A surprising allomorphic span in Cypriot Greek

Jason Merchant and Natalia Pavlou

University of Chicago, February 2016

In Cypriot Greek, the element *thα* is an allomorph of the present tense copula and a selected subordinating element; it marks the future in negative contexts. In this paper, we document the distribution of this item for the first time, and we then present an analysis in Distributed Morphology that analyzes *thα* as a portmanteau morpheme realizing two heads in the context of negation. This analysis requires that we admit a more expansive definition of span (or targets of Fusion) than the previous literature has adopted.

1 The future and negation in Cypriot Greek

1.1 The periphrastic future: copula+*na*

Cypriot Greek, like its standard Greek sister spoken to its northwest across the Mediterranean (which we will refer to as ‘Mainland Greek’ for convenience), has a copula verb that inflects for person, number, and tense, but shows no number distinction in the third person (the past plural forms *itan/ito* are in free variation; the difference does not index number):

(1) Cypriot Greek copula *ime* ‘be’

<table>
<thead>
<tr>
<th></th>
<th>present</th>
<th>past</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td><em>ime</em></td>
<td><em>imun</em></td>
</tr>
<tr>
<td>2</td>
<td><em>ise</em></td>
<td><em>isun</em></td>
</tr>
<tr>
<td>3</td>
<td><em>en</em></td>
<td><em>itan/ito</em></td>
</tr>
<tr>
<td></td>
<td><em>imasten</em></td>
<td><em>imastan</em></td>
</tr>
<tr>
<td></td>
<td><em>isaste</em></td>
<td><em>isastan</em></td>
</tr>
</tbody>
</table>

*Acknowledgments to be added.*
This verb is used with adjectival, nominal, and prepositional predicates, just as in English (we illustrate only the first case).

(2) Ta pethkja en arosta.
\( \text{the child.NEUT.pl be.NONPAST.3 sick.NEUT.pl} \)
‘The children are sick.’

Cypriot Greek has a two-way tense distinction morphologically represented: past and non-past (see Holton et al. 1997 for more detail on the identical standard Greek system), and two aspects: imperfective and perfective. It marks the future\(^1\) periphrastically, using a construction that combines the (here invariant) 3rd person copula with a clause headed by the subordinating particle \( \text{na} \)\(^2\) (traditionally called ‘subjunctive’, and which we will assume is a complementizer for the time being) and a verb in the non-past:

(3) En na pao.
\( \text{be.NONPAST.3 na go.PERF.NONPAST.1sg} \)
‘I will go.’

This \( \text{na} \) is found in a range of contexts, as the head of selected control and raising CPs—Cypriot Greek, like standard Greek, lacks infinitives, and uses \( \text{na} \)-clauses in their stead:

(4) Thelo na pao.
\( \text{want.IMPERF.NONPAST.1sg na go.PERF.NONPAST.1sg} \)
‘I want to go.’

(5) Thelo na ton dho.
\( \text{want.IMPERF.NONPAST.1sg na him see.PERF.NONPAST.1sg} \)
‘I want to see him.’ (Terzi 1999b:230)

\(^1\)It is traditional to call this form the future, though close examination of its uses shows that it has non-temporal, modal uses as well; see Giannakidou and Mari (2014) for references and extensive discussion of the standard Greek. Pending a similar semantic investigation of the Cypriot Greek, we will continue to call this form the future for convenience. The semantic issues raised by a complex, multi-head construction being interpreted as a single modal operator are the usual ones for idioms, idiom chunks, collocations, and ‘constructions’ generally; compare e.g. \textit{must} with \textit{have to}.

\(^2\)Analyses vary somewhat in whether they take \( \text{na} \) to be a complementizer, as Philippaki-Warburton 1982 and Agouraki 1991 do, or to head a MoodP or other IP-internal functional projection, possibly selected by a null C or otherwise embedded under a C, as Roussou 2001 and Giannakidou 2009 do. All that is important for us is that \( \text{na} \) signals the presence of a CP layer, for reasons that will become clear below. To sidestep these labeling issues, we will therefore gloss \( \text{na} \) simply as ‘\( \text{na} \)’.
Each of the elements in *en na*, in other words, has an independent use: *en* as a verb and *na* as a (clausal) subordinator. The null hypothesis for the syntactic structure of the periphrastic future which combines these two elements, therefore, is to put them together in the usual way, with the verb *en* selecting a CP headed by *na*, as illustrated in the following simplified structure (where we suppress the representation of verb movement to T, as well as of other elements possibly present in the clause):

(6) TP
   T    VP
      V    CP
        en  C
          CP
            T    VP
              (pro\_1\_sg) V
                  pao

The *na* that occurs in this locution has the usual properties found with *na*-clauses elsewhere in the language: it triggers the dependent nonpast form of the verb following it, and occurs only with the negator *men* (see Chatzopoulou 2012). We illustrate the latter property here:

(7) Akusa oti en na men pais.
    hear.PERF.PAST.1sg that be.NONPAST.3 na NEG go.PERF.NONPAST.2sg

‘I heard that you will not go.’

