Abstract
The present paper investigates the use of *oti*, *na* and negation in wh-question production in L1 Greek. Children’s preference is explored for use of *oti* and *na*, and for use of the negation markers *den* and *min*. These elements have been extensively studied from a theoretical perspective, yet they remain poorly investigated from an acquisition perspective, hence the present study. In long-distance wh-questions *na* is predicted to be preferred over *oti* due to its stronger entrenchment as clause-introducing element and as mood marker; in short-distance questions, however, *na* is predicted to be less preferred than the indicative due to the enriched modal semantics it carries in matrix clause environments. In negative matrix questions *den* is expected to be the preferred choice, since *min* occurs with *na*, which carries an extra semantic/pragmatic load. To test these predictions, a group of ninety four-to-seven-year-old Greek children participated in elicited production tasks designed mainly along the methodological principles of Crain and Thornton (1998). The results were generally in line with initial expectations. Children resorted mostly to *na* in long-distance contexts and to the semantically simpler indicative questions in short-distance contexts. With negative questions, higher accurate use rates were attested for target *den* than for target *min*, reflecting the simpler semantics associated with the former. Overall, these findings provide evidence that children opt for economy, with semantic factors contributing to their economy-based choices.

Keywords
wh-questions; L1: production; clause-introducing elements; negation; Greek

1. Introduction
The aim of the present article is to investigate the use of *oti*, *na* and of negation *den* and *min* in Greek children’s L1 acquisition of wh-interrogatives. A core underlying hypothesis is that access to long-distance (LD) movement occurs from around the age of three onwards (e.g. Thornton 1990, Thornton and Crain 1994, de Villiers et al. 1990). An additional hypothesis is that economy considerations seem to be in play during the acquisition process; as a result, such considerations should dictate the least marked choices between *oti*, *na*, *den* and *min* when it comes to the production of LD and short-distance (SD) wh-questions.
2. Theoretical Background

In the present section a brief presentation will be provided of basic properties of *oti*, *na* and negation markers in Greek.

2.1. Long-Distance Movement in Greek wh-Interrogatives

Here are examples of Greek LD wh-questions.

(1) Ti ipes oti ðjavasun ta peðja?
what said-2SG that will read-3PL the children

‘What did you say that the children will read?’

(2) Ti ipes na ðjavasun ta peðja?
what said-2SG to read-3PL the children

‘What did you tell the children to read?’

As shown in (1) and (2), Greek allows wh-extraction out of a complement clause in the presence of the elements *oti* and *na*; in general, *oti* introduces embedded tense-dependent clauses while *na* tense-independent ones (Mackridge 1985). The categorial status of *oti* and *na* has been the subject of much discussion in the literature.

To begin with *oti*, in terms of interpretability, it seems to be the least specified complementiser in Greek in the sense that it is a morpheme introducing declarative clauses (Mastropavlou and Tsimpli 2011). *oti* just presents facts (Philippaki-Warburton and Veloudis 1984) without attributing any additional semantic connotations to the clause it introduces (Roussou 2006). Under a revised version of Rizzi’s (1997) Split-CP hypothesis, Roussou (2000) claims that the C domain in Greek has three basic C positions: the higher one is specified for subordination (*C*), the middle one for clause-typing (*C*<sub>OP</sub>) and the lower one for modality features (*C*<sub>M</sub>). In this proposal, *oti* is taken to be directly merged in the second higher one, i.e. in *C*<sub>OP</sub>. Being merged there, *oti* spells out features for subordination and clause typing.

Similarly to other complementisers across languages (e.g. English, Italian), under certain circumstances *oti* can optionally be given a null spellout at PF. This pattern is mostly restricted to contexts where *oti* denotes opinion or belief and complements verbs like *nomizo*, *pistevo*, *elpizo*, *leo*, etc. which can acquire a parenthetical meaning (e.g. Drachman 2003, Drachman 2004, Mastropavlou and Tsimpli 2011, Roussou 2006). Besides, omission of *oti* in such contexts is not

---

1) See Drachman (2004) for examples.
precluded in the presence of negation (Drachman 2003). An illustration of this pattern is provided through the following examples from Drachman (2004, 5).

(3) *Pjo palto nomizis (oti) ða tu pai?
which coat think-2SG that will him suit-3SG
‘Which coat do you think (that) will suit him?’

(4) Pjo palto nomizis (oti) ðen ða tu pai?
which coat think-2SG that not will him suit-3SG
‘Which coat do you think (that) will not suit him?’

On the whole, *oti* is considered to be a typical complementiser in Greek, with one of its functions being to introduce an embedded clause out of which wh-extraction occurs.

Turning to *na*, its categorial identity has been strongly questioned. On one view, *na* is considered to be a complementiser, just like *oti* (e.g. Agouraki 1991, Tsoulas 1993). On another view, *na* is not a complementiser but merely the marker of subjunctive mood in the language (e.g. Philippaki-Warburton 1992, Philippaki-Warburton and Veloudis 1984, Roussou 2000). Arguments in support of this claim come, for example, from the fact that *na*, unlike *oti*, can co-occur in the surface syntax with a wh-phrase moved in the embedded SpecCP, which is the typical complementiser position (Philippaki-Warburton 1992, Roussou 2000).

(5)* Den ðimate pjon oti ða ði.
not remember-3SG who that will meet-3SG
‘S/he does not remember who s/he is meeting.’

(6) Den kseri pjon na pari mazi tu.
not know-3SG who to take-3SG with him
‘He doesn’t know who to take with him.’ (from Philippaki-Warburton 1992, 261)

Moreover, as opposed to typical complementisers like *oti* ((7)), *na* can also surface in matrix clauses ((8)) (Roussou 2000, 2006) encoding modality (Pavlidou 1991, Rouchota 1993, Tsangalidis 2000), while it attributes a note of uncertainty or non-fact to its complement clause ((9) vs. (10)) (Philippaki-Warburton and Veloudis 1984, Roussou 2006, Varlokosta 1994).

(7)* Oti fijis.
that leave-2SG
‘Leave!’

(8) Na fijis!
to leave-2SG
‘Leave!’ (from Roussou 2006, 37)
(9) O Janis elpizi oti θa erθi i Maria.
the John hope-3SG that will come-3SG the Maria

'John hopes that Maria will come.'

(10) O Janis elpizi na erθi i Maria.
the John hope-3SG to come-3SG the Maria

'John hopes that Maria will come.'

As regards its position in the clause structure, *na* is argued to move from the lower *C_m* position to the higher *C_op*, and thus to spell out features of modality and clause typing (Roussou 2000).

In the present paper, a rather compromising approach to *na* will be adopted. Given that ‘diachronically *na* is the residue of the complementiser *ina* which introduced the subjunctive mood in classical Greek’ (Varlokosta 1994, 97), *na* is viewed as a morpheme with the double function of subjunctive/modality marker (carrying the modality feature) and clause-introducing element. In other words, *na* seems to oscillate between its subjunctive marker and complementiser status. As a consequence, it seems to have wider distribution than *oti*, which is a typical complementiser in Greek (Mastropavlou and Tsimpli 2011, Roussou 2006). This observation will be of great significance for the discussion below.

In summary, LD wh-questions in Greek are formed by extracting a wh-phrase out of an embedded finite clause introduced by *oti* or *na*. In terms of structure, the position of *oti* and *na* is represented as follows (adapted from Roussou 2000, 79):

\[
[C \ [\text{Topic/Focus} \ [C \ [\text{OP} \ [\text{oti/na} \ [\text{Neg} \ \text{den/min} \ [\text{C_m} \ \text{t_na} \ [I \ldots]]]]]]]]
\]

### 2.2. Negation and Greek Interrogatives

When a question is negated, this is expressed by means of a preverbal negation marker, which is actually the marker of all sentential negation in Greek. Depending on mood distinctions, the form of this negative marker varies. Specifically, there is claimed to be a one-to-one correspondence between indicative/subjunctive mood and the negative markers *den/min* respectively (Philippaki-Warburton 1994, Roussou 2000, Tsimpli and Roussou 1991). Here are examples of Greek negative wh-questions of both types.

(11) Pjos *den* pije sti θulia simera?
who not went-3SG to the work today

‘Who didn’t go to work today?’
(12) Ti na min po otan ton do?
what to not say-1SG when him see-1SG

‘What should I not say when I see him?’

As shown in the above examples, both negative markers need an overt host, and this host is always the verb raising to C. Due to its sensitivity to mood, negation in Greek has been claimed to project a NegP in the C domain. Under Roussou’s (2000) proposal that there are three C positions in Greek (see section 2.1), NegP is taken to be located between the two lower C heads. This system yields the right order with *den but not with min (i.e. ‘*min na’). As stated in the discussion of na in the previous section, na raises to the higher C_{OP}, and this gives rise to the correct order na min.

