Non-sentential utterances (NSUs), utterances that lack an overt verbal (more generally predicative) constituent, are common in adult speech. This paper presents the results of a corpus study of the emergence of certain classes of NSUs in child language, based primarily on data from the Manchester Corpus from CHILDES. Our principal finding is the \textit{late short query effect}: the main classes of non-sentential queries (NSQs) are acquired much later than non-sentential answers (NSAs). At a stage when the child has productive use of sentential queries, and has mastered elliptical declaratives and the polar lexemes ‘yes’ and ‘no’, non-sentential questions are virtually absent. This happens despite the fact that such questions are common in the speech of the child’s caregivers and that the contexts are ones which should facilitate the production of such NSUs. We argue that these results are intrinsically problematic for analyses of NSUs in terms of a single, generalized mechanism of phonological reduction, as standard in generative grammar. We show how to model this effect within an approach of \textit{dialogue-oriented constructionism}, wherein NSUs are grammatical words or constructions whose main predicate is a contextual parameter resolved in a manner akin to indexical terms, the relevant aspect of context being the discourse topic. We sketch an explanation for the order of acquisition of NSUs, based on the different timing of their emergence.

---

[1] Earlier versions of this paper were presented at the 2002 LAGB meeting in Manchester, at the 2003 Georgetown Roundtable in Washington and at CLRF’04 at Stanford, and at seminars in Durham, Essex, and Edinburgh. We wish to thank audiences there for many useful comments and others who have discussed issues related to the paper, in particular Bob Borsley, Penelope Brown, Harald Clahsen, Eve Clark, Robin Cooper, Raquel Fernández, Pat Healey, Ruth Kempson, Shalom Lappin, Elena Lieven, and David Schlangen. This work has been supported by grant number RES-000-23-0065 from the Economic and Social Research Council of the United Kingdom, by a Research Leave Grant AH/E50440X from the Arts and Humanities Research Council of the United Kingdom, and by the EU project Dealing with Uncertainty in Dialogue (DEAWU). Many thanks to Raquel Fernández and David Schlangen for corpus processing software and other computational assistance and to Marina Kolokode for help with an initial pilot study. We would also like to thank two anonymous reviewers for the \textit{Journal of Linguistics} for many useful suggestions, and Ewa Jaworska and an external copy-editor for the \textit{Journal of Linguistics} for their assiduous reading of the manuscript. Thanks also to Ewa for her help and patience in getting the paper to press.
on a notion which combines accessibility of contextual parameters and complexity of content construction.

1. Introduction

Language acquisition is often viewed as a process correlated with an increase in the mean length of utterance (MLU) by the child (going back to Brown 1973). Nonetheless, one of the important components of adult linguistic competence is actually the ability to use short utterances when context so allows. This competence is perhaps at its most pronounced in the case of Non-sentential Utterances. Non-sentential Utterances (NSUs) are intuitively complete utterances that lack a verbal (more generally predicative) constituent. NSUs include 'short answers', and reprise utterances used to acknowledge or request clarification of prior utterances. Examples of these are provided in boldface in (1):

(1) A: Wasn’t he refused the chair in Oxford?
    B: Who?
    A: Skeat. Wasn’t he refused
    B: That’s Meak.
    A: Oh Meak, yes. [London-Lund Corpus S.1.9, p. 245]

Estimates of the frequency of NSUs are somewhat variable, depending on the classificational criteria applied. de Waijer (2001) provides figures of 40%, 31%, and 30% respectively for the percentage of one-word utterances in the speech exchanged between adults and infant, adult and toddler, and among adults in a single Dutch-speaking family consisting of two adults, one toddler and one baby over two months. Fernández (2006) cites a figure of 9% for the percentage of utterances lacking a verbal predicate, based on random sampling from (by and large) adult speech in the British National Corpus (BNC; Burnard 2000).

All approaches to first language acquisition, even strongly nativist ones, acknowledge the importance of the input to which the child is exposed. Given this and the ubiquity of NSUs in the input, it is noteworthy that there is hardly any work on NSUs in the area of language acquisition. The few studies that do exist (e.g. Bloom, Rocissano & Hood 1976, McTear 1985, Ninio & Snow 1996) do not provide detailed analyses of their syntactic and semantic properties. This possibly reflects the fact that, with a few isolated

---

[2] Fernández does not include in her tally fragmentary utterances resulting from interruptions or overlaps and formulaic uses such as greetings and partings.
[3] Bloom et al. (1976) use the term linguistically contingent for a child utterance that exhibits, in their words, a 'structural continuity' with the preceding adult utterance. Such utterances can be either sentential (non-elliptical) or elliptical. (The latter are not
exceptions (e.g. Barton 1988, Ginzburg & Sag 2000, Merchant 2002, Schlangen 2003, Stainton & Elugardo 2004, Stainton 2005, Fernández 2006), there is a dearth of work even in the area of adult NSUs. Although the formal grammatical study of dialogue is an emerging field in theoretical and computational linguistics (see e.g. work presented at the SEMDIAL conferences (1997–) and the SIGDIAL conferences (2000–), and work in the EU project TRINDI), there has to date been no work applying the newly developed analytical tools of dialogue modelling to ontogenetic investigation.

In this paper we present the results of a corpus study of the emergence of certain classes of NSUs. The main source of data are four English speaking children; a smaller-scale study provides evidence from the speech of one Greek child. Our principal finding is the late short query effect: in both English and Greek subjects, the main classes of non-sentential queries (NSQs) are acquired much later than non-sentential answers (NSAs). At a stage when the child has productive use of sentential questions, and has mastered elliptical declaratives and the polar lexemes ‘yes’ and ‘no’, non-sentential queries are virtually absent. This happens despite the fact that such questions are common in the speech of the child’s caregivers and that the contexts are ones which should facilitate the production of such NSUs.

We will argue that these results have significant implications for the grammatical analysis of NSUs and, more generally, for the issue of how context needs to be integrated with grammar. We will argue that these results provide support for a particular approach to the explication of NSU constructions, dialogue-oriented constructionism (e.g. Ginzburg & Sag 2000, Ginzburg 2009) over the approach of underlying sententialism, adopted in most generative linguistic work on ellipsis, wherein NSUs are underlyingly canonical sentences that (under conditions of contextual recoverability) involve unpronounced material. (See e.g. Ross 1969, Morgan 1973, Chung, Ladusaw & McCloskey 1995, Merchant 2002.) In fact, we suggest that the late short query effect is problematic for any attempt to account for NSUs in terms of a generalized mechanism of phonological reduction (e.g. Schwarzschild 1999), be it syntactically or semantically driven (the latter as in accounts derived from higher-order unification (Dalrymple, Pereira & Shieber 1991, Pulman 1997).

The paper is structured as follows: in section 2, we briefly survey the distribution of NSU constructions in adult English and possible grammatical

systematically categorized, which makes it difficult to draw an explicit comparison with the NSU categories we posit in this paper.) The authors claim that ‘asking a linguistically contingent question appears to be more difficult than producing a linguistically contingent statement or asking a non-contingent question’ (i.e. a question that does not rely on information in a preceding adult utterance) (pp. 538–539). Though it is not clear from the data they cite whether their linguistically contingent questions do indeed include sluices and clarification queries, their overall conclusion is a priori consistent with the main findings of our paper.
explications thereof. In section 3, we describe results concerning the acquisition of NSUs, based on a corpus study of the Manchester Corpus of Child English. We discuss some possible pragmatic explanations for this data in section 4, explanations we believe do not fully explain the observed effect. We also briefly describe results concerning NSUs based on a corpus study of the Stephany Corpus of child Greek, which replicate closely the results for the Manchester Corpus. Finally, in section 5, we explain how the late Query Effect can be modelled grammatically and sketch an explanation based on a metric which combines accessibility of contextual parameters and complexity of content construction.