There are two pieces of evidence that make it clear that the sequence *en na* is not a fixed or lexicalized expression of a single head, tense or otherwise.\(^3\) First, the

\(^3\)Precisely the same reasoning applies to the English modal expression *be to*, which has undergone a different path of development and has slightly different properties, but which bears obvious similarities to the Cypriot Greek locution. See Huddleston et al. 2002:113 and Kayne 2013 for brief discussion. The fact that Greek also has such an expression shows that Kayne’s proposed generalization about the distribution of this construction is wrong: he analyzes *Abby is to appear tonight* as *Abby is MEANT FOR to appear tonight* and claims that only a language like English that has a case-assigning complementizer like *for* will be able to have this construction. Greek lacks any such case-assigning complementizer. Kayne’s reasoning is based on the idea that there
copula can appear independently in the past tense, as discussed in Pavlou 2014, and illustrated in (8):

(8) Itan na pao ekso. (Pavlou 2014 (2a))

be.PAST.3 na go.PERF.NONPAST.1sg outside

‘I was going to go out.’

This is entirely expected if en is merely the nonpast 3rd person copula: itan, also found as ito (Vassiliou 2002), is the regular past tense.

Second, we find coordinated na-clauses under en (and under itan, shown in (27) below):

(9) a. En na pao che na
be.NONPAST.3 na go.PERF.NONPAST.1sg and na
kathariso.
clean.PERF.NONPAST.1sg
‘I will go and clean.’

b. En na mairepsis i na
be.NONPAST.3 na cook.PERF.NONPAST.2sg or na
katharisis?
clean.PERF.NONPAST.2sg
‘Are you cooking or cleaning?’

(Or: ‘Are you cooking, or are you cleaning?’)

c. En na mairepsis oksa na
be.NONPAST.3 na cook.PERF.NONPAST.2sg or na
katharisis?
clean.PERF.NONPAST.2sg
‘Which of the two are you doing: cooking, or cleaning?’

If en na were a single head or fixed lexical expression, then coordination of such na-clauses alone should not be possible.4

4As in standard Greek, we do not find coordinated VPs or Vs under na: na is a proclitic that cannot cliticize onto the first element of a coordinate structure.

4 is no way to assign the attested modal semantics to be to in its individual parts (and therefore the modality must come from unpronounced elements). But this reasoning strikes us as unfounded: the same question arises with any multi-morphemic idiom or partial compositional phrase, including call up ‘muster’, call out ‘challenge’, spill the beans ‘reveal the secret’, etc. Any solution to the issues raised for compositionality in such phrases (see e.g. Kobele 2012 and Harley 2014) will apply to be to and en na as well.
The verbal form selected by *na* in the periphrastic future also behaves as expected for such forms in Cypriot Greek with respect to the placement of clitics. Though object clitics in Cypriot Greek are enclitic in a number of matrix and embedded clauses (see Terzi 1999b), as illustrated in (10), clitic arguments of dependent verbs under *na* are proclitic, as seen in (11).

(10) Ida ton extes.
    *see.PERF.PAST.1sg him.CL yesterday*
    ‘I saw him yesterday.’

(11) a. Thelo na ton do avrio.
    *want.1sg na him.CL see.PERF.NONPAST.1sg tomorrow*
    ‘I want to see him tomorrow.’

The periphrastic future construction also triggers proclisis, as expected if the construction contains a regular *na*-clause:

(12) a. En na ton do avrio.
    *be.NONPAST.3 na him.CL see.PERF.NONPAST.1sg tomorrow*
    ‘I will see him tomorrow.’

Historical work on the development of the future *tha* in standard mainland Greek also suggests that more complex structures were involved, similar to the one we argue for here. Joseph and Pappas 2002 and Roberts and Roussou 1999 argue that mainland Greek *tha* involves a redeployment of the volitional verb *thelo* ‘want’ and the infinitive, the latter later replaced by the head *ina*. A use of the volitional verb and the *na*-clause to give a future reading in Cypriot Greek has also been documented for medieval Cypriot Greek in Aerts 1983:

(13) I tis theli na mini as mini.
    *PRT someone want.3sg na stay let stay.3sg*
    ‘If someone will stay, let him stay.’

Given such developments, it comes as no surprise that the Cypriot Greek future periphrastic construction is built on a biclausal structure, involving a *na*-clause.

5
The Cypriot Greek periphrastic en na future and the mainland Greek tha future also share some of the restrictions on their appearance, which is to be expected if the functions or meanings of the two are largely the same. For example, neither can appear as marking the main predicate in clauses with Cypriot Greek dixa ‘without’ (in (14)) or mainland Greek xoris ‘without’ (in (15)):

(14) a. *Dixa en na pi tipotis, efien.
   without be.3 na say.PERF.NONPAST.3s nothing, left.PERF.PAST.3s
   ‘Without saying anything, she left.’

   b. Dixa na pi tipotis, efien.
   without na say.PERF.NONPAST.3s nothing, left.PERF.PAST.3s
   ‘Without saying anything, she left.’

(15) a. *Xoris tha pi tipota, efige.
   without tha say.PERF.NONPAST.3s nothing, left.PERF.PAST.3s
   ‘Without saying anything, she left.’

   b. Xoris na pi tipota, efige.
   without na say.PERF.NONPAST.3s nothing, left.PERF.PAST.3s
   ‘Without saying anything, she left.’