3. Literature Review

3.1. Acquisition of Complementation

The acquisition of LD questions by children is closely related to the acquisition of complementation. On this ground, an overview of child complementation studies is essential.

Findings from English demonstrate that children acquire complementation between the ages of two and three. For instance, in Bloom et al.’s (1989) longitudinal study with two-to-three-year-old children, it was shown that wh-complementisers were the most frequent to be used, while that was rarely used to introduce the embedded clausal complement. Furthermore, Radford (1994) reports that after their second year of life, children start producing sentential complements introduced by complementisers like if, that and for as well as by other words that in the adult language do not function as complementisers. Similar findings have been observed by Diessel (2004), with wh-complements appearing earlier than that and if’ complements. Besides, the early English findings reported above gain further support from other early languages. For example, in Sesotho, embedded CP complements start appearing between ages 2;8 and 3;0 and are introduced with well-formed complementisers from the beginning (Demuth 1995). In some studies, however, it has been noted that children do not use overt complementisers in their sentential complements (e.g. Josefsson and Hakansson 2000, Meisel and Müller 1992, Poeppel and Wexler 1993). Yet, as noted by Poeppel and Wexler (1993, 20), “the lack of overt complementisers ... does not conclusively prove the absence (or unavailability) of the linguistic category in question,” given that young children’s omission of the introducing complementiser in subordinate clauses is not systematic.

Regarding Greek, Baslis (1993) studied spontaneous data from Greek children up to the age of three and noted that na appeared near the end of the
second year of life and earlier than *oti*. Similarly, Stephany (1997) observed that Greek children start producing complement clauses introduced by *na* at around 1;10, and the use of *na* becomes productive by 2;4. On the contrary, *oti* subordinate clauses appear later, at around the age of 2;10. As noted in Rousou (2006), *oti* is the latest to be acquired. In a more recent study recording a girl's spontaneous speech in the age range of 1;8 to 4 years, Katis and Stam-ppouliadou (2009) attribute the earlier systematic use of *na*- compared to *oti*-complement clauses to a usage-based motivation: *na* does not occur only in embedded clauses; from an early age, it also occurs in simple clauses as a subjunctive marker. On the contrary, *oti* may only occur in embedded clauses, and hence it is more weakly entrenched than *na*. Finally, Mastropavlou and Tsimpi (2011) studied the spontaneous speech of three-to-six-year-old children. Similarly to previous studies, *na* was observed to be acquired earlier and to be more frequently used than *oti*. In Mastropavlou and Tsimpli’s (2011) approach, the acquisition and use of *oti* and *na* are associated with the notion of semantic richness, i.e. with the formal LF-interpretable features that they encode. More concretely, the higher frequency in use of *na* as opposed to *oti* was attributed to the dual function of *na* as the carrier of the [modality] feature and as the marker of clause typing. In other words, *na* is semantically richer than *oti*, and hence children seem to acquire it earlier and more easily. Finally, omissions of *na* and of *oti* have also been reported, with the latter often restricted to contexts of acceptable drop (Mastropavlou and Tsimpli 2011).

To summarise, it is clear that complementation is acquired at a very early age across languages. Sentential complements may be introduced by either an overt complementiser or an omitted one; irrespective of the complementiser spell-out, however, the significant conclusion is that subordinate complement clauses are present in child grammar by age three or so. What this implies, in turn, is that children by that age are in principle able to produce complex questions involving a subordinate embedded clause.

### 3.2. Acquisition of Negation

Another aspect that has been of considerable interest across child language studies is the acquisition of questions containing negation. Specifically, the presence of matrix scope negation has been examined in child negative questions, with the aim to see what role negation plays in children’s acquisition of wh-interrogatives.

Child studies have shown that the acquisition of negative questions may differ from one language to the other. As regards child English, Guasti et al. (1995) tested the production of matrix questions by 3;8-to-4;7-year-old children who were prompted to pose questions to a puppet about himself and about sce-
narios acted out with props. What they found was that affirmative questions were mostly grammatical, whereas the reverse held for negative questions. Few adult-like forms were produced, while the two usual types of errors reported were firstly auxiliary doubling, and secondly negation and auxiliary doubling. In both error types, the doubled constituent(s) surfaced in C and in IP. Using the same elicitation technique, Thornton (1995) also observed that children aged between 4;1 and 5;4 produced negative questions with a doubled auxiliary. In a similar study Hiramatsu (2003) found that 3;10-to-4;11-year-old children’s affirmative questions were mostly grammatical, as opposed to their negative questions which were an even mix of grammatical and ungrammatical auxiliary doubling questions. Similarly to Thornton (1995) but unlike Guasti et al. (1995), in Hiramatsu’s (2003) study children did not produce any negation and auxiliary doubling questions. What is interesting to mention is that Hiramatsu’s children also participated in a grammaticality judgment experiment where they had to indicate whether a puppet’s produced questions were grammatical or not. The results showed that children rejected the auxiliary doubling questions that they had produced during the elicited production experiment. According to Hiramatsu, the grammaticality judgment data reflect what children know about questions, whereas the production data are ‘not the result of a non-adult grammar but rather the result of a performance phenomenon based on incorrect lexical information regarding constituent negation’ (Hiramatsu 2003, 99).

However, Hiramatsu’s explanation does not seem to gain wide applicability, since the production of negative questions, as indicated earlier, does not follow a uniform pattern across languages. Guasti (1996), for instance, reports that the English pattern outlined above is not attested in Italian. Given that any verb can move to C in Italian, the Italian counterpart of the early English auxiliary doubling interrogative would be a main verb doubling structure. Nevertheless, in her elicited production experiment replicating the one by Guasti et al. (1995) for English, Guasti (1996) found that Italian children of ages 3;1 to 4;8 produced adult affirmative and negative questions, while verb doubling or negation and verb doubling were never attested. This asymmetry between early English and early Italian has been attributed to the hypothesis that children initially decide that the Neg-Criterion must be satisfied in the projection in which the verb checks its features, including the negation feature (Guasti 1996, Guasti et al. 1995). This V-related projection is different in English and in Italian: in English only auxiliaries and modals can raise to C, whereas in Italian all verbs can raise. As a consequence, CP is the V-related projection only in languages like Italian but not in ones like English. In the latter type of languages C cannot host an inflected verb, and hence IP is the only candidate for the V-related projection (Guasti 1996). On these grounds, English children retain negation inside IP, which accounts for the ungrammaticality of their
produced questions, while Italian children raise it to CP and thus produce adult interrogatives (Guasti et al. 1995).

To sum up with respect to the acquisition of negative questions by children, it can be concluded that the structural position that negation occupies across child languages can determine the grammaticality of the produced negative interrogatives. As a result, negation is an important constraint which is operative in the acquisition of interrogatives by children.

4. The Study

The predictions stated with reference to oti, na and negation during the acquisition of interrogatives in Greek, and aimed to be evaluated in the present empirical study, are based on the relevant cross-linguistic findings outlined in the previous section and on some Greek-specific properties which can modify the formulation of the predictions.

Prediction 1: With regard to LD question production, it is predicted that the Greek children’s produced LD questions will mostly contain an embedded clause introduced by na rather than one with oti. This prediction is based on the assumption that na may be more prototypical than oti; its dual status as a clause-typing element and as a subjunctive marker makes it more prototypical either because of its stronger entrenchment (Katis and Stampouliadou 2009) or of its greater semantic richness (Mastropavlou and Tsimpli 2011) compared to oti.

Prediction 2: As for negation, it is predicted to be present in target negative questions, and this follows from its attested presence from early on in child grammar (e.g. Déprez and Pierce 1993, Stephany 1997); as a matter of fact, accurate use of target ḍen is expected to be higher than accurate use of target min in all groups, due to the fact that min, as opposed to ḍen, presupposes the presence of na, which carries an extra semantic load of modality when appearing in SD questions (Pavlidou 1991, Rouchota 1993, Tsangalidis 2000).

4.1. Method

In order to test the above predictions concerning early L1 Greek, a study was designed in which the methodology of research was one of elicited production, adapted mainly from Crain and Thornton (1998). What follows is a detailed description of the participating subjects, of the experimental tasks set and of the procedure followed, as well as of the measurement steps employed for the analysis of the collected data.
4.2. Participants

The study group consisted of ninety normally developing children aged 4;0 to 7;0. These children were divided into three equivalent subgroups A, B and C. Group A included thirty children between four and five (mean age range: 4;6), group B thirty children between five and six (mean age range: 5;5) and group C thirty children between six and seven years old (mean age range: 6;7). Group A and B children were in their first and second year in kindergarten respectively, while group C children attended the first grade in primary school. The sixty children of groups A and B were selected from four state kindergartens, while the thirty children of group C were recruited from two state primary schools. All state kindergartens and primary schools are located in Trikala, Greece.