2. Adult NSUs in dialogue

<table>
<thead>
<tr>
<th>NSU class</th>
<th>Example</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plain Acknowledgement</td>
<td>A:  _  B: mmh</td>
<td>599</td>
</tr>
<tr>
<td>Short Answer</td>
<td>A: Who left? B: Bo</td>
<td>188</td>
</tr>
<tr>
<td>Affirmative Answer</td>
<td>A: Did Bo leave? B: Yes</td>
<td>105</td>
</tr>
<tr>
<td>Repeated Acknowledgement</td>
<td>A: Did Bo leave? B: Bo, hmm.</td>
<td>86</td>
</tr>
<tr>
<td>Clarification Ellipsis</td>
<td>A: Did Bo leave? B: Bo?</td>
<td>79</td>
</tr>
<tr>
<td>Rejection</td>
<td>A: Did Bo leave? B: No.</td>
<td>49</td>
</tr>
<tr>
<td>Factive Modifier</td>
<td>A: Bo left. B: Great!</td>
<td>27</td>
</tr>
<tr>
<td>Repeated Affirmative Answer</td>
<td>A: Did Bo leave? B: Bo, yes.</td>
<td>26</td>
</tr>
<tr>
<td>Helpful Rejection</td>
<td>A: Did Bo leave? B: No, Max.</td>
<td>24</td>
</tr>
<tr>
<td>Check Question</td>
<td>A: Bo isn’t here. Okay?</td>
<td>22</td>
</tr>
<tr>
<td>Filler</td>
<td>A: Did Bo ... B: leave?</td>
<td>18</td>
</tr>
<tr>
<td>Bare Modifier Phrase</td>
<td>A: Max left. B: Yesterday.</td>
<td>15</td>
</tr>
<tr>
<td>Propositional Modifier</td>
<td>A: Did Bo leave? B: Maybe.</td>
<td>11</td>
</tr>
<tr>
<td>Conjunction + Fragment</td>
<td>A: Bo left. B: And Max.</td>
<td>10</td>
</tr>
</tbody>
</table>

Total dataset 1283

Table 1

NSUs in a sub-corpus of the BNC

Fernández & Ginzburg (2002) developed a taxonomy of NSUs in the BNC. As reported in Fernández (2006), this taxonomy was used to manually classify a randomly selected section of 200 speaker-turns from 54 BNC files. The examined sub-corpus contains 14,315 sentences. A total of 1,299 NSUs were found. Of these, 1,283 were successfully labelled according to the taxonomy, as illustrated below, the remaining 16 instances did not fall in any of the categories of the taxonomy and were labelled as ‘Other’. Table 1 shows...
Our focus in studying the emergence of NSUs will be on five subclasses of NSUs, \textit{short answers}, (plain and repeated) \textit{affirmative answers}, (plain and helpful) \textit{rejection answers}, \textit{clarification ellipsis}, and \textit{sluicing}, which constitute the majoritarian classes of NSUs used to query and answer. For subsequent comparative reference, table 2 lists the proportion of each class in this restricted sub-corpus.

<table>
<thead>
<tr>
<th>NSU class</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affirmative + Rejection</td>
<td>204 (41%)</td>
</tr>
<tr>
<td>Short Answer</td>
<td>188 (38%)</td>
</tr>
<tr>
<td>Clarification Ellipsis</td>
<td>79 (16%)</td>
</tr>
<tr>
<td>Sluice</td>
<td>24 (5%)</td>
</tr>
<tr>
<td><strong>Total dataset</strong></td>
<td><strong>495</strong></td>
</tr>
</tbody>
</table>

Table 2
Query/Answer NSUs

the distribution of NSU categories in the aforementioned sub-corpus of the BNC.

Our focus in studying the emergence of NSUs will be on five subclasses of NSUs, \textit{short answers}, (plain and repeated) \textit{affirmative answers}, (plain and helpful) \textit{rejection answers}, \textit{clarification ellipsis}, and \textit{sluicing}, which constitute the majoritarian classes of NSUs used to query and answer. For subsequent comparative reference, table 2 lists the proportion of each class in this restricted sub-corpus.

**Short answers**

We use this term to designate elliptical answer-conveying responses to \textit{wh}-questions, though the question can also potentially be elliptical, as in (2b):

(2) (a) A: Who’s that?
    B: My Aunty Peggy \textbf{[last or full name]}. My dad’s sister. [G58, 33–35]\footnote{This notation indicates the BNC file (G58) together with the sentence numbers (33–35).}
    (b) A: What’s plus three time plus three?
    B: Nine.
    A: Right. And minus three times minus three?
    B: Minus nine. [J91, 172–176]

**Affirmative + Rejection**

The typical context of this superclass is a polar question, as in (3a). We include also answers to \textit{implicit} polar questions, e.g. clarification ellipsis questions like (3b) and responses to assertions (3c). We do not include acknowledgements or acceptances such as (3d):

(3) (a) A: Did you bring the book I told you?
    B: Yes./No.
    (b) A: That one?
    B: Yeah. [G4K, 106–107]
We mention below some additional facts about the distribution and use of sluicing and clarification ellipsis, the two NSU types whose delayed emergence will be the central topic of the paper.

**Sluicing**

As indicated in table 3, sluices make up approx. 33% of all *wh*-interrogative queries in the BNC. Although there is a large generative literature on sluicing, it has almost without exception ignored the fact that bare *wh*-phrases are systematically ambiguous between at least three uses. In such cases the antecedent of the sluice is a *wh*-phrase:

(i)  *Larna: We’re gonna find poison apple and I know where that one is.*

   *Charlotte: Where? [KD1, 2371]*

We will not discuss this significantly rarer use here; see Ginzburg (2009) for an account.

---

**Table 3**

<table>
<thead>
<tr>
<th></th>
<th>what</th>
<th>why</th>
<th>who</th>
<th>where</th>
<th>when</th>
<th>which</th>
<th>how</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sluice</td>
<td>3045</td>
<td>1125</td>
<td>491</td>
<td>350</td>
<td>107</td>
<td>175</td>
<td>50</td>
<td>5343</td>
</tr>
<tr>
<td>Sentential interrogative</td>
<td>7965</td>
<td>1328</td>
<td>1696</td>
<td>1877</td>
<td>420</td>
<td>225</td>
<td>1640</td>
<td>15151</td>
</tr>
</tbody>
</table>

---

[5] These figures need to be treated with some caution, deriving as they do from an automated search using a small number of templates in a search engine, Matt Purver’s SCoRE (Purver 2001).

[6] The examples cited here are from Fernández, Ginzburg & Lappin (2007), who show that this ambiguity is reliably coded by human subjects and develop an automatic classifier for the ambiguity, using machine learning techniques. Fernández et al. (2007), in fact, suggest the existence of a fourth use, which they dub the *wh*-anaphoric use. In such cases the antecedent of the sluice is a *wh*-phrase:
ANSWERS WITHOUT QUESTIONS

(4) (a) A: Can I have some toast please?
    B: Which sort? [KCH, 104–105]
(b) Caroline: I’m leaving this school.
    Lyne: When? [KP3, 538]

Reprise

The utterer of the sluice cannot understand some aspect of the previous utterance which the previous speaker assumed as presupposed:

(5) (a) Geoffrey: What a useless fairy he was.
    Susan: Who? [KCT, 1753]
(b) Pat: You might find something in there actually.
    Carole: Where? [KBH, 1817]

Repetition

The sluice is used to ask for repetition of the previous utterance as a whole.

(6) June: Only wanted a couple weeks.
    Ada: What?
    June: Only wanted a couple weeks. [KB1, 3312]

2.1.1 Clarification ellipsis (CE)

This category is employed to classify reprise fragments which are used to clarify an utterance that has not been fully comprehended. Clarification requests (CRs) constitute approx. 4% of all dialogue moves in the BNC (Purver, Ginzburg & Healey 2001) and this proportion is somewhat higher in task-oriented dialogue – Rodriguez & Schlangen (2004) cite a figure of 5.8% for the Bielefeld Corpus dealing with model plane construction. Clarification ellipses make up approx. 29% of all clarification requests in the BNC (Purver et al. 2001). As with sluicing, CEs are ambiguous between a number of readings, the two most prominent being the clausal confirmation reading and the intended content or constituent reading, illustrated in (7):

(7) (a) Marsha: Yeah that’s it, this, she’s got three rottweilers now and ...
    Sarah: Three? (=clausal confirmation reading: Are you saying she’s got THREE rottweilers now?)
    Marsha: yeah, one died so only got three now.
(b) Tim: Those pink things that af after we had our lunch.
    Dorothy: Pink things? (=intended content reading: What do you mean by saying ‘Pink things’?)
    Tim: Yeah. Er those things in that bottle.
2.2 Explicating NSUs: with and without constructions

How should NSUs be incorporated into grammatical analysis? This depends to a large extent on how one resolves the issue of whether NSUs are to be assimilated to some other grammatical phenomenon such as phonological reduction or anaphora, or whether they are actually in some significant way *sui generis*. Theories that follow the former route we will dub *UNITARIAN* theories; in such theories ellipsis resolution is associated with a single, typically extra-grammatical mechanism. Theories that follow the latter route are *CONSTRUCTIONIST* theories, in which NSUs are incorporated in the grammar as distinct constructions, which specify among other things the contextual characteristics which govern their use.

Within the unitarian approach is included the commonest generative approach, which we dub *UNDERLYING* (NSU) *SENTENTIALISM*. On this view NSUs such as sluicing and short answers are analyzed as underlyingly *SENTENTIAL* constructions. (For sluicing see e.g. Ross 1969, Chung et al. 1995, Merchant 2002; for short answers see Morgan 1973, Merchant 2004.) On this view, which has attracted a variety of implementations, the grammar does not need to reify NSUs as such – no new rules or principles need to be posited for individual NSU constructions. A non-grammatical module serves – under appropriate conditions – to recover material that is phonologically redundant and remains unpronounced. This type of approach accounts for parallelism phenomena that are exhibited by NSUs, such as the case parallelism between source and target pointed out by Ross for sluicing and by Morgan for short answers, using grammatical principles that apply to non-elliptical sentences.