1.2 Sentential negator: en

Sentential negation in Cypriot Greek is expressed in indicative clauses with the preverbal element en (compare standard Greek dhen), homophonous with the 3rd person nonpast form of the copula seen above:

(16) En pao.
    NEG go.IMPERF.NONPAST.1sg
    ‘I am not going.’

As expected, this negator can co-occur with the copula, both in the present and past:

(17) a. Ta pethkja en en arosta.
    the children NEG be.NONPAST.3 sick
    ‘The children are not sick.’

   b. Ta pethkja en itan arosta.
    the children NEG be.PAST.3 sick
    ‘The children were not sick.’
The expectation, then, is that the negator *en* and the copula with *na* in the periphrastic future should be able to combine. And this is true, when the copula is in the past:

(18) En itan na pao.

\[ \text{NEG be.PAST.3 na go.PERF.NONPAST.1sg} \]

‘I wasn’t going to go.’

It therefore comes as a surprise that sentential negation and the nonpast copula *en* in the periphrastic future cannot co-occur:

(19) *En en na pao.

\[ \text{NEG be.NONPAST.3 na go.PERF.NONPAST.1sg} \]

(Intended: ‘I will not go.’)

The fact that *en en* is not found in (19) is not due to mere haplology (even syntactically conditioned haplology, as in Salzmann 2013 and Merchant 2014), given that (17a) is well-formed.

What we find instead for the negated future is surprising. It has been observed before that “Cypriot Greek lacks the future particle [tha] of Standard Greek” (Terzi 1999a:110 fn 24), as shown in (20).

(20) *Tha tha pao.

\[ \text{th a go.PERF.NONPAST.1sg} \]

(‘I will go.’)

But precisely this *tha* that is otherwise absent from the language is the form that appears in the negated periphrastic future:

(21) En tha pao.

\[ \text{NEG th a go.PERF.NONPAST.1sg} \]

‘I will not go.’

The appearance and identity of this *tha* is the puzzle to be solved.

---

5 The example in (20) is of course acceptable in standard mainland Greek: it is in fact the usual way to express the future (see Tsangalidis 1998, Giannakidou and Mari 2014). The particle is only acceptable among Greek Cypriot speakers when the high variety of standard mainland Greek is used in code-switching or code-mixing (Tsiplakou 2009, 2014).

6 Similarly to *na*, we gloss this element simply as *tha*, in order not to prejudge its category or function. (Analyses differ on where the standard mainland Greek *tha* is introduced: in a lower C modality projection (Roussou 2000), in a FutP above TP (Spyropoulos and Philippaki-Warburton 2001), or as the realization of T *tout simple.*)
2 Spanning: Vocabulary Items that realize more than one node

The element *tha* is in complementary distribution with the usual future marking strategy *en na*. The most parsimonious analysis therefore is to take the two sets of elements to be allomorphs of the same nodes, one realized under negation, and the other otherwise.\(^7\)

Such an analysis can be implemented in a theory such as Distributed Morphology by positing the following Vocabulary Items:

\[
\begin{align*}
(22) & \quad \textbf{a. } \sqrt{\text{BE}} \leftrightarrow \text{en} / \text{T[pres]} \\
& \quad \textbf{b. } C \leftrightarrow \text{na} \\
& \quad \textbf{c. } \text{T[pres]} \sqrt{\text{BE}} C \leftrightarrow \text{tha} / \text{NEG} \\
\end{align*}
\]

By the Elsewhere Principle, insertion of the Vocabulary Items in (22a,b) will be preempted by the more narrow context of insertion in (22c).

The unusual property of the entry in (22c) is that it realizes three nodes. This apparent difficulty is overcome in a theory that countenances spanning (Svenonius 2012; Merchant 2015; Bennett et al. 2015).\(^8\) The resulting structure is as follows:

\(\quad \)

---

\(^7\)This distribution has obvious parallels to negative polarity items; on one reading, we can claim that *tha* is an NPI in Cypriot Greek. Such a claim does not immediately help with understanding its morphosyntactic distribution, of course. *Tha*, unlike the nonemphatic items in Greek, is not licensed by higher negation, questions, conditionals, disjunctions, imperatives, or any of the other environments that non-emphatic NPIs appear in in Greek (see Giannakidou 1998).

\(^8\)An equivalent result could be achieved in a theory that posits an operation of Fusion, subject to the same locality conditions that spanning is subject to. See Embick 2010, Arregi and Nevins 2012, and Bobaljik 2015.
We note here that it is not crucial that the root of the copula be categorized as V, as we have done here for the sake of simplicity, or that the copula has not head-moved to T. Either of the routine assumptions that the root is categorized by a separate V node or that V moves to T are compatible with our analysis; they simply require slight reformulations of the Vocabulary Items in (22). Indeed, we could leave the copula entirely in situ and make the presence of en vs. itan contingent on whether the neighboring Tense node is specified as present or past, as in (24). This kind of analysis merely extends the logic of spanning from the insertion of Vocabulary Items to their conditioning environments (as argued for in Merchant 2015).\footnote{Another possibility would be to analyze tha as only realizing the copular, with a null variant of C preempting na there: while such an analysis may capture the facts by positing mutual allomorphy, it essentially makes the codependency an accident and thus fails to make such codependency follow from the lexical entries.}

\[(24) \quad \sqrt{BE} \ C \leftrightarrow \text{tha} / \text{NEG T[pres]} \]

Whichever implementation is ultimate to be preferred, this kind of local allomorphic analysis makes a prediction: if the adjacency between en and na is disrupted, tha should fail to appear. There are in fact two such cases, both of which bear out this prediction of our analysis.