All child consultants who participated in the study group were selected with the criterion that they were native speakers of Greek and that they did not show any serious language difficulties or signs of psychosocial and neurological disorders. Their typical linguistic and social developmental profile was established by information obtained from schoolteachers. In addition, efforts were made to select an equal number of boys and girls, the goal being to alleviate differences in performance potentially attributed to sex variation. Finally, with regard to the setting, each child was tested separately in a room next to their classroom during their daily school program.

4.3. Materials

What follows is a detailed presentation of the experimental tasks set to the child consultants.

Task 1: Production of Long-Distance Questions with *oti*

Specific Rationale

In task 1 the aim was the production of LD questions with *oti* by children. The rationale evolved around the hypothesis that in cases of LD production, most of them would contain a clausal complement introduced by non-target *na* rather than target *oti*, given the pragmatically (Katis and Stampouliadou 2009)/semantically (Mastropavlou and Tsimpli 2011) more prototypical nature of the marker *na*.

Materials

For the elicitation of LD questions with an embedded *oti* clause, a guessing game was conducted for the purposes of which a puppet called ‘Astrulis’ that
had come from another planet was introduced to the child. This game included nine prompts of the type *Rotise ton Astruli X mandevi aftos* (‘Ask Astrulis [i.e. the puppet] X he guesses’), where X stood for the respective wh-element.

Presented just below is a sample of the game protocol used and its accompanying prompt.

The experimenter presents five toys: one car, one pink and one blue comb, and one red and one yellow ball. She asks both the child and the puppet to cover their eyes while she is hiding each of these toys in different numbered boxes. Then she asks the child and the puppet to uncover their eyes, and the game proceeds. After the experimenter has elicited the child's guess, she prompts the child to elicit the puppet's guess. One set of stimulus sentences she uses is the following:

**Experimenter:** Sto kuti 3 ekripsa mia xtena. Ja mandepse pja.

‘In box 3 I hid one of the combs. Guess which one.’

(*the child says his/her guess …*)

**Experimenter:** Esi mandevis oti sto kuti 3 ekripsa tin ... xtena. Rotise ton Astruli pja mandevi aftos.

‘You are guessing that in box 3 I hid the ... comb. Ask Astrulis which one he is guessing.’

**Target question:** Pja xtena mandevis oti ekripsi sto kuti 3?

‘Which comb are you guessing that she (i.e. the experimenter) hid in box 3?’

**Procedure**

The duration of this game was about fifteen minutes. The game was similar to that used by Thornton (1996) and proceeded as follows: both the child and the puppet covered their eyes while the experimenter hid objects in small numbered boxes. The child guessed what was hidden in each box and s/he was then prompted to ask the puppet about his guess. In this way, the child was prompted in a random order to produce the target LD questions. In case the child did not react to the experimenter’s prompt, the prompt (and the relevant part of the game) was repeated twice; if still no question was elicited, the procedure continued with the presentation of the next part of the game and the corresponding prompt.

**Task 2: Production of Long-Distance Questions with na**

**Specific Rationale**

The rationale behind task 2 was the same as in task 1. The aim was to see whether children would produce target LD questions; contrary to task 1, however, target LD question production was more prominent as an expectation given that *na* is pragmatically (Katis and Stampouliadou 2009)/semantically (Mastropavlou and Tsimpi 2011) more prototypical than *oti.*
Materials
The experimental task set for the elicitation of LD questions with an embedded *na* clause consisted of six prompts, which were of the form Roti se ton Astruli X protimai aftos (‘Ask Astrulis [i.e. the puppet] X he prefers’), where X stood for the respective wh-element.

A sample set of the task and the stimulus sentences used is given here:

The experimenter presents four toy characters: one rabbit, one dog and two horses. She explains to the child that three of these toy characters have to be matched with certain actions.

Experimenter: Kapjo alogaki tha pai volta. Roti se ton Astruli pjo protimai aftos.
‘One of the two horses is going for a walk. Ask Astrulis which one he prefers.’

Target question: Pjo alogaki protimas na pai volta?
‘Which horse do you prefer that it go for a walk?’

Procedure
The present experimental task was very short in duration, namely about five minutes long. As regards the procedure, it was based on an idea from Thornton (1996): that is, the child and the puppet were presented with some toy characters and some possible actions they could do. The child was prompted to ask the puppet to decide which action each toy character would do. Thus, through this game, the child was prompted to produce the target LD questions. In case the child did not respond to the prompt, the latter (and the relevant part of the game) was repeated twice; if no question was elicited, the procedure proceeded with the provision of the next prompt.

Task 3: Production of Short-Distance Questions in the Indicative

Specific Rationale
Task 3 aimed at the production of SD questions in the indicative. The rationale underlying this experiment was that SD question production would not posit any serious problems to children. Specifically, children’s performance was tested on target affirmative vs. target negative questions. Expectations leaned towards highly accurate performance, since negation is found to be present from early on in child grammar (e.g. Déprez and Pierce 1993, Stephany 1997).

Materials
The puppet called ‘Astrulis’ was again employed for the purposes of the present experiment. In total, nineteen prompts were designed on the basis of scenarios and short stories close to children’s everyday life. Fourteen of these prompts were true test items, each of which ended on the phrase Roti se ton Astruli X
(‘Ask Astrulis [i.e. the puppet] X’), where X stood each time for the respective wh-element. The remaining five prompts targeted yes/no question production and were included simply for the provision of smooth transitions within the stories.

Here is a sample test story and the relevant stimulus sentences:

(English translation)

Stimulus context 1: Three smurfs are roller-skating in the forest when they meet two horses. While playing with the two horses, a bear suddenly appears and tramples the brown horse.

Experimenter: I arkuða patise kapjo alogaki. Rotise ton Astruli pjo.

‘The bear trampled one of the horses. Ask Astrulis which one.’

Target question: Pjo alogaki patise i arkuða?

‘Which horse did the bear trample?’

Stimulus context 2: At night, the three smurfs go to sleep. The two of them get covered with the blanket but the third one does not because it cannot sleep.

Experimenter: Kapjo strumfaki ðen skepazete me tin kuverta. Rotise ton Astruli pjo.

‘One smurf does not get covered with the blanket. Ask Astrulis which one.’

Target question: Pjo strumfaki ðen skepazete me tin kuverta?

‘Which smurf does not get covered with the blanket?’

Procedure

This task lasted about twenty minutes. For the elicitation of questions in the indicative, the child was first prompted to pose questions to the puppet about himself and his friends and about life on his planet (e.g. Thornton 1996). Then, stories acted out with props or accompanied by pictures were presented to both the child and the puppet. At the end of each story, using lead-in statements the experimenter prompted the child to ask the puppet the target questions so as to see whether the puppet understood what had happened in the story. In case the child did not react to the experimenter’s prompt, the prompt (and the story where necessary) was repeated twice; if no question was elicited, the procedure continued with the presentation of the next story and its accompanying prompt.

Task 4: Production of Short-Distance Questions with na

Specific Rationale

The rationale behind task 4 was exactly the same as in task 3 with regard to the expected lack of difficulty in producing SD questions. Again, the parameter examined was the affirmative/negative, with the only difference being that the focus was on the production of questions featuring na and not indicative. On
the basis of previous acquisition data, expectations were high with regard to the production of *na* given its presence from early on in child grammar (e.g. Stephany 1997, Katis and Stampouliadou 2009, Mastropavlou and Tsimpli 2011), but quite lower with regard to the production of the negated form *na min* (e.g. Stephany 1997).

Materials
This experimental task consisted of nineteen prompts which, as a whole, constituted a big story. All prompts included the phrase *Rotise ton Y X* (‘Ask Y X’), where X stood each time for the respective wh-element and Y for a character from the story.

Here is a sample part of the test story and relevant prompts:

*(English translation)*

**Stimulus context 1:** Jim goes shopping. He wants to buy a ball. He goes in a shop and sees two balls, a red one and a yellow one.

**Experimenter:** O politis lei sto Ðimitri na min pari kapja bala. Rotise ton pja. ‘The salesman tells Jim not to take one specific ball. Ask him which one.’

