case, gender, number) that define them. Thus, it might be argued, the sequence téēs téēs is not phonologically ill-fomed but morpho-syntactically ill-formed because the full set of features that identifies each morpheme is repeated: [gen fem sg] [gen fem sg]. The prohibition in question could then be given as in (53), where $\square F$, $\square G$, $\square H$ are the feature-value pairs that define the morphemes in question.

(53) *$[\square F \square G \square H] [\square F \square G \square H]$

The prohibition is ad hoc but does make the correct prediction for the case at hand, *$[\text{gen fem sg}] [\text{gen fem sg}]$. But consider now the full set of articles in Greek.
Although the prohibited *tées tées consists of two identical morpho-syntactic feature matrices, only some of the cases of prohibited *tóu tóu and *tóon tóon are featurally identical. This is because the phonological shape tóu may be either [gen masc sg] or [gen neu sg]; and the phonological shape tóon may be [gen fem pl], [gen masc pl] or [gen neu pl]. So although (53) would rule out all cases of *tées tées, it would rule out only some of the *tóu tóu cases (the cases in which both articles are either masculine or neuter) and only some of the *tóon tóon cases (the case in which both articles are either feminine, or masculine or neuter) as shown below.

(55) Morpho-syntactically identical (=) vs. unattested (*) pairs

<table>
<thead>
<tr>
<th></th>
<th>Nom</th>
<th>Gen</th>
<th>Dat</th>
<th>Acc</th>
<th>Nom</th>
<th>Gen</th>
<th>Dat</th>
<th>Acc</th>
</tr>
</thead>
<tbody>
<tr>
<td>fem</td>
<td>hee</td>
<td>tées</td>
<td>tééi</td>
<td>téén</td>
<td>hai</td>
<td>tóon</td>
<td>táis</td>
<td>taás</td>
</tr>
<tr>
<td>masc</td>
<td>ho</td>
<td>tóu</td>
<td>tóoi</td>
<td>tón</td>
<td>hoí</td>
<td>tóon</td>
<td>tóis</td>
<td>toús</td>
</tr>
<tr>
<td>neut</td>
<td>tó</td>
<td>tóu</td>
<td>tóoi</td>
<td>tó</td>
<td>tó</td>
<td>tóon</td>
<td>tóis</td>
<td>tóis</td>
</tr>
</tbody>
</table>

But the phonological sequences *tóu tóu and *tóon tóon are prohibited regardless of their morpho-syntactic featural composition. Thus an analysis in terms of morpho-syntactic features is unable to capture the classes of non-occurring pairs of articles.\(^{11}\)
It might be possible to devise an underspecified morpho-syntactic representation in which phonologically identical articles are always featurally identical. The genitive masculine and neuter singular τόu might then be simply [gen, -fem, -pl]. One would then need to locate some principle of grammar (analogous to the OCP) that forbade such a sequence; I am aware of no such principle in the literature. In any case, a morphosyntactic analysis cannot work with the morphological cases of antihomophony in Greek (section 3), in Japanese or in Italian (section 4) because the items that violate antihomophony in each of these cases are indisputably morphosyntactically distinct. I turn now to the first of these cases.

3 Antihomophony and morphology
Antihomophony holds both within lexical phonological words and within postlexical phonological words or clitic groups.\textsuperscript{12} We’ve already considered what happens when a syntactic construction (a center-embedded NP) violates it: the construction is blocked and another equivalent syntactic structure (NP with a postposed possessor) takes its place. We turn now to the effects of antihomophony on morphology. Here there are two cases, one involving morphological haplology (2.1), the other involving allomorphy of a negative particle (2.2).

3.1 Reduplicative e- and past tense e-
The first case is discussed by Stemberger (1981) as an instance of morphological haplology. Consider the reduplicated perfect and pluperfect stems graph ‘write’ and phthi ‘decline’:

(56) Stem | Perfect | Pluperfect  
---|---|---  
graph- | gé-graph-a | e-ge-gráph-ee  
phthi- | é-phthi-mai | e-phthí-meen

Roots like graph-, which reduplicate the initial consonant with a fixed vowel [e], take the past tense prefix e- in the Pluperfect. But stems like phthi-, whose reduplicant consists solely of e-, fail to take the past tense marker e- in the Pluperfect. Stemberger points out that morphological haplology is obligatory before a morpheme e- and never occurs with a non-morphemic stem-initial /e/. That is, only heteromorphemic e’s that are themselves morphemes delete, in conformity with (44). The deletion can be understood as a response to the ranking of prosodic constraints over morphological constraints in Greek:

