Journal of Child Language (2025), 1–26
doi:10.1017/S0305000924000588

CAMBRIDGE UNIVERSITY PRESS

REGISTERED REPORT

Null subject comprehension and production revisited: a look at English and Italian

Juliana Gerard[1], Muskaan Singh[2], Giulia Bencini[3] and Virginia Valian[4,5]

[1]School of Communication and Media, Ulster University, Belfast, Northern Ireland; [2]School of Computing, Engineering and Intelligent Systems, Ulster University, Derry~Londonderry, Northern Ireland; [3]Department of Linguistics and Comparative Cultural Studies, Ca' Foscari University of Venice, Italy; [4]The Graduate Center, City University of New York, New York City, NY, USA and [5]Department of Psychology, Hunter College, New York City, NY, USA
Corresponding author: Juliana Gerard; Email: j.gerard@ulster.ac.uk

(Received 20 November 2023; revised 11 September 2024; accepted 23 September 2024)

Abstract
This study will investigate how children acquire the option to drop the subject of a sentence, or null subjects (e.g., "Tickles me" instead of "He tickles me"). In languages that do not permit null subjects, children produce sentences with null subjects from 1 to 3 years of age. This non-adultlike production has been explained by two main accounts: first, the null subject sentences may accurately reflect the children's linguistic knowledge, that is, a competence account. Alternatively, they may result from immature processing resources, therefore underestimating children's competence, that is, a performance account. We will test the predictions of these accounts by using a central fixation preference procedure and elicited imitation to measure children's comprehension and production, respectively, in monolingual 19- to 28-month-olds acquiring English (a non-null subject language) and Italian (a null subject language). The results will shed light on acquisition across languages, and the features that provide evidence to a learner.

Keywords: null subject; comprehension; production; looking time; imitation

Introduction

Children acquire language quickly – within a few short years, they advance from babbling to first words to complex sentences. Meanwhile, the early success of language acquisition may be illuminated precisely by the contexts where it appears to be unsuccessful. Importantly, these contexts are not random and can isolate specific processes involved in acquisition.

One such context is the option to drop a subject, that is, to use a null subject, as in (1):

(1) a. Vedo un aereo.
see–1SG an airplane
"I see an airplane" (Italian; from Guasti, 2002)

b. Kanjian ta le.
see him ASP
"He saw him" (Chinese; from Huang, 1984)

The availability of null subjects varies across languages (Biberauer, 2008; Biberauer et al., 2009; Camacho, 2013, 2016) – for example, languages like English, German, and French do not permit null subjects in declarative contexts like (1), and require an overt subject even in expletive contexts like (2):

(2) a. It is raining. (English)
b. Il pleut. (French)

In contrast, the same contexts in null subject languages require a null subject (Rizzi, 1982):

(3) Piove. (Italian)
rain-3SG
"It rains"

Children must, therefore, acquire the appropriate optionality based on their linguistic input. Moreover, this acquisition is more complicated than choosing between an overt subject language as in (2), or a null subject language as in (3). Rather, this optionality varies across languages depending on features of discourse, morphosyntax, and lexical frequency (Allen, 2000; Allen & Schröder, 2003; Allen et al., 2008; Shlonsky, 2009; Bondaruk, 2001; Brennan, 2008; Travis & Torres Cacoullos, 2021; Erker et al., 2023). Therefore, the acquisition problem must involve identifying those specific contexts which require an overt subject, and the contexts in which subjects may be omitted.

However, children initially appear to acquire the incorrect optionality, particularly when their language does not allow null subjects in declarative contexts like (1) (for reviews see Hyams, 2011; Valian, 2016a). For these languages, children produce sentences with null subjects between the ages of 1 and 3 years, as in (4) (from Guasti, 2002):

(4) a. Tickles me. (Adam, 3;6)
c. Mange du pain. (Grégoire, 2;1, French)
eat–3SG some bread.

b Se, blomster har. (Jens, 2;2, Danish)
look, flowers have/has.
"Look, (I/you/she/we) have/has flowers"

Different explanations have been proposed for this "null-subject stage" in non-null subject languages. For example, this stage may be accounted for by the temporary acquisition of a grammar that generates null subjects, that is, a competence account (Hyams, 1986, 1992; Rizzi, 2018; Hyams & Wexler, 1993; Orfitelli & Hyams, 2012; Wexler, 2013). Non-adult competence is similarly posited by usage-based approaches, although in isolated contexts depending on input frequency (Bybee, 2002; Brown & Rivas, 2012; Brown & Shin, 2022). However, null subject sentences may also be produced with a non-null subject grammar if limitations in processing resources cause the subject to be

https://doi.org/10.1017/S0305000924000588 Published online by Cambridge University Press

dropped, that is, a performance account (Bloom, 1989, 1990, 1993; Valian, 1991; Valian et al., 1996; Valian & Aubry, 2005).

In the context of null subjects, competence and performance accounts have largely been formulated based on children's production, that is, subjects omitted from children's own sentences. Meanwhile, variation in null subject contexts across languages must be acquired from the linguistic input; this acquisition depends on children's comprehension of the input. If children's comprehension accurately reflects the input, then the acquisition process will consist primarily of identifying the relevant contexts. This process becomes more complicated, however, if children's comprehension of the input does not match the input itself, for example, if children who produce sentences with null subjects also comprehend sentences in the input with null subjects (Orfitelli & Hyams, 2012). Competence and performance accounts make different predictions about children's representation of the input in different contexts (Bloom, 1989, 1990, 1993; Valian, 1991; Valian et al., 1996). In this study, we test these predictions for children's comprehension of null subjects and consider further implications for the timeline of acquisition and relevant learning mechanisms.

Competence and performance accounts

Children's observable behavior is the output of various unobservable cognitive systems (Omaki & Lidz, 2015). These include linguistic knowledge, processing mechanisms, and discourse pragmatics, among others. Competence and performance accounts attribute children's non-adultlike behavior to different processes; this, in turn, entails different explanations for the transition from non-adultlike to adultlike behavior. Competence and performance accounts therefore have different implications for language acquisition, both for the specific context of null subjects and more broadly. In the following sections, we discuss competence and performance accounts of children's null subjects in production, and their predictions for comprehension.

Competence accounts

Under a competence account, the source of children's null subject sentences is a grammar that generates null subjects (Hyams, 1986, 1992; Rizzi, 2018; Hyams & Wexler, 1993; Orfitelli & Hyams, 2012; Wexler, 2013). For children who are acquiring a non-null subject language, this *non*-adult null subject grammar must be discarded in favour of the adult non-null subject grammar; competence accounts therefore require an explanation of how the non-adult grammar becomes adultlike. For null subjects, this transition is triggered by evidence from the linguistic input (Hyams, 2011).

Different competence accounts have involved different types of null subject grammars. For example, adult grammars that generate null subjects include grammars with *pro* subjects (Hyams, 1986), as well as grammars with topic drop (Hyams, 1992). These accounts predict that the contexts of children's null subjects will match those of the respective adult grammar. Meanwhile, null subjects are also predicted with a non-adult grammar with root truncation, which predicts additional non-adult elements depending on the truncated structure (Hyams, 2011; Rizzi, 2018, 2005).

Importantly, all competence accounts predict that children will produce null subject sentences and that the source is grammatical, because of a qualitatively different grammar from the adult grammar. Therefore, this requires a qualitative change to children's

https://doi.org/10.1017/S0305000924000588 Published online by Cambridge University Press

linguistic knowledge – from a grammar that generates null subjects to a grammar that does not. Given the cross-linguistic variation in subject use, this change necessarily involves evidence from the linguistic input. Finally, the evidence must be robust to the null subject grammar. For example, sentences with overt subjects do not constitute evidence for a non-null subject grammar, since overt subjects are also acceptable in null subject languages. Rather, evidence for a non-null subject language must be impossible to represent with a null subject grammar. One such type of evidence is expletive subjects (Hyams, 1986) as in (2), repeated below as (5):

(5) a. It is raining. (English)
b. Il pleut. (French)

This predicts that adult grammar is only acquired once children encounter sufficient instances of expletive subjects (Yang, 2002). In addition, children would be predicted to have non-adult grammar until this time.

The null subject stage is observed across languages, on similar acquisition timescales. The most parsimonious account of this transition will, therefore, involve the same type of evidence across languages. The relevant evidence need not be restricted to expletive subjects; however, the same predictions apply with respect to the onset of the adult grammar – null subjects are predicted *before* children acquire the adult grammar, but not *after*. This contrasts with the predictions of performance accounts.