**Target question:** Pja bala na min pari o Ðimitris? ‘Which ball should Jim not take?’

**Stimulus context 2:** Maria is at Jim’s party and is playing with the cat. They are playing with two toy cars, a big one and a small one.

**Experimenter:** I gata lei sti Maria na tis ðosi kapjo aftokinitaki. Rotise tin pjo. ‘The cat asks Maria to give her one car. Ask her which one.’

**Target question:** Pjo aftokinitaki na su ðosi i Maria? ‘Which car should Maria give you?’

Procedure
The duration of the present experimental task was about twenty-five minutes. For the elicitation of questions with *na*, a big story was acted out in front of the child and the puppet, with the latter playing this time the role of a passive observer. The child was prompted at frequent intervals to pose SD questions to the toy characters of the story, so as to find out how the story continued. In cases that the child did not produce a question as a response to the experimenter’s prompt, the prompt (and the story where necessary) was repeated twice; if still no question was elicited, the procedure went on with the continuation of the story and the presentation of the relevant prompt.

4.4. Analyses and Measurements
The data collected from all four tasks were inserted into a database using the statistical software SPSS 17.0 for Windows. Mean ratings were then calculated
for each participant; these ratings were tested statistically with mixed ANOVA analyses which provided within- and between-subject comparisons. These, in turn, enabled the checking for significance of the effect of category and group on the ratings. Apart from the group-level analyses, we also carried out individual-subject analyses in order to provide a more fine-grained picture of the results obtained.

On the whole, the analyses of the data carried out at group and individual-subject level were the following: a. correct responses with regard to the prediction tested, and b. type of errors. For analyzing the collected data, we followed common practice in the sense that we included only those responses that were relevant each time to the condition tested; responses that were irrelevant to the specific condition were removed from the analysis in question.

4.5. Evaluation of Methodology

Overall, the tasks used were based both on older and on newly-created protocols which were adjusted to the purposes of the present study. In all cases, however, it was ensured that the scenarios and stories included in the tasks were adapted to children’s age and school experience. On the whole, the method of elicited production provided meaningful results indicating its experimental validity, as in previous studies that employed a similar method. Besides, no serious procedural problems were encountered, since children reacted mostly as expected to the given prompts; elicitation of no response whatsoever was very rare across all children tested.

5. Results

On the whole, the results obtained from tasks 1, 2, 3 and 4 were in line with the predictions outlined in section 4. A detailed presentation of the relevant results at a group and at an individual-subject level is provided below.

5.1. Group Analysis

In this section, the results of the main study will be presented at a group level, structured around the formulated predictions. Within- and between-group comparisons will be drawn at each stage.

5.1.1. Prediction 1: Preference for Use of na

As expected, in LD question production na was indeed easier to use than oti for the introduction of the embedded clause, and this difference was very
prominent in the younger groups. A detailed analysis of these results is provided below.

As an introductory note, it is important to mention that overall LD question production remained at a rather low frequency rate in all groups. With regard to prediction 1, two observations are of relevant interest here. Firstly, in the instances of accurate use of LD question production, the ones with an embedded target *na* clause were more frequent than those with an embedded target *oti* clause in all groups, as shown in Table 1 below.

<table>
<thead>
<tr>
<th>Condition</th>
<th>GROUP A</th>
<th>GROUP B</th>
<th>GROUP C</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Target oti</strong></td>
<td>11/35 31.43%</td>
<td>42/54 77.78%</td>
<td>96/118 81.35%</td>
</tr>
<tr>
<td><strong>Target na</strong></td>
<td>39/39 100%</td>
<td>71/72 98.61%</td>
<td>89/102 87.26%</td>
</tr>
</tbody>
</table>

1 The number of contexts in both conditions in each group corresponds to the total number of LD questions produced with target *oti* and with target *na*. There is no issue of missing cases here.
2 GA= Group Analysis
3 StdD= Standard Deviation

For the two younger groups, accurate production rates of target *na* were much higher than the respective rates of target *oti*, this difference being alleviated in the older group C. As for StdD scores, they suggest some variance in group A's performance with regard to target *oti*; in all other cases, children's accurate production of *oti* and *na* showed consistency, as indicated by the relatively low StdD scores. Besides, it is interesting to note that when it came to LD questions with target *oti*, children often produced a non-target *na* instead or a wh-element in the medial COMP position, or they left this position altogether empty. However, these error patterns were scarcely attested in LD questions with target *na*, except for group C. The relevant error analysis is provided in Table 2.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Error Type</th>
<th>GROUP A</th>
<th>GROUP B</th>
<th>GROUP C</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Target oti</strong></td>
<td><em>na</em></td>
<td>16/24 66.67%</td>
<td>0/12 0%</td>
<td>15/22 68.18%</td>
</tr>
<tr>
<td></td>
<td>'pos'</td>
<td>1/24 4.17%</td>
<td>1/12 8.33%</td>
<td>0/22 0%</td>
</tr>
<tr>
<td></td>
<td>Null COMP</td>
<td>7/24 29.17%</td>
<td>8/12 66.67%</td>
<td>1/22 4.55%</td>
</tr>
<tr>
<td></td>
<td>Wh-element</td>
<td>0/24 0%</td>
<td>3/12 25%</td>
<td>6/22 27.27%</td>
</tr>
<tr>
<td><strong>Target na</strong></td>
<td><em>oti</em></td>
<td>0/0 0%</td>
<td>1/1 100%</td>
<td>8/13 61.54%</td>
</tr>
<tr>
<td></td>
<td>Null COMP</td>
<td>0/0 0%</td>
<td>0/1 0%</td>
<td>1/13 7.69%</td>
</tr>
<tr>
<td></td>
<td>Wh-element</td>
<td>0/0 0%</td>
<td>0/1 0%</td>
<td>4/13 30.77%</td>
</tr>
</tbody>
</table>

1 GA= Group Analysis
Examples of all these error types in both conditions are provided just below.

Condition: target *oti*

**na**

(13) Ti nomizis esi na ine eki mesa?  
what think-2SG you to be-3SG there inside  
target: Ti mandevis oti ekripse sto kuti 3? (‘What are you guessing that she hid in box 3?’)  
(A: 7)

(14) Esi Astruli pjo pjo aftokinikaki mandevis na ine, to kitrino i to you Astruli which which car guess-2SG to be-3SG the yellow or the kokino?  
red  
target: Pjo aftokinikaki mandevis oti ine sto kuti 3? (‘Which car are you guessing to be in box 3?’)  
(C: 16)

**pos**

(15) Astruli esi ti nomizis pos ine?  
Astruli you what think-2SG that be-3SG  
target: Ti mandevis oti ine sto kuti 1? (‘What are you guessing to be in box 1?’)  
(A: 26)

(16) Astruli pjo pjo nomizis pos ine?  
Astruli which which think-2SG that be-3SG  
target: Pja bala mandevis oti ine sto kuti 5? (= ‘Which ball are you guessing to be in box 5?’)  
(B: 12)

**null COMP**

(17) Astruli pu mandevis ine to alogaki?  
Astruli where think-2SG be-3SG the horse  
target: Pu mandevis oti ine to alogaki? (‘Where are you guessing the horse to be?’)  
(A: 10)

(18) Ti mandevis ine sto kuti to 1?  
what guess-2SG be-3SG in the box the 1  
target: Ti mandevis oti ine sto kuti 1? (‘What are you guessing to be in box 1?’)  
(C: 8)

**wh-element**

(19) Pu mandevis esi pja ine i kafe bala?  
where guess-2SG you which be-3SG the brown ball  
target: Pu mandevis oti ine i kokini bala? (‘Where are you guessing the red ball to be?’)  
(B: 1)

---

2) Such coding in all the examples cited corresponds to the reference code of the child producing the respective question. Specifically, what is indicated is the age group that the child belongs to (A, B or C) and his/her position (out of 30 children) in the sequence of testing.
(20) Pjo mandevis esi Astruli o-\textit{pu} ine to kokino i to kitrino?
which guess-2SG you Astruli where be-3SG the red or the yellow

\quad target: Pjo aftokinitaki mandevis oti ine sto kuti 3? (‘Which car are you guessing to be in box
3?’) (C: 20)

\textbf{Condition: target \textit{na}}

\textit{oti}

(21) Pjo protimas \textit{oti} \text\check{a} pai volta, to kafe i to mavro?
which prefer-2SG that will go-3SG walk the brown or the black

\quad target: Pjo alogaki protimas na pai volta? (‘Which horse do you prefer to go for a walk?’) (B: 1)

(22) Astruli pjo alogaki nomizis \textit{oti} \text\check{a} pai volta?
Astruli which horse think-2SG that will go-3SG walk