First, the copula allows for a missing complement: dropping the CP sister to the copula will bleed the rule in (22c). Under our analysis, the environment
for the insertion of *tha* is not met (due to the lack of C), and we expect to find instead an emergence of the unmarked effect (see McCarthy and Prince 1994). The following data show that this expectation is met:

(25) a. O Yannis itan na pai ekso
    *the Giannis.NOM be.PAST.3s na go.PERF.NONPAST.3s out*
    extes, ala i Maria en itan.
    *yesterday but the Maria.NOM NEG be.PAST.3s*
    ‘Giannis was going to go out yesterday, but Maria wasn’t.’ (=going to go out yesterday)

b. *O Yannis en na pai ekso*
    *the Giannis.NOM be.NONPAST.3s na go.PERF.NONPAST.3s out*
    avrio, ala i Maria en tha.
    *tomorrow but the Maria.NOM NEG FUT*
    (Intended: ‘Giannis will go out tomorrow, but Maria will not.’)

We can see that the missing *na*-clause is due to ellipsis, not Null Complement Anaphora, because it is possible to extract from the missing complement. In (25a), the preverbal subject *i Maria* in the matrix clause has moved there by an A-movement from the embedded clause (see section 2.1 below). A′-dependencies can also reach into the missing CP (unlike what is possible in cases of Null Complement Anaphora).\(^{10}\)

(26) To aftokinito itan na plini o Kostas, che tin motora
    *the car was na wash.3s the Kostas.NOM and the motorbike.ACC*
    itan i Maria (na plini).
    *was the Maria.NOM na wash.3s*
    ‘The car, Kostas will wash, and the motorbike, Maria will (wash).’

The fact that ellipsis blocks allomorphy that depends on a triggering element internal to the ellipsis site is a simple ordering effect: ellipsis bleeds the allomorphic rule in (22c) by removing part of the structural description of the rule. This interaction is thus fully parallel to those studied in Bennett et al. 2015 for Irish, for example.

The second piece of evidence that the appearance of *tha* involves locally conditioned allomorphy over a span comes from the behavior of *en na* and *tha* in

\(^{10}\)It is immaterial here whether the ellipsis in question is of the CP or of the VP, with verb-raising having moved the copula out of the target of ellipsis (as in V-stranding VP-ellipsis of the kind studied by McCloskey 1991 and many since).
coordinations. As seen in (9) above and in (27) here, *na*-clauses can be coordinated under the copula, with the futurate meaning maintained for both conjuncts.

(27) Itan (che) na mairepso supa che na
    be.PAST.3s both na cook.PERF.NONPAST.1sg soup and na
    kathariso to domatio avrio.
    clean.PERF.NONPAST.1sg the room tomorrow
‘I was (going) (both) to cook soup and to clean the room tomorrow.’

But in the negated present, we find an asymmetry: the conjunct closest to negation shows the expected replacement of *en na* by *tha*, but any nonlocal conjunct does not. In other words, when the span targeted by (22c) is interrupted, the *tha* allomorph is impossible in the conjunct that isn’t adjacent to Neg—instead, we find the regular *na* again:

(28) a. En tha mairepso che na kathariso
    NEG tha cook.PERF.NONPAST.1sg and na clean.PERF.NONPAST.1sg
    avrio.
    tomorrow
‘I will not cook and clean tomorrow.’

b. *En tha mairepso che tha
    NEG tha cook.PERF.NONPAST.1sg and tha
    kathariso avrio.
    clean.PERF.NONPAST.1sg tomorrow
‘I will not cook and clean tomorrow.’

We assume that the coordination of the CP complements to *en* is a balanced coordination, represented here for convenience as a ternary branching structure.11 Vocabulary Insertion targets a span of nodes that includes the C head of the closest conjunct, in an apparent violation of the Coordinate Structure Constraint:

---

11The copula *en* itself cannot head a conjunct under negation: there is no VP coordination under negation in any variety of Greek, presumably for the same reason that coordination under *tha* and *na* is impossible. The negator is a proclitic and cannot cliticize into a conjunct, though whether this is a cause or an effect, we cannot determine on the basis of these data.
Note that an analysis based on head movement followed by Fusion would have difficulty accounting for these data: head movement is subject to the Coordinate Subject Constraint, so the requisite complex head consisting of T, $\sqrt{BE}$, and $C_1$ (which would form the input to the Fusion operation) cannot be formed in the syntax.

If this analysis is correct, we must countenance an expanded domain for spans, allowing T-V-C in (23) and (29) to count as a span (or merely V-C, if the rule in (24) is correct). This conclusion is at odds with the definitions of spans that restrict spans to extended projections, such as the following, from Merchant (2015):

(30) Let $T$ be an ordered n-tuple of terminal nodes $\langle t_1, \ldots, t_n \rangle$ such that for all $t \in T$, $t = t_1$ or $t$ is an element of the extended projection of $t_1$.
   a. For all $k = 1 \ldots n$, $t_k$ is a span. (Every node is a trivial span.)
   b. For any $n > 0$, if $t_k$ is a span, then $\langle t_k, \ldots, t_{k+n} \rangle$ is a span.

