The morphosyntax of the periphrastic future under negation in Cypriot Greek

Jason Merchant and Natalia Pavlou
University of Chicago, 2017*

To appear in the Journal of Greek Linguistics

Abstract

In Cypriot Greek, the negated future is marked by an element, *tha*, which appears instead of the expected present tense copula and a selected subordinating element. This paper documents the distribution of this item for the first time, and presents an analysis in Distributed Morphology that analyzes *tha* as a portmanteau morpheme realizing two heads in the context of negation. This analysis requires that spans (or targets of Fusion) can include a verb and the head of its C complement.

1 The future and negation in Cypriot Greek

1.1 The periphrastic future: copula+*na*

Cypriot Greek, like its Standard Modern Greek sister spoken primarily to its northwest across the Mediterranean, has a copula verb that inflects for person, number, and tense, but shows no number distinction in the third person:

*Our thanks to the reviewers for the Journal of Greek Linguistics, to Karlos Arregi, Anastasia Giannakidou, Marika Lekakou, and Andrew Nevins for valuable feedback on the paper, as well as to audiences at the 12th International Conference on Greek Linguistics and at the University of Chicago, where earlier versions of this work were presented. All Cypriot Greek data reported in this paper without further attribution come from the second author, who is a native speaker from Limassol. Our thanks also to the ten additional speakers who provided judgments in a written questionnaire.
This verb is used with adjectival, nominal, and prepositional predicates:

(2) a. *Ta mora en arosta.*
the child.NEUT.pl be.NONPAST.3 sick.NEUT.pl
‘The children are sick.’

b. *Ta mora en kali mathites.*
the child.NEUT.pl be.NONPAST.3 good.MASC.pl pupils.MASC.pl
‘The children are good students.’

c. *Ta mora en sto dhomatio.*
the child.NEUT.pl be.NONPAST.3 in.the room
‘The children are in the room.’

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 periphrastically, using a construction that the (here invariant) 3rd person copula with a clause headed by the subordinating particle *na* (traditionally called ‘subjunctive’, and which we will assume is a complementizer for the time being) and a verb in the non-past:

---

1. The past plural forms *itan/ito/itun* are in free variation; the difference does not index number.

2. 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 to appear 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 regarding the challenge of compositionality for idioms, idiom chunks, collocations, and ‘constructions’ generally; compare e.g. *must* with *have to*, and *will* with *be going to*, among many other such examples.

3. Analyses vary somewhat in whether they take *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 2000, 2001 and Giannakidou 2009 do. All that is important for us is that *na* signals the presence of a CP layer, for reasons that will become clear below. To sidestep these labeling issues, we will therefore gloss *na* simply as ‘na’.
(3)  *En na pao.*  
    be.NONPAST.3 na go.PERF.NONPAST.1sg  
    ‘I will go.’

The particle *na* is found in a range of contexts, including as the head of selected control and raising CPs; Cypriot Greek, like standard Greek, lacks infinitives, and uses *na*-clauses in their stead:

(4)  *Thelo na pao.*  
    want.IMPERF.NONPAST.1sg na go.PERF.NONPAST.1sg  
    ‘I want to go.’

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

As also pointed out in Roussou and Tsangalidis (2010), it is unclear whether *en na* is a single lexical item or not. We claim that each of the elements in *en na* has an independent use: *en* as a verb[^4] 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):