(57)

<table>
<thead>
<tr>
<th></th>
<th>PHON</th>
<th>MORPH</th>
</tr>
</thead>
<tbody>
<tr>
<td>e-e-phthí-meen</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>&lt;e-&gt;e-phthí-meen</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

The phonological constraint in question is antihomophony and the morphological constraint is EXPONENCE (Prince & Smolensky 1993), a constraint which requires that morphemes be overt. Antihomophony forces violation of the latter constraint, suggesting that this phonological constraint outranks a morphological constraint.
3.2 meé, meé and ou

Other data from Greek show antihomophony at work in a startling way: in these data, antihomophony forces the appearance of a contextually inappropriate allomorph of a negative particle. In this case, phonology does not pick between competing structures as it does with center-embedded and postposed NPs; instead, it forces a violation in the morphology, yielding a realis negative in an irrealis context.

The data concerns the homophones meé ‘not’ and meé ‘lest’. The first is a negative particle that occurs in irrealis clauses, the second a complementizer that governs irrealis clauses. The fact to be explained is that meé 'not' is replaced with the realis negative ou 'not' in irrealis clauses just in those cases where the clause is governed by meé 'that'. I argue that this is another instance of antihomophony.14 meé is used in irrealis contexts primarily governing optative (58), subjunctive (59) and imperative (60) verb forms.

(58) eè meé dzoó-ieen
     or not live₁ opt
     ‘or may I not live’                      Ar. Eq. 833

(59) meè phóo-men
     not say₁p subj
     ‘shall we not say?’                       P. Rep.554b

(60) meè még-a lég-e
     not big₁s:n:p speak₂ imp
     ‘don’t boast’                              P. Ph.95b

The other negative particle, ou (ouk before a vowel-initial word) is used in realis contexts primarily governing indicative forms of the verb (61).

(61) ouk en-no-óó
The string meé also occurs as a subordinating conjunction (‘lest’) introducing subordinate clauses like the object clauses used with verbs of fearing (62-63).

(62) dé-doi-ka meè... epilatho-tha t-ées oik-ades homeward
redupfear1pf lest lose1p subj theg:f roadg:f
‘I fear we may forget the way home’ X. A.3.2.25

(63) édei-s-an h-oi héllon-es meè forth-lead3p opt
fear3paor thenm:p Greeknm:p lest ‘The Greeks feared they might advance...’ X. A.1.10.9

Depending on the tense/aspect of the matrix verb, the verbs of these subordinate clauses are marked for either subjunctive (62) or optative (63) mood. Recall now that the normal negative for subjunctives and optatives is (irrealis) meé rather than (realis) ou. When (62) and (63) are negativised with meé, we expect to find two adjacent instances of meé (‘lest’ and ‘not’); in precisely these cases, the second meé is replaced by ou, despite the fact that the clause in which the negative occurs is irrealis. Consider the negative object clause in (64), comparing it to the non-negative object clauses in (62-63).

(64) dé-di-men meè ou bébaï-oï ée-te
redupfear1p indic lest not be2p subj
‘We fear you are not to be depended on’ T. 3.57

In (64) the verb of the subordinate clause (éete ‘you are’) appears in the subjunctive, but the negative that goes with it is (realis) ou rather than the expected (irrealis) meé. The conditioning factor is the preceding meé ‘lest’. The prosodic constituency of (64) is given in (65): meé, ou and éete ‘be’ are all function words and do not form phonological words of their own. Note that
replacing realis \textit{ou} with (the contextually appropriate) irrealis \textit{meé} would result in adjacent homophones internal to a phonological word (66).

\begin{enumerate}
\item[(65)] contextually inappropriate negative: grammatical
\begin{center}
\begin{tikzpicture}
  \node {dédimen \textit{meé ou} bébaioi éete};
\end{tikzpicture}
\end{center}

\item[(66)] contextually appropriate negative: ungrammatical
\begin{center}
\begin{tikzpicture}
  \node {dédimen \textit{meé meé} bébaioi éete};
\end{tikzpicture}
\end{center}
\end{enumerate}