Extra-syntactic (performance; executive function) accounts

Extra-syntactic accounts attribute children's null subjects to limited processing resources (Bloom, 1989, 1990, 1993; Valian, 1991; Valian et al., 1996; Valian & Aubry, 2005) or limited development in other, extra-syntactic, components of the grammar, such as phonology and prosody. These accounts predict null subjects in production, even after the adult grammar is acquired. Such accounts therefore involve an explanation of how limitations in other parts of the grammar or cognition result in unexpressed subjects and also involve a specification of how the adult grammar is acquired (Valian, 2016a). In contrast with competence accounts, performance accounts place the acquisition of the requirement of overt subjects in English before the consistent use of such subjects. If that account is correct, it has significant implications for the relevant evidence, which must be accessible in the linguistic input well before children exit the null subject "stage."

Because processing resources are required to produce an overt subject, processing accounts predict that the subject will be dropped with insufficient resources, that is, there is a production bottleneck (Valian, 2016b). The availability of these resources may be modulated by internal factors (e.g., processing capacity) or external factors (e.g., sentence complexity). Importantly, variation in the contexts where children omit subjects is not random under a processing account. Rather, subject omission is highly principled depending on the type of extra-syntactic pressure (Lidz, 2022), meaning that processing accounts make clear predictions about where subjects will and will not be omitted by children in the null subject stage.

For example, null subjects vary across individuals by mean length of utterance (MLU), a production measure of language development (Valian, 1991; Valian et al., 1996). In addition, further variation is observed in the production of null subjects depending on the sentence context, with systematic variation across contexts: null subjects are more likely

https://doi.org/10.1017/S0305000924000588 Published online by Cambridge University Press

in contexts that place greater pressure on the production system. This effect is observed most consistently with the trade-off between overt subjects and verb phrase (VP) length, with a greater likelihood of a null subject, the longer the VP (Bloom, 1989, 1990; Valian, 1991).

Based on this variation in the likelihood of a null subject, processing accounts make a key prediction, in (6):

(6) Null subjects are less likely in contexts with a lower processing load. Therefore, ***decreasing*** the processing load of a given context should ***decrease*** the likelihood of a null subject.

This prediction was tested for children's production by Valian and Aubry (2005), who used a double imitation paradigm to reduce the processing load of the test sentence in a second imitation attempt. Pronominal and expletive subjects were less likely to be dropped in the second attempt, consistent with the prediction in (6). They note, however, that decreasing the processing load may involve various factors. For example, a second imitation attempt may alleviate a production bottleneck, decreasing the pressure on a limited production system. However, a second attempt may also reduce the processing load for comprehension, with a similar result of a more accurate imitation (Valian et al., 1996; Valian & Aubry, 2005).

This comprehension option has important implications for children's acquisition of null subjects, which necessarily involves linguistic input. If children's comprehension of the input does not match the input itself, then this limits the possible options for evidence in the input. In the next section, we consider the prediction in (6) for children's comprehension.

Comprehension of null subjects

Children's null subjects have primarily been investigated with production measures. Children's comprehension has been tested in just one study to date, by Orfitelli and Hyams (2012). We discuss this study in the following section, followed by the predictions for the current study.

Orfitelli and Hyams (2012)

In a study on children's comprehension, Orfitelli and Hyams (2012) tested the predictions of competence and performance accounts of null subjects. If children's null subjects are because of a null subject grammar, then a competence account predicts that they should also accept null subjects in comprehension. In contrast, Orfitelli & Hyams hypothesized under a performance account that children with adult grammar should not accept null subjects in comprehension (see also Lutken et al., 2020).

In a modified truth value judgment task, younger children (who produced sentences with null subjects) and older children (who no longer did) judged sentences with null subjects like (7) in declarative and imperative contexts:

(7) Play with blocks.
   a. declarative: description of children playing with blocks
   b. imperative: command for children to play with blocks

https://doi.org/10.1017/S0305000924000588 Published online by Cambridge University Press

Importantly, the imperative context is acceptable in adult grammar, while the declarative context is only possible with the null subject grammar.

Consistent with the competence account, the children who produced null subjects did accept the declarative contexts, while the older children only accepted the imperative. Moreover, Orfitelli and Hyams (2012) provide an explanation for the role of the parser under this competence account, which must select between declarative and imperative interpretations for a grammar that still allows this. In particular, they note that "when children exit the [null subject] stage, their grammar changes and the processor is no longer faced with the problem of ambiguity resolution."

This qualitative change from a grammar that generates both imperative and declarative interpretations (i.e., a non-adult grammar) to one that generates only the declarative (i.e., adult grammar) suggests that extra-syntactic factors do not interact with the adult grammar to generate an imperative interpretation. However, just as for production, processing limitations may also influence children's comprehension (Trueswell et al., 1999; Conroy et al., 2009; Lidz et al., 2017; Gerard et al., 2018). If limited processing resources can cause children to drop the subject, then these (or similar) limitations may also cause children to accept null subjects in comprehension. The prediction in (6) for production may, therefore, be extended to comprehension, in (8):

(8) If children accept null subjects as grammatical because of limited processing resources, then ***decreasing*** the processing load of a given context should ***decrease*** the likelihood of accepting null subjects as grammatical.

We test this prediction in the current study.

Current study

In this study, we compare children's comprehension of sentences with null subjects across languages. We focus on English – a non-null subject language – and Italian – a null subject language. To assess children's comprehension during the null subject stage, we use the central fixation paradigm (Maye et al., 2002; Shi et al., 2006). We compare this comprehension measure with children's production, using an imitation paradigm.

The Central Fixation Paradigm

For the current study's comprehension measure, we use a central fixation paradigm that measures children's attention to audio and visual stimuli. In the paradigm, children are presented with a continuous animation paired with a set of sentences (e.g., null subject sentences). Different conditions include different sentences (e.g., sentences with null subjects vs. overt subjects), and the dependent measure is the time spent looking at animations in each condition. This paradigm differs from methods like the preferential-looking paradigm, which measures children's interpretations with respect to a visual context. Instead, a constant visual stimulus is presented without any relation to the auditory stimuli and is only used to capture children's attention. The central fixation paradigm may be deployed with a wide age range, depending on the type of visual stimuli (Li et al., 2023). For the target age range of 19–28 months, we will first validate this methodology to identify when children become sensitive to central fixation manipulation within this age range (discussed further below in the Methods section).

https://doi.org/10.1017/S0305000924000588 Published online by Cambridge University Press

Importantly, if children have acquired the adult grammar for null subjects, then different preferences will be observed across languages, depending on the acceptability of null subjects in the language. In the central fixation paradigm, infants may exhibit either a familiarity preference or a novelty preference (Hunter & Ames, 1988; Kidd et al., 2012; Kosie et al., 2023). For example, a familiarity preference predicts that children will exhibit longer looking times for the more natural subject form in their respective language, that is, overt subjects in English, and null subjects in Italian; the reverse preferences are predicted by a novelty preference.

In previous looking time studies, children have generally preferred grammatical sentences over ungrammatical sentences, that is, a familiarity preference (Santelmann & Jusczyk, 1998; Hohle et al., 2006; Perkins & Lidz, 2021; but see Saffran et al., 1996). However, in previous studies on null subjects in production, children have exhibited delayed reaction times for unnatural sentences (e.g., with null subjects in English); this behavior is consistent with a novelty preference (Valian et al., 1996; Valian & Aubry, 2005). As a result, we consider the possibilities of either a novelty or a familiarity preference for the current study's central fixation paradigm. Children's preferences may also vary depending on internal or external factors (Hunter & Ames, 1988; Kidd et al., 2012; Kosie et al., 2023), discussed further in the Analysis section.

In the previous study by Orfitelli and Hyams (2012), children who produced null subjects also accepted null subject sentences in comprehension. This result is consistent with a competence account of children's null subjects; however, the same pattern is predicted by a performance account if children's comprehension was influenced by processing limitations, for example, the processing demands of the task. If so, then lowering these demands should yield a different result.

The central fixation paradigm imposes lower demands in two ways. First, the test sentences are not interpreted with respect to an external discourse context. Rather, the sentences are unambiguously declarative (as opposed to, e.g., imperative or interrogative), with minimal discourse context. As a result, children's responses are based on the form of the sentence itself rather than the acceptability of the form within an external context, thus reducing the processing cost (Gerard, 2022; Gerard et al., 2018). Next, the central fixation paradigm uses looking time for each condition for measuring children's comprehension. Compared with an explicit behavioral measure, this implicit measure involves fewer steps between the presentation of the test sentence and the child's response. While an explicit response involves planning and producing a response to the test sentence, these additional steps are not required for the looking time measure, further reducing the processing cost (Adani & Fritzsche, 2015). In sum, the central fixation paradigm should more accurately reflect children's grammatical knowledge than a task with higher demands. In particular, for children with a non-null subject grammar, this predicts lower acceptability for sentences with null subjects. In contrast, these sentences should be acceptable for children with a null subject grammar. These contrasts in acceptability should then be reflected in the central fixation paradigm as inverse patterns of looking behavior, that is, preferences for opposite subject forms.