A significantly distinct unitarian approach to the description of ellipsis constructions is the higher-order unification (HOU) approach pioneered in Dalrymple et al. (1991) for VP ellipsis, and extended to short answers in Pulman (1997) and to corrections in Gardent, Kohlhase & van Leusen (1996). On this view an ellipsis construction involves the recovery of a property from context such that when it is applied to the content of the ellipsis target, it resolves the interpretation of the elliptical construction. The property is recovered from context by using higher order unification to solve an equation that relates the target with its antecedent. In contrast to *UNDERLYING SENTENTIALISM*, HOU approaches make no appeal to syntactic spelling out of linguistic material. However, in common with *UNDERLYING SENTENTIALISM*, NSUs are viewed as arising from the application of a single generalized extra-grammatical mechanism. The HOU-based account offers a single mechanism that uniformly accounts for VP ellipsis, focus constructions and certain NSU constructions such as short answers and corrections. Another

---

[7] Stainton (2005) argues that, in fact, an approach such as that of Merchant (2004) does end up introducing new devices into the grammar to account for the data.
commonality is the absence of an explicit spelling out of an interface between grammar and context: neither account offers an explicit proposal about how the information required for contextual resolution actually emerges in context, how it is accessed, etc.

An alternative to such unitarian approaches that has emerged in recent years, inspired by developments in construction grammar (see e.g. Sag 1997, Fillmore & Kay 1999) and in the modelling of dialogue context (Ginzburg 1996, 2009; Larsson 2002; Asher & Lascarides 2003) is the approach of DIALOGUE-ORIENTED CONSTRUCTIONISM (see e.g. Ginzburg & Sag 2000, Schlangen 2003, Ginzburg & Cooper 2004, Purver 2004, Fernández 2006). On this view NSUs are grammatical words (e.g. ‘yes’, ‘no’, ‘pardon’) or constructions (e.g. short answers, sluicing), where semantic composition is driven primarily by context rather than syntax. Concretely, in such words or constructions the main predicate is a contextual parameter resolved in a manner akin to indexical terms (‘I’, ‘you’, ‘here’, ‘now’) – in this case the relevant aspect of context is the discourse topic, known also as the maximal element in the contextual repository of Questions-Under-Discussion (QUD; see Van Kuppevelt 1995, Ginzburg 1996, Roberts 1996). One motivation for pursuing a dialogue-oriented constructionist approach is the existence of various syntactic and semantic divergences between NSUs and their putative sentential correlates. (For such data concerning short answers and sluicing, see Ginzburg 1999, Ginzburg & Sag 2000, Stainton 2005; concerning CE see Ginzburg & Copper 2004; for some counterarguments concerning short answers and sluicing see Merchant 2004, with further arguments pro in Ginzburg 2009). A particularly strong case in this respect can be built for NSUs oriented to metacommunicative uses such as CE. A syntactically sententialist analysis of the INTENDED CONTENT reading of (8) is difficult to conceive of, whereas an analysis of the CLAUSAL CONFIRMATION reading would presuppose some version of the performative hypothesis (Levinson 1983):

(8) A: Did Bo leave?
   B: Bo? (=Who do you mean ‘Bo’? Or: Are you asking if BO of all people left?)

In similar vein, Ginzburg & Cooper (2004) point out that an HOU-based analysis can be set up to derive the CLAUSAL CONFIRMATION reading of CE, but there is no apparent way of accommodating the INTENDED CONTENT reading.

Incorporating dialogue context, including the context that arises in meta-communicative interactions such as grounding (Clark 1996) and clarification

[8] Schlangen’s account, formulated in Segmented Discourse Representation Theory (SDRT; Asher & Lascarides 2003) differs somewhat from the other accounts mentioned here – it does not appeal to QUD in resolution, nor is context integrated into the grammar as such: NSU resolution, in common with anaphora resolution, is effected as a consequence of the general process of inference used to compute the rhetorical relation which connects an NSU to its antecedent.
interaction (Schegloff, Jefferson & Sacks 1977, Ginzburg & Cooper 2004), in the grammatical component has enabled the development of analyses which cover the majority of classes of NSUs postulated in the taxonomy of Fernández & Ginzburg (2002). Whether unitarian approaches like UNDERLYING SENTENTIALISM and HOU can scale up to cover these phenomena, particularly metacommunicative interaction-oriented NSUs, remains an open issue.

We turn now to consider some patterns of the acquisition of NSUs, which could offer an additional means of choosing between unitarian and dialogue-oriented constructionist approaches to NSUs. Before actually doing so, we note one key entailment of unitarian approaches to NSUs. This is stated in (9) in two slightly different versions:

(9) PUNCTUAL NSU emergence:

(a) UNDERLYING SENTENTIALISM version: Once a canonical (sentential) structure is acquired (declarative or interrogative sentential structures) and a mechanism for phonological elision exists, the corresponding NSUs (short answers, sluicing) are available to the grammatical system.

(b) HIGHER-ORDER UNIFICATION version: Once the relevant semantic content types for antecedents have been acquired (wh-questions, quantified propositions) and the (higher order unification) resolution mechanism exists, the corresponding NSUs (short answers, sluicing, CE) are all available to the grammatical system.

Whereas PUNCTUAL NSU emergence seems to be an intrinsic characteristic of both unitarian approaches, albeit in subtly distinct versions, it is not expected and would, therefore, be difficult to account for within DIALOGUE-ORIENTED CONSTRUCTIONISM. We will argue below that there are some indications that PUNCTUAL NSU emergence in fact fails. This means that neither unitarian approach can be used to formulate a theory of the acquisition of NSUs. In contrast, we will use a version of DIALOGUE-ORIENTED CONSTRUCTIONISM to sketch such an account.

3. AN ASYMMETRY IN THE ACQUISITION OF NSUS

3.1 NSU distribution for two-year-olds

We conducted a study of NSUs in child English based on the Manchester Corpus in CHILDES (Theakston et al. 2001, Roland et al. 2003), a longitudinal study of 12 monolingual English-speaking children between the ages...
of approximately two and three years. The children were engaged in normal
game activities with their mothers at home and the hourly (audio) recordings
took place on two separate occasions every three weeks. Our data comes
primarily from the manual coding of a sub-corpus of the Manchester Corpus
consisting of 25,060 turns (111,969 words). We studied the transcript files of
four children: Aran, Becky, Carl, and Dominic, including at least six files per
child (average size of file: 830 turns). The manual coding initially involved six
files per child (Aran: files 12b, 13b, 16b, 18b, 20b, 22b covering age span
2;3–2;6; Becky: files 12b, 14b, 16b, 18b, 20b, 22b covering age span 2;3–2;7;
Carl: files 16b, 20a, 25a, 28a, 31a, 34b covering age span 2;1–2;8; Dominic:
files 16a, 20a, 24b, 28a, 32a, 34a covering age span 2;4–2;10.). In each file the
following categories were manually tagged and their numbers tallied:
sentential answers (declaratives used to answer questions), sentential interrogatives (wh-interrogatives, polar and intonation interrogatives) and the
two classes of NSUs listed in table 2 above.

The period covered by our discussion in the current section is one where
three of the four children (Aran, Becky, Carl) were at Brown’s stage III
(MLU 2.5–2.99) during the first half of this period and attained stage IV/V
(MLU 3+) subsequently; Dominic correspondingly was at stage II (MLU
2-2.49) and subsequently attained stage III (see table 2 in Rowland et al.
2003). As table 4 demonstrates, there is significant evidence that short

<table>
<thead>
<tr>
<th>Subject</th>
<th>Period</th>
<th>Declarative yes/no</th>
<th>Total Non-sentential</th>
<th>Total Answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aran</td>
<td>2;3–2;6</td>
<td>125 (29%)</td>
<td>207</td>
<td>307 (71%)</td>
</tr>
<tr>
<td>Becky</td>
<td>2;3–2;7</td>
<td>88 (25%)</td>
<td>210</td>
<td>265 (75%)</td>
</tr>
<tr>
<td>Carl</td>
<td>2;1–2;8</td>
<td>99 (50%)</td>
<td>41</td>
<td>99 (50%)</td>
</tr>
<tr>
<td>Dominic</td>
<td>2;4–2;10</td>
<td>85 (25.5%)</td>
<td>130</td>
<td>238 (74.5%)</td>
</tr>
</tbody>
</table>

Table 4
Distribution of senential/non-sentential answers for four children aged 2;1–2;10 in the Manchester Corpus

[10] The longer time span for Carl and Dominic is due to the fact that they exhibit a less
advanced stage of acquisition than Aran and Becky. Aran and Becky were studied more
closely, as discussed in section 4, with an additional set of three files each manually coded,
covering the period up to 2;10.

[11] The coding was performed by the two authors: all of Aran and Becky’s files were coded by
both authors using the annotation tool MMAX (see http://mmax.eml-research.de; Müller
& Strube 2001). We calculated \( k \) for one file per child: (aran16b: \( k = 0.85 \), becky16b: \( k = 0.82 \)). Carl and Dominic’s files were coded by the second author on two separate oc-
casions, separated by more than one year. Disagreements were checked and reconciled. All
manually coded data cited here is the reconciled data.
answers are by now acquired by all four subjects. Three of the four subjects show a preference for using NSUs over sentential answers, though the sentential option is clearly a robust option.