A computer search reveals that adjacent instances of heteromorphemic \textit{meé} are completely unattested in the Greek corpus.\textsuperscript{15}

That the conditioning factor is \textit{meé} can also be seen in negative clauses introduced by conjunctions other than \textit{meé}, in which irrealis negative \textit{meé} surfaces with subjunctive and optative verbs as expected. Compare irrealis \textit{meé} in (67) and (68) with contextually inappropriate realis \textit{ou} in (60).

\begin{enumerate}
\item[(67)] hína \textit{meé} t-à zdeúg-ee heem-óon strateeg-éei
so that not the\textsuperscript{n:n:p} wagon\textsuperscript{n:n:p} us\textsuperscript{g:m:p} command\textsuperscript{3 sg subj}
‘so that our wagons not command us’ X.A.3.2.27

\item[(68)] hópoos \textit{meé} apothán-eei
so that not die\textsuperscript{3 sg subj}
‘so that he will not die’ X.M.2.10.2
\end{enumerate}

The unexpected occurrence of the realis negative particle \textit{ou} in irrealis contexts in which the negative particle follows \textit{meé} ‘lest’ is another clear case of
antihomophony. The prohibition against adjacent homophonous function words captures a significant generalization between two otherwise puzzling facts: (i) adjacent homophonous articles are systematically avoided and (ii) the realis particle ou is used in irrealis contexts only when it occurs in a clause headed by the complementiser meé.

The important difference between these two cases lies in the nature of the structures that antihomophony chooses from. The tableaux in (69-70) summarize the analysis. Consider first the center-embedded NP case (69). The first candidate violates antihomophony; the third violates a morphological constraint on agreement (tōu is masculine or neuter, póleoos is feminine); the second violates neither and is the only grammatical candidate.

(69)

<table>
<thead>
<tr>
<th></th>
<th>SYN</th>
<th>PHON</th>
<th>MORPH</th>
</tr>
</thead>
<tbody>
<tr>
<td>(tées tées póleoos) (arkhées)</td>
<td></td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>(tées arkhées) (tées póleoos)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(tées tōu póleoos) (arkhées)</td>
<td></td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

In the meé meé case (70) antihomophony chooses between a contextually appropriate form (irrealis meé) and a contextually inappropriate form (realis ou) and forces the appearance of the latter.

(70)

<table>
<thead>
<tr>
<th></th>
<th>SYN</th>
<th>PHON</th>
<th>MORPH</th>
</tr>
</thead>
<tbody>
<tr>
<td>(meè meé bêbaioi éete)</td>
<td></td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>(meè ou bêbaioi éete)</td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>
This shows that antihomophony is ranked higher than the constraint that selects appropriate allomorphs for various morpho-syntactic environments.

4 Other cases

The argument presented above rests on the claim that prosodic constraints never force syntactic ill-formedness (but may force morphological ill-formedness). In this section I review all the cases I know of that bear on this issue and conclude that they are consistent with this claim.

4.1 Antihomophony

I’m aware of two additional cases of antihomophony in the literature. Both show the same thing: when the syntax threatens a violation of antihomophony, the morphology, never syntax, is the component that yields.

4.1.1 Japanese

Japanese has three syntactic particles—genitive, nominal, copula— with the same phonological shape [no]. As Poser (1984) has shown, however, the language does not generally tolerate adjacent occurrences of no. Where we would expect two instances of no we find only one:

(71) Japanese *no no

a. no (genitive) + no (nominal )

*Zyon no no
John GEN NOM
‘John’s (thing)’

Zyon no
‘John’s (thing)’
b. no (copula) + no (nominal)

\[ \text{utyooteN no no wa Hanako da} \]
\[ \text{ecstatic COP NOM TOP Hanako be} \]
\[ \text{utyooteN na no no wa Hanako da} \]
\[ \text{‘The one who is ecstatic is Hanako.’} \]

Poser proposes the following deletion rule for Japanese\(^6\):

(72) No-Haplology: no > Ø / ___ no

The alignment constraints governing prosodic phrasing in Japanese are given in (73), following Terada (1986), Selkirk (1990) and others.