[^4]: We know of no compelling reason to suspect that *en* in *en na* was ever anything but a form of the copula verb historically. We thus agree with Symeonides 2006:240, who writes that “I have the impression that this category of the future is based on ἔνι (ἐν) and *en=*einai, prokeitai” (translation ours), where *einai* is the 3rd person nonpast unreduced copula verb and *prokeitai* is another paraphrastic future (roughly, ‘be going to’). While it is not our aim here to try to trace the history of this construction, nor is it in any way relevant to our synchronic analytical goals, we may also point out that historical accounts that do not tie *en na* to the copula encounter numerous difficulties. In particular, we find the speculation of Chatziioannou 1999:92 that this *en* somehow derives from an earlier *thelo* (“thelo na > thel na > then na > enna”) to be highly unlikely: such a derivation (which lacks any basis in the historical record we are aware of) would require an idiosyncratic sound change to account for the loss of the initial interdental voiceless fricative θ. (It is in precisely this last, implausible, step that Chatziioannou’s proposal diverges from the well established history of standard modern Greek *tha*, as argued in Joseph and Pappas 2002 and Roberts and Roussou 1999. We do not rule out the possibility of a parallel diachronic development for Cypriot Greek *tha*, but we have not conducted the relevant historical research needed to establish the point.) See section 3.4 below for additional reasons to reject a derivation of *en* from *the.*
The *na* that occurs in this location has the usual properties found with *na*-clauses elsewhere in the language. First, it can select for what we follow Holton et al. 1997 in calling the dependent form of the verb; the dependent verb form is the perfective nonpast, and it is always selected for—it occurs only under certain particles, like *na*, and in certain embedded constructions (see Giannakidou 2009 for the claim that the dependent verbal form in the perfective nonpast cannot be identified with the utterance time of the context). This was illustrated for the usual *na* in (4) and (5) above, where *pao* and *dho* are respectively the dependent forms (the perfective nonpast), and in (3) for *en na*.

Second, as with all *na*-clauses in Cypriot Greek, when the embedded predicate in the *na*-clause is negated, the negator that appears under *na* is the ‘subjunctive’ negator *men*, as in (7), glossed NEG.SUBJ (see Chatzopoulou 2012 for the history of the Standard Modern Greek cognate *min*), not the indicative *en* (see section 1.2 for more on *en*).

(7)  *Thelo*  *na men*  *pais.*
    want.1sg  *na*  NEG.SUBJ  go.PERF.NONPAST.2sg
    ‘I want for you not to go.’

(8)  *Akusa*  *oti en na men*
    hear.PERF.PAST.1sg  that.be.NONPAST.3  *na*  NEG.SUBJ
    *pais.*
    go.PERF.NONPAST.2sg
    ‘I heard that you will not go.’

The appearance of *men* in this context is expected if the *na* that we see in (8) is the usual *na* found elsewhere the language, such as that in (7).
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. First, the copula part can appear independently in the past tense as illustrated in (9):

(9) *Itan na pao ekso.*
    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) or *itun*, is the regular 3rd person past tense of the copula.

Second, we find coordinated *na*-clauses under *en* (and under *itan*, shown in (25) below):

(10) 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 mairopsis i na*
    be.NONPAST.3 na cook.PERF.NONPAST.2sg or *na katharis?*
    clean.PERF.NONPAST.2sg
‘Are you cooking or cleaning?’

---

5Precisely 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 cannot be correct: 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 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 partially 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.
c. *En na mairepsis oksa na
cook.PERF.NONPAST.2sg or na katharisis?
clean.PERF.NONPAST.2sg
‘Which of the two are you doing: cooking, or cleaning?’
(Or: ‘Are you cooking, or are you cleaning?’)

If *en na were a single head or fixed lexical expression, then coordination of such na-clauses alone should not be possible.6

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; Mavrogiorgos 2013, to appear among others), as illustrated in (11), clitic arguments of dependent verbs under na are proclitic, as seen in (12).

(11) Idha ton extes.
see.PERF.PAST.1sg him.CL yesterday
‘I saw him yesterday.’

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

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

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

6As 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.
1.2 Sentential negator *en* and the periphrastic future

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

(14)  \textit{En pieno.}  
\texttt{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:

(15)  \begin{itemize}
  \item \textit{Ta mora \textit{en} \textit{en} arosta.}  
        \texttt{the children NEG be.NONPAST.3 sick}  
        ‘The children are not sick.’
  \item \textit{Ta mora \textit{en} \textit{itan} arosta.}  
        \texttt{the children NEG be.PAST.3 sick}  
        ‘The children were not sick.’
\end{itemize}

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:

(16)  \textit{En \textit{itan na} pao.}  
\texttt{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:

(17)  \textit{*En \textit{en na} pao.}  
\texttt{NEG be.NONPAST.3 na go.PERF.NONPAST.1sg}  
(\texttt{Intended: ‘I will not go.’})

The fact that *en* *en* is not found in (17) is not due to mere haplology (even syntactically conditioned haplology, as in Salzmann 2013 and Merchant 2014), given that (15a) is well-formed.