To exemplify this, (1) contains data from a single subject at 2;4 and 2;5, showing his mastery of short answers, which he uses with a variety of distinct wh-phrase antecedents, as well as yes/no; (2) exemplifies some sentential answers.

(10) (a) Mother: He’s been eating the what? Aran: the bad food. (aran20b)
(b) Mother: Who ate it? Aran: Pippin. (aran20b)
(c) Mother: Where is it? Aran: In the box. (aran20b)
(d) Mother: What colour is that one? Aran: Green. (aran18b)
(e) Mother: What is it? Aran: A hen. (aran16b)
(f) Mother: Did you have one of these this morning? Aran: Yeah.
(g) Mother: Shall we move it? Aran: No. (aran20b)

(11) (a) Investigator: What is it? Aran: That is Bumbo. (aran18b)
(b) Mother: Where is the other shoe? Aran: It is lost. (aran16b)
(c) Mother: What have you got there? Aran: I have got a hair in my mouth. (aran16b)
(d) Mother: Oh is he climbing right to the top? Aran: He can’t. (aran20b)

The data for sentential/non-sentential queries is entirely different: all subjects throughout this period use sentential queries productively, primarily wh-interrogatives. According to the criteria of Rowland et al. (2003), all four children had acquired the wh-words ‘what’ and ‘where’ by the beginning of the period we studied, and two of the children (Aran and Becky) had also acquired ‘who’. Some examples from Aran are given in (12):

(12) (a) Where is Dolly? (aran16b)
(b) What is that? (aran16b)
(c) What did you have this morning? (aran20b)
(d) What does he eat? (aran20b)

Nonetheless, as table 5 reveals, for three of the four subjects NSQs are almost entirely absent. The different sentential/non-sentential distributions for answers vs. questions is of high statistical significance (Aran: \( p \leq .0001, \chi^2 = 86 \); Becky: \( p \leq .0001, \chi^2 = 241 \)).

To what extent does this distribution reflect the input of NSUs available to the children? We manually coded the input provided by the children’s mothers, based on the first and last file in the examined period (files 12b, 22b).

---

[12] We should emphasize that by NSQs in this paper we refer specifically only to sluicing and CE. We make no claims about other kinds of NSUs which are semantically questions.
for Aran and Becky; 16b, 34b for Carl; 16a, 34a for Dominic). Tables 6 and 7 provide average figures per conversation for mothers and children. One obvious difference, a clear consequence of the fact that the mothers are in charge of keeping the conversation moving along, is that the mothers ask significantly more than they answer. Two of the mothers’ answers are less NSU-oriented than the children’s, Becky’s and Aran’s, though this is significant only in the case of Becky ($\chi^2 = 10.53, p < .001$), not Aran ($\chi^2 = 1.25$, two-tailed $p = .72$). With Dominic and Carl’s mothers the answer distribution is a bit harder to gauge given their sparseness, in both sentential and non-sentential forms.

However, strikingly, for questions there is what seems to be a categorical difference: all the mothers use NSQs productively, whereas NSQs are almost

### Table 5

Distribution of sentential/non-sentential questions for four children in the Manchester Corpus (same period as table 4)

<table>
<thead>
<tr>
<th>Subject</th>
<th>Sentential</th>
<th>Total</th>
<th>Sentential</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-sentential</td>
<td>Answers</td>
<td>Non-sentential</td>
<td>Answers</td>
</tr>
<tr>
<td>Aran</td>
<td>49 (96%)</td>
<td>51</td>
<td>2 (4%)</td>
<td>20</td>
</tr>
<tr>
<td>Becky</td>
<td>204 (92%)</td>
<td>222</td>
<td>14 (8%)</td>
<td>54</td>
</tr>
<tr>
<td>Carl</td>
<td>52 (96%)</td>
<td>54</td>
<td>2 (4%)</td>
<td>20</td>
</tr>
<tr>
<td>Dominic</td>
<td>17 (85%)</td>
<td>20</td>
<td>2 (15%)</td>
<td>20</td>
</tr>
</tbody>
</table>

### Table 6

Averaged distribution for answers per conversation of four children and their mothers in the Manchester Corpus (average over six conversations for the children, two for the mothers)

<table>
<thead>
<tr>
<th>Subject</th>
<th>Total</th>
<th>Total</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Declarative</td>
<td>yes/no</td>
<td>Short-answer</td>
</tr>
<tr>
<td>Aran</td>
<td>51 (71%)</td>
<td>72</td>
<td>6 (29%)</td>
</tr>
<tr>
<td>A’s mother</td>
<td>10 (62%)</td>
<td>16</td>
<td>6 (38%)</td>
</tr>
<tr>
<td>Becky</td>
<td>9 (76%)</td>
<td>59</td>
<td>15 (25%)</td>
</tr>
<tr>
<td>B’s mother</td>
<td>12 (38%)</td>
<td>32</td>
<td>20 (62%)</td>
</tr>
<tr>
<td>Carl</td>
<td>17 (52%)</td>
<td>33</td>
<td>16 (48%)</td>
</tr>
<tr>
<td>C’s mother</td>
<td>5 (63%)</td>
<td>8</td>
<td>3 (37%)</td>
</tr>
<tr>
<td>Dominic</td>
<td>40 (74%)</td>
<td>54</td>
<td>14 (26%)</td>
</tr>
<tr>
<td>D’s mother</td>
<td>12 (93%)</td>
<td>13</td>
<td>1 (7%)</td>
</tr>
</tbody>
</table>
entirely absent from Aran’s, Carl’s and Dominic’s productions and rare for Becky (statistical significance for Becky vs. her mother: $\chi^2 = 2.55$, two-tailed $p = .07$).

Table 8 shows the average NSU production in these conversations. It reveals that in fact all the mothers deviate quite substantially from the distribution evinced in our sub-corpus of the BNC (table 2 above), producing substantially higher proportions of NSQs.

Table 7
Averaged distribution for the questions per conversation of four children and their mothers in the Manchester Corpus (average over six conversations for the children, two for the mothers)

<table>
<thead>
<tr>
<th>Subject</th>
<th>wh</th>
<th>Polar</th>
<th>Total</th>
<th>Sentential</th>
<th>Sluice</th>
<th>CE</th>
<th>Non-sentential</th>
<th>Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aran</td>
<td>8</td>
<td>1</td>
<td>9 (100%)</td>
<td>0</td>
<td>0</td>
<td>0 (0%)</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>A’s mother</td>
<td>40</td>
<td>68</td>
<td>108 (86%)</td>
<td>8</td>
<td>10</td>
<td>18 (14%)</td>
<td>126</td>
<td></td>
</tr>
<tr>
<td>Becky</td>
<td>18</td>
<td>16</td>
<td>34 (92%)</td>
<td>2</td>
<td>1</td>
<td>3 (8%)</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>B’s mother</td>
<td>27</td>
<td>44</td>
<td>71 (78%)</td>
<td>6</td>
<td>14</td>
<td>20 (22%)</td>
<td>91</td>
<td></td>
</tr>
<tr>
<td>Carl</td>
<td>8</td>
<td>1</td>
<td>9 (100%)</td>
<td>0</td>
<td>0</td>
<td>0 (0%)</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>C’s mother</td>
<td>30</td>
<td>17</td>
<td>47 (83%)</td>
<td>4</td>
<td>6</td>
<td>10 (17%)</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>Dominic</td>
<td>2</td>
<td>0</td>
<td>2 (100%)</td>
<td>0</td>
<td>0</td>
<td>0 (0%)</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>D’s mother</td>
<td>37</td>
<td>31</td>
<td>68 (64%)</td>
<td>25</td>
<td>13</td>
<td>38 (36%)</td>
<td>106</td>
<td></td>
</tr>
</tbody>
</table>

Table 8
Averaged distribution of NSUs per conversation for four children and their mothers in the Manchester Corpus, compared with proportion in BNC sub-corpus (average over six conversations for the children, two for the mothers)

<table>
<thead>
<tr>
<th>Subject</th>
<th>yes/no</th>
<th>Short-answer</th>
<th>Sluice</th>
<th>CE</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aran</td>
<td>34 (67%)</td>
<td>17 (32%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>51</td>
</tr>
<tr>
<td>A’s mother</td>
<td>8 (29%)</td>
<td>2 (7%)</td>
<td>8 (29%)</td>
<td>10 (35%)</td>
<td>28</td>
</tr>
<tr>
<td>Becky</td>
<td>35 (75%)</td>
<td>9 (19%)</td>
<td>2 (5%)</td>
<td>1 (1%)</td>
<td>47</td>
</tr>
<tr>
<td>B’s mother</td>
<td>8 (25%)</td>
<td>4 (12%)</td>
<td>6 (19%)</td>
<td>14 (44%)</td>
<td>32</td>
</tr>
<tr>
<td>Carl</td>
<td>7 (41%)</td>
<td>10 (59%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>17</td>
</tr>
<tr>
<td>C’s mother</td>
<td>3 (20%)</td>
<td>2 (13%)</td>
<td>4 (27%)</td>
<td>6 (40%)</td>
<td>15</td>
</tr>
<tr>
<td>Dominic</td>
<td>22 (55%)</td>
<td>18 (45%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>40</td>
</tr>
<tr>
<td>D’s mother</td>
<td>11 (22%)</td>
<td>1 (2%)</td>
<td>25 (50%)</td>
<td>13 (26%)</td>
<td>50</td>
</tr>
<tr>
<td>BNC pct</td>
<td>41%</td>
<td>38%</td>
<td>5%</td>
<td>16%</td>
<td>100%</td>
</tr>
</tbody>
</table>

654
3.2 Interim summary: two generalizations about NSU emergence

We can summarize the NSU data from the four children aged 2–2;10 in the Manchester Corpus as follows:

(13) (a) Empirical Generalization 1: With short answers and sentential interrogatives acquired, non-sentential queries (sluicing and clarification ellipsis) are, on the whole, still absent.