\quad target: Pjo alogaki protimas na pai volta? (‘Which horse do you prefer to go for a walk?’) (C: 27)

\textbf{null COMP}

(23) Pu protimas Astruli \text\check{a} pai o lagos?
where prefer-2SG Astruli will go-3SG the hare

\quad target: Pu protimas na pai to kunelaki? (‘Where do you prefer that the rabbit go?’) (C: 29)

\textbf{wh-element}

(24) Pjos protimas \textit{pjos} \text\check{a} pai ja ipno?
who prefer-2SG who will go-3SG for sleep

\quad target: Pjos protimas na pai ja ipno? (‘Who do you prefer to go to sleep?’) (C: 6)

On the basis of the above examples, it is interesting to note in passing that errors were mainly grammatical (cf. (13), (17), (18), (19), (20), (21), (23), (24)), while a small minority merely reflected contextual inadequacy and not ungrammaticality (cf. (14), (15), (16), (22)). Contextual inadequacy refers to instances where the child used an element to introduce the embedded clause that was non-target but still grammatical; the verb \textit{mandevo}, for instance, can combine with either \textit{oti} or \textit{na}, so the child’s use of \textit{na} in (14) is not ungrammatical but simply non-target. Likewise, the use of \textit{nomizo pos} instead of \textit{nomizo oti} in (15)–(16), and of \textit{nomizo oti} instead of \textit{protimo na} in (22) is grammatical but non-target. However, this difference between grammaticality and contextual inadequacy errors expanded beyond the aims tested, and hence it was not taken into consideration in statistical analyses.
Various error types aside, the following figure offers a schematic representation of the accurate production rates of target *oti* and target *na*.

![Figure 1. Accurate Use of Target *oti* vs. Target *na* in Long-Distance Question Production](image)

A two-way mixed ANOVA showed a significant main effect within subjects of target marker \[F(1,38)=28.308, p=.000\], and of target marker by group interaction \[F(2,38)=7.981, p=.001\]. The main between-subjects effect of group, however, was not significant \[F(2,38)=2.293, p=.066\]. Pairwise, the difference in frequency of target *oti* and target *na* was statistically significant within the two younger groups A (\(p=.000\)) and B (\(p=.057\)); between-groups, the burst in the use of target *oti* occurred between groups A and B (significant comparisons: A-B: \(p=.017\) and A-C: \(p=.004\)), while in the use of target *na* no significant comparisons were found; accurate production rates were extremely high in all three groups.

As for the second observation concerning production, it was attested that even in cases where a LD question was produced instead of a target SD one,\(^3\) *na*

---

\(^3\) It is interesting to note that the overproduction of LD questions was a surprising, unexpected finding. Instances of LD overuse primarily employed verbs like *thelo* ‘want’ and *boro* ‘can’ in the matrix verb position, and they mostly pertained to target negative questions or to target *na* questions. Given its occurrence in these specific contexts, LD overproduction is not considered to be
was used to introduce the embedded clause. *oti* clauses appeared scarcely only in the oldest children’s questions, as presented in the table and the accompanying figure below.

**Table 3. Overuse of Long-Distance Question Production (GA)**

<table>
<thead>
<tr>
<th>Group</th>
<th>No</th>
<th>Mean</th>
<th>StdD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preference for <em>oti</em> A</td>
<td>0/148</td>
<td>0%</td>
<td>0.00</td>
</tr>
<tr>
<td>B</td>
<td>0/208</td>
<td>0%</td>
<td>0.00</td>
</tr>
<tr>
<td>C</td>
<td>3/260</td>
<td>1.15%</td>
<td>5.19</td>
</tr>
<tr>
<td>Preference for <em>na</em> A</td>
<td>148/148</td>
<td>100%</td>
<td>0.00</td>
</tr>
<tr>
<td>B</td>
<td>208/208</td>
<td>100%</td>
<td>0.00</td>
</tr>
<tr>
<td>C</td>
<td>257/260</td>
<td>98.85%</td>
<td>5.19</td>
</tr>
</tbody>
</table>

1 The number of contexts in each group corresponds to the total number of LD questions produced instead of target SD ones. No missing cases are involved here.
2 GA = Group Analysis
3 StdD = Standard Deviation

an artifact of the experiment. Instead, it is assumed here that the underlying trigger for this pattern seems to be lexicalization of modality. For a more detailed analysis, see section 6.3.
The overuse of LD questions is illustrated through the following examples.

\textit{oti}

(25) Pjos ipes \textit{oti} ðen ða pai puðena simera?
who said-2SG that not will go-3SG anywhere today
target: Pjos na min pai puðena simera? (‘Who should not go anywhere today?’) (C: 23)

\textit{na}

(26) Ti ipes \textit{na} pao na fero?
what said-2SG to go-1SG to bring-1SG
target: Ti na feri mesa apo to spiti? (‘What should he bring from the house?’) (B: 2)

(27) Astruli pjo pjo ðen bori \textit{na} pari o Petros?
Astruli which which not can-3SG to take-3SG the Peter
target: Pjo pragma ðen ða pari mazi tu o Petros? (‘Which thing won’t Peter take with him?’) (C: 2)

The two observations presented just above with regard to the LD production data are suggestive of the earlier emergence and greater preference for \textit{na} as opposed to \textit{oti} in child grammar. Besides, even the error analysis presented in table 2 converges to the same conclusion. In the two younger groups A and B, errors were mainly limited to target \textit{oti} questions and pertained mostly to resort to non-target \textit{na}; on the contrary, in group C errors were found with the use of both target \textit{oti} and target \textit{na}. These error patterns also indicate that the use of \textit{na} is less problematic than the use of \textit{oti} from the youngest test age. When \textit{oti} starts becoming more productive in use (i.e. around group C age), then errors occur with both \textit{oti} and \textit{na} until their acquisition process is completed.

Before ending the discussion on prediction 1, it is interesting to mention that contrary to LD question production, the preference for \textit{na} was not evident in SD question production. Specifically, in the set of the produced SD questions in target SD contexts, the questions with target \textit{na} were less frequent than those with target indicative in all groups. Inaccurate use implies use of \textit{na} instead of target indicative and use of indicative instead of target \textit{na}.

Table 4. Accurate Use of Short-Distance Question Production (GA)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Group</th>
<th>No</th>
<th>Mean</th>
<th>StdD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target indicative questions</td>
<td>A</td>
<td>168/172</td>
<td>97.67 %</td>
<td>5.00</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>232/234</td>
<td>99.15 %</td>
<td>16.52</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>231/234</td>
<td>98.72 %</td>
<td>17.37</td>
</tr>
<tr>
<td>Target \textit{na} questions</td>
<td>A</td>
<td>190/238</td>
<td>79.83 %</td>
<td>28.95</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>345/386</td>
<td>89.38 %</td>
<td>26.93</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>349/412</td>
<td>84.71 %</td>
<td>22.98</td>
</tr>
</tbody>
</table>
The number of contexts in both conditions in each group corresponds to the total number of SD questions produced with target indicative and with target *na*. There is no issue of missing cases here.

GA= Group Analysis

StdD= Standard Deviation

This is clearly shown in the following figure as well.

Within subjects, the main effect of target mood was significant \( F(1,87)=18.220, p=.000 \) but that of its interaction with group was not \( F(2,87)=.058, p=.944 \). Between subjects, the main effect of group was not significant either \( F(2,87)=.219, p=.803 \). Pairwise, within all groups accurate use of target indicative was significantly higher than accurate use of target *na* (A: \( p=.009 \), B: \( p=.031 \), C: \( p=.013 \)), while between groups no significant comparisons were found for either condition; accurate production rates were high across groups.

As a consequence, it becomes clear that the choice of mood in SD questions differs from the choice of mood for the embedded clause in LD questions. Use of *na* predominates in the latter, thus contrasting with use of indicative in the former. To a certain extent, this difference might be attributable to the enriched modal semantics that *na* carries in matrix environments (e.g. Pavlidou 1991, [Figure 3. Accurate Use of Target Indicative vs. Target *na* in Short-Distance Question Production])
Rouchota 1993, Tsangalidis 2000). Also, it might be related to the use of specific ‘complement taking predicates’ in LD questions. That is, LD questions often involved the matrix verbs protimo, boro and ëelo, which are routinely followed by a na complement clause (Katis and Stampouliadou 2009). Further and more elaborate explanations of this difference will be provided in section 6.1.

5.1.2. Prediction 2: Negation Effects

Turning to Prediction 2, negation effects were predicted to be operative in the Greek child data. Table 5 below presents the average accurate performance rates of all groups with respect to the production of negation in target negative contexts, irrespective of accurate SD/overuse of LD extraction counts.