In addition, we use an imitation paradigm to compare children's comprehension of null subjects with their production. As in previous studies with an imitation paradigm, we expect that children will produce null subject sentences. We also expect a higher rate of null subjects for test sentences with a pronominal subject than with a full lexical noun phrase (NP) (Gerken, 1991; Valian et al., 1996; Valian & Aubry, 2005). Importantly, for children who produce null subject sentences, competence and performance accounts

https://doi.org/10.1017/S0305000924000588 Published online by Cambridge University Press

make different predictions about children's comprehension, as well as the correspondence between comprehension and production.

Predictions

Under a competence account, children's null subjects are due to a grammar that generates null subject sentences. This predicts that children in the null subject stage will also accept null subject sentences in comprehension, and will do so as long as they produce null subject sentences (Orfitelli & Hyams, 2012). For the central fixation paradigm, if children have a familiarity preference, then Italian children will prefer the videos with null subject sentences over the videos with overt subject sentences. Similarly, English children who drop the subject in the imitation task should also exhibit a preference for null subject sentences. The reverse pattern is predicted for a novelty preference: children who drop the subject in the imitation task should prefer overt subject sentences in the central fixation paradigm. Importantly, under a competence account, the same behavior should be observed in comprehension and production, with English children who drop the subject in the imitation task exhibiting the same preference as Italian children in the central fixation paradigm.

Under a performance account, children's null subjects are because of limited processing resources. This predicts that lowering the processing load of the context will reduce the likelihood that children with a non-null subject grammar will accept null subjects as grammatical. If the processing load is sufficiently reduced for the central fixation paradigm, then English and Italian children will distinguish between null and overt subject conditions by exhibiting *opposite*-looking behaviors – with a familiarity preference, English children will prefer overt subjects and Italian children null subjects, and vice versa with a novelty preference. *In addition*, null subjects in the imitation task should not predict a children's preferences in the central fixation paradigm: English children who drop the subject at higher rates in the imitation task should exhibit the same preference in the central fixation paradigm as children with low/adultlike rates of subject omission. These predictions are spelled out in Table 1.

Importantly, the English and Italian children serve as controls for each other: a preference in one language alone cannot support either account; rather, the competence account predicts the same preference for both Italian children and English children who omit the subject, while the performance account predicts opposite preferences.

Table 1. Predictions for the central fixation and imitation task

	Account	
Prediction	Competence	Performance
Central fixation familiarity preference	English: Null subjects Italian: Null subjects	English: Overt subjects Italian: Null subjects
Central fixation novelty preference	English: Overt subjects Italian: Overt subjects	English: Null subjects Italian: Overt subjects
Comprehension matches production?	Yes	No

https://doi.org/10.1017/S0305000924000588 Published online by Cambridge University Press

Experiment 1: central fixation paradigm validation

In this study, we aim to investigate the source of children's null subjects in production, by testing the predictions of competence and performance accounts for comprehension and production. Comprehension will be assessed using a central fixation paradigm in an unmoderated online context, which has not been used previously to test children's acquisition of null subjects. While children in the target age range of 19–28 months are expected to produce null subjects, the novel use of the central fixation paradigm introduces some uncertainty regarding the age at which they will start to differentiate between null and overt subjects for this methodology. To address this uncertainty, we will first validate children's behavior on the central fixation paradigm for the full 19–28 month age range to confirm the target range, such that a familiarity (or novelty) preference is observed throughout this range.

Participants

For the initial validation of the central fixation paradigm, we will collect data from children between the ages of 19–28 months in English and Italian. This age range is determined based on two factors. First, previous studies have reliably observed higher rates of subject omission within this age range, both in naturalistic studies (Valian, 1991; Aronoff, 2002; Lorusso et al., 2005), and in experimental paradigms – including the elicited imitation task to be replicated in the current study (Gerken, 1991; Valian et al., 1996; Valian & Aubry, 2005). Second, the lower end of the age range is based on looking time evidence for children's sensitivity to missing arguments (Seidl et al., 2003; Gagliardi et al., 2016; Perkins & Lidz, 2020, 2021), suggesting that children have the requisite syntactic knowledge to track null subjects in the input by 19 months.

The participants will be children raised in a monolingual English or monolingual Italian environment. They will be recruited via social media or Children Helping Science, a platform for remote data collection created by the 2023 merger of Lookit (Scott et al., 2017; Scott & Schulz, 2017) and Children Helping Science (Sheskin et al., 2020). Participants will be compensated with a £5 or $5 (English) or €5 (Italian) gift card for completing the study.

Design

The test trials of the central fixation paradigm will include sentences with null subjects and with overt subjects as in (9), with person and number held constant in the third person singular:

(9) The girl is eating some soup.
   a. She spills the bowl.
   b. ∅ spills the bowl.

Importantly, both conditions will include a context sentence, which is required for both types of subjects in (9). The pragmatic context is therefore kept constant for both conditions of the independent variable, subject form (null/overt).

Additionally, the third-person singular inflection on the verb in (9) supports an unambiguous declarative interpretation. This contrasts with other present verb forms which are homophonous with the imperative form "spill." Meanwhile, these

https://doi.org/10.1017/S0305000924000588 Published online by Cambridge University Press

Table 2. Design for the central fixation paradigm

	Subject form (following "The girl is eating some soup …")	
Language	Overt subject sentences "… She spills the bowl"	Null subject sentences "… ∅ spills the bowl"
English (non-null subject)	Natural	Unnatural
Italian (null subject)	Unnatural	Natural

interpretations can also depend on the frequency of the imperative at the lexical level: verbs that occur more often in the imperative are more likely to be interpreted as such (Bybee, 2002; Brown & Rivas, 2012). Therefore, to minimize the likelihood of the imperative interpretation, the central fixation paradigm sentences will include verbs that are less likely to occur as imperatives in a speech to children.[1] The full set of materials is available at the project OSF site at https://osf.io/86tu4/.

The null and overt conditions (within subjects) will be presented in both English and Italian (between subjects), a second independent variable. For the context in (9), a different subject form is acceptable for each language – null for Italian, overt for English. This design is spelled out in Table 2.

Audio passages will be constructed with corresponding items in each condition, with matched timing for corresponding null and overt sentences. Each passage will contain four sentences with a basic structure, as in (10) (adapted from Santelmann and Jusczyk (1998)):

(10) In the kitchen, the mom is talking on the phone. {She/∅} laughs at a joke. The girl is eating some soup. {She/∅} spills the bowl.

Each child will hear passages with overt subjects and passages with null subjects, with passages lasting an average of 15 seconds and with the same overt subject form used throughout a given passage (e.g., "she" in (10); Tincoff et al., 2000). The passages will alternate between conditions, with the first condition counterbalanced between null and overt subjects. The videos with the first two passages will serve as "familiarization" trials, while the remaining passages serve as "test" trials (with items randomly assigned to the familiarisation phase). While looking times are expected to be longer for the familiarization phase than the test phase, the same contrast between natural and unnatural sentences is predicted within each phase. To determine the appropriate number of "test" trials for this central fixation paradigm design with the target age range, we will pilot the procedure starting with 12 trials (i.e., 6 natural and 6 unnatural), and adjust the trials based on how many are completed during piloting.

Each trial will start with an attention-getter to direct the child's attention to the screen. The experiment then advances automatically to the test trial (Figure 1), where children will hear the passage with four sentences while viewing a repeating animation (Li et al., 2023). Parents are instructed not to interfere or influence their child's behavior, and if possible to wear headphones for the duration of the study.

Children's looking behavior will be recorded via a laptop webcam, and each frame will be coded manually from the trial onset for looks at the screen. Preference for null or overt

[1]Identified with imperative coding from the childesdb package in R (Sanchez et al., 2019) of all CHILDES English corpora.

https://doi.org/10.1017/S0305000924000588 Published online by Cambridge University Press

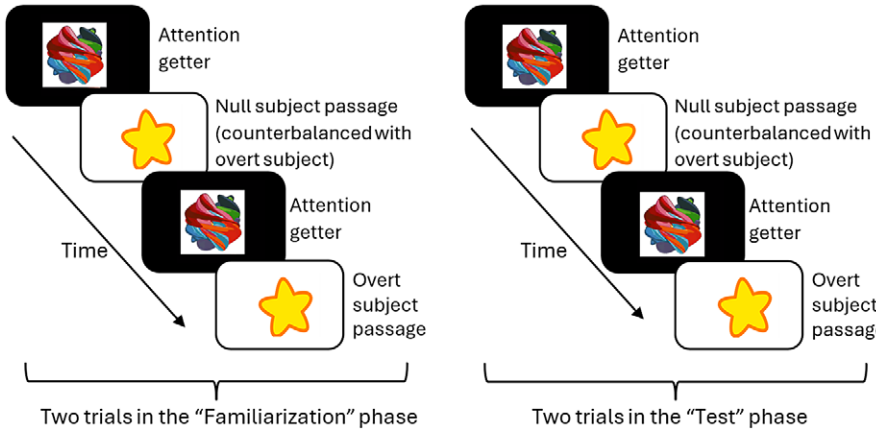

Figure 1. Example of the trial structure of the central fixation, with the familiarization phase starting with a natural condition, followed by alternating conditions. The first condition will be counterbalanced between natural and unnatural conditions, which will in turn determine the conditions of subsequent trials.

subjects will be determined based on looking times to the passages in each condition, which will be used to identify the target age range for the comparison with children's production in Experiment 2. The identification criteria are discussed further in the following analysis section.