(b) Empirical Generalization 2: NSQs are common in the speech of the child’s caregivers.

In what follows, we will refer to (13a), Empirical Generalization 1, as the LATE SHORT QUERY EFFECT, given that it indicates that NSQs are delayed with respect to a system that already provides for declarative NSUs and sentential interrogatives. The LATE SHORT QUERY EFFECT is, as the reader may already suspect, our evidence against PUNCTUAL NSU EMERGENCE. Significant work remains, however, in order to demonstrate that the LATE SHORT QUERY EFFECT is not merely epiphenomenal.

4. Establishing the Late Short Query Effect

The LATE SHORT QUERY EFFECT can be explained in two ways, either pragmatically or mechanistically. A pragmatic explanation sees the absence of NSQs as a consequence of two factors: (a) the CONTEXTUAL UNAVAILABILITY of antecedents for NSQs and/or (b) the SEMANTIC UNDESIRABILITY of the contents that NSQs express in the type of conversational interaction recorded in this corpus. A mechanistic explanation ties the absence of NSQs to the absence of appropriate linguistic/conversational competence. In this section we attempt to show that a pragmatic explanation is not sufficient. We need, therefore, to refute the two components of the pragmatic explanation:

(14) CONTEXTUAL AVAILABILITY: Do children get the appropriate contexts for NSQs? Given an appropriate context, how likely is an NSQ to be produced?

(15) SEMANTIC DESIRABILITY: Do NSQs encode meanings children want to convey in the current setting?

SEMANTIC DESIRABILITY is less of an issue for direct sluices since these convey questions which on the whole are no different from ones typically conveyed by sentential queries. But SEMANTIC DESIRABILITY is an issue for NSUs expressing clarification requests, since prima facie it is not clear that children express clarification requests at this stage or, in principle, in the kind of interaction recorded in this corpus.

Thus, there are some clear empirical desiderata we need to establish in order to demonstrate the need for a mechanistic explanation of the LATE...
SHORT QUERY EFFECT. What is needed is evidence for at least one of the following:

(16) (a) The child ignores the presence of triggering context for NSQs. Moreover, this happens with a likelihood which is significantly distinct from the corresponding behaviour of adult speakers.

(b) The child expresses clarification requests using non-elliptical constructions.

The following two sub-sections provide the requisite evidence. The final subsection provides an additional supporting piece of evidence: data from a distinct corpus in a different language, the Stephany Corpus of Modern Greek, where the LATE SHORT QUERY EFFECT is replicated.

4.1 Ignoring triggering context

There is clear evidence for the existence of contexts in which the child could produce an NSQ but does not. For each child we manually searched six files (21a–23b, Aran), (20a–22b, Becky), and (32a–34b, Carl and Dominic), collating all questions for both child and mother. From these we categorized as a potential NSQ context any case where either an NSQ was produced or in which a phonologically reduced version of the question that was actually produced could have been realized as an NSQ without significantly increasing ambiguity. (17a) is an example of a potential but unrealized NSQ, whereas (17b) is a case we did not categorize as a potential NSQ since the resulting NSQ (‘What?’) would have been very ambiguous:


(b) Child: I can build a church. Mother: You want to build what? → What?

We start by discussing Aran and Becky’s data. Examples from potential but unrealized NSQ contexts are given in (18) and (19):

(18) (a) Mother: Well where is the ship? Aran: I don’t know where it is. (aran21a)

→ I don’t know where.

(b) Mother: Oh somebody has dressed dolly nicely, haven’t they?

Aran: Who did it? (aran23b)

→ Who?

(c) Mother: He has got a hat on as well, hasn’t he? Aran: Got a hat on.

Aran: What has he got a hat on for? (aran23a)

→ What for?
**ANSWERS WITHOUT QUESTIONS**

<table>
<thead>
<tr>
<th>Subject</th>
<th>Potential NSQ contexts</th>
<th>NSQs produced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aran</td>
<td>6</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>A’s mother</td>
<td>35</td>
<td>28 (80%)</td>
</tr>
<tr>
<td>Becky</td>
<td>16</td>
<td>4 (25%)</td>
</tr>
<tr>
<td>B’s mother</td>
<td>36</td>
<td>30 (83%)</td>
</tr>
</tbody>
</table>

**Table 9**

Rate of NSQ production given triggering context in files (21a–23b, Aran); (20a–22b, Becky)

(a) Mother: He is crying. Becky: Why is he crying? (becky20b)  → Why?

(b) Mother: I think you can put this one wherever you like, Becky in the picture. Mother: Put him next to that other mouse. Becky: Where does he go? (becky20b)  → Where?

(c) Mother: There is a bag of something there, Becky. Becky: Where is bag? (becky20b)  → Where?

(d) Mother: All gone. Becky: Are they all gone now? (becky20b)  → All gone?

(e) Mother: Somebody has made something already, look. Mother: Haven’t they? Becky: Who was that, Mum? (becky21b)  → Who?

(f) Mother: It goes anywhere. Becky: Does it go there? (becky 21b)  → There?

Note that the test used here is a conservative one that underestimates the potential for producing NSQs: it evaluates this potential merely from the pool of questions that were actually produced, ignoring other contexts where an NSQ could possibly have been produced. Moreover, it considers only the most direct phonologically reduced variant as a potential NSU in such a context. The results obtained are given in table 9. Aran avoids producing NSQs completely despite various opportunities; Becky’s rate of exploitation is non-zero, although at least two of the NSQs she produces are repetitions – we cannot be certain that these are genuine cases of CE, particularly in light of the fact that in later files, discussed further below, where she has clearly mastered NSQ production, she eschews this construction. In contrast, both mothers exploit their opportunities for NSQs with a probability higher than 50%.

[13] Here the distribution of Becky compared with her mother is statistically significant ($\chi^2 = 14.18$, $p \leq .0001$). With Aran vs. his mother, applying $\chi^2$ is not possible given the null
We also searched for unexploited triggering contexts in Carl and Dominic’s data. Given that they were generally far less advanced than Aran and Becky we used the final six files of the corpus (32a–34b). However, neither child provided evidence that would support or refute the existence of unexploited triggering contexts. Carl has three cases of repetition ‘What?’; one ‘What?’ which is simply used to indicate his presence, an effect that could equally well be attained by saying ‘Yes’; and one ‘Where?’, which might almost be dismissed as a false start. Dominic has one repetition ‘What?’ and one ‘What?’ used to indicate his presence. There are no other relevant sentential or non-sentential cases.\(^{14}\)

(b) Mother: Where did you go yesterday? Carl: What? (carl 32a)
(c) Mother: Well I’ll stick him under and you can dry him. There you go. Carl: Where? Where is he going? Mother: Here, Carl. Carl: Where? (carl 33a)
(d) Mother: Carl? Carl: What? (carl 33a)
(e) Mother: Where did you go yesterday? Carl: What? (carl 33a)
(g) Mother: Dominic. Dominic: What? (dominic 34a)

4.2 Semantic desirability

The data presented in the previous section makes it clear that the lack of NSQs is not a consequence of the unavailability of contexts in which an NSQ would be appropriate. The other component of an exclusively pragmatic explanation of the late short query effect is semantic desirability – whether the contents communicated by an NSQ is something that the child might have the desire to express. This is certainly an issue since the majoritarian function of NSQs is the posing of clarification requests (CRs). There is no question that children aged 2;6 or thereabouts comprehend and respond appropriately to CRs (Garvey 1979, Gallagher 1981). However, their own interest in and actual production of these is more in question, as is the influence on CRs of the task in which they are engaged during the conversations exemplified in the corpus.

\[^{14}\] Dominic and Carl’s mothers, on the other hand, exhibit patterns of triggering context exploitation akin to Aran and Becky’s mothers: Carl’s mother has 17 triggering contexts altogether in these files and uses NSUs in 15 (88%) of these cases, whereas Dominic’s mother has 52 triggering contexts and uses NSUs in 47 (90%) of these cases.
We address the issue of *semantic desirability* in two ways. First, there is evidence that CRs are realized by non-NSQ utterances in the period we have been considering:

(21) (a) Investigator: That is the cow’s udders. Becky: What is that? Investigator: Udders. (becky16b)
(b) Mother: In there? Becky: Where is xxx there. (becky22b)
(d) Mother: I think you can put this one wherever you like, Becky in the picture. Mother: Put him next to that other mouse. Becky: Where does he go? (becky22b)
(e) Mother: Let us put panda in his car, should we? Aran: Where is his car? (aran 23b)
(f) Mother: Has he eaten you? Aran: What is he call [*]? Mother: He is called a tiger, isn’t he? (aran21b)

Our second argument relates to data from somewhat later productions of Aran and Becky. Table 10 contains the distribution of Aran and Becky’s questions, as recorded in the files (aran28b,29b,30b (2;8–2;9), becky28b,29b,30b (2;9–2;10)).