Table 5. Accurate Production of Target Negative Questions (GA)

<table>
<thead>
<tr>
<th>Group</th>
<th>No</th>
<th>Mean</th>
<th>StdD</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>194/319</td>
<td>60.82%</td>
<td>37.68</td>
</tr>
<tr>
<td>B</td>
<td>397/478</td>
<td>83.05%</td>
<td>17.33</td>
</tr>
<tr>
<td>C</td>
<td>464/517</td>
<td>89.75%</td>
<td>23.48</td>
</tr>
</tbody>
</table>

As indicated in this table, the presence of negation was not avoided in target negative questions; on the contrary, accuracy rates were above 50% from the youngest test age, with significant differences attested between groups A and B (p=.002), and between groups A and C (p=.000). This suggests that a significant burst in the production of negation occurred between A and B, and the accuracy rates increased gradually thereafter.

In more detail with respect to the two negative markers ðen and min, table 6 below presents children’s accurate use rates of each of the two negative markers separately.

Table 6. Accurate Use of Negative Markers (GA)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Group</th>
<th>No</th>
<th>Mean</th>
<th>StdD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target ðen</td>
<td>A</td>
<td>89/93</td>
<td>95.70%</td>
<td>19.92</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>184/184</td>
<td>100%</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>215/215</td>
<td>100%</td>
<td>0.00</td>
</tr>
<tr>
<td>Target min</td>
<td>A</td>
<td>54/101</td>
<td>53.47%</td>
<td>46.85</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>166/213</td>
<td>77.93%</td>
<td>37.88</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>209/249</td>
<td>83.94%</td>
<td>29.49</td>
</tr>
</tbody>
</table>

1 The total number of negative test contexts was 570 in each group. Missing cases correspond to instances of isolated wh and irrelevant questions.
2 GA= Group Analysis
3 StdD= Standard Deviation

1 The total number of test contexts was 270 in the first condition and 300 in the second condition in each group. Missing cases correspond to instances of isolated wh and irrelevant questions.
2 GA= Group Analysis
3 StdD= Standard Deviation
In the production of target negative questions, the accurate use of target *den* was very high in all groups, as opposed to the accurate use of target *min* that was quite low in the youngest group but increased gradually with age. As for StdD scores, they were high throughout the accurate use of target *min*, and especially in the youngest group. This indicates that the youngest children were highly inconsistent in their use of target *min*. The following figure illustrates the accurate use of target *den* and target *min* across groups.

![Figure 4. Accurate Use of Target *den* vs. Target *min* in Negative Question Production](image)

The main effect of type of negative marker within subjects was significant \( [F(1,82)=35.225, p=.000] \), but the effect of its interaction with group was not \( [F(2,82)=2.851, p=.064] \). Yet, the main between-subjects effect of group was found to be significant \( [F(2,82)=6.511, p=.002] \). This pattern indicates that there is some heterogeneity between groups in terms of their use of the two negative markers *den* and *min*, which will be elaborated further in the individual-subject analysis section. Pairwise, within-group comparisons revealed that accurate use of target *den* was significantly better than accurate use of target *min* in all groups (A: \( p=.000 \), B: \( p=.004 \), C: \( p=.038 \)). As for the between-group analysis, in the accurate use of target *den* no significantly different comparisons were observed as its rates were very high from the youngest test age. In the accurate use of target *min* the significant comparisons attested were the
A-B (p=.02) and A–C (p=.004) ones, which point to a significant burst between groups A and B.

As for instances of inaccurate use of negative markers, these correspond in the first condition to use of min instead of target δen and in the second condition to use of δen instead of target min. Here are examples of the inaccurate use of both types.

inaccurate min

(28) Pjos na min pai na voιθisi ti mama?  
who to not go-3SG to help-3SG the mum  
(target: Pjos δen pai na voιθisi ti mama? (‘Who doesn't go to help mum?’) (A: 1)

inaccurate δen

(29) Pjo δen έa pai ja poδilato?  
which not will go-3SG for bicycle  
(target: Pjo peδi na min pai ja poδilato? (‘Which child should not go for a ride?’) (A: 1)

Irrespective of the negative marker used, it is interesting to mention, as a final note, that the negative questions produced by the Greek children in the present study were grammatical in their overwhelming majority. Instances of negation doubling were not attested, while instances of verb doubling were scarce and amounted to only six across groups: five in group A and one in group C; as a matter of fact, all these six instances involved doubling not only of the verb but also of the wh-element. The rest of the negative questions produced across groups were grammatical, with two of them actually instantiating the use of a cleft construction. A detailed discussion of the lack of negation and verb doubling phenomena in the produced negative questions will be provided in section 6.2 below.

On the whole with regard to prediction 2, negation effects were present in the child grammar from the youngest test age; accurate use of target δen was more prominent than accurate use of target min for marking of negation. In other words, prediction 2 is fully confirmed.

5.2. Individual-Subject Analysis

In this section, an individual-subject analysis will be presented of the above group results. On a par with a common tradition in experimental studies (e.g. Griffin et al. 2006, Montrul and Bowles 2008, O’Grady 2000), this analysis was motivated by the high StdD scores in certain measurements.

From a methodological point of view, certain specifications were made from the beginning so as to avoid random measurements. To begin with, in mea-
surements concerning the distinction between accurate and inaccurate performance, chance performance was established at the percentage of 50% out of the total. Below and above chance performance was then defined accordingly. Furthermore, another specification that was necessary to establish concerned all measurements that involved preference. A child’s preference for a certain pattern was recorded when this pattern was observed in 2/3 of the child's total number of utterances, i.e. at a percentage of 66.67%. Chance performance in these preference cases indicates that none of the elicited patterns amounted to 66.67% but that the child resorted to more than one of these patterns without showing preference for a specific one. On the whole, in both accurate/inaccurate and preference cases, a percentage was taken to denote consistent performance when it amounted to 66.67% (i.e. 2/3 of the total).

Taking these specifications into account, the presentation will proceed with the individual-subject results. Predictions 1 and 2 are examined in turn, while correlations with the corresponding group results are drawn at each stage. The aim is to provide a more fine-grained analysis of the results of the present study at an individual-subject level.

5.2.1. Prediction 1: Preference for Use of na

Group results revealed a limited number of produced LD questions overall. In target use of LD question production, target na seemed to be easier and more frequent to use than target oti in the two younger groups for the introduction of the embedded clause. In the oldest group, the frequency in use of target oti and target na was similar. These tendencies are reflected in the individual-subject results presented in table 7.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Performance</th>
<th>GROUP A</th>
<th></th>
<th>GROUP B</th>
<th></th>
<th>GROUP C</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>Mean</td>
<td>No</td>
<td>Mean</td>
<td>No</td>
<td>Mean</td>
</tr>
<tr>
<td>target oti</td>
<td>Chance</td>
<td>0/8</td>
<td>0 %</td>
<td>0/12</td>
<td>0 %</td>
<td>1/21</td>
<td>4.76%</td>
</tr>
<tr>
<td></td>
<td>Below chance</td>
<td>6/8</td>
<td>75 %</td>
<td>2/12</td>
<td>16.67%</td>
<td>3/21</td>
<td>14.29%</td>
</tr>
<tr>
<td></td>
<td>Above chance</td>
<td>2/8</td>
<td>25 %</td>
<td>10/12</td>
<td>83.33%</td>
<td>17/21</td>
<td>80.95%</td>
</tr>
<tr>
<td>target na</td>
<td>Chance</td>
<td>0/8</td>
<td>0 %</td>
<td>0/12</td>
<td>0 %</td>
<td>1/21</td>
<td>4.76%</td>
</tr>
<tr>
<td></td>
<td>Below chance</td>
<td>0/8</td>
<td>0 %</td>
<td>0/12</td>
<td>0 %</td>
<td>2/21</td>
<td>9.52%</td>
</tr>
<tr>
<td></td>
<td>Above chance</td>
<td>8/8</td>
<td>100 %</td>
<td>12/12</td>
<td>100 %</td>
<td>18/21</td>
<td>85.71%</td>
</tr>
</tbody>
</table>

1 The numbers of children refer to those children out of 30 in each group that produced accurate LD questions with target oti and target na.
2 ISA= Individual-Subject Analysis

The production of target oti was consistently below chance in the youngest group and consistently above chance for the two older groups. The production
of target *na*, however, was consistently above chance from group A onwards. Table 8 below provides an analysis at an individual-subject level of the error patterns to which children resorted in place of target *oti* and *na*.