What we find instead for the negated future is surprising and shows an interesting fact of syntactic micro-variation between two closely related varieties. It
has been observed before that “Cypriot Greek lacks the future particle \textit{tha} of Standard Greek” (Terzi 1999a:110 fn 24), as shown in (18).\footnote{The example in (18) is acceptable in Standard Modern Greek: it is in fact the usual way to express the future (see Tsangalidis 1998, Giannakidou and Mari to appear). The particle is only acceptable among Greek Cypriot speakers when the high variety of Standard Modern Greek is used in code-switching or code-mixing (Tsiplakou 2009, 2014). Apart from the second author’s own judgments, the data presented here concerning both the use of \textit{en na} and the appearance of \textit{tha} in negated sentences were confirmed by the results of a questionnaire, administered in the summer of 2016 to ten speakers of Cypriot Greek.}

\begin{align*}
(18) \quad & \text{{\textit{tha pao}.}} \\
& \text{\textit{tha go.PERF.NONPAST.1sg}} \\
& \text{\textquoteleft I will go.\textquoteright} \\
\end{align*}

But precisely this \textit{tha} that is otherwise absent from the language is the form that appears in the negated periphrastic future:

\begin{align*}
(19) \quad & \text{{\textit{en tha pao}.}} \\
& \text{\textit{NEG tha go.PERF.NONPAST.1sg}} \\
& \text{\textquoteleft I will not go.\textquoteright} \\
\end{align*}

The appearance and identity of this \textit{tha} is the puzzle to be solved.

\section{Spanning: Vocabulary Items that realize more than one node}

The element \textit{tha} is in complementary distribution with the usual future marking strategy \textit{en na}. The most parsimonious analysis therefore is to take the two sets of elements to be differing, competing realizations of the same nodes, one realized under negation, and the other otherwise: allomorphs, in other words, expanding the use of that term slightly.\footnote{Similarly to \textit{na}, we gloss this element simply as \textit{tha}, in order not to prejudge its category or function. (Analyses differ on where the Standard Modern Greek \textit{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 \textit{T tout simple}.)}

\footnote{This distribution has obvious parallels to negative polarity items (NPIs); on one reading, we can claim that \textit{tha} is an NPI in Cypriot Greek. Such a claim does not immediately help with understanding its morphosyntactic distribution, of course. \textit{Tha}, unlike the nonemphatic items in Greek, is not licensed by higher negation, questions, disjunctions, imperatives, or any of the other environments that non-emphatic NPIs appear in in Greek (see Giannakidou 1998), with the exception of conditionals; for these latter, see section 3.4 below.}
Such an analysis can be implemented in a theory such as Distributed Morphology (see Arregi and Nevins 2012 and Bobaljik 2015 and Embick and Noyer 2007) in which a single generative system is responsible both for word structure and phrase structure and derivation of complex objects is syntactic. Distributed Morphology (Halle and Marantz 1993) incorporates hierarchical structure into morphology by positing that the input to morphology is syntactic structure. Traditional features (or feature bundles) are distributed over nodes, which in turn are subject to Vocabulary Insertion, the process that replaces or realizes morphosyntactic featurally specified nodes as particular morphemes. In our analysis, we posit the following Vocabulary Items (abstracting away from person features here for simplicity):

(20)  
   a. $\sqrt{BE} \leftrightarrow \text{en} / T[\text{pres}]$  
   b. $C \leftrightarrow \text{na}$  
   c. $T[\text{pres}] \sqrt{BE} C \leftrightarrow \text{tha} / \text{NEG}$

By the Elsewhere Principle, insertion of the Vocabulary Items in (20a,b) will be preempted by the more narrow context of insertion in (20c).

