Sentence processing in bilingual children: evidence from garden-path sentences

Article

Accepted Version

Pontikas, G., Cunnings, I. ORCID: https://orcid.org/0000-0002-5318-0186 and Marinis, T. (2024) Sentence processing in bilingual children: evidence from garden-path sentences. Linguistic Approaches to Bilingualism. ISSN 1879-9272 doi: https://doi.org/10.1075/lab.22104.pon Available at https://centaur.reading.ac.uk/115605/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1075/lab.22104.pon

Publisher: John Benjamins

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Sentence Processing in Bilingual Children: Evidence from Garden-Path Sentences

George Pontikas[a], Ian Cunnings[a] and Theodoros Marinis[b]

[a] School of Psychology and Clinical Languages Sciences, University of Reading, Reading, UK
[b] Department of Linguistics, University of Konstanz, Konstanz, Germany

ABSTRACT

Research in sentence processing in bilingual children is emergent but incomplete as very few studies examine the processing of structurally complex sentences or bilingual children's real-time interpretation of sentences. One underexplored linguistic feature which can offer insights in this direction are garden-path sentences, i.e., sentences with temporary syntactic ambiguity. These are difficult to process for monolingual children as incremental processing results in an initial misinterpretation and the need for reanalysis. Studies on bilingual children's processing of garden-path sentences have used paradigms with limited ecological validity and which are not informative about one's interpretation while listening. This study bridges this gap by investigating the processing of garden-path sentences in bilingual children with the visual-world eye-tracking paradigm. It further explores the role of referential context in the visual stimuli to aid disambiguation. Monolingual and bilingual children aged 8-11 years completed a task similar to Trueswell et al. (1999). The results showed similar difficulty with revising garden-path sentences as evidenced by comprehension accuracy for both groups but only the monolinguals showed real-time garden-path effects in the gaze data. We interpret these findings as a manifestation of slower sentence processing in bilingual children. Both groups made limited use of the referential context to facilitate processing.

Keywords: sentence processing, bilingual children, eye-tracking, visual-world paradigm, garden-path sentences, psycholinguistics

1. INTRODUCTION

A hallmark of incremental sentence processing is garden-path effects, where an initial misinterpretation of a local syntactic ambiguity needs to be revised once the entire sentence has been heard (e.g. *"Put the frog on the napkin on the table"*). Research has shown that 5-8-year old monolingual children experience garden-path effects similarly to adults (Trueswell et al., 1999). However, they have greater difficulty recovering from these effects and do not benefit from contextual cues to avoid misinterpretation in the same way as adults do. How bilingual children process such sentences remains an empirical question.

Studies on sentence processing in bilingual children have used mainly grammatical violation paradigms (Chondrogianni et al. 2015; Chondrogianni & Marinis, 2012, 2016; Vasić & Blom, 2011). These studies have showed overall slower processing for bilingual than monolingual children – even when there is similar comprehension accuracy - with qualitatively similar patterns. While these findings have provided useful insights into children's knowledge of morphosyntactic features, they cannot elucidate bilingual children's ability to revise initial interpretations when the sentences do not involve grammatical violations, for example in garden-path sentences.

Sentence processing in monolingual and bilingual children may differ for a number of reasons. One reason may be related to their linguistic experience; the input bilingual children receive may be qualitatively and quantitatively different (e.g., amount of exposure, age of onset). This may result in lower proficiency in the additional language and/or slower processing speed. Evidence for slower processing comes from studies on processing phrase-level ungrammaticality (Chondrogianni & Marinis, 2012) but also for complex morphosyntactic structures such as object which-questions (Pontikas et al., 2023). Slower processing could influence sentence comprehension in different ways. Firstly, qualitative processing differences could be the by-product of slower processing if complex syntactic representations are not computed (or re-computed in the case of local ambiguity resolution) quickly enough in real time, resulting in reduced garden-path effects. An alternative possibility is that the same computational processes implicated in understanding a sentence take place, but occur over a longer time period. This study teases these issues apart by examining how bilingual children aged 8-11 years process garden-path sentences in English in real time.

One empirical question in bilingual language processing is to what extent bilingual children can utilise various sources of information to aid processing. Trueswell et al. (1999) showed that monolingual children aged 5 to 8 years could not utilise contextual cues to aid sentence processing. However, this has not been tested in bilingual children. A secondary aim of our study is thus to explore whether bilingual children can utilise referential cues to overcome garden-path effects.

If bilingual children do show slower processing and computation, this may also impede their integration of (non)linguistic information, such as contextual information in the visual field. Available evidence on the use of morphosyntactic cues in bilingual children points to a more nuanced and potentially reduced use of cues (Lemmerth & Hopp, 2019; Lew-Williams, 2017; Roesch & Chondrogianni, 2016). Some studies suggest bilingual children aged 6 to 10 years old can use a preceding linguistic referential context to aid interpretation of ambiguity (Chondrogianni et al., 2015, van Dijk et al., 2022a). Integration of visual contextual cues may however place greater demand on attentional control as the human parser needs to integrate both linguistic and non-linguistic cues in real time processing. Similarly to Trueswell et al. (1999) for monolingual children, this study examines the use of visual referential context as a facilitatory cue in sentence processing for bilingual children.