(73) \( X \frac{}{\sqcup} = \frac{}{\sqcup} \) (X a lexical category)
\( X \frac{}{\sqcup} = \frac{}{\sqcup} \)

This yields the following prosodic constituencies for the phrases in (71):

(74) Prosodic Phrasing
a. (Zyon no no)\( \frac{}{\sqcup} \) (71a)

b. (utyooteN no no wa)\( \sqcup \) (Hanako da)\( \sqcup \) (71b)

The haplology in (71) can thus be attributed to antihomophony, exactly as in Greek. The different outcomes in the two languages are attributable to the syntax: Japanese has no syntactic alternatives to the no no cases but Greek does have syntactic alternatives to the tées tées, etc. cases, namely postposition of the embedded NP. A tableau for (71b) is given below, in which \(<\text{no}>\) marks the occurrence of an EXPONENCE violation.

(75)
4.1.2 Italian

The second case of antihomophony outside of Greek comes from Italian. The syntax of Italian should produce a number of homophonous clitic sequences, but these are not allowed. Instead, one finds substitutions of other clitics for one of the offending homophones (Lepschy & Lepschy 1989, 212; Bonet 1991). The expected sequence *si si (reflexive + impersonal) in (76) surfaces instead as *ci si.

(76) *quando si si sgevlia presto, si si alza volentieri
when refl imp wake early refl imp rise willingly

quando ci si sgevlia presto, ci si alza volentieri
‘when one wakes up early, one gets up willingly’

Here again, the syntax offers no other order of clitics, so antihomophony is not selecting among two or more equal syntactic candidates. Rather, antihomophony forces the mis-selection of a morphological exponent for the reflexive: reflexive *si is realized as *ci along side impersonal *si for the same reason that Greek *meé ‘not’ is realized as ou alongside meé ‘lest’—avoidance of homophones within a phonological word is more important than exact morphological identity:

(77)

<table>
<thead>
<tr>
<th></th>
<th>SYN</th>
<th>PHON</th>
<th>MORPH</th>
</tr>
</thead>
<tbody>
<tr>
<td>(si si alza) (volentieri)</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(ci si alza) (volentieri)</td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

4.2 Heaviness

Two additional cases of prosodic influence on syntax need to be examined here. Both have been recently discussed in Zec & Inkelas (1990).
4.2.1 English

The first involves heavy NP shift in English:

(78) a. He threw into the wastebasket the letter which he had not decoded
    b. * He threw into the wastebasket the letter
    c. * He threw into the wastebasket it

(78b, c) are ungrammatical according to Zec and Inkelas because the postposed NPs the letter and it are not prosodically heavy, i.e., do not contain branching phonological phrases. But both postposed and in situ constuctions are available in English, as (79) shows:

(79) a. He threw the letter which he had not decoded into the wastebasket
    b. He threw the letter into the wastebasket
    c. He threw it into the wastebasket

The prosodic constraint here rules out one of two available syntactic structures, exactly as in the Greek data discussed in section 2:

(80)

<table>
<thead>
<tr>
<th></th>
<th>SYN</th>
<th>PHON</th>
<th>MORPH</th>
</tr>
</thead>
<tbody>
<tr>
<td>(he threw it) (into the wastebasket)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(he threw) (into the wastebasket) (it)</td>
<td></td>
<td></td>
<td>*!</td>
</tr>
</tbody>
</table>

4.2.2 Serbo-Croatian

Zec and Inkelas’ second case involves topicalization in Serbo-Croatian: a topicalised constituent must consist of a branching phonological phrase. Again, the heaviness constraint does not place any direct conditions on syntactic
constructions—it merely chooses between two already available constructions, one with a preposed topic, the other with the topic in situ.

5 Alternative accounts
In the preceding sections I’ve tried to show that the only effect prosody ever has on syntax is to choose between structures which are equally well-formed syntactically. The constraint-based analysis in (3) predicts exactly this because it is couched within a theory of grammar (OT) whose only role for dominated constraints is to select among representations which are not ranked by higher ranked constraints. (3) does not allow prosody to block a syntactic construction if it is the only syntactic construction available: the very nature of Optimality Theory guarantees an output, even if that output violates a constraint.

On the other hand, syntactic influence on prosody is almost total: the phrasing algorithms for phonological words and phrases are strictly determined by (but not reducible to) syntactic bracketing and labeling. Again, (3) has this as a natural result: alignment of prosodic and syntactic categories brings prosodic constituency in line with syntactic constituency, not the reverse.