<table>
<thead>
<tr>
<th></th>
<th>Total Sentential</th>
<th>Sluicing CE</th>
<th>Total Non-sentential</th>
<th>Total Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aran</td>
<td>57</td>
<td>23</td>
<td>1</td>
<td>81 (82.5%)</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>0</td>
<td>17 (17.5%)</td>
<td>98</td>
</tr>
<tr>
<td>Becky</td>
<td>106</td>
<td>60</td>
<td>8</td>
<td>174 (77%)</td>
</tr>
<tr>
<td></td>
<td>52</td>
<td>0</td>
<td>52 (23%)</td>
<td>226</td>
</tr>
</tbody>
</table>

Table 10
Distribution of sentential/non-sentential questions for Aran and Becky at (2;8–2;9) and (2;9–2;10), respectively

We address the issue of *semantic desirability* in two ways. First, there is evidence that CRs are realized by non-NSQ utterances in the period we have been considering:

(21) (a) Investigator: That is the cow’s udders. Becky: What is that? Investigator: Udders. (becky16b)
(b) Mother: In there? Becky: Where is xxx there. (becky22b)
(d) Mother: I think you can put this one wherever you like, Becky in the picture. Mother: Put him next to that other mouse. Becky: Where does he go? (becky22b)
(e) Mother: Let us put panda in his car, should we? Aran: Where is his car? (aran 23b)
(f) Mother: Has he eaten you? Aran: What is he call [*]? Mother: He is called a tiger, isn’t he? (aran21b)

Our second argument relates to data from somewhat later productions of Aran and Becky. Table 10 contains the distribution of Aran and Becky’s questions, as recorded in the files (aran28b,29b,30b (2;8–2;9), becky28b,29b,30b (2;9–2;10)).

In terms of raw frequency this demonstrates significant production of NSQs by Aran and Becky, some examples of which are in (22) and (23):

(22) (a) Mother: She has got such a long fringe now. Mother: There we are. Becky: What fringe? (becky29b)
going to speak to you on the telephone. Becky: Who? Mother: The cow. Mother: This one. (becky29b)

(23) (a) Mother: Get me that top from over there please. Aran: Yes. Aran: Course. Aran: Which top? (aran28b)
(b) Mother: Well we’re not going to play this game if you’re going to be silly. Aran: Why? (aran28b)
(c) Mother: It has come off. Aran: Which one? Mother: Off this one. (aran28b)

It is also worth noting the make-up of these NSQs, as exemplified and tabulated in table 11. More than half of Becky’s relatively high frequency of NSQs (23%, see table 10) – higher proportionwise than three of the four adults recorded in table 7 – derives from ‘repetition ‘What?’’ and another third is made up of ‘why’ sluices. As we discuss in section 5, the contextual resolution of these sluices is more straightforward than either reprise or direct sluices. With Aran the distribution is more even. Most interestingly, it is worth noting that although both children do produce reprise sluices, which express CRs, there is complete absence of the CE construction. We discuss the relevance of this to the issue of how well parental input is correlated with order of acquisition in section 5.3.

The data from the somewhat later productions of Aran and Becky indicates that there is nothing specific to the nature of the interactions occurring in this corpus which biases against the production of NSQs by the children. This suggests that the earlier absence of NSQs is due, in part at least, to insufficiently developed competence.

Finally, one additional argument can be made here, deriving from a comparison of the distribution of NSAs for Aran and Becky at this same stage vs. earlier stages of the same children. Table 12 shows that the distribution of NSAs over this period has not changed significantly, rising slightly for Aran by comparison with the period covered earlier ($\chi^2 = 1.6$, Fisher exact probability one-/two-tailed $p = .1$, $p = .18$, respectively), and falling slightly for Becky ($\chi^2 = .98$, Fisher exact prob one-two-tailed $p = .16, p = .31$, respectively). This suggests that the increasing frequency of NSQs should not be attributed to a recently developed predilection for phonologically reduced forms.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Repetition</th>
<th>Why sluice</th>
<th>Reprise sluice</th>
<th>Direct sluice</th>
<th>CE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aran</td>
<td>0</td>
<td>6</td>
<td>4</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Becky</td>
<td>30</td>
<td>17</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 11
Distribution of non-sentential questions for Aran and Becky at (2;8–2;9) and (2;9–2;10), respectively
In this section we suggest that the delayed emergence of elliptical queries is not due to language-specific factors: there is prima facie evidence that NSQs emerge late also in child Greek. This provides evidence that the late short query effect is both language- and task-independent of the specific conditions existing in the Manchester set-up.

Our study of NSUs in child Greek is based on the Stephany Corpus in CHILDES (Stephany 1995): four monolingual Greek-speaking children in free conversation with their caregiver and/or investigator observed at 1;9–1;11, 2;3–2;5, and 2;9. We mention here solely the data from the child Mairi, the most intensely studied child in the corpus. Tables 13 and 14 provide data about her productions of NSAs and NSQs, respectively, at 21, 27, and 33 months. The productions at 27 months are illustrative of the late short query effect: Mairi has mastery of NSAs and of sentential interrogatives, but produces virtually no NSQs. By 33 months, the situation has changed significantly: NSQs now constitute already 14% of her questions. This is another indication that the effect is developmentally grounded, not situation-/task-dependent. We also include data about Mairi’s mother from the pre-NSQ period (21 and 27 months), data that indicates that Mairi’s mother uses NSQs productively.

### 5. The Late Short Query Effect: Theoretical Explanation

In section 4 we sought to demonstrate that the late short query effect cannot be explained away pragmatically: the absence of NSQs from young children’s productions is not merely a consequence of the contextual unavailability of antecedents for NSQs and/or the semantic undesirability of such questions.

[15] In each case this data is based on coding the first file for each period of investigation: respectively, files mai21a1 (948 turns), mai27b1 (1,128 turns), mai33c1 (1,275 turns). The coding was done by the second author.
of the contents NSQs express. In so far as this demonstration is successful, this suggests that a mechanistic explanation is required. It is not possible to give such an explanation within theories of NSUs such as UNDERLYING
SENTENTIALISM or HIGHER-
ORDER UNIFICATION given that, as we saw in section 2.2, each of these enforces a version of PUNCTUAL NSU EMERGENCE: ellipsis resolution is pinned to a single mechanism, which, given the presence of requisite inputs (wh-interrogative sentences, quantified propositions and
wh-phrase denotations), cannot prevent the corresponding NSUs (e.g. sluicing) being available in the linguistic system, contrary to what is desired. In what follows, we provide an account of the LATE SHORT QUERY EFFECT using a dialogue-oriented construction grammar. Such an account comprises two components, given in (24).
ANSWERS WITHOUT QUESTIONS

(24) (a) **Propositional NSUs and sentential interrogatives without NSQs:** The specification of a grammar and contextual system in which short answers and sentential interrogatives co-exist without NSQs.

(b) **The order of NSU acquisition:** An explanation for the emergence of NSAs before NSQs.

Before turning to these in order, we provide a brief sketch of the grammatical and contextual analysis of NSUs which we presuppose.

5.1 *Grammar and context for NSUs*

As mentioned earlier, in *dialogue-oriented constructionism*, NSU words and constructions are akin to other indexical words and constructions, with the relevant contextual parameter being the discourse topic, or more technically the **MAX-QUD**. For certain NSU constructions, an additional contextual parameter is an antecedent sub-utterance which specifies *inter alia* syntactic and/or phonological parallelism conditions for the fragment utterance. Intuitively, this parameter provides a partial specification of the focal sub-utterance, and hence it is dubbed the **FOCUS ESTABLISHING CONSTITUENT** (FEC).\(^{[16]}\) The FEC plays a role somewhat akin to the role played by the **PARALLEL ELEMENT** in higher-order unification-based approaches to ellipsis resolution.