Experiment 1 analysis plan

The analysis of children's looking time in the central fixation paradigm will aim to identify the earliest age at which children distinguish between null and overt subject forms in comprehension. Following data pre-processing, we will analyze children's looking time across the 19–28 month age range in English and Italian.

Data preprocessing and criteria for exclusion

Video data from the test trials will be coded without the corresponding audio by the research team and trained research assistants. Test trial videos will be isolated and exported to ELAN (ELAN (Version 6.8) [Computer Software], 2024)), developed by Max Planck Institute for Psycholinguistics for media annotation. With ELAN, test trial videos will be coded to indicate whether the child was looking at the screen, with the following exclusion criteria:

- The child cried or fell asleep during the trial
- The parent interfered during the child's response or paused the trial
- The child's response could not be scored, either because of blurriness or moving out of frame

A participant's data from the central fixation paradigm will not be included in the analysis if all trials in one of the conditions (natural/unnatural) are excluded based on these criteria. The total looking time will be determined for each trial by summing over

https://doi.org/10.1017/S0305000924000588 Published online by Cambridge University Press

coded frames, to produce the dependent measure for the central fixation paradigm analysis.

Statistical model

To analyze children's looking times in the central fixation paradigm, we will use a mixed effects linear regression with the continuous dependent variable of total looking time at the screen for each trial (determined by the manual coding by video frame). The fixed effects will be the central fixation paradigm subject form (overt/null, within-subjects) and language (English/Italian), as well as age in 2-month bins to identify the target age range. Random effects will include participants and the central fixation paradigm items.

The planned model will include the maximal random effects structure, with terms removed as necessary in case of non-convergence. The model will be fit using the lmerTest package in R (Kuznetsova et al., 2020), and will be specified as:

(11) looking_time ~ cfp_subject_form * language * age_2months +
(1 + cfp_subject_form | participant) +
(1 + cfp_subject_form * language * age_2months | item)

For each language, we expect an effect of subject form for children who distinguish between these conditions; ultimately, we expect an interaction between language and subject form because of a reverse preference for English versus Italian; however, a grammaticality account predicts the same preference for children who omit the subject in production. Finally, to identify the age at which an effect of looking time on a subject form is first observed in each language, we will conduct post hoc tests by language and age for the model in (11). This target age range will then be used in Experiment 2.

A key assumption for this analysis is that at least one of the bins will yield a preference (either familiarity or novelty), especially for the older ages. However, if no preference is observed across the age range, this would be problematic as we then have no target range for Experiment 2. In the case of this outcome, we would first need to identify why no preference has been observed for this instance of the central fixation paradigm before proceeding to Experiment 2. We would then adjust the parameters of the paradigm accordingly in Experiment 2.

Power analysis

To determine the sample size for Experiment 1, we simulated 500 experiments with each of 24, 64, 120, and 200 participants in each language (the simulation code is available at https://osf.io/86tu4/). Under this approach, we identify the proportion of experiments in which the relevant effects are significant ($p < .05$), using the model in (12) with random intercepts only to minimize non-convergence:

(12) looking_time ~ cfp_subject_form * language * age_2months +
(1 | participant) + (1 | item)

The experiments were simulated by generating children's looking times for the central fixation paradigm from a normal distribution truncated at 0 (i.e., no negative looking times). The parameters for the distribution were based on the means and standard

https://doi.org/10.1017/S0305000924000588 Published online by Cambridge University Press

deviations reported by Santelmann and Jusczyk (1998) for the natural and unnatural conditions.[2] Importantly, the parameters for natural and unnatural conditions were crossed with the language conditions for central fixation paradigm subject form; that is, we used the mean and standard deviation for the natural condition to generate looking times for the overt subject condition for English, and the null subject condition for Italian. As a result, we expect a subject form to interact with language.

Meanwhile, the primary effect of interest for this power analysis is the effect of subject form, for each 2-month bin within each language. With an age range of 10 months (19–28 months), the simulations included five 2-month bins, with the same generation procedure and parameters used for each bin.[3] To isolate the effect of subject form in each language, we included post hoc tests on the model in (12) by age and language. These tests were included for each experiment in the power analysis, using the joint_tests function from the emmeans package in R (Lenth et al., 2022).

The results of the simulations are in Table 3, which presents the proportion of experiments where the post hoc tests on subject form were significant. Because of the simulation procedure with different parameters for English and Italian, this effect for English would be because of longer looking times for the overt subject condition, while the significant effects for Italian are because of longer looking times for the null subject condition. Overall, we achieved 80% power with 120 participants in each language, that is, 24 participants for each 2-month age bin. This sample size is therefore planned for Experiment 1, to identify the optimal age range for the central fixation paradigm to investigate children's comprehension of sentences with null and overt subjects.

Table 3. Simulations of central fixation paradigm looking times in English and Italian: proportion of experiments in which the effect of subject form was significant (80% power with 120 participants, in bold)

		Number of participants in each language			
Language	Age bin	24	64	120	200
English	1	0.734	0.78	**0.772**	0.884
	2	0.694	0.736	**0.816**	0.872
	3	0.692	0.726	**0.802**	0.876
	4	0.686	0.74	**0.796**	0.862
	5	0.7	0.716	**0.806**	0.884
Italian	1	0.734	0.78	**0.772**	0.884
	2	0.694	0.736	**0.816**	0.872
	3	0.692	0.726	**0.802**	0.876
	4	0.686	0.74	**0.796**	0.862
	5	0.7	0.716	**0.806**	0.884

[2]While the paradigm used by Santelmann and Juscyk was a headturn preference paradigm rather than central fixation, the looking times are comparable to those observed for central fixation (Shi et al., 2006).

[3]We expect that this sensitivity will be observed at all bins following the earliest 2-month bin (e.g., if children are sensitive to the contrast at 23–24 months, then we also expect to observe the contrast at 25–26

https://doi.org/10.1017/S0305000924000588 Published online by Cambridge University Press

Results

To be added.

Experiment 2: comprehension and production

Experiment 2 will address the core aim of the current study: to investigate the source of children's null subjects in production, by testing the predictions of competence and performance accounts for comprehension and production. Comprehension will be assessed using the central fixation paradigm, replicating the effect with the age range identified in Experiment 1 and comparing this with children's production via an elicited imitation task. Both tasks will be conducted in an unmoderated online context as in Experiment 1, with an online adaptation of previous in-person instances of the imitation task (Valian et al., 1996; Valian & Aubry, 2005). To confirm the validity of this online adaptation, we will also measure children's MLU, which reliably predicts children's null subjects in production.

Participants

The participants will be children raised in a monolingual English or monolingual Italian environment, with the age range based on the results of Experiment 1. As for Experiment 1, participants will be recruited via social media or Children Helping Science, a platform for remote data collection created by the 2023 merger of Lookit (Scott et al., 2017; Scott & Schulz, 2017) and Children Helping Science (Sheskin et al., 2020). Participants will be compensated with a $10 or £10 (English) or €10 gift (Italian) gift card for completing the study (both tasks and the MLU measurement).

Design

The study will use two methodologies – the central fixation paradigm and imitation task – to test children's comprehension and production of null subjects, respectively. The central fixation paradigm design will be identical to the design in Experiment 1 and will be directly compared with the measure for children's production (the imitation task).

Imitation task design

The imitation task will be a replication of Valian et al. (1996), with the same participants as the central fixation paradigm. Children will be prompted to repeat sentences with pronominal and lexical NP subjects, as in (13):

(13) a. She spills the bowl
b. The girl spills the bowl

In addition to this manipulation of subject form, the sentences will vary in VP length, with a higher likelihood of dropping the subject expected for longer VPs:

and 27–28 months). For the purposes of the power analysis, however, we focus on each bin individually in order to identify the sample size needed to detect children's sensitivity within this age range.

https://doi.org/10.1017/S0305000924000588 Published online by Cambridge University Press

Table 4. Design for the imitation task

	VP length	
NP type	Short VP	Long VP
Lexical subject	The girl spills the bowl (n = 6)	The girl spills the bowl on the floor (n = 6)
Pronominal subject	She spills the bowl (n = 6)	She spills the bowl on the floor (n = 6)

(14) a. She spills the bowl
b. She spills the bowl on the floor

Children will be presented with six sentences in each condition, for a total of 24 test sentences (Table 4).