How might the models based on the premises in (1) and (2) account for this? Let’s begin with the first type, repeated below as (81):

(81) Syntax derivationally precedes phonology
(81) is subject to both a weak and a strong interpretation. Myers (1991) offers the strong interpretation, claiming that constraints on phonological representations can block only phonological rules. If the antihomophony analysis of the Greek facts above is correct, this must be wrong.
A weaker interpretation of (81) would include constraints that can in principle filter out unwanted syntactic structures. Such a model would allow a prosodic constraint like antihomophony to filter out the offending cases. The problem in this case is that such filters are clearly too strong. (81) can handle the center-embedding facts by generating both center-embedded and post-posed possessives and filtering out offending cases with antihomophony. Given such an analysis, however, (81) becomes too weak to rule out another possibility. Imagine a language Greek such that the syntax only produces center-embedded possessors; a grammar with rules and filters has no principled way of keeping antihomophony from blocking these structures too, with the result that there would be no grammatical way to express a possessed NP with homophonous articles. The attested Greek facts do not warrant such a strong role for filters.

Similarly for the English and Serbo-Croatian cases: each shows a prosodic constraint that rules out one of two competing syntactic expressions. There are no cases where a prosodic constraint rules out the only available syntactic expression. Such cases are easily imaginable: eg, a dialect of German in which the strings die die and das das are ungrammatical or a dialect of Italian in which ‘one gets up willingly’ is not expressible.

Now let us consider Zec and Inkelas’s (1990) bidirectional model (2), repeated below as (82):

(82) Syntax constrains prosody and vice versa

Such a model has no difficulty with the Greek data presented here, of course. The problem is that (82) predicts the existence of grammars in which prosody plays a role in determining syntactic well-formedness: in particular, (82) is
compatible with a grammar in which prosodic structure and labeling dictate syntactic structure and labeling. So in addition to cases like (83a), where the constituency and labeling (XP, X₀) of the syntax determine the prosodic structure, we should have cases like (83b) where the constituency and labeling ([], P) determine the syntactic structure.

(83)  a.  S determines P  b.  P determines S

(83b) allows constraints on the syntax of the following type:

(83) Possible constraints in a bidirectional model
    (i) any branching phonological phrase constitutes a verb phrase
    (ii) any non-branching phonological phrase constitutes a noun phrase

But such prosodic constraints on syntactic constituency and labeling do not occur in natural language. This shows that the Zec and Inkelas model is too permissive, as it has no principled way of excluding constraints of this type.
5 Conclusion

All the data of which I’m aware suggest that prosodic constraints play a very limited, but nonetheless important role in filtering out possible syntactic representations. Essentially, prosody can block syntactic constructions only if there are other structures that can fill in for the blocked ones. This special type of blocking scenario is predicted by the model presented in (3). Derivational models (1) and models in which prosody interacts with syntax bidirectionally (2) are unable to capture this directly.
Notes

* This work would not have been possible without extensive free use of the computerized corpus Thesaurus Linguae Graecae. I would like to thank D. J. Mastronarde for helping me get access to this incredible source of data. I would also like to thank Cheryl Chan, Stan Hoffer, Sharon Inkelas, Michael Israel, Paul Kiparsky, Geoff Pullum, John Rickford, Ivan Sag, Tom Wasow, Richard Wiese and Moira Yip for helpful input. A large debt of gratitude to two anonymous NLLT reviewers, two anonymous Phonology reviewers and to Ellen Kaisse, whose careful comments and suggestions were invaluable. All errors or inconsistencies are of course my own.

1 Zec and Inkelas (1990) assume that all syntax-phonology interactions are mediated by prosodic structure. This allows prosody some interaction with syntax but denies it to segmental and sub-segmental phonological information. Thus the term ‘prosody’ rather than ‘phonology’ in (2).

2 The constraint seems to generalize to any string of homophonous function words, at least for some speakers. In my speech, for instance, phrases like *the star in In Harm’s Way, *he went to To Live and Die in L.A., *the video of Of Thee I Sing are ungrammatical without a pause between the offending homophones.

3 The second and third homophones may be separated by an adverb:
   (a) die die gestern die Blumen gekauft haben
       those who yesterday the flowers bought have
       ‘those who bought the flowers yesterday’
   (b) das das wirklich das Problem ist
       that this really the problem is
       ‘that this really is the problem’
Hayes 1989 was circulated in 1984 as a manuscript; hence the funny chronology of the citations.