Table 8. Inaccurate Performance in Target Long-Distance Question Production: Error Analysis (ISA)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Preferred Error Type</th>
<th>GROUP A</th>
<th>GROUP B</th>
<th>GROUP C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>Mean</td>
<td>No</td>
</tr>
<tr>
<td>target <em>oti</em></td>
<td>Na</td>
<td>2/7</td>
<td>28.57%</td>
<td>0/6</td>
</tr>
<tr>
<td></td>
<td>'pos'</td>
<td>0/7</td>
<td>0%</td>
<td>1/6</td>
</tr>
<tr>
<td></td>
<td>Null COMP</td>
<td>4/7</td>
<td>57.14%</td>
<td>2/6</td>
</tr>
<tr>
<td></td>
<td>Wh-element</td>
<td>0/7</td>
<td>0%</td>
<td>3/6</td>
</tr>
<tr>
<td></td>
<td>Chance performance</td>
<td>1/7</td>
<td>14.29%</td>
<td>0/6</td>
</tr>
<tr>
<td>target <em>na</em></td>
<td><em>Oti</em></td>
<td>–</td>
<td>–</td>
<td>1/1</td>
</tr>
<tr>
<td></td>
<td>Null COMP</td>
<td>–</td>
<td>–</td>
<td>0/1</td>
</tr>
<tr>
<td></td>
<td>Wh-element</td>
<td>–</td>
<td>–</td>
<td>0/1</td>
</tr>
<tr>
<td></td>
<td>Chance performance</td>
<td>–</td>
<td>–</td>
<td>0/1</td>
</tr>
</tbody>
</table>

1 The numbers of children in each group correspond to those children that produced LD questions in target LD contexts but did not always use the target complementiser *oti* or *na* to introduce the embedded clause.

2 ISA= Individual-Subject Analysis

In the first condition, the picture is not uniform across groups. Specifically, in the youngest group the most common error pattern was preference for a null complementiser. In group B use of a wh-element and a null complementiser in the embedded SpecCP were the mostly preferred patterns, while in group C preference for non-target *na* and for a wh-element predominated. In the second condition, however, the picture is much clearer. No child of the youngest group produced any error pattern, while only one showed preference for non-target *oti* in group B. Preference for non-target *oti* was also the most common error pattern in group C, with preference for a wh-element following.

As an additional note on the frequency of use of *na* as opposed to *oti*, group results revealed that in the overuse of LD questions, *na* was used to introduce the embedded clause rather than *oti*. The relevant individual-subject results in table 9 below show that all children within all groups unexceptionally showed consistent preference for *na*.

Table 9. Overuse of Long-Distance Question Production (ISA)

<table>
<thead>
<tr>
<th>Performance</th>
<th>GROUP A</th>
<th>GROUP B</th>
<th>GROUP C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
<td>Mean</td>
<td>No</td>
</tr>
<tr>
<td>Preference for <em>oti</em></td>
<td>0/28</td>
<td>0%</td>
<td>0/30</td>
</tr>
<tr>
<td>Preference for <em>na</em></td>
<td>28/28</td>
<td>100%</td>
<td>30/30</td>
</tr>
</tbody>
</table>

1 In groups A and C, 28 out of 30 children showed overuse of LD question production.

2 ISA= Individual-Subject Analysis
Although the preference for use of na over oti was clearly evident in the case of LD question production, the group analysis showed that in SD question production the higher frequency of na was not that evident. More concretely, in the set of the produced target SD questions, accurate use of target na was high but significantly lower than accurate use of target indicative; this difference compared to LD questions was attributed to the enriched modal semantics of na in matrix contexts (Pavlidou 1991, Rouchota 1993, Tsangalidis 2000) and to the occurrence of specific na-complement-taking matrix verb predicates in LD questions (Katis and Stampouliadou 2009). At an individual-subject level, the corresponding results from target SD questions are provided just below.

Table 10. Accurate Use of Short-Distance Question Production (ISA)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Performance</th>
<th>GROUP A</th>
<th>GROUP B</th>
<th>GROUP C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>Mean (%)</td>
<td>No</td>
</tr>
<tr>
<td>Target indicative</td>
<td>Chance</td>
<td>0/30</td>
<td>0%</td>
<td>0/30</td>
</tr>
<tr>
<td></td>
<td>Below chance</td>
<td>0/30</td>
<td>0%</td>
<td>0/30</td>
</tr>
<tr>
<td></td>
<td>Above chance</td>
<td>30/30</td>
<td>100%</td>
<td>30/30</td>
</tr>
<tr>
<td>Target na</td>
<td>Chance</td>
<td>0/30</td>
<td>0%</td>
<td>0/30</td>
</tr>
<tr>
<td></td>
<td>Below chance</td>
<td>4/30</td>
<td>13.33%</td>
<td>3/30</td>
</tr>
<tr>
<td></td>
<td>Above chance</td>
<td>26/30</td>
<td>86.67%</td>
<td>27/30</td>
</tr>
</tbody>
</table>

1 ISA= Individual-Subject Analysis

In both conditions, all groups tested were consistently above chance in their performance with respect to use of accurate mood in target SD questions, and especially so in accurate use of target indicative. As a consequence, these findings are in line with the relevant group results.

On the whole, the individual-subject analysis showed that consistency in the use of target oti was evident in groups B and C. As for na, consistency in target use and overuse was attested in all three groups.

5.2.2. Prediction 2: Negation Effects

Prediction 2 was that negation effects will be present in children’s production of questions. The group results have shown that accurate performance was high in production of negation in target negative contexts. The corresponding individual-subject analysis is provided just below.
Table 11. Accurate Production of Target Negative Questions (ISA)

<table>
<thead>
<tr>
<th>Performance</th>
<th>GROUP A</th>
<th>GROUP B</th>
<th>GROUP C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
<td>Mean</td>
<td>No</td>
</tr>
<tr>
<td>Chance</td>
<td>3/29</td>
<td>10.34%</td>
<td>0/30</td>
</tr>
<tr>
<td>Below chance</td>
<td>7/29</td>
<td>24.14%</td>
<td>1/30</td>
</tr>
<tr>
<td>Above chance</td>
<td>19/29</td>
<td>65.52%</td>
<td>29/30</td>
</tr>
</tbody>
</table>

1 In group A, 29 out of 30 children were included in the measurement. These 29 children were those that produced negative questions; the remaining one was restricted to the production of isolated wh and irrelevant questions.

2 ISA= Individual-Subject Analysis

A clear pattern is observed: the vast majority of children in each group tested were above chance with regard to accurate performance; actually, only the youngest children are slightly more random in their performance, but still, however, the percentage of above-chance performers is marginally consistent (i.e. 65.52%). Thus it can be concluded that, on the whole, children of all ages performed consistently well in the production of negation in target negative questions.

As for a more detailed examination of the production of target negative questions, the group analysis revealed differences in the use of the two target negative markers den and min. Specifically, accurate use of target den was found to be significantly better than accurate use of target min in all groups; the former was very high throughout, while for the latter a significant burst was attested by group B. At an individual-subject level, the analysis of the corresponding data looks as follows.

Table 12. Accurate Use of Negative Markers (ISA)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Performance</th>
<th>GROUP A</th>
<th>GROUP B</th>
<th>GROUP C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
<td>Mean</td>
<td>No</td>
<td>Mean</td>
</tr>
<tr>
<td>Target den</td>
<td>Chance</td>
<td>0/25</td>
<td>0%</td>
<td>0/29</td>
</tr>
<tr>
<td></td>
<td>Below chance</td>
<td>0/25</td>
<td>0%</td>
<td>0/29</td>
</tr>
<tr>
<td></td>
<td>Above chance</td>
<td>25/25</td>
<td>100%</td>
<td>29/29</td>
</tr>
<tr>
<td>Target min</td>
<td>Chance</td>
<td>0/25</td>
<td>0%</td>
<td>0/29</td>
</tr>
<tr>
<td></td>
<td>Below chance</td>
<td>10/25</td>
<td>40%</td>
<td>6/29</td>
</tr>
<tr>
<td></td>
<td>Above chance</td>
<td>15/25</td>
<td>60%</td>
<td>23/29</td>
</tr>
</tbody>
</table>

1 The numbers of children reported here correspond to those children out of 30 in each group that produced negative questions with the target negative markers den and min in the first and in the second condition respectively.

2 ISA= Individual-Subject Analysis

From the above table it becomes evident that the accurate use of target den was consistently above chance from the youngest test age, and this pertained to all children across groups that produced den questions. As for target min, its
accurate use was again above chance in all groups, yet consistently only in B and C. However, even in these two groups above chance rates for target min were lower than the respective rates for target den. This pattern shows a higher degree of consistency in use of target den than in use of target min overall.