Children will be prompted to imitate the sentences in an imitation game. In the game, sentences to be imitated are first read off of a card. After the child attempts an imitation, the card is "eaten" by an animated basket. For the online adaptation of this game, an on-screen character will be introduced to read the sentences on each card. The sentences will be pre-recorded for consistency across children, and cards will be presented with simple pictures rather than the sentences themselves to avoid an additional presentation of the sentence from the parent. If no imitation attempt is made initially, the parent may select a prompt for a second attempt, by presenting the recorded sentence a second time.

To assess the validity of this online adaptation of the elicited imitation task, we will measure children's MLU. This measure has predicted children's rate of subject use in in-person contexts, and replicating this correlation will allow for greater confidence in the production measure as a predictor of children's looking behavior in the central fixation paradigm.

MLU measurement

Children's MLU will be measured by recording a 20-minute conversation between a child and their parent before their participation in the central fixation paradigm and imitation task. The conversation will be facilitated by providing an online copy of Richard Scarry's Best Word Book Ever, a picture book with minimal text that prompts children to talk about the content of the pictures.

The recordings will be transcribed by native speakers of English and Italian for the respective languages, and MLU calculated as the mean number of words in a child's utterance. Note that MLU in English is commonly based on the mean number of morphemes; however, MLU in words will allow for greater comparability between English and Italian, while still being a reliable predictor of subject use (Valian, 1991).

This recording will be completed at the start of the online study session and may be followed by a break as needed. Next, children will complete the imitation task or the central fixation paradigm (counterbalanced), followed by a second break and the second of the two tasks.

https://doi.org/10.1017/S0305000924000588 Published online by Cambridge University Press

Experiment 2 analysis plan

The analysis aims to determine whether children's comprehension matches their production by addressing the predictions in Table 1. Following the data preprocessing, we will first analyze children's production to assess the reliability of the online imitation task. Next, we will analyze children's behavior in the central fixation paradigm by using the production data to predict children's comprehension.

Data preprocessing and criteria for exclusion

As in Experiment 1, the central fixation paradigm will involve manual coding to produce children's null subject measures, with the same preprocessing procedure and exclusion criteria. Coding in ELAN will require blinding to the experimental condition; therefore, an index will be constructed before exporting the videos to link the video coding with other participant measures across tasks.

Meanwhile, manual coding for the imitation task is based on criteria described in Valian et al. (1996): each trial will be scored based on (a) whether an imitation was attempted, and (b) whether the subject was included in the imitation attempt or omitted.[4] Trials with no attempt will be excluded, and participants with four or fewer imitation attempts will not be included in the analysis.

Imitation task model

For the production analysis, we will use a mixed effects logistic regression with the binary dependent variable of a null (0) or overt (1) subject of the imitated sentence, fixed effects of language (English/Italian), NP type (pronoun/lexical NP), and VP length (short/long), and random effects of participants and items.

The planned model will include the maximal random effects structure, with terms removed as necessary in case of non-convergence. The model will be fit using the lmerTest package in R (Kuznetsova et al., 2020), and will be specified as:

(15) subject_included ~ language * np_type * vp_length
(1 + np_type * vp_length | participant) +
(1 + language * np_type * vp_length | item)

For the imitation task, we expect to replicate previous effects observed in in-person contexts; these include the following:

- A main effect of language: more subjects will be produced in English than in Italian (Grinstead, 2004; Valian, 1991)
- A main effect of NP type: more subjects will be produced with a lexical subject than with a pronominal subject (Valian, 1991; Valian et al., 1996; Valian & Aubry, 2005)
- A main effect of VP length: more subjects will be produced with a short VP than a long VP (Bloom, 1990; Valian, 1991; Valian & Aubry, 2005)

[4]Children's attempts may also deviate from the target if any additional content is omitted, or if the content itself is changed (e.g., if children use a different verb). For the imitation task coding, we will not separate these errors from the main analysis based on subject omission; however, an exploratory analysis of error type will be included if they are prevalent in the data.

https://doi.org/10.1017/S0305000924000588 Published online by Cambridge University Press

- A correlation with MLU: children with a greater MLU will be more likely to produce the subject (Valian, 1991; Valian et al., 1996)

While these effects are reliable in previous studies, they have received different explanations under the processing and competence accounts. For the purposes of the current study, we aim to replicate these effects to confirm the validity of the imitation task for providing a measure of children's subject omissions in production. The lack of these effects may require modifications to the imitation task, particularly for NP type and the correlation with MLU.

While null subjects are predicted at different rates across conditions on the imitation task, the analysis across tasks will consider just one rate for each participant because of the complexity of the cross-task comparison model (discussed in the following section). Therefore, the cross-task analysis will include only the pronominal condition of the NP type factor, as this condition is likely to have a broader distribution of subject omissions than the lexical condition. Meanwhile, the inclusion of the VP length condition will depend on whether the effect of VP length is replicated in the imitation task: if a significant difference is observed between short and long VPs, then the cross-task analysis will include only long VPs; however, if this effect is not replicated then the cross-task analysis will include responses from both VP types (collapsed across conditions).

Cross-task analysis

For the analysis across tasks, we will use a mixed effects linear regression with the continuous dependent variable of total looking time at the screen for each trial (determined by the manual coding by video frame), as in Experiment 1. The fixed effects will be the central fixation paradigm subject form (overt/null), language (English/Italian), and the proportion of subjects included in the imitation task (continuous), with participants and the central fixation paradigm items as random effects.

As above, the planned model will include the maximal random effects structure, with terms removed as necessary in case of non-convergence. The model will be fit using the lmerTest package in R and will be specified as:

(16) looking_time ~ cfp_subject_form * language * subjects_included +
(1 + cfp_subject_form | participant) +
(1 + cfp_subject_form * language | item)

Potential results and implications. Under both the competence and performance accounts, adultlike comprehension will be realized as a preference – either familiarity (natural condition) or novelty (unnatural condition). As these conditions are different for English and Italian, the inverse pattern of behavior is predicted, regardless of preference type (familiarity vs. novelty): either (a) English children will look longer for the overt subject passages and Italian children will look longer at the null subject passages, that is, a familiarity preference, or (b) English children will look longer for the null subject passages and Italian children will look longer at the overt subject passages, that is, a novelty preference. In the context of the central fixation paradigm regression model, from either of these patterns, we would expect a 2-way interaction between subject form and language.

https://doi.org/10.1017/S0305000924000588 Published online by Cambridge University Press

Importantly, the competence and performance accounts make different predictions about ***when*** this 2-way interaction between subject form and language will emerge. This is because of the difference between competence and performance accounts in the onset of adultlike knowledge: under performance accounts, adultlike knowledge is acquired early; under competence accounts, adultlike knowledge is acquired late. Since the interaction between subject form and language reflects adultlike knowledge, a performance account predicts that this interaction will be observed early – before children start to produce multi-word utterances. Therefore, the interaction between subject form and language in the central fixation paradigm should ***not*** change with the proportion of subjects produced in the imitation task (Figure 2).

In contrast, a competence account predicts that the interaction between subject form and language will not be observed until after English children's rate of subjects produced has converged with the rate of adults. Before this, that is, while the English children are in the null subject stage, the lower rate of subject production is because of a null subject grammar under a competence account. The resulting pattern predicted by the competence account is as follows:

- English children with low rates of subject production will exhibit the same looking preference as Italian children, who also have a null subject grammar
- in contrast, English children with higher subject production rates will exhibit the opposite preference (Figure 3).

That is, the emergence of the 2-way interaction between subject form and language (as an indicator of adult grammar) will depend on children's null subjects in the imitation

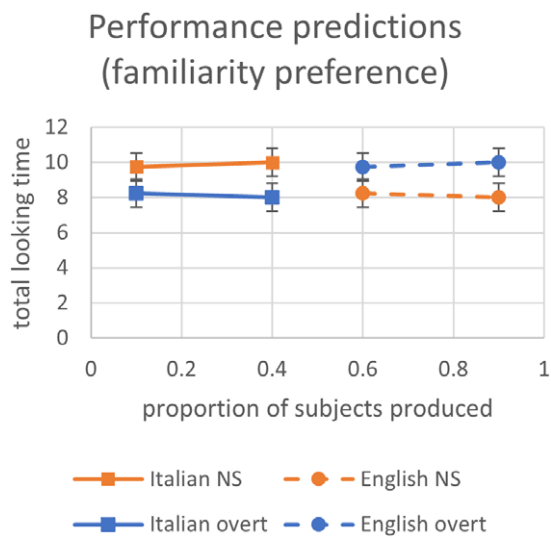

Figure 2. Predictions of a performance account, assuming a familiarity preference – Italian children look longer for null subject (NS) passages, while English children look longer for overt subject passages: a 2-way interaction between the language (English/Italian) and the subject form of the passage (null/overt). A novelty preference predicts reverse-looking patterns, but the same 2-way interaction.

https://doi.org/10.1017/S0305000924000588 Published online by Cambridge University Press