The unusual property of the entry in (20c) is that it realizes three nodes, not just one. It is for this reason that we adopt a theory that allows Vocabulary Insertion to target multiple nodes with a single rule: these nodes must form what is known as a span (Svenonius 2012; Merchant 2015; Bennett et al. 2015), and our innovation here is to claim that tha is a portmanteau morpheme that is inserted for T, the copula root, and C simultaneously.\(^\text{10}\) The resulting structure is as follows:

\(^{10}\)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, nor is it crucial that the copula has not 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 (20). 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 (22). 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): such an alternative would need to take the sequence of nodes that constitute the conditioning environment together, forming themselves a span of nodes.\(^{11}\)

\[(22) \quad \sqrt{BE} \ C \leftrightarrow \text{tha / NEG T[pres]} \]

Whichever implementation is ultimately 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.

\[^{11}\text{Another possibility would be to analyze *tha* as only realizing the copular, with a null variant of C preempting *na* there: while such an analysis may capture the facts by positing mutual allomorphy, it essentially makes the codependency of the realizations of the copula and of the subordinating C in this context an accident and thus fails to capture the complementary distribution in the usual way.}\]
First, the copula allows for a missing complement: dropping the CP sister to the copula will bleed the rule in (20c). Under our analysis, the environment for the insertion of *tha* is not met (due to the lack of *C*) (compare 23b and 23c), and we expect to find instead an emergence of the unmarked effect (McCarthy and Prince 1994). The following data show that this expectation is met:

(23) 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 pai.
  tomorrow but the Maria.NOM NEG FUT go.PERF.NONPAST.3
  ‘Giannis will go out tomorrow, but Maria will not.’

c. *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 (23a), the preverbal subject *i Maria* in the matrix clause has moved there by an A-movement from the embedded clause (see section 3.1 below). A’-dependencies can also reach into the missing CP (unlike what is possible in cases of Null Complement Anaphora; see Merchant 2013).12

(24) *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

12It 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); see Merchant 2016 for an argument that standard modern Greek has the latter.
‘The car, Kostas was going to wash, and the motorbike, Maria was (going
to 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 allomor-
phic rule in (20c) 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 con-
ditioned allomorphy over a span comes from the behavior of en na and tha in
 coordinations. As seen in (10) above and in (25) here, na-clauses can be coordi-
nated under the copula, with the futurate meaning maintained for both conjuncts.

(25) Itan
  (che) na mairepso
  supb a 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 con-
junct does not. In other words, when the span targeted by (20c) is interrupted, the
tha allomorph is impossible in the conjunct that isn’t adjacent to Neg—instead,
we find the regular na again:

(26) 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
(symmetrical) coordination with the expected morphology on both conjuncts (see
Johannessen 1993), represented here for convenience as a ternary branching structure.\(^{13}\) 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:

\(\text{(27)}\)

\[
\begin{array}{c}
\text{NegP} \\
\text{Neg} \\
\text{en} \\
\text{T} \\
\sqrt\text{BE} \\
\text{CP} \\
\text{Conj} \\
\text{C}_1 \\
\text{TP} \\
\text{che} \\
\text{C}_2 \\
\text{TP} \\
\end{array}
\]

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, \(\sqrt{\text{BE}}\), and \(\text{C}_1\) (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 (21) and (27) to count as a span (or merely V-C, if the rule in (22) 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:

\(\text{(28)}\) 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 \(n > 0, \text{if } t_k\) is a span, then \(\langle t_k, \ldots, t_{k+n} \rangle\) is a span.

\(^{13}\)The 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.
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 (27): the CSC is a constraint on syntactic (or semantic) representations, not on strings. The definition in (30) is much weaker than that in (28), since it eliminates the requirement that the elements be in an extended projection.\textsuperscript{14}

\textsuperscript{14}This move is presaged in part by a similarly weaker definition offered in Abels and Muriungi (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{15}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.