Summing up the individual-subject findings about prediction 2, it is clear that children of all ages showed consistent sensitivity to negation effects at the production level.

6. Discussion and Conclusions

In this section we will discuss the results of the main study along the two predictions stated in section 4. Comparisons will be drawn between the main study findings and cross-linguistic ones for every prediction, with the aim to detect potential similarities and/or differences between Greek and other child languages. In addition, a more general interpretation will be attempted of the Greek findings with implications about the broader domain of language acquisition.

6.1. Prediction 1: Preference for Use of na

As noted in the relevant analysis in section 5, LD question production was limited in all test groups. Children's produced LD questions were expected to contain embedded clauses introduced more frequently by na than by oti. This expectation was mainly based on the greater degree of pragmatic/semantic richness associated with na as compared to oti (Katis and Stampouliadou 2009, Mastropavlou and Tsimpi 2011).

The higher frequency in use of na as opposed to oti in LD questions was evident from two main sources. Firstly, in the set of the accurately produced target LD questions, questions with an embedded target na were more frequent than questions with a target oti across groups (significantly in groups A and B). As a matter of fact, accurate use of target na was high throughout, whereas accurate use of target oti increased significantly at a later stage, i.e. by group B. Secondly, in instances of overuse of LD question production, use of na predominated, with use of oti being limited to only three occurrences in the oldest group. The main underlying reason for the higher frequency of na in accurate use and overuse of LD questions may be its stronger entrenchment, namely its wider distribution not only as an element introducing an embedded clause but also as the marker of the subjunctive mood in the language (Katis and Stampouliadou 2009, Mastropavlou and Tsimpi 2011, Roussou 2000, 2006). In fact, the dual use of na as a clause-introducing element and as a subjunctive marker has been attested in early acquisition data (e.g. Katis and Stampouliadou 2009,
On the whole, then, prediction 1 was confirmed.

Before closing this section, it is important to mention that in target SD question production, accurate use of target indicative was found to be significantly higher than accurate use of target na in all groups. On the basis of this observation, it is clear that the choice of mood differed between SD questions and the embedded clause of LD questions, with na being the preferred choice in the latter but not in the former. As suggested in previous sections (e.g. 3.1, 5.1.1, 5.2.1), this difference may, on the one hand, be attributed to the employment by children of specific na-complement-taking matrix verbs in LD question production (Katis and Stampouliadou 2009). On the other hand, an additional reason for the difference between LD and SD questions in children’s preference for na may be the following. When used in LD questions, na is associated with semantic neutrality; still, its use in SD questions encodes modality, thus adding a note of desirability or potentiality to the question. On this ground, SD na questions carry pragmatically enriched semantics, thus being in semantic contrast with SD non-na questions (Pavlidou 1991, Rouchota 1993, Tsangalidis 2000). As a result, SD questions in the indicative encoded simpler semantics and were thus easier for children to produce than SD questions with na, while the reverse was attested for LD contexts.

6.2. Prediction 2: Negation Effects

In question production, the effect of negation was predicted to be evident through the non-avoidance of use of negation in target negative contexts; at a more fine-grained level, accurate use rates were expected to be higher for target ðen than for target min on the basis of modality considerations involved in the latter as opposed to the former.

The early Greek data showed that the presence of negation was generally not avoided in target negative questions, inasmuch as accurate performance rates had become consistently high in group B. A closer look to the production of negative markers revealed that accurate use of target ðen was significantly higher than that of target min in all groups, with the corresponding rates in the former amounting to consistently top performance from group Â and in the latter increasing significantly in group B. This lag between ðen and min is well justified on the ground that min is selected by the particle na while this is not the case with ðen. As noted in the previous section, na in matrix environments carries an extra semantic/pragmatic load of modality (Pavlidou 1991, Rouchota 1993, Tsangalidis 2000). Due to this extra load, min is more difficult and thus less preferred by children compared to ðen in order to mark negation.

Another point that deserves a comment is the fact that almost all negative questions produced by the Greek children were grammatical. At a crosslinguis-
tic level, early English data has shown that children’s negative questions often involve some type of doubling, like doubling of the auxiliary or doubling of the auxiliary along with negation (Guasti et al. 1995, Hiramatsu 2003, Thornton 1995). On the contrary, early Italian negative interrogatives are adult in form, with doubling occurrences never being attested (Guasti 1996). Given these two trends in the production of negative questions, early Greek seems to pattern together with Italian, since the only observed instances of doubling included wh and verb doubling and amounted to only six across groups. Drawing on the claim that children initially assume that negation must stay in a V-related projection (Guasti 1996, Guasti et al. 1995), it is argued here that Greek children, on a par with their Italian but unlike their English peers, seem to hypothesise correctly that the Neg-Criterion must be satisfied inside CP and not IP. Unlike English, all verbs can raise to C in Greek and Italian, and hence CP is the V-related projection where the negation feature is checked. For this reason, Greek children locate negation inside CP and produce adult-like negative interrogatives.

To summarise, the negation effects that were tested under prediction 2 were found to be prominent in all test ages.

6.3. Generalised Findings and Implications

In the previous section, it was shown that predictions 1 and 2 were confirmed to a great extent. In the present section, more generalised conclusions will be drawn about language acquisition in its broader sense.

Overall, the early Greek data presented in this paper shows that children seem to opt for the most economical choices during their productions. One piece of evidence converging to this conclusion is the following: children resorted to embedded *na* clauses in LD wh-questions but to indicative (non-*na*) questions in SD contexts. As explained in section 5.1.1, the produced LD questions involved mainly verb predicates subcategorizing for a *na* complement;

---

4) It is important to mention, however, that this correct hypothesis concerning negation refers here to Greek children in age groups that go well beyond the initial stages of language acquisition. In other words, it is not implied here that children make this correct hypothesis right from start of the acquisition process.

5) Contrary to what Greek children do, Italian children have to *raise* negation to CP in order to produce grammatical negative questions. This difference between Greek and Italian stems from the fact that negation occupies a higher structural position in the former compared to the latter language. That is, in Greek NegP is located between the two lower C heads C<sub>OP</sub> and C<sub>M</sub> within a split-CP domain (Roussou 2000) (see section 2.1 for more details), while in Italian negation is located within IP (Guasti 1996, Guasti et al. 1995). As a result of this difference, adult-like negative interrogatives in Greek, unlike Italian, do not involve raising to but merely retaining of negation inside CP.
besides, *na* is more strongly entrenched than *oti* (see Katis and Stampouli-adou 2009). In addition, it introduces semantically neutral verb complements in LD structures, while it carries an extra semantic load of modality in SD contexts (Pavlidou 1991, Rouchota 1993, Tsangalidis 2000). Therefore, it is clear that children resort to the semantically less complex choices, and this points to a semantically-based economy hierarchy.

Further confirmation of an economy-based parser at a semantic level comes from negative questions. The higher accurate use rates for target *ðen* than for target *min* are grounded on the fact that *min*, as opposed to *ðen*, occurs with *na*, which encodes enriched modal semantics in matrix environments (Pavlidou 1991, Rouchota 1993, Tsangalidis 2000). *ðen* is semantically simpler than *min* and hence more preferred by children.

Finally, even the attested instances of overuse of LD question production point to a lag in semantics during acquisition. As mentioned in section 5.1.1, the instances of LD overuse pertained mostly to target negative or to target *na* SD questions. The underlying trigger for this pattern seems to be lexicalization of modality. Children often chose to express the enriched modal semantics associated with negation and matrix *na* questions in a more analytic way that involved increase in the syntactic complexity of the produced question. This choice for greater syntactic complexity essentially contradicts the Derivational Complexity Hypothesis (Jakubowicz 2005), which assumes that working memory constraints always induce the production of derivationally simpler constructions. On the contrary, working memory constraints may also lead to semantically less packed constructions, as in the case at hand here.

To summarise, all these findings converge on a common point: in child grammar an economy-based hierarchy of choices seems to be in play. What is essential to underline is that, as suggested in the above discussion, semantic factors seem to affect the formation of this hierarchy. On the basis of this suggestion, it could be argued that a lag in the acquisition of the syntax-semantics mapping seems to be in play. Children may have problems not at the level of syntactic processing (as suggested by Jakubowicz’s 2005 Derivational Complexity Hypothesis) but at the level of semantic integration. In this respect, a continuity account of language acquisition is supported for syntactic computations (e.g. Hyams 1986, Pinker 1984), with maturation pertaining to semantically-based problems.

References