(31) **Spanning Insertion Hypothesis**: A span and only a span can be targeted for Vocabulary Insertion.

Instead, it would appear at first glance that we must define a span to include a contiguous string of elements after Linearization. By locating the point of the derivation that spanning is sensitive to after Linearization, this proposal also explains why no Coordinate Structure Constraint violation is registered in (29): the CSC is a constraint on syntactic (or semantic) representations, not on strings. The
definition in (32) is much weaker than that in (30), since it eliminates the requirement that the elements be in an extended projection.\textsuperscript{12}

(32) Let $T_D$ be the unique totally ordered n-tuple of terminal nodes $\langle t_1, \ldots, t_n \rangle$ over the elements in a derivation $D$ that satisfy the Linearization statements generated by $D$.\textsuperscript{13}

a. For all $k = 1 \ldots n$, $t_k$ is a span. (Every node is a trivial span.)

b. For any $n > 0$, if $t_k$ is a span, then $\langle t_k, \ldots, t_{k+n} \rangle$ is a span.

That the conditions on insertion of Vocabulary Items must in some cases be made sensitive to linear adjacency is a conclusion argued for independent grounds by Arregi and Nevins 2012 and Ostrove 2015 (though see Moskal and Smith 2016 for an opposing view).

At this point in our investigation, we have little evidence that mere adjacency is not the best model for these data, though allowing such string adjacency to be the sole conditioning factor for allomorphy is widely thought to overgenerate (see Svenonius 2012 for discussion). Another possibility for analyzing the coordinate structures would be to assimilate the spanning insertion condition to that found for closest conjunct agreement (as in McCloskey 1986, Munn 1999, Haegeman and van Koppen 2012, and Villavicencio et al. 2005, among many others). These in turn could privilege the first conjunct for structural reasons: it could be that the first conjunct CP is the true and only complement to the copula, and noninitial conjuncts are mere adjunct CPs to the first conjunct. These latter would show na, therefore, because they are not in the extended projection. This move raises a number of obvious difficulties (for extraction, subcategorization, and agreement), but these are the usual difficulties in handling conjunction to begin with.

Nevertheless, given the wide range of predicted but absent phenomena that mere linear adjacency would allow, it seems most prudent to find a middle ground. What we need is to define spans as consisting of all tuples of nodes that are in the set of adjacent terminal nodes and that stand in a (possibly transitive) selection relation. This is easily done:

\textsuperscript{12}This move is presaged in part in by a similarly weaker definition offered in Abels and Muringi (2008:719), who propose a version of a span (which they call a ‘stretch’) that includes the selectional requirement but jettisons the requirement that the heads be in an extended projection: “We suggest that a morpheme can realize a stretch of functional heads; by a stretch we mean one or more heads that select each other’s maximal projections.”

\textsuperscript{13}On some theories, such an ordered tuple is the output of Linearization; on others, it can be generated by the transitive closure over the Linearized pairs. Arregi and Nevins 2012 argue that the output of Linearization maintains hierarchical information as well.
Let $T_{≺≺}$ be the unique set of ordered pairs of terminal nodes $⟨t_i, t_j⟩$ over the elements in a derivation $D$ such that $t_i$ immediately precedes $t_j$. Let $T_S$ be the set of all pairs of nodes in $D$ $⟨t_i, t_j⟩$ such that $t_i$ selects $t_j$.

a. For all $k = 1 \ldots n$, $t_k$ is a span. (Every node is a trivial span.)

b. For any $n > 0$, $⟨t_1, \ldots, t_n⟩$ is a span iff for all $1 \leq i \leq n$, $⟨t_i, t_{i+1}⟩ \in T_{≺≺} \cap T_S$

This definition picks out a subset of those nodes that select other nodes: the subset of such nodes that also stand in the immediate precedence relation. This means that when a verb selects a CP and immediately precedes it, $⟨V, C⟩$ will be a span. Likewise when a $v$ selects and immediately precedes a $V$. But when a $v$ selects a specifier to its left (say, a DP), $⟨v, D⟩$ will not form a span: although $v$ selects $D$, it does not precede it. Likewise for potential selection relations that hold between a selecting head and a head remote from the selecting head, as was the case in the coordinations in (29) above: $√\text{BE}$ selects $C_2$, but does not immediately precede it. Since $⟨√\text{BE}, C_2⟩ \notin T_{≺≺}$, these two elements do not form a span.