\begin{equation}
\text{(29) Spanning Insertion Hypothesis: A span and only a span can be targeted for Vocabulary Insertion.}
\end{equation}

\begin{equation}
\text{(30) Let } T_D \text{ be the unique totally ordered n-tuple of terminal nodes } \langle t_1, \ldots, t_n \rangle \text{ over the elements in a derivation } D \text{ that satisfy the Linearization statements generated by } D.\textsuperscript{15}
\end{equation}

\begin{enumerate}
\item For all } k = 1 \ldots n \text{, } t_k \text{ is a span. (Every node is a trivial span.)}
\item For any } n > 0 \text{, if } t_k \text{ is a span, then } \langle t_k, \ldots, t_{k+n} \rangle \text{ is a span.}
\end{enumerate}

That the conditions on insertion of Vocabulary Items must in some cases be made sensitive to linear adjacency is a conclusion argued for on 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, Villavicencio et al. 2005, and Haegeman and van Koppen 2012, 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:

(31) Let $T_{\prec\prec}$ be the unique set of ordered pairs of terminal nodes $\langle t_i, t_j \rangle$ 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 $D$ $\langle t_i, t_j \rangle$ 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 $n > 0$, $\langle t_1, \ldots, t_n \rangle$ is a span iff for all $1 \leq i \leq n$, $\langle t_i, t_{i+1} \rangle \in T_{\prec\prec} \cap T_S$

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, $\langle V, C \rangle$ 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), $\langle v, D \rangle$ 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 (27) above: $\sqrt{BE}$ selects $C_2$, but does not immediately precede it. Since $\langle \sqrt{BE}, C_2 \rangle \notin T_{\prec\prec}$, 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:

(32) a. *En tha ute mairepso supb a ute na

\[
\text{NEG tha neither cook.PERF.NONPAST.1sg soup nor na}
\]

\[16\]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 (31), but does not include the crucial adjacency requirement.
kathariso to dhomatio avrio.
clean.PERF.NONPAST.1sg the room tomorrow

b. *En ute tha mairepso sup'h a ute na
   NEG neither tha cook.PERF.NONPAST.1sg soup nor na
kathariso to dhomatio avrio.
clean.PERF.NONPAST.1sg the room tomorrow

(33)   En en ute na mairepso sup'h a ute na
   NEG be.NONPAST.3 neither na cook.PERF.NONPAST.1sg soup nor na
kathariso to dhomatio avrio.
clean.PERF.NONPAST.1sg the room tomorrow
‘I will neither cook soup nor clean the room tomorrow.’

The example in (33) 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 (33), 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 (33), the allomorphic rule in (20c) will not apply, and instead we
find the usual exponents of these morphemes.

3 Additional issues and questions

3.1 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 target-
ing a position associated with agreement in their own clause (see Alexiadou and
Anagnostopoulou 2002 for raising out of na-clauses, and Spyropoulos and Re-
vithiadou 2007 for a discussion of preverbal subjects).

For example, subject idiom chunks from clausal idioms like the one in (34) do
not appear in control clauses (35a) or as cross-clausal topicalized phrases (35b):
(34) \textit{Efkalen malja i ylossa mu.} \\
\textit{grow.PAST.PERF.3s hair.ACC the.NOM tongue.NOM my} \\
‘I talked a lot about the same thing; I wore myself out talking.’ (\textit{lit.} ‘My tongue grew hair.’)

(35) a. *I ylossa mu eprospathise na \\
\textit{the.NOM tongue.NOM my try.PAST.PERF.3s na} \\
fkali malja. \\
\textit{grow.NONPAST.PERF.3s hair.ACC} \\
(\textit{Intended:} ‘I tried to talk a lot’, \textit{lit.} ‘My tongue tried to grow hair.’)

b. *[I ylossa tu], se proidhopiisa oti \textit{t2 fkali} \\
\textit{the.NOM tongue.NOM his you.ACC warned.1s that grows} \\
\textit{malja.} \\
hair \\
(\textit{Intended:} ‘I warned you that he talks a lot’, \textit{lit.} ‘I warned you that his tongue grows hair’)

But these idiom chunks can appear before the copula, as seen in (36a,b) for the present and past copulas, respectively.