Tonally dependent words are best analyzed as bearing floating high tones (Sauzet 1989, Golston 1989). There is no segmental, metrical or epigraphic evidence for a prosodic constituent E-Word, as argued for in Steriade 1988.

I use NP as a theory-neutral cover term for DP, N”, etc.

Glosses are marked for case, gender and number in that order, using the following abbreviations:

\[ \begin{align*}
    a &= \text{accusative} & g &= \text{genitive} & d &= \text{dative} \\
    f &= \text{feminine} & m &= \text{masculine} & n &= \text{neuter} \\
    p &= \text{plural}
\end{align*} \]

Nominative case and singular number are left unmarked in the glosses.

The following abbreviations are used for authors and their works.

\[
\begin{aligned}
A &\quad \text{Aristotle} \\
\text{Poet} &\quad \text{Ars Poetica} \\
M &\quad \text{Metaphysics} \\
Ar &\quad \text{Aristophanes} \\
\text{Eq} &\quad \text{Equites} \\
I &\quad \text{Isocrates} \\
L &\quad \text{Lysias} \\
P &\quad \text{Plato} \\
\text{A} &\quad \text{Apology} \\
\text{Euth} &\quad \text{Euthyphro} \\
\text{Ph} &\quad \text{Phaedo} \\
\text{Pol} &\quad \text{Politicus (‘Statesman’)} \\
\text{Rep} &\quad \text{Republic} \\
S &\quad \text{Sophist} \\
\text{Symp} &\quad \text{Symposium} \\
S &\quad \text{Sophocles} \\
\text{OC} &\quad \text{Oedipus at Colonus} \\
T &\quad \text{Thucydides} \\
\text{Peloponnesion Wars} &\quad \text{Peloponnesion Wars} \\
X &\quad \text{Xenophon} \\
\text{A} &\quad \text{Anabasis} \\
\text{M} &\quad \text{Memorabilia} \\
\text{Oec} &\quad \text{Oeconomicus}
\end{aligned}
\]

The corpus contains one possible exception, from Aristotle’s Metaphysics:
In Aristotle’s time, tone and word-division were not marked. What is written here as τόu τόu was merely TOUTOU; this means that it could also be read as the genitive singular deictic τοúτου ‘this’:

τοúτ-ου οντ-ος τ-ο mèn tó-de éinai
this gen being the a:nt PRT this be inf
‘of this being of being this’

The text is found in 4 medieval manuscripts, all of which indicate both tone and word-division: two manuscripts record τόu τόu ‘the the’ and two record τοúτου ‘this’. The single instance of adjacent homophonous articles in Greek is thus textually suspect and extremely isolated.

10 Ellen Kaisse (p.c.) points out that the Greek dual provides an additional testing ground for the proposal here. Forms of the dual articles are τόo (nominative, accusative) and τόin (genitive, dative). The TLG corpus contains no sequences of τόin τόin, as predicted; but it contains no sequences of τόo τόin either, due to the low occurrence of duals in Greek. The lack of τόin τόin sequences cannot therefore be taken as additional evidence for antihomophony.

11 Ivan Sag (p.c.) suggests that antihomophony might be due to processing factors rather than grammar. He points out that center-embedded sentences in English such as That that John left surprised me worried Mark are noticeably more difficult to process than similar sentences in which the complementizers are not homophonous: If that John left surprised me had worried Mark, he would have
said so. Something about the identity of the complementizers adds to the parsing problem.

The point is well taken. Indeed, if there were any evidence that center-embedded NPs in Greek were difficult to parse, one would want to treat them in a manner analogous to center-embedded sentences in English. But exactly the opposite seems to be the case: whereas (even) putatively easy center-embedded sentences like *That that John left surprised me worried Mark* are practically non-existent in English texts, center-embedded NPs are the normal way of constructing possessed NPs in Greek. Furthermore, whereas triple center-embedding seems to be wholly impossible in English (*The rat the cat the dog chased ate died*), it is clearly not so in Greek, as (4) and (5) demonstrate.

The fact that enter-embedding occurs as frequently as it does in Greek suggests that the prohibition against it in English is not due to the processing abilities of humans but to the grammar of English. The only alternative is to posit that Greek speakers processed language differently than we do today, an unlikely possibility.