An additional prediction is made by the present analysis: because T-V-C must form an uninterrupted span to surface as tha under negation, the presence of a marker on the left conjunct should make tha impossible. This is the case in balanced coordinations, such as those involving $ute \ldots ute$ ‘neither ...nor’ under negation (see Giannakidou 2007 for further discussion of the properties of $ute$). In such a situation, the unmarked $na$ should appear. These predictions are also borne out:

(34) a. *En tha ute mairepso supa ute na
   NEG tha neither cook.PERF.NONPAST.1sg soup nor na
   kathariso to domatio avrio.
   clean.PERF.NONPAST.1sg the room tomorrow

b. *En ute tha mairepso supa ute na
   NEG neither tha cook.PERF.NONPAST.1sg soup nor na
   kathariso to domatio avrio.
   clean.PERF.NONPAST.1sg the room tomorrow

\textsuperscript{14}Svenonius 2012:2 fn 3 considers the possibility that spans may include heads from across multiple extended projections, writing that “c-selection essentially turns a selected complement into part of the extended projection, at least for the purposes of lexical insertion”. This presages part of our definition in (33), but does not include the crucial adjacency requirement.
The example in (35) is particularly significant: it shows that the appearance of *tha* for *en na* under negation is not due to some incompatibility of negation with *en na*, either syntactic or semantic: in (35), the regular sentential marker *en* does occur with *en na*, yielding the expected meaning. This remarkable reappearance of *en na* under negation is predicted by our analysis: because the copular *en* is not adjacent to *na* in (35), the allomorphic rule in (22c) will not apply, and instead we find the usual exponents of these morphemes.

2.1 On the properties of the copula

The nature of the interactions between the higher verb, realized as *en* or *itan*, and the surface subject, are tangential to our concerns in this paper, so we will confine ourselves to only a few remarks.

It appears that apparent preverbal subjects in the matrix clause are moved there from the embedded clause via one of the strategies that underlie the robust word order permutations that Greek enjoys: the movement that derives much of the attested variation appears to have A-movement-like properties despite not targeting a position associated with agreement in their own clause (see Alexiadou and Anagnostopoulou 1999 for raising out of *na*-clauses, and Spyropoulos and Revithiadou 2007 for a discussion of preverbal subjects).

For example, subject idiom chunks from clausal idioms like the one in (36) do not appear in control clauses (37a) or as cross-clausal topicalized phrases (37b):

(35) En en ute na mairepsos supa ute na
NEG be.NONPAST.3 neither na cook.PERF.NONPAST.1sg soup nor na
kathariso to domatio avrio.
clean.PERF.NONPAST.1sg the room tomorrow
‘I will neither cook soup nor clean the room tomorrow.’

(36) Efkalen malja i glosa mu.
grow.PAST.PERF.3s hair.ACC the tongue.NOM my
‘I talked a lot about the same thing; I wore myself out talking.’ (lit. ‘My tongue grew hair.’)

(37) a. *I glosa mu eprospathise na fkali
the tongue.NOM my try.PAST.PERF.3s na grow.NONPAST.PERF.3s
malja.
hair.ACC
(Intended: ‘I tried to talk a lot.’, lit. ‘My tongue tried to grow hair.’)
b. * [I glosa tu], se proidhopisa oti t2 fkali malja.
the tongue.NOM his you.ACC warned.1s that grows hair
(Intended: ‘I warned you that he talks a lot.’)

But these idiom chunks can appear before the copula, as seen in (38a,b) for the present and past copulas, respectively.

(38) a. I glosa mu en na fkali
the tongue.NOM my be.NONPAST.3s na grow.NONPAST.PERF.3s
malja.
hair.ACC
‘I will talk a lot about the same thing.’ (lit. ‘My tongue will grow hair’)

b. I glosa mu itan na fkali
the tongue.NOM my be.PAST.3SG na grow.NONPAST.PERF.3s
malja.
hair.ACC
‘I was going to talk a lot about the same thing.’ (lit. ‘My tongue would grow hair’)

Similarly, active/passive synonymy is maintained under en na: (39a) and (39b) are synonymous.

(39) a. O jatros en na eksetasi
the doctor.NOM be.NONPAST.3SG na examine.NONPAST.PERF.3SG
ton Kosta.
the Kostas.ACC
‘The doctor will examine Kostas.’

b. O Kostas en na eksetasti pu ton
the Kostas.NOM be.NONPAST.3s na examine.PASSIVE.3s by the
jatro.
doctor.ACC
‘Kostas will be examined by the doctor.’

If this reasoning is correct, we have a case of A-movement out of an embedded finite clause, a conclusion in line with that of Alexiadou and Anagnostopoulou 1999 for certain aspectual predicates. Two additional tests from Alexiadou and Anagnostopoulou 1999 yield the same result: as they discuss, pseudocleft-like
structures as in (40) do not allow the $na$-clause complement of raising verbs to be the pivot, and nominative anaphors as in (41) allow backward binding into subjects of raising, but not control, predicates (see Landau 2013 for extensive additional discussion and diagnostics).

(40) *Afto pu itan i Maria itan na pai ekso.

\[
\text{\textit{that which be.PAST.3s the Maria was na go.3s outside}}
\]
\[
\text{\textit{\textquoteleft What Maria was was to go outside.\textquoteright}}
\]

(41) O eaftos tu en na tu aresi.

\[
\text{\textit{the self.NOM his be.PRES.3s na him.GEN appeal.3s}}
\]
\[
\text{\textit{\textquoteleft He will like himself.\textquoteright}}
\]

Two other typical tests that distinguish A- from A'-movement, the presence of weak crossover effects and the licensing of parasitic gaps, yield less than reliable results in Cypriot Greek. While these subjects do not trigger weak crossover effects with respect to pronouns inside matrix adjuncts, from which we would normally conclude that they reach their surface positions by A-movement, it is difficult to interpret the results of this diagnostic since it is not clear whether Greek has weak crossover effects for topicalized phrases from embedded clauses in all cases in the first place (this effect is not dependent on the presence of an object clitic, unfortunately). It is also unclear whether Greek has a set of gaps with the profile of parasitic gaps in languages like English.