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

b. I ylossa mu itan na fkali \\
\textit{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.’ (\textit{lit.} ‘My tongue would grow hair’)

Similarly, active/passive synonymy is maintained under \textit{en na}: (37a) and (37b) are synonymous.
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 2002 for certain aspectual predicates. Two additional tests from Alexiadou and Anagnostopoulou 2002 yield the same result. First, as they discuss, pseudocleft-like structures as in (38) do not allow the na-clause complement of raising verbs to be the pivot.

(38) *Afto pu itan i Maria itan na pai ekso.  
that which be.PAST.3s the Maria was na go.3s outside  
(Intended: ‘What Maria was going to do was go outside’, lit. ‘What Maria was was to go outside.’)

Second, nominative anaphors as in (39) allow backward binding into subjects of raising, but not control, predicates (see Landau 2013 for extensive additional discussion and diagnostics). This shows that the anaphor can be reconstructed to and bound in its base position. In (39), o eafos tu is in the subject position of the main clause and it is bound by the pronoun in the na-clause; this is consistent with the supposition that o eafos tu A-moves out of the embedded clause to the matrix clause.

(39) O eafos tu en na tu aresi.  
the self.NOM his be.PRES.3s na him.GEN appeal.3s  
‘He will like himself.’

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, and so we do not report the data here. Briefly, quantificational subjects of en na clauses appear to be able to bind pronouns that are
inside matrix adjunct. In other words, such subjects do not trigger weak crossover effects. We would normally conclude that they reach their surface positions by A-movement (over matrix adjuncts), expanding their A-binding domains. But it is difficult to interpret the results of this diagnostic because it is not clear whether even clearly topicalized phrases from embedded clauses trigger weak crossover effects in all cases in the first place. 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).

(40) a. Ἐγώ ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐν ἐ

While unusual, this would not be the only verb in Greek to show this behavior: impersonal verbs like prepi ‘must’ and bori ‘is possible’ have a similar pattern. There are two obvious analytical paths for accounting for the lack of agreement on en in (40).

First, the locality domains for Agreement and for movement could be different (Potsdam and Polinsky 2008, Keine 2016): 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 that 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 and Martinović
2015, who study such opaque interactions in detail: the movement feature on the head that agrees would precede the agreement feature: \( \langle \bullet \text{D}, \ast \phi \ast \rangle \). 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/titan and the aspectual verbs studied by Alexiadou and Anagnostopoulou 2002, 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, the interaction of default values with Agreement failures in Preminger 2014). It is not important for our purposes to choose between these alternatives, and we leave adjudication between them to future work.

### 3.2 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 (such as Giannakidou 2009, Philippaki-Warbuton 1994, and others) 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/tatan in these structures selects for MoodP directly, bypassing the CP layer, and that this truncated complement phrase is the realization of a 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.
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.

(42) *Thelo ton Yanni na ton want.IMPERF.NONPAST.1sg the.ACC Yannis.ACC na him apolisun avrio.
fire.PERF.NONPAST.3pl tomorrow
‘I want them to fire Yannis 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*:\(^\text{17}\)

---

\(^\text{17}\)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 Yanni.ACC na him fire.IMPERF.NONPAST.3pl tomorrow
‘They are firing Yanni tomorrow.’

21
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 (21) 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 $\phi$-features via Agree with the embedded subject.

### 3.3 On the *the na* construction

Greek Cypriot speakers also find examples like (44) 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.

(44)  
\[ E(n) \text{ the na } pao \text{ popse.} \]  
\[ \text{NEG the na go.PERF.NONPAST.3SG tonight} \]  
’I am not going tonight.’

This use of *the* + *na* is also reported in Roussou and Tsangalidis 2010 as a reduced form that maintains the *na*-complement. Markopoulos 2008 notes the emergence of *the na* in the 14th century AD 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 Modern 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.