12 See Zec (1993), who replaces Hayes’ (1989) term Clitic Group with the postlexical phonological word.


14 Or was. Brett Kessler (p.c.) points out that other particles may intervene between meé ‘lest’ and ou ‘not’. That is, the realis negative is chosen even when the complementizer is not string-adjacent to the negative. This suggests that the use of realis ou ‘not’ in irrealis clauses headed by meé 'lest' derives from a period of Greek in which complementizer and negative were always adjacent.
Alternatively, morphologically induced violations of the OCP may simply span longer distances than syntactically induced violations.

15  Adjacent homomorphemic instances of meé, on the other hand, do occur, as predicted.

    (i) meè meè meé m’anéree tís éimi
        not not not me ask who I:am
        ‘Don’t! don’t! don’t ask me who I am’ S. OC. 210

An anonymous reviewer points out that such cases probably involve distinct phonological phrases. Other such cases occur in Sophocles’ Ajax (1191) and in Aristophanes’ Wasps (1418) and Peace (927); a similar but post-classical case occurs in Matthew 5.37.

16  Deletion is blocked when an NP-boundary (#) separates the two no:

    (i) no (nominal) + no (genitival )
        akai no # no futa
        red NOM GEN lid
        *akai no futa
        ‘the red one’s lid’

Poser adds this as a condition to the rule of No-haplology, but we need not stipulate this if the NP-boundary induces a new phonological phrase, as we’d expect from the general phrasing algorithms.

17  ci is otherwise a clitic that means (i) ‘us’, (ii) ‘with her/him’, (iii) ‘here, there’.

18  Or clitic group. Nespor & Vogel (1986, 147ff) argue for the latter, noting that, with respect to phonological rules, Italian clitics behave neither as part of a word nor as totally independent words.
References


* This work would not have been possible without extensive free use of the computerized corpus Thesaurus Linguae Graecae. I would like to thank D. J. Mastronarde for helping me get access to this incredible source of data. I would also like to thank Cheryl Chan, Stan Hoffer, Sharon Inkelas, Michael Israel, Paul Kiparsky, Geoff Pullum, John Rickford, Ivan Sag, Tom Wasow, Richard Wiese and Moira Yip for helpful input. A large debt of gratitude to two anonymous NLLT reviewers, two anonymous Phonology reviewers and to Ellen Kaisse, whose careful comments and suggestions were invaluable. All errors or inconsistencies are of course my own.

1 Zec and Inkelas (1990) assume that all syntax-phonology interactions are mediated by prosodic structure. This allows prosody some interaction with syntax but denies it to segmental and sub-segmental phonological information. Thus the term ‘prosody’ rather than ‘phonology’ in (2).

2 The constraint seems to generalize to any string of homophonous function words, at least for some speakers. In my speech, for instance, phrases like *the star in In Harm’s Way, *he went to To Live and Die in L.A., *the video of Of Thee I Sing are ungrammatical without a pause between the offending homophones.
The second and third homophones may be separated by an adverb:
(a) die die gestern die Blumen gekauft haben
   those who yesterday the flowers bought have
   ‘those who bought the flowers yesterday’
(b) das das wirklich das Problem ist
   that this really the problem is
   ‘that this really is the problem’

Hayes 1989 was circulated in 1984 as a manuscript; hence the funny
chronology of the citations.

Tonally dependent words are best analyzed as bearing floating high tones
(Sauzet 1989, Golston 1989). There is no segmental, metrical or epigraphic
evidence for a prosodic constituent E-Word, as argued for in Steriade 1988.

I use NP as a theory-neutral cover term for DP, N”, etc.

Glosses are marked for case, gender and number in that order, using the
following abbreviations:

a = accusative  g = genitive  d = dative
f = feminine  m = masculine  n = neuter
p = plural

Nominative case and singular number are left unmarked in the glosses.

The following abbreviations are used for authors and their works.