Given the parallels to raising predicates, then, it comes as something of a surprise that in these constructions, the copular verb always appears in the 3rd person (recall from (1) that the 3rd person shows no number distinction in this verb).

(42) a. Ego en na pao.

\[
\text{\textit{I be.NONPAST.3 na go.PERF.NONPAST.1sg}}
\]
\[
\text{\textit{\textquoteleft I will go.\textquoteright}}
\]

b. *Ego ime na pao.

\[
\text{\textit{I be.NONPAST.1sg na go.PERF.NONPAST.1sg}}
\]
\[
\text{\textit{\textquoteleft I will go.\textquoteright}}
\]

While unusual, this would not be the only verb in Greek to show this behavior: impersonal verbs like \textit{prepi} ‘must’ and \textit{bori} ‘is possible’ have a similar pattern. There are two obvious analytical paths for accounting for the lack of agreement on \textit{en} in (42).
First, the locality domains for Agreement and for movement could be different: in Greek, apparently, the embedded CP na-clause, which does not assign nominative case, allows for EPP-driven (∗D∗-driven) A-movement from its subject to the higher subject position. The ability of agreement to probe into the CP could be limited, perhaps because CP itself has φ-features. The necessity for default inflection would follow on this analysis from the fact the higher T node, where the probing φ-features are located, is separated from the potential controller of agreement by a clause boundary, CP. Movement of the DP into the higher domain does not alter this fact, since the probe only agrees downward.

The second possibility again would take it that Agree to value the φ-features on the higher T can only probe T’s c-command domain, but would derive the lack of agreement from an interaction of movement and Agree. If movement occurs before Agree, then movement of the subject above T would bleed Agree. Such an analysis can be implemented in a system like that of (Georgi 2014), who studies such opaque interactions in detail: the movement feature on the head that agrees would precede the agreement feature: ⟨•D•, ∗φ∗⟩. The advantage of this approach is that it could code on individual lexical items a differing order of Move and Agree triggers, allowing us to understand the difference between en/itan and the aspectual verbs studied by Alexiadou and Anagnostopoulou 1999, which do agree with their derived subjects, as a lexical difference, specified in the lexical entries for the different verbs.

In either case, any theory of default values should capture the fact that what surfaces is 3rd person (for example, Preminger 2014’s account of Agreement failures). It is not important for our purposes to choose between these alternatives, and we leave adjudication between them to future work.

2.2 Is clause union an alternative?

There is in principle another possibility for analyzing the allomorphic competition between en na and tha, one that would preserve the definition of span built on the notion of extended projection, and not weaken it to mere transitive selection (modulo the immediate precedence condition). This alternative would require two changes to our analysis: first, we could adopt the position of many researchers (Giannakidou 2009; Philippaki-Warbuton 1994; Roussou 2010) that na is in a lower clause-internal projection, such as Mood, and not in C, and does not require embedding under a C. Second, we could claim that the copula en/itan in these structures selects for MoodP directly, bypassing the CP layer, and that this truncated complement phrase is the realization of the restructuring context.
As in Germanic and Romance restructuring phenomena (Wurmbrand 2004), the selecting V and the lower V would be in the same domain for certain purposes, including allomorphy (a phenomenon familiar from German for example, where one subset of restructuring contexts licenses the Infinitivus Pro Participio, a locally conditioned variant realization of a verb under the perfect auxiliary). On this analysis, na and en would not be in different domains, and our definition of span requires no revision along those lines: instead, we could analyze the entire V-na-V complex as involving only one extended projection (see also Grano 2012 for an in-depth discussion of some Greek restructuring verbs).

On this analysis, we would have structures such as the following.

(43) NegP
    |       Neg
    |       TP
    |       en T VP
    |       MoodP
    |       Mood TP
    |       tha T VP
    |       (pro_{1sg}) pao

In this case, we would therefore not find a CP layer, even though na-clauses in other contexts project CPs, as can be seen by the appearance of a clitic-left-dislocated DP between a selecting verb and na in the following examples.

(44) Thelo ton Gianni na ton
    want.IMPERF.NONPAST.1sg the.ACC Giannis.ACC na him
    apolisun avrio.
    fire.PERF.NONPAST.3pl tomorrow
    ‘I want them to fire Giannis tomorrow.’

The en na construction, however, appears to have an embedded CP layer: there is a landing site of at least marginal acceptability for clitic-left dislocated phrases.
between the matrix verb and na: 15

(45) ?Itan ton Gianni na ton apolisun avrio.
    be.PAST.3s the Giannis.ACC na him fire.PERF.NONPAST.3pl tomorrow

‘They were going to fire Giannis tomorrow.’

But a significant remaining question on such an analysis is why the higher T cannot agree with the lower subject. If a clause boundary intervenes (a CP, as in (23) above), standard theories of locality of agreement will correctly predict that the embedded subject only triggers agreement on the embedded verb. If we collapse the clauses, it would seem, ceteris paribus, that the matrix T would now be in the same phase as the lower subject, and hence able to value its φ-features via Agree with the embedded subject.