(45)  
\[ An \text{ the na mbun apu tin tripan, t’ aloga apothe na ta } \]  
\[ \text{if ‘the’ na enter.3pl from the hole the horses from.where na them mbasomen?} \]  
\[ \text{pass.through.1pl} \]  
‘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 (46a) 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 (46b), 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).

\[(46) \hspace{1em} \text{a. } En \text{ thelo } na \text{ pao} \]
\[
\begin{align*}
\text{NEG} & \text{ want.NONPAST.IMPERF.1sg na go.NONPAST.IMPERF.1sg} \\
\text{avrio, alla en na pao.} \\
\text{tomorrow, but be.3sg na go.NONPAST.IMPERF.1sg} \\
\text{I don’t want to go tomorrow, but I will go.}
\end{align*}
\]

\[
\begin{align*}
\text{b. } #\text{En the na pao avrio, alla en na} \\
\text{NEG ‘the’ na go.NONPAST.IMPERF.1SG tomorrow, but be.3sg na pao.} \\
\text{go.NONPAST.IMPERF.1SG} \\
\text{‘I will not go tomorrow, but I will go.’}
\end{align*}
\]

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.

3.4 An additional point of comparison with Standard Modern Greek tha

The assumption that the Cypriot Greek future periphrastic construction is built on a biclausal structure, involving a na-clause, comes as no surprise when considering the development of the future particle tha in Standard Modern Greek. Historical work on the development of the future particle tha 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 Standard Modern Greek tha involves a redeployment of the volitional verb thelo ‘want’ and the infinitive, the latter replaced by the head ina. A use of the volitional verb and the na-clause
to give a future reading has also been documented for medieval Cypriot Greek in Aerts 1983 in (47) and a similar development to today’s use of en na has also been previously assumed (Chatziioanou 1999:92, Menardos 1969).

(47) I tis theli na mini as mini.

PRT someone want.3sg na stay let stay.3sg
‘If someone will stay, let him stay.’

The Cypriot Greek periphrastic en na can profitably be compared to the Standard Modern Greek future particle tha (see Tsangalidis 1998 and Giannakidou and Mari to appear for extensive discussion of Standard Modern Greek tha). In all contexts in which we saw en na appear above, Standard Modern Greek would use the particle tha: like en na, tha takes both perfective and imperfective nonpast verbal complements in its futurate use (while also taking the past in its epistemic use). Both en na and tha can appear, for example, in protases of conditionals:

(48) a. An en na vreksi, en na
    if be.NONPAST.3 na rain.PERF.NONPAST.3s be.NONPAST.3 na
    minume esso.
    stay.PERF.NONPAST.1pl home
    ‘If it rains, we will stay home.’
    [Cypriot Greek]

b. An tha vreksi, en na
    if FUT rain.PERF.NONPAST.3s be.NONPAST.3 na
    minume esso.
    stay.PERF.NONPAST.1pl home
    ‘If it rains, we will stay home.’
    [Cypriot Greek]

(49) An tha vreksi, tha minume spiti.
    if FUT rain.PERF.NONPAST.3s FUT stay.PERF.NONPAST.1pl home
    ‘If it rains, we will stay home.’
    [Standard Modern Greek]

The fact that tha can appear in the Cypriot Greek conditional protasis in (48b) is unexpected, unless the conditional head an in Cypriot Greek bears the relevant conditioning feature NEG that the Vocabulary Insertion rule in (20c) above requires. We are not in a position to pursue this further here, other than to note the cross-linguistically variable appearance of so-called ‘expletive’ negation in conditionals (see Yoon 2010 for discussion).

This comparison of the uses of Standard Modern Greek tha and Cypriot Greek en na serves strengthen the supposition that these two elements play parallel roles and have parallel distributions in the two language varieties.
4 Conclusion

The 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 clause-mate 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. In one sense, *tha* is an allomorph of *en na*.

Cypriot Greek *tha* therefore represents a morpheme whose environment for Vocabulary Insertion cannot be stated within a single extended projection. For this reason, we 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 immediately 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!

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.


Merchant, Jason. 2016. Does Greek have VVPE? Yes, but it’s easy to overlook. Paper presented at the 5th annual Midwest Greek Linguistics Workshop, University of Chicago.


Symeonides, Charalampos P. 2006.*Istoria tis Kypriakis dialektou. [The history of the Cypriot dialect].*
of the Cypriot dialect]. Lefkosia: Kentro Meleton Ieras Monis Kikkou.