A Aristotle
   Poet Ars Poetica
M Metaphysics
Ar Aristophanes
   Eq Equites
I Isocrates
L Lysias
P Plato
   A Apology
   Euth Euthyphro
   Ph Phaedo
   Pol Politicus (‘Statesman’)
   Rep Republic
   S Sophist
   Symp Symposium
The corpus contains one possible exception, from Aristotle’s *Metaphysics*:

\[
tóu \; tóu \; ónt-os \; t-ò \; mèn \; tó-de \; éinai
\text{the} \; \text{the} \; \text{being} \; \text{the} \; \text{PRT} \; \text{this} \; \text{be}_{\text{inf}}
\]
\‘of the (property) of the being of being this’  A.M.2.1089a 14

In Aristotle’s time, tone and word-division were not marked. What is written here as \(tóu \; tóu\) was merely TOUTOU; this means that it could also be read as the genitive singular deictic toútou ‘this’:

\[
tóu \; tóu \; ónt-os \; t-ò \; mèn \; tó-de \; éinai
\text{the} \; \text{the} \; \text{being} \; \text{the} \; \text{PRT} \; \text{this} \; \text{be}_{\text{inf}}
\]
\‘of this being of being this’

The text is found in 4 medieval manuscripts, all of which indicate both tone and word-division: two manuscripts record \(tóu \; tóu\) ‘the the’ and two record toútou ‘this’. The single instance of adjacent homophous articles in Greek is thus textually suspect and extremely isolated.

Ellen Kaisse (p.c.) points out that the Greek dual provides an additional testing ground for the proposal here. Forms of the dual articles are toó (nominative, accusative) and toín (genitive, dative). The TLG corpus contains no sequences of toín toín, as predicted; but it contains no sequences of toó toín either, due to the low occurrence of duals in Greek. The lack of toín toín sequences cannot therefore be taken as additional evidence for antihomophony.
Ivan Sag (p.c.) suggests that antihomophony might be due to processing factors rather than grammar. He points out that center-embedded sentences in English such as *That that John left surprised me worried Mark* are noticeably more difficult to process than similar sentences in which the complementizers are not homophonous: *If that John left surprised me had worried Mark, he would have said so*. Something about the identity of the complementizers adds to the parsing problem.

The point is well taken. Indeed, if there were any evidence that center-embedded NPs in Greek were difficult to parse, one would want to treat them in a manner analogous to center-embedded sentences in English. But exactly the opposite seems to be the case: whereas (even) putatively easy center-embedded sentences like *That that John left surprised me worried Mark* are practically non-existent in English texts, center-embedded NPs are the normal way of constructing possessed NPs in Greek. Furthermore, whereas triple center-embedding seems to be wholly impossible in English (*The rat the cat the dog chased ate died*), it is clearly not so in Greek, as (4) and (5) demonstrate.

The fact that center-embedding occurs as frequently as it does in Greek suggests that the prohibition against it in English is not due to the processing abilities of humans but to the grammar of English. The only alternative is to posit that Greek speakers processed language differently than we do today, an unlikely possibility.

See Zec (1993), who replaces Hayes’ (1989) term Clitic Group with the postlexical phonological word.

Or was. Brett Kessler (p.c.) points out that other particles may intervene between meé ‘lest’ and ou ‘not’. That is, the realis negative is chosen even when the complementizer is not string-adjacent to the negative. This suggests that the use of realis ou 'not' in irrealis clauses headed by meé 'lest' derives from a period of Greek in which complementizer and negative were always adjacent. Alternatively, morphologically induced violations of the OCP may simply span longer distances than syntactically induced violations.

Adjacent homomorphemic instances of meé, on the other hand, do occur, as predicted.

(i) meè meè meè m’ anéree tís éimi
    not not not me ask who I:am
     ‘Don’t! don’t! don’t ask me who I am’ S. OC. 210

An anonymous reviewer points out that such cases probably involve distinct phonological phrases. Other such cases occur in Sophocles’ Ajax (1191) and in Aristophanes’ Wasps (1418) and Peace (927); a similar but post-classical case occurs in Matthew 5.37.

Deletion is blocked when an NP-boundary (#) separates the two no:

(i) no (nominal) + no (genitival )

    akai no # no futa
    red NOM GEN lid

*akai no futa
  ‘the red one’s lid’
Poser adds this as a condition to the rule of No-haplology, but we need not stipulate this if the NP-boundary induces a new phonological phrase, as we’d expect from the general phrasing algorithms.

17 *ci* is otherwise a clitic that means (i) ‘us’, (ii) ‘with her/him’, (iii) ‘here, there’.

18 Or clitic group. Nespor & Vogel (1986, 147ff) argue for the latter, noting that, with respect to phonological rules, Italian clitics behave neither as part of a word nor as totally independent words.