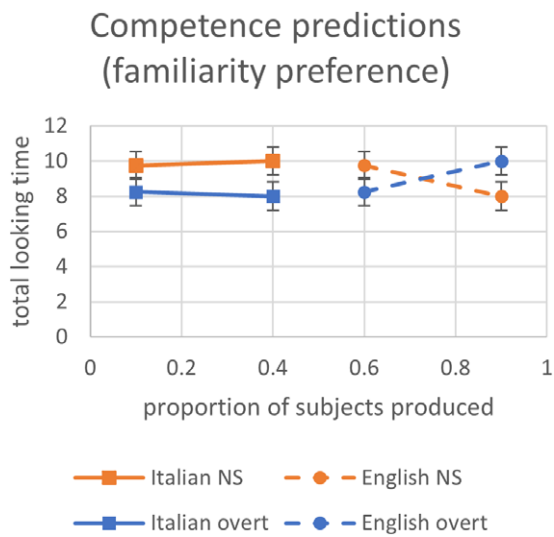

Figure 3. Predictions of a competence account, assuming a familiarity preference – Italian children look longer for null subject (NS) passages, as do English children who produce *fewer* subjects. In contrast, English children who produce more subjects look longer for overt subject passages: a 3-way interaction between the language (English/Italian), the subject form of the passage (null/overt), and the proportion of subjects produced. A novelty preference predicts reverse-looking patterns, but the same 3-way interaction.

task. In the context of the central fixation paradigm regression model, this predicts a 3-way interaction between subject form, language, and subject omissions.

This prediction by the competence account for the same preference in Italian children and English children in the null subject phase is what motivates the use of total looking times rather than the looking time difference between overt and null subject passages: if the same non-zero preference is predicted in both languages, then a model with difference times does not disambiguate between this non-zero preference and a pattern with no preference in either language: in both patterns, the same preference is exhibited in across languages, meaning that for both patterns there is no main effect of language. To differentiate between a zero preference and a non-zero preference across languages, an additional statistical test against zero would be required for each language. We therefore use total looking times to account for both outcomes in a single model with subject omissions.

It is also important to note that the above predictions all involve a preference for one subject form condition over another, rather than no preference between conditions. This is because a no preference result is consistent with both accounts before children have acquired either grammar. However, this result would also be expected if the central fixation paradigm were too demanding, thus motivating the initial validation of the paradigm in Experiment 1. By restricting the age range to a population that is expected to exhibit a preference in the central fixation paradigm *and* to omit subjects in production, we avoid this ambiguous outcome with respect to the competence and performance accounts. However, the cross-task analysis for Experiment 2 depends on a replication of the looking time preferences observed in Experiment 1. If these preferences are not replicated in Experiment 2, then we would carry out a more fine-grained analysis of the looking time differences between the two experiments to better understand the conflicting

https://doi.org/10.1017/S0305000924000588 Published online by Cambridge University Press

results. This analysis can consider children's responses to the imitation task, which may still predict looking times despite the absence of group-level effects.

Variation in novelty and familiarity preferences. Finally, children's preferences for novel or familiar sentences may vary, depending on various factors (Hunter & Ames, 1988; Kidd et al., 2012; Kosie et al., 2023). In the context of the cross-task comparison, the same effects are predicted regardless of whether children exhibit a novelty preference or a familiarity preference; however, these predictions depend on a consistent preference throughout the sample. This consistency may be challenged by two key factors, namely, the broad age range of the sample and any unintended variation in task complexity.

To address the issue of variation in preference by age, exploratory analysis will be conducted for children's preference by age in English and Italian. For each language, children's preferences for the novel and familiar conditions will be analysed by age. If children's preferences are predicted by age, then the cross-task analysis will account for age in the looking time measure.

Similarly, children's responses may be affected by the complexity of the task, which may vary over the course of the experiment session. If children's preferences in the familiarization phase are significantly different from their preferences in a test phase, then each phase will be analysed separately to minimize the risk of including different preferences in the cross-task analysis.

Power analysis

To determine the sample size for Experiment 2, we used the same simulation-based approach as for Experiment 1 to identify the proportion of experiments in which the relevant effects were significant ($p < .05$). We simulated 500 experiments with each of 24, 64, 120, and 200 participants in each language (The simulation code is available at https://osf.io/86tu4/), using the model in (17) with random intercepts only to minimize non-convergence:

(17) looking_time ~ cfp_subject_form * language * subjects_included + (1 | participant) + (1 | item)

The experiments were simulated by generating the following distributions:

- For the imitation task: the proportion of subjects produced, from a truncated normal distribution between 0 and 1
  - For English children in the null subject stage, generated from the means and standard deviations reported by Valian et al. (1996)
  - For English children with adultlike rates of subject use, also generated from the means and standard deviations reported by Valian et al. (1996)
  - For Italian children, generated from the means and standard deviations reported by Valian (1991)
- For the central fixation paradigm: looking times in seconds, from a normal distribution truncated at 0 (i.e., no negative looking times)
  - Looking times for a natural condition, generated from the means and standard deviations reported by Santelmann and Jusczyk (1998)
  - Looking times for an unnatural condition, also from Santelmann and Jusczyk (1998)

https://doi.org/10.1017/S0305000924000588 Published online by Cambridge University Press

The simulations were repeated for the predictions of a performance account – in which children's underlying knowledge is adultlike – and for a competence account – where children who omit the subject do so because of having a null subject grammar. To distinguish between children with a null subject grammar versus a non-null subject grammar, we used the manipulation of the subject form: for children with a null subject grammar, looking times generated for the natural condition were linked with "null subject" passages, while unnatural condition looking times were linked with "overt subject" passages; the reverse links were made for a non-null subject grammar (i.e., assuming a familiarity preference; the same result is obtained for the power analysis with a novelty preference).

The key difference between the two sets of simulations was in the looking time preference for English children in the null subject stage: under a performance account, these children have a non-null subject grammar; therefore, in this simulation, the natural

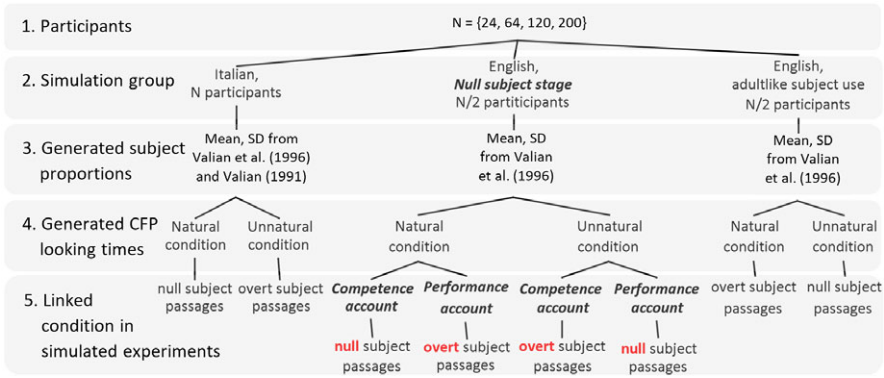

Figure 4. Simulation procedure – the proportions of subjects produced in the imitation task (3) were generated from the means and standard deviations reported in the studies by Valian (1991) and Valian et al. (1996), while the looking times for the central fixation paradigm (4) were generated from the means and standard deviations reported by Santelmann and Jusczyk (1998) (Table 5 and 6).

Table 5. Simulations with a competence account (the null subject stage is because of a null subject grammar): proportion of experiments in which the fixed effects and interactions were significant; predicted effects in bold

	Number of participants in each language			
Model term	24	64	120	200
cfp_subject_form	**0.782**	**0.894**	**0.948**	**0.994**
language	0.028	0.03	0.046	0.032
subjects_included	0.042	0.05	0.056	0.048
cfp_subject_form : language	0.626	0.658	0.768	0.856
cfp_subject_form : subjects_included	0.718	0.748	0.838	0.91
language : subjects_included	0.032	0.044	0.054	0.054
cfp_subject_form : language : subjects_included	**0.648**	**0.746**	**0.824**	**0.89**

https://doi.org/10.1017/S0305000924000588 Published online by Cambridge University Press

Table 6. Simulations with a performance account (the null subject stage is because of performance limitations on the deployment of the adult grammar): proportion of experiments in which the fixed effects and interactions were significant; predicted effects in bold

	Number of participants in each language			
Model term	24	64	120	200
cfp_subject_form	0.524	0.534	0.578	0.524
language	0.028	0.03	0.046	0.032
subjects_included	0.042	0.05	0.056	0.048
cfp_subject_form : language	**0.698**	**0.762**	**0.84**	**0.906**
cfp_subject_form : subjects_included	0.642	0.572	0.608	0.566
language : subjects_included	0.032	0.044	0.054	0.054
cfp_subject_form : language : subjects_included	0.588	0.588	0.62	0.582

condition was linked to "overt subject" passages. This contrasts with the predictions of a competence account, where children in the null subject stage have a null subject grammar; the natural condition was therefore linked to "null subject" passages. This procedure is spelled out in Figure 4, and the results of the simulations (the proportion of experiments in which the effects from (17) were significant) are in Table 5 (for the predictions of a competence account) and Table 6 (for the predictions of a performance account).