3 Conclusion

The allomorphic variation of the Cypriot Greek periphrastic future en na∼tha presents a puzzle for standard spanning theory, where spans are restricted to extended projections. We analyzed en na as being just what it seems to be: the copula followed by the ‘subjunctive’ subordinating particle na; tha replaces both of these under clausemate negation. In other words, tha is a portmanteau form realizing the copula and the embedded head that normally surfaces as na: concretely, tha realizes a (T-)V-C span.

Cypriot Greek tha therefore represents a morpheme whose environment for Vocabulary Insertion cannot be stated within a single extended projection. We therefore revised the definition of span to make it sensitive to mere selection (including across a clause boundary). The facts from coordination, in particular the ability of the closest conjunct to license tha across its edge, and the fact that CP ellipsis bleeds tha, required a further, more radical revision: spans are sequences of selecting heads that also immediate precede one another.

Constraints on possible morphemes can come from the kinds of linearizations that are possible, from cyclic constraints on the generation of those statements, or elsewhere, but one conclusion is inescapable: eppur si spane!

15 For reasons that are unclear to us at present, cases in the present tense are much worse:

(1) *En ton Yanni na ton apolisun avrio.
    be.NONPAST.3 the.ACC John.ACC na him fire.IMPERF.NONPAST.3pl tomorrow

‘They are firing John tomorrow.’
A A note on the the na construction

Greek Cypriot speakers also find examples like (46) acceptable. At first sight, this seems to involve a contracted form of the volitional verb thelo (to the; see Joseph and Pappas 2002 and Pappas 2001 for relevant discussion); but this the is not found anywhere else in the language.

(46) E(n) the na pao popse.
    NEG the na go.PERF.NONPAST.3SG tonight
    ‘I am not going tonight.’

Markopoulos 2008 notes the emergence of the na in the 14th c. as a construction used for expressing volition-related meaning, and he treats the as the product of the loss of the unstressed word-final /i/ of the volitional verb theli ‘want’ following the loss of intervocalic /l/. Similarly, the 2nd person singular in standard mainland Greek also exhibited a similar reduction, from thelis to thes, in the same period, a fact suggesting that it belonged to the same pattern as the the na construction.

(47) An the na mbun apu tin tripan, t’ aloga apothe na ta
    if the na enter.3pl from.the hole, the horses from.where na them.CL
    mbasomen?
    pass-through.3pl
    ‘If they will go through the hole, where will we pass the horses through?’
    (Cypriot Greek, Mahairas, 509)

Synchronically, the is hardly used in any other context and speakers do not necessarily associate it in the aforementioned construction with a volitional reading. As the example in (48a) shows, it is no contradiction to not want to do something but to nevertheless predict or plan for it to happen; the fact that the in precisely the same context gives rise to a contradiction, as indicated in (48b), shows that the is not merely a reduced form of thelo and is compatible with the indicating futurity in the same way as en na (though we will have to leave a close investigation of their semantic differences to future work).

(48) a. En thelo na pao
    NEG want.NONPAST.IMPERF.1sg na go.NONPAST.IMPERF.1sg
    avrio, alla en na pao.
    tomorrow; but be.3sg na go.NONPAST.IMPERF.1sg
    ‘I don’t want to go tonight, but I will go.’
Note finally that the fact that the na can be felicitously embedded under negation makes it unlikely that the failure of en na to similarly appear under negation is due to its meaning or to something idiosyncratic about na in these contexts: it is due to the fact that there is a more specific morpheme, tha, which is competing for precisely the copular+na span under negation, and which, by the Elsewhere Principle, pre-empts the appearance of en na.

References


Bennett, Ryan, Emily Elfner, and James McCloskey. 2015. Prosody, focus, and ellipsis in Irish. Ms., Yale, University of British Columbia, and University of California, Santa Cruz.


Merchant, Jason. 2014. The Afrikaans final negative particle is a clitic on VP. Ms., University of Chicago.


Spyropoulos, Vassilios, and Anthi Revithiadou. 2007. Subject chains in Greek
and PF processing. In Workshop on Greek Syntax and Semantics, MIT, 20–22.
Terzi, Arhonto. 1999a. Clitic combinations, their hosts and their ordering. Natural
language & linguistic theory 17:85–121.
Terzi, Arhonto. 1999b. Cypriot Greek clitics and their positioning restrictions. In
Studies in Greek syntax, ed. Artemis Alexiadou, Geoffrey Horrocks, and Melita
future. Thessaloniki: University Studio Press.
varieties: Establishing the Blueprint. International Journal of the Humanities
6:49–65.
Tsiplakou, Stavroula. 2014. How mixed is a ‘mixed’ system? The case of Cypriot
Vassiliou, Erma. 2002. The word order of Medieval Cypriot. Doctoral Disserta-
tion, La Trobe University.
Villavicencio, Aline, Louisa Sadler, and Doug Arnold. 2005. An HPSG account of
closest conjunct agreement in NP coordination in Portuguese. In Proceedings of
the HPSG05 conference, ed. Stefan Müller. Stanford, Calif.: CSLI Publications.
114:991–1014.