Given this, the grammar and contextual processing of NSUs involves specifications such as described below.\(^{[17]}\)

• **Polar lexemes**

  **Context:** For propositional lexemes such as ‘yes’ and ‘no’, the **MAX-QUD** is a polar question \( p? \), which is identical with the (content of the) interrogative utterance or derived from an assertion \( p \).\(^{[18]}\)

\[
\begin{align*}
(a) \quad & A: \text{Did Bo phone?} \quad B: \text{Yes/No} \, (= \text{Bo phoned/didn’t phone}) \\
(b) \quad & A: \text{Bo phoned.} \quad B: \text{Yes/No} \, (= \text{Bo phoned/didn’t phone}) \\
(c) \quad & \text{MAX-QUD: } \text{?Phone}(b, t)
\end{align*}
\]

---

\(^{[16]}\) Somewhat less illuminatingly referred to as the **SALENTI-UTTERANCE** in Ginzburg & Sag (2000).

\(^{[17]}\) The treatment informally sketched here originates in Ginzburg & Sag (2000), Ginzburg & Cooper (2004). However, it takes into account a modified semantic analysis, formulated in Fernández (2006). Ginzburg (2009). These latter works jettison unification-based semantic operations used in the earlier treatment in favour of a semantics rooted in \( \lambda \)-calculus operations (\( \lambda \)-abstraction, function application, etc.), based on the semantic framework of Type Theory with Records (Cooper 2006). This allows, among other things, for a simpler characterization of the semantic combinatorial operations used in NSU constructions, a key component in our explication of the order of acquisition of NSUs in section 5.3.

\(^{[18]}\) This follows the assumption that an assertion \( p \) gives rise to the question \( p? \) in QUD (Ginzburg 1996, 2009).
Grammar: ‘Yes’ is lexically specified as having a content identical with the proposition queried in MAX-QUD. ‘No’ is used to negate when answering positive polar questions and to affirm when answering negative polar questions.19 Hence, ‘no’ is lexically specified as denoting either the queried proposition in MAX-QUD or its negation, whichever is negatively marked.

- Short answers

Context: The MAX-QUD is a wh-question; the FEC in such a case is the sub-utterance of the wh-phrase:

(26) (a) A: Who phoned? B: Bo (\(=\) Bo phoned).
(b) MAX-QUD: \(\lambda x.\) Phone\((x, t)\); FEC = ‘Who’.

Grammar: Short answers are an instance of a construction in which MAX-QUD is the predicate and the bare fragment is the argument; the construction enforces syntactic parallelism between the FEC and the fragment.20

- Direct sluicing

Context: The MAX-QUD is a quantified polar question, entering into the context via assertion or querying; the FEC is the sub-utterance corresponding to the widest scoping quantificational NP:

(27) (a) A: A student phoned. B: Who? (\(=\) Which student phoned?)
(b) MAX-QUD: \(\exists x.\) Phone\((x, t)\); FEC: ‘A student’.

Grammar: Direct sluices are an instance of a question-denoting construction.22 The domain of the question is provided by the fragment whP.

---

19 This dual nature of ‘no’ is one which English-speaking children seem aware of very early on:

(i) Mother: shall we move it? Aran: No. (\(=\) negation of: (We shall move it.)) Mother: All right. Mummy won’t [: will-’nt] move it.
(ii) Mother: Aren’t you going to put it back together again? Aran: No. (\(=\) affirmation of: (Aran) is not going to put it back together again.) Mother: Why not?

(both examples from aran20b)

These uses of ‘no’ need to be distinguished from ‘refusal no’, used to refuse to perform an action. This latter use has a distinct semantics and is acquired at an earlier stage.

20 This approach to short answers has a long history in work on interrogative semantics; see e.g. Hausser (1983).

21 Why’ sluices (‘Why?’), which are among the commonest sluices in adult speech, arguably warrant a simpler treatment. They invariably occur without an antecedent quantificational NP, i.e. an FEC. Their contextual background is, therefore, similar to that of propositional lexemes. Grammatically, they can be analyzed as functors that take the contextually provided proposition as an argument.

22 Questions denote propositional abstracts: functions from domains partially specified by the wh-phrase into propositions.
ANSWERS WITHOUT QUESTIONS

whereas the range of the question arises by a substitution operation on (the proposition queried in) max-qud: the fragment whP’s variable is substituted in place of the FEC’s variable. The construction enforces syntactic parallelism between the FEC and the fragment.  

• Reprise sluicing/CE24,25

Context: The context for questions used in CRs arises via an accommodation process (a clarification context update rule) whose input is a partially understood utterance (Ginzburg & Cooper 2004, Purver 2004, Ginzburg 2009). For instance, the accommodation operation of parameter focussing gives rise to the contextual specification in (28a):

(28) (a) Input utterance: A: Did Bo leave?
   (b) Context assuming the reference of ‘Bo’ cannot be fully resolved: max-qud: \( \lambda x.\text{Ask}(A,?\text{Leave}(x)) \) (Who is A asking if I left?); FEC: ‘Bo’.
   (c) Bo? (=Are you asking if BO [of all people] left?)/Who? (=Who are you asking if they left?)

Grammar: An instance of the same construction as short answers, in combination with the construction used to form questions from propositional denotations.

The semantic composition and contextual reference in the above NSU constructions can be summarized in tables 15 and 16. With this in hand, we can turn to describing the two components of our explanation of the LATE SHORT QUERY EFFECT.

5.2 Propositional NSUs and sentential interrogatives without NSQs

Ginzburg & Sag (2000) specify a grammar which accommodates inter alia sentential declaratives, sentential interrogatives, and the various NSU constructions discussed above – polar lexemes, short answers, direct and reprise sluices, and CE. In this grammar, as in other work in the Construction Grammar tradition, constructions are organized into a type hierarchy. This enables higher-level types to be posited which allows relevant cross-cutting generalizations to be captured. For our current purposes it is worth mentioning two relevant pieces of the construction hierarchy, that which concerns interrogatives and that which concerns NSUs.

[23] The fact that such a parallelism requirement exists for short answers and sluicing is captured in the type hierarchy of constructions, as further discussed below.

[24] For reasons of space, we restrict attention here to the clausal confirmation reading of CE.

[25] ‘Repetition what’ can be analyzed in a somewhat simpler fashion: in common with other words that fulfill a similar function (‘pardon’, ‘eh’, etc.), it can be lexically specified to express a question that queries the form of a contextually provided utterance.
The hierarchy for interrogatives is displayed in figure 1.

The most general interrogative type is the type *inter-cl*, whose defining characteristic is that such clauses denote questions. ‘Canonical filler-gap’ *wh*-interrogatives are governed by the type *wh-int-cl*, whose defining characteristic is that such clauses denote a question involving at least one abstraction, which emanates from the dislocated filler phrase. Direct sluices and reprise sluices are sub-types of a distinct branch of the interrogative hierarchy. *slu-int-cl*, the type on the basis of which direct sluices are analyzed, specifies how their semantics arises in the terms discussed above. It also

<table>
<thead>
<tr>
<th>NSU type</th>
<th>Content specification</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>max-quad.prop (= (1))</td>
<td>max-quad’s proposition</td>
</tr>
<tr>
<td>No sub/neg-q</td>
<td>NegProp ∈ {(1), ¬(1)}</td>
<td>Selection of negative proposition from polar answers</td>
</tr>
<tr>
<td>Short-answer</td>
<td>max-quad(fragment.content)</td>
<td>Function application of max-quad to fragment’s content</td>
</tr>
<tr>
<td>Direct sluice</td>
<td>λ fragment.domain (max-quad.prop(FEC.content) → fragment.content)</td>
<td>λ-abstraction of fragment’s domain from (3), where (3) = Substitution in (1) of fragment’s content for FEC’s content</td>
</tr>
<tr>
<td>Reprise sluice</td>
<td>λ fragment.domain (max-quad(fragment.content))</td>
<td>λ-abstraction of fragment’s domain from (2)</td>
</tr>
</tbody>
</table>

Table 15
Semantic composition for NSUs in a dialogue-oriented constructionist grammar

<table>
<thead>
<tr>
<th>NSU type</th>
<th>Contextual parameters involved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes, no, short answer</td>
<td>max-quad</td>
</tr>
<tr>
<td>Direct sluice</td>
<td>max-quad, FEC</td>
</tr>
<tr>
<td>CE, reprise sluice</td>
<td>max-quad, FEC via accommodation triggered by partially understood utterance</td>
</tr>
</tbody>
</table>

Table 16
Contextual parameters involved in NSU contents
specifies the semantic and syntactic parallelism between the fragment and the FEC; the syntactic part of this specification is shared with other NSU constructions and is captured in the NSU type hierarchy to be discussed shortly. Reprise sluices are instances of the type \textit{dir-is-int-cl}, a type whose defining characteristic is the generation of question contents from daughters of proposition denoting heads and which also accommodates ‘intonation questions’ and non-reprise \textit{in situ} \textit{wh}-interrogatives.

The crucial point about a grammar organized in terms of a hierarchy like figure 1 is that its acquisition can be modelled in a straightforward monotonic way, as pointed out originally by Green (2000) (see e.g. Goldberg 1995, Tomasello 2003). In our case a grammar without sluicing involves an interrogative hierarchy with only the circled branches. Acquiring the additional constructions involves a monotonic refinement of the hierarchy.