For the 3-way interaction predicted by a competence account between subject form, language, and the proportion of subjects produced in the imitation task, we achieved 82.4% power with 120 participants in each language group. Similarly, we achieved 84% power with 120 participants for the 2-way interaction predicted by a performance account between subject form and language.

To confirm that these interactions reflect the predicted effects for each account, we included post hoc tests for each simulation alongside the mixed effects model. The post hoc tests also support a sample of 120 participants (Tables 7 and 8):

Table 7. Post hoc tests of simulations with a competence account, by language: proportion of experiments in which the post hoc tests by language were significant; predicted effects in bold

		Number of participants in each language			
Language	Model term	24	64	120	200
English	cfp_subject_form	0.676	0.77	0.844	0.914
	subjects_included	0.034	0.046	0.056	0.044
	cfp_subject_form : subjects_included	**0.752**	**0.87**	**0.964**	**0.992**
Italian	**cfp_subject_form**	**0.714**	**0.84**	**0.92**	**0.972**
	subjects_included	0.04	0.056	0.046	0.05
	cfp_subject_form : subjects_included	0.6	0.594	0.584	0.614

https://doi.org/10.1017/S0305000924000588 Published online by Cambridge University Press

Table 8. Post hoc tests of simulations with a performance account, by language: proportion of experiments in which the post hoc tests by language were significant; predicted effects in bold

		Number of participants in each language			
Language	Model term	24	64	120	200
English	**cfp_subject_form**	**0.734**	**0.842**	**0.956**	**0.986**
	subjects_included	0.034	0.046	0.056	0.044
	cfp_subject_form : subjects_included	0.554	0.568	0.588	0.576
Italian	**cfp_subject_form**	**0.712**	**0.838**	**0.92**	**0.972**
	subjects_included	0.04	0.056	0.046	0.05
	cfp_subject_form : subjects_included	0.608	0.596	0.584	0.618

- Competence account:
  - English: subject form * subjects included, 96.4% power
  - Italian: subject form, 92% power
- Performance account:
  - English: subject form, 95.6% power
  - Italian: subject form, 92% power

Results

To be added.

Timeline for completion of the study

The Stage 1 manuscript was submitted for review on 20 November 2023. Between March and August 2024, we addressed reviewer comments and constructed the study on the Children Helping Science platform. We will submit the study for approval on Children Helping Science and begin piloting in September, with data collection projected to begin in October 2024 for Experiment 1. Data collection is expected to be completed for Experiment 1 by December 2024, with data analysis completed by March 2025. Data collection will begin for Experiment 2 in January 2025, to be completed in June 2025, with data analysis to be completed in August 2025. We aim to submit the Stage 2 manuscript by December 2025.

Following Stage 1 in principle acceptance, we agree to register the approved protocol on the OSF.

Discussion

To be added.

Competing interest. The authors declare none.

https://doi.org/10.1017/S0305000924000588 Published online by Cambridge University Press

References

Adani, F., & **Fritzsche, T.** (2015). On the relation between implicit and explicit measures of child language development: Evidence from relative clause processing in 4-year-olds and adults. In *Proceedings of the 39th annual Boston university conference on language development*, Boston, MA: Cascadilla Press.

Allen, S. (2000). A discourse-pragmatic explanation for argument representation in child Inuktitut. *Linguistics*, **38**(3). https://doi.org/10.1515/ling.38.3.483

Allen, S., & **Schröder, H.** (2003). Preferred argument structure in early inuktitut spontaneous speech data. In J. W. Du Bois, L. E. Kumpf, & W. J. Ashby (Eds.), *Studies in discourse and grammar* (Vol. **14**, pp. 301–338). John Benjamins Publishing Company. https://doi.org/10.1075/sidag.14.14all

Allen, S., **Skarabela, B.**, & **Hughes, M.** (2008). Using corpora to examine discourse effects in syntax. In H. Behrens (Ed.), *Trends in language acquisition research* (Vol. **6**, pp. 99–137). John Benjamins Publishing Company. https://doi.org/10.1075/tilar.6.07all

Aronoff, J. M. (2002). *Null subjects in child language: Evidence for a performance account.* /paper/Null-subjects-in-child-language-%3A-evidence-for-a-Aronoff/5448c986b6d3c2bce03f205c57d96209e2effc6a

Biberauer, T. (2008). *The limits of syntactic variation*. John Benjamins Publishing.

Biberauer, T., **Holmberg, A.**, **Roberts, I.**, & **Sheehan, M.** (2009). *Parametric variation: Null subjects in minimalist theory*. Cambridge University Press.

Bloom, P. (1989). Why do children omit subjects? *Papers and Reports on Child Language Development*, **28**, 57–64.

Bloom, P. (1990). Subjectless sentences in child language. *Linguistic Inquiry*, **21**(4), 491–504.

Bloom, P. (1993). Grammatical continuity in language development: The case of subjectless sentences. *Linguistic Inquiry*, **24**(4), 721–734.

Bondaruk, A. (2001). Pro-drop in Polish and Irish. *Poznań Studies in Contemporary Linguistics: PSiCL.*, **37**, 35–52.

Brennan, J. (2008). Irish prepositions: Agreement and impoverishment. In *Proceedings of the 26th West Coast Conference on formal linguistics, Sous La Dir. de Charles B. Chang and Hannah J. Haynie* (pp. 105–113).

Brown, E. L., & **Rivas, J.** (2012). Grammatical relation probability: How usage patterns shape analogy. *Language Variation and Change*, **24**(3), 317–341.

Brown, E. L., & **Shin, N.** (2022). Acquisition of cumulative conditioning effects on words: Spanish-speaking children's [subject pronoun + verb] usage. *First Language*, **42**(3), 361–382. https://doi.org/10.1177/01427237211067574

Bybee, J. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*, **14**(3), 261–290.

Camacho, J. (2013). *Null subjects*.

Camacho, J. (2016). The null subject parameter revisited. In *The Morphosyntax of Portuguese and Spanish in Latin America* (pp. 27–48).

Conroy, A., **Takahashi, E.**, **Lidz, J.**, & **Phillips, C.** (2009). Equal treatment for all antecedents: How children succeed with principle B. *Linguistic Inquiry*, **40**(3), 446–486.

ELAN (Version 6.8) [Computer software]. (2024). Max Planck Institute for Psycholinguistics, The Language Archive. https://archive.mpi.nl/tla/elan

Erker, D., **Adli, A.**, **Bayley, R.**, **Guy, G.**, & **Beaman, K.** (2023, October 13). *Cognition and the human speech community: Large-scale evidence of crosslinguistic similarities in variable subject pronoun expression*. In *Relating rates and constraints in subject pronominal variation*. New ways of analyzing variation 51, Queens College, CUNY.

Gagliardi, A., **Mease, T. M.**, & **Lidz, J.** (2016). Discontinuous development in the acquisition of filler-gap dependencies: Evidence from 15- and 20-month-olds. *Language Acquisition*, **23**(3), 234–260.

Gerard, J. (2022). The extragrammaticality of the acquisition of adjunct control. *Language Acquisition*, **29**(2), 107–134. https://doi.org/10.1080/10489223.2021.1971231

Gerard, J., **Lidz, J.**, **Zuckerman, S.**, & **Pinto, M.** (2018). The acquisition of adjunct control is colored by the task. *Glossa: A Journal of General Linguistics*, **3**(1). https://doi.org/10.5334/gjgl.547

Gerken, L. (1991). The metrical basis for children's subjectless sentences. *Journal of Memory and Language*, **30**(4), 431–451.

Grinstead, J. (2004). Subjects and interface delay in child Spanish and Catalan. *Language*, **80**(1), 40–72.

Guasti, M. T. (2002). *Language acquisition: The growth of grammar*. MIT Press.

Hohle, B., **Schmitz, M.**, **Santelmann, L. M.**, & **Weissenborn, J.** (2006). The recognition of discontinuous verbal dependencies by German 19-month-olds: Evidence for Lexical and structural influences on children's early processing capacities. *Language Learning and Development*, **2**(4), 277–300. https://doi.org/10.1207/s15473341lld0204_3

Huang, C. J. (1984). On the distribution and reference of empty pronouns. *Linguistic Inquiry*, 531–574.

Hunter, M. A., & **Ames, E. W.** (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in Infancy Research*, **5**, 69–95.

Hyams, N. (1986). *Language acquisition and the theory of parameters* (Vol. **3**). Springer Netherlands. https://doi.org/10.1007/978-94-009-4638-5

Hyams, N. (1992). A reanalysis of null subjects in child language. In *Theoretical issues in language acquisition: Continuity and change in development* (p. 249).

Hyams, N. (2011). Missing subjects in early child language. In *Handbook of generative approaches to language acquisition* (pp. 13–52). Springer.

Hyams, N., & **Wexler, K.** (1993). On the grammatical basis of null subjects in child language. *Linguistic Inquiry*, 421–459.

Kidd, C., **Piantadosi, S. T.**, & **Aslin, R. N.** (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLoS ONE*, **7**(5), e36399.

Kosie, J., **Zettersten, M.**, **Abu-Zhaya, R.**, **Amso, D.**, **Babineau, M.**, **Baumgartner, H. A.**, **Bazhydai, M.**, **Belia, M.**, **Benavides-Varela, S.**, **Bergmann, C.**, **Berteletti, I.**, **Black, A. K.**, **Borges, P.**, **Borovsky, A.**, **Byers-Heinlein, K.**, **Cabrera, L.**, **Calignano, G.**, **Cao, A.**, **Cox, C. M. M.**, … **Lew-Williams, C.** (2023). *ManyBabies 5: A large-scale investigation of the proposed shift from familiarity preference to novelty preference in infant looking time.* PsyArXiv. https://doi.org/10.31234/osf.io/ck3vd

Kuznetsova, A., **Brockhoff, P. B.**, **Christensen, R. H. B.**, & **Jensen, S. P.** (2020). *lmerTest: Tests in linear mixed effects models* (Version 3.1–3) [Computer software]. https://cran.r-project.org/web/packages/lmerTest/index.html

Lenth, R. V., **Buerkner, P.**, **Giné-Vázquez, I.**, **Herve, M.**, **Jung, M.**, **Love, J.**, **Miguez, F.**, **Riebl, H.**, & **Singmann, H.** (2022). *emmeans: Estimated marginal means, aka least-squares means* (Version 1.8.2) [Computer software]. https://CRAN.R-project.org/package=emmeans

Li, D., **Zhong, S.**, & **Schuler, K.** (2023). *Proceedings of the 47th annualBoston University Conference on Language Development*, ed. **Paris Gappmayr** and **Jackson Kellogg**, 499–504. Somerville, MA: Cascadilla Press.

Lidz, J. (2022). Parser-grammar transparency and the development of syntactic dependencies. *Language Acquisition*, **30**, 1–12. https://doi.org/10.1080/10489223.2022.2147840

Lidz, J., **White, A. S.**, & **Baier, R.** (2017). The role of incremental parsing in syntactically conditioned word learning. *Cognitive Psychology*, **97**, 62–78.

Lorusso, P., **Caprin, C.**, & **Guasti, M. T.** (2005). Overt subject distribution in early Italian children. In *Proceedings of the twenty-ninth Boston university conference on language development*, 7.

Lutken, C. J., **Legendre, G.**, & **Omaki, A.** (2020). Syntactic creativity errors in children's Wh-questions. *Cognitive Science*, **44**(7), e12849. https://doi.org/10.1111/cogs.12849

Maye, J., **Werker, J. F.**, & **Gerken, L.** (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, **82**(3), B101–B111.

Omaki, A., & **Lidz, J.** (2015). Linking parser development to acquisition of syntactic knowledge. *Language Acquisition*, **22**(2), 158–192.

Orfitelli, R., & **Hyams, N.** (2012). Children's grammar of null subjects: Evidence from comprehension. *Linguistic Inquiry*, **43**(4), 563–590.

Perkins, L., & **Lidz, J.** (2020). Filler-gap dependency comprehension at 15 months: The role of vocabulary. *Language Acquisition*, **27**(1), 98–115. https://doi.org/10.1080/10489223.2019.1659274

Perkins, L., & **Lidz, J.** (2021). Eighteen-month-old infants represent nonlocal syntactic dependencies. *Proceedings of the National Academy of Sciences*, **118**(41), e2026469118.

Rizzi, L. (1982). Negation, wh-movement, and the null subject parameter. *An Annotated Syntax Reader*, **169**. https://as.nyu.edu/content/dam/nyu-as/linguistics/documents/Kayne%20Leu%20Zanuttini%200520%20-2013-proofs%20paper.pdf#page=181

Rizzi, L. (2018). Early null subjects and root null subjects. In *Syntactic Theory and First Language Acquisition*, Psychology Press. (pp. 151–176).

Rizzi, L. (2005). Grammatically-based target-inconsistencies in child language. In *Proceedings of the inaugural conference of GALANA* (pp. 19–49).

Saffran, J. R., **Aslin, R. N.**, & **Newport, E. L.** (1996). Statistical learning by 8-month-old infants. *Science*, **274**, 13.

Sanchez, A., **Meylan, S. C.**, **Braginsky, M.**, **MacDonald, K. E.**, **Yurovsky, D.**, & **Frank, M. C.** (2019). childes-db: A flexible and reproducible interface to the child language data exchange system. *Behavior Research Methods*, **51**(4), 1928–1941. https://doi.org/10.3758/s13428-018-1176-7

Santelmann, L., & **Jusczyk, P.** (1998). Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. *Cognition*, **69**(2), 105–134. https://doi.org/10.1016/S0010-0277(98)00060-2

Scott, K., **Chu, J.**, & **Schulz, L.** (2017). Lookit (part 2): Assessing the viability of online developmental research, results from three case studies. *Open Mind*, **1**(1), 15–29. https://doi.org/10.1162/OPMI_a_00001

Scott, K., & **Schulz, L.** (2017). Lookit (part 1): A new online platform for developmental research. *Open Mind*, **1**(1), 4–14. https://doi.org/10.1162/OPMI_a_00002

Seidl, A., **Hollich, G.**, & **Jusczyk, P. W.** (2003). Early understanding of subject and object Wh-questions. *Infancy*, **4**(3), 423–436.

Sheskin, M., **Scott, K.**, **Mills, C. M.**, **Bergelson, E.**, **Bonawitz, E.**, **Spelke, E. S.**, **Fei-Fei, L.**, **Keil, F. C.**, **Gweon, H.**, **Tenenbaum, J. B.**, **Jara-Ettinger, J.**, **Adolph, K. E.**, **Rhodes, M.**, **Frank, M. C.**, **Mehr, S. A.**, & **Schulz, L.** (2020). Online developmental science to foster innovation. Access, and impact. *Trends in Cognitive Sciences*, **24**(9), 675–678. https://doi.org/10.1016/j.tics.2020.06.004

Shi, R., **Werker, J. F.**, & **Cutler, A.** (2006). Recognition and representation of function words in English-Learning infants. *Infancy*, **10**(2), 187–198. https://doi.org/10.1207/s15327078in1002_5

Shlonsky, U. (2009). Hebrew as a partial null-subject language. *Studia Linguistica*, **63**(1), 133–157. https://doi.org/10.1111/j.1467-9582.2008.01156.x

Tincoff, R., **Santelmann, L.**, & **Jusczyk, P.** (2000). Auxiliary verb learning and 18-month-olds' acquisition of morphological relationships. In *Proceedings of the 24th annual Boston university conference on language development*, **2**, Cambridge MA: Cascadilla Press. (pp. 726–737).

Travis, C. E., & **Torres Cacoullos, R.** (2021). Categories and frequency: Cognition verbs in Spanish subject expression. *Languages*, **6**(3), 126.

Trueswell, J. C., **Sekerina, I.**, **Hill, N. M.**, & **Logrip, M. L.** (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition*, **73**(2), 89–134.

Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition*, **40**, 21–81. https://doi.org/10.1016/0010-0277(91)90046-7

Valian, V. (2016a). Null subjects. In *The Oxford handbook of developmental linguistics*. https://doi.org/10.1093/oxfordhb/9780199601264.013.17

Valian, V. (2016b). When children don't say what they know. In D. Barner & A. S. Baron (Eds.), *Core knowledge and conceptual change* (pp. 261–276). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190467630.003.0015

Valian, V., & **Aubry, S.** (2005). When opportunity knocks twice: Two-year-olds' repetition of sentence subjects. *Journal of Child Language*, **32**(03), 617. https://doi.org/10.1017/S0305000905006987

Valian, V., **Hoeffner, J.**, & **Aubry, S.** (1996). Young children's imitation of sentence subjects: Evidence of processing limitations. *Developmental Psychology*, **32**(1), 153.

Wexler, K. (2013). A new theory of null-subjects of finite verbs in young children. In M. Becker, J. Grinstead, & J. Rothman (Eds.), *Language acquisition and language disorders* (pp. 325–356). John Benjamins Publishing Company. https://benjamins.com/catalog/lald.54.14wex

Yang, C. D. (2002). *Knowledge and learning in natural language*. Oxford University Press on Demand.

Cite this article: Gerard, J., Singh, M., Bencini, G., & Valian, V. (2025). Null subject comprehension and production revisited: a look at English and Italian. *Journal of Child Language* 1–26, https://doi.org/10.1017/S0305000924000588