This applies equally to the hierarchy of NSUs, shown in figure 2. A general type here is \textit{hd-frag-ph}, which specifies the afore-mentioned fragment-FEC parallelism. Short answers are instances of the type \textit{decl-frag-cl}, which specifies the semantic construction combining \textit{MAX-QUD} and the fragment. Sluicing is an instance of the type \textit{slu-int-cl}, as mentioned above. A grammar for NSUs without direct sluicing involves merely the circled branch in figure 2.

\textit{Figure 1}

\textit{Sketch of interrogatives type hierarchy from Ginzburg & Sag (2000)
This concludes the first part of our explanation of the **late short query effect**: the specification of a grammar and contextual system in which short answers and sentential interrogatives co-exist without NSQs.

**5.3 The order of NSU acquisition**

We can now consider how semantic and contextual processing might play a role in the observed order of acquisition of NSUs, most specifically the fact that NSAs emerge before NSQs. We pursue an account of this in terms of a notion that combines semantic and contextual complexity, as will be explained shortly. There have been many accounts of the order of acquisition of lexical items and constructions that have appealed to some notion of complexity (see e.g. Brown & Hanlon 1970, Bloom, Merkin & Wooten 1982). There have also been recent alternative accounts of such phenomena in terms of input frequency (see e.g. Theakston et al. 2001, Rowland et al. 2003), and references therein). A glance at table 8 above indicates that at least the most simplistic input factor, the distribution of caregiver NSUs, is not well correlated with the order of NSU acquisition: the frequency of Aran’s mother’s NSUs – admittedly on the basis of a small sample – (CE $\geq$ sluicing $\geq$ yes/no $\geq$ short-answer) is precisely the opposite of the order of acquisition, while Becky’s mother’s (CE $\geq$ yes/no $\geq$ sluicing $\geq$ short-answer) is not much better. The most striking discorrelation concerns CE: as we noted in section 4.2, although both children do produce reprise sluices, which express CRs, there is complete absence of the CE construction. This contrasts sharply with the distribution evinced by their mothers: table 7 indicates that on average Becky’s mother was producing 20 NSQs per conversation of which 14 were CE, whereas Aran’s mother was producing 18 NSQs of which 10 were CE. We certainly do not rule out the...
existence of other possible factors relating to input which could play a major role, for instance the frequency of NSU triggering contexts presented to the child – though this is somewhat tricky to assess since, in principle, any utterance is a potential context for CE, which is acquired last of all NSUs.

When considering the semantic complexity of NSUs, there is a rather obvious notion that is worth considering, namely the complexity of the semantic operations involved in constructing the content, the semantic composition complexity. However, to the extent that this is a relevant factor, it clearly tells us only half the story since access to the contextual parameters also plays a crucial role in the semantic processing. Hence, our notion of complexity should also make reference to contextual accessibility. There are a number of ways that these notions could be combined. We assume a notion of complexity that gives priority to contextual complexity; only if the contextual complexity is equal is an appeal made to semantic composition complexity:

\[
C_1 \preceq C_2 \text{ iff either } C_1 \text{ has a simpler contextual background than } C_2, \text{ or } C_1 \text{ and } C_2 \text{ have equivalent contextual backgrounds but } C_1 \text{ is combinatorially simpler than } C_2. 
\]

Let us examine the notion of semantic composition complexity. In our case, the operations involved are: function application, \( \lambda \)-abstraction, proposition extraction, substitution, negation, and polar-questioning.

We evaluate this in terms of the operations involved in each construction type. Let us consider some examples: the content of ‘yes’ is specified in table 15 as arising by extracting the proposition from \( \text{max-qud} \). The complexity of constructing ‘yes’, therefore involves one proposition extraction operation. Similarly, given the specification of short answer content in table 15, its complexity involves one function application operation. A direct sluice involves a propositional extraction, a \( \lambda \)-abstraction and a substitution. Its complexity involves, therefore, 1 proposition extraction + 1 substitution + 1 abstraction.

Table 17 lists the associated semantic composition complexity of the various NSU classes we have considered.

If we use tables 17 and 16 to extract, respectively, semantic composition complexity and contextual parameter accessibility, we can explain the fundamental order of emergence we have demonstrated in this paper:

\[
\text{Non-sentential answers} \preceq \text{Non-sentential questions}; \text{ Short answers and yes/no have fewer contextual parameters than direct sluices and are also simpler combinatorially; short answers and yes/no have simpler contextual backgrounds than reprise sluices and CE.}
\]

\[26\] In so listing the operations, we are taking a somewhat theory-neutral view of semantic combinatorics, shows how to explicate proposition extraction, substitution, negation, and polar-questioning in terms of the logical formalism of Type Theory with Records.
6. C O N C L U S I O N S

Although non-sentential utterances are a pervasive feature of adult speech, their acquisition has to date not been the focus of much attention. In this paper we have presented findings about the order of acquisition of certain major types of NSUs, demonstrating the existence of the LATE SHORT QUERY EFFECT, the absence of NSQs (sluicing and clarification ellipsis) at an acquisition stage where a mechanism for NSUs is well established – short answers and yes/no are used productively – and sentential interrogatives have been acquired. We have attempted to demonstrate that this effect is not merely a pragmatic epiphenomenon, one that can be explained on the basis of either the CONTEXTUAL UNAVAILABILITY OR SEMANTIC UNDESIRABILITY of NSQs. Moreover, we have shown that the emergence of NSQs is not correlated with increased use of other types of NSUs, as might be expected if the absence of NSQs were a consequence of the gradual emergence of a generalized mechanism of phonological reduction. Although our data is primarily based on English speaking children, we have presented some evidence suggesting that our conclusions hold also for Greek.

We believe that the LATE SHORT QUERY EFFECT provides useful evidence in the ongoing debate concerning the nature of NSU constructions, and more generally the integration of context in grammar. It is not possible to give an account of this effect within theories of NSUs such as UNDERLYING SENTENTIALISM OR HIGHER-ORDER UNIFICATION given that each of these enforces a version of PUNCTUAL NSU EMERGENCE. In these theories ellipsis resolution is pinned to a single mechanism, which means that later emerging NSU types

<table>
<thead>
<tr>
<th>NSU type</th>
<th>Content specification</th>
<th>Semantic composition operations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>max-quad.prop (= (1))</td>
<td>1 proposition extraction</td>
</tr>
<tr>
<td>No</td>
<td>NegProp ∈ {(1), ¬(1)}</td>
<td>1 proposition extraction + 1 negation + 1 choice-from-2</td>
</tr>
<tr>
<td>Short-answer</td>
<td>max-quad(fragment.content)</td>
<td>1 function application</td>
</tr>
<tr>
<td>Direct sluice</td>
<td>λ fragment.domain (max-quad.prop(FEC.cont ↦ fragment.content))</td>
<td>1 λ-abstraction + 1 substitution + 1 proposition extraction</td>
</tr>
<tr>
<td>CE_confirmation</td>
<td>?max-quad(fragment.content)</td>
<td>?-op + function application</td>
</tr>
<tr>
<td>Reprise sluice</td>
<td>λ fragment.domain (max-quad(fragment.content))</td>
<td>1 λ-abstraction + 1 function application</td>
</tr>
</tbody>
</table>

Table 17
Semantic composition complexity for NSUs in a dialogue-oriented constructionist grammar

670
are available in the linguistic system given the presence of requisite inputs (wh-interrogative sentences, quantified propositions and wh-phrase denotations).

We have shown how the late short query effect can be modelled in an approach to grammar which integrates phonological/syntactic/semantic information with detailed specification of dialogue context. NSUs are treated as constructions, which inter alia carry a specification of the contextual conditions regulating their resolution. As discussed earlier, within this type of approach to the description of NSUs, detailed formal analysis of a very high proportion of NSU constructions has been attained. We have used these analyses as the basis of an explanation for the order of emergence of certain classes of NSU constructions, using a notion which combines semantic composition complexity and the accessibility to the contextual parameters. Scaling up this kind of explanation to the full range of NSU constructions is a challenge which we are currently pursuing.

REFERENCES


Green, Georgia. 2000. Modelling grammar growth: Universal grammar without innate principles or parameters. Revised and expanded version of a paper presented at GALA 97, University of Edinburgh.


Müller, Christoph & Michael Strube. 2001. MMAX: A tool for the annotation of multimodal corpora. 2nd DICAI Workshop on Knowledge and Reasoning in Practical Dialogue Systems, 45–50. Seattle, WA.


Authors’ addresses: (Ginzburg)
Department of Computer Science, King’s College London, Strand, London WC2R 2LS, UK
jonathan.ginzburg@kcl.ac.uk

(Kolliakou)
School of English Literary and Linguistic Studies, Newcastle University, Newcastle upon Tyne NE1 7RU, UK
Dimitra.Kolliakou@ncl.ac.uk