1.1 Garden-path sentence processing in monolinguals

Garden-path sentences have been shown to be difficult for children to process in numerous studies (Kidd, Stewart & Serratrice, 2011; Snedeker & Trueswell, 2004; Trueswell et al., 1999). The core difficulty appears not to be building a syntactic representation, but revising an initial misinterpretation and building a new one. For example, in Trueswell et al. (1999), children aged 5 years old listened to either locally ambiguous sentences or their unambiguous counterparts with a full RC exemplified in (1). At the same time, they looked at four objects and needed to act-out the instruction in the sentence. A camera recorded their gaze while they heard the sentence, and act-out accuracy and looks over time were analysed.

(1) Put the frog (that's) on the napkin in the box

(1) is ambiguous when there is no full RC because the PP "*on the napkin*" may be a modifier to the NP "*the frog*" or the destination argument of the verb "put". Participants saw a series of objects; a) the target entity to be moved (i.e., a frog on a napkin), b) the incorrect destination – what they initially misinterpreted as the destination (i.e., an empty napkin), c) the correct destination – the destination in the correct interpretation of the sentence (i.e., the box), and d) a distractor object (either another frog or different animal). Both adult and child participants appeared to be garden-pathed in that there were more looks to the incorrect destination in ambiguous compared to control unambiguous sentences. For the subsequent act-out task, child participants moved the frog to the incorrect destination (onto the empty napkin) more often in the ambiguous condition than in the unambiguous condition, but this was not observed for the adults. This suggests that, although both adults and children were garden-pathed and initially misinterpreted the sentence, only the adults were able to revise the initial misinterpretation (see also Choi & Trueswell, 2010; Hurewitz et al., 2000; Kidd et al., 2011; Meroni & Crain, 2003).

Trueswell et al. (1999) manipulated the number of referents in the scene in order to test how participants integrate contextual information. Participants saw combinations of objects with either one referent (a frog in the aforementioned example) or two referents. In the two-referent scene, one frog was on the napkin while the other was not. It was expected that in the two-referent condition, the second frog would guide the listener to interpret the PP "on the napkin" as a modifier to the NP "the frog", as this identifies which frog in the visual scene is being referred to. In contrast, in the one-referent condition, there is only one frog in the scene which, as such, will not act as a disambiguating cue. Both groups misinterpreted the sentences in the one-referent condition but differed in the two-referent condition. The adults utilised the presence of the second referent and, thus, did not experience garden-path effects. However, the children were garden-pathed in both one- and two-referent conditions. This was interpreted as reflecting qualitatively similar processing patterns between adults and children in terms of incrementality, while the differences lie in the children's (in)ability to integrate visual cues to disambiguate during parsing. So far, no studies have investigated how bilingual children process this type of ambiguous sentence.

1.2 Garden-path sentence processing in bilinguals

While there is ample research on ambiguous sentence processing in bilingual adults (Dussias & Cramer Scaltz, 2008; Hopp, 2015), the available literature on bilingual children is substantially scarcer. Examining ambiguity resolution in bilingual children however is crucial to our understanding of sentence processing in bilingual populations more broadly, beyond the existing literature on bilingual adults.

For bilingual adults, evidence suggests that adult second language (L2) learners process syntactic ambiguities in a manner similar to monolingual adults. Moreover, like monolingual adults, adult L2 learners can use a variety of cues to facilitate processing (see e.g. Dussias & Cramer Scaltz 2008; Hopp, 2015; Pan & Felser, 2011; Roberts & Felser, 2011; Williams, Möbius & Kim, 2001). Differences between adult L2 and monolingual sentence processing, primarily in relation to increased difficulty in reanalysis (Pozzan & Trueswell, 2016), have been largely attributed to individual differences in working memory (McDonald, 2006; Roberts, 2012), ability to integrate syntactic information (Hopp, 2015), or decoding and processing speed (McDonald, 2006). Note that this literature on adult bilingualism has focused on late, sequential L2 learners who have mostly not been exposed to the L2 in a naturalistic setting, and has been typically phrased around the distinction between native and non-native speakers. This dichotomy is less applicable to bilingual children who tend to have an earlier age of acquisition and qualitatively more naturalistic exposure to the additional language. As such, it is unclear whether and to what extent findings from the adult literature are to be replicated in bilingual children.

Only a handful of studies have so far examined real time processing of sentences that do not involve a grammatical violation in bilingual children (Bentea & Marinis, 2022; Bosch & Foppolo, 2023; Meir et al., 2020; Papangeli & Marinis, 2010; Pontikas et al., 2023, van Dijk et al., 2022a; 2022b). Only two studies examined local syntactic ambiguity processing in bilingual children. Papangeli & Marinis (2010) examined locally ambiguous sentences with a subject-object ambiguity in L2 Greek-speaking children aged 9-11 years with L1 Russian using a self-paced listening task. Participants heard sentences with optionally transitive verbs ("*While Mary dressed the baby slept*") in segments and then had to judge the sentences' grammaticality. Garden-path effects were expected in the form of increased reaction times at the segment of the second VP where subject-verb agreement precluded the ambiguous NP from being the object of the first verb, but instead meant it could only be a subject for the second one. Papangeli & Marinis (2010) manipulated intonation as a cue; in one condition, the intonation of the sentence was flat, providing no cues towards either interpretation; in the other condition, the sentences were read with natural intonation as a cue which guides towards the correct interpretation from the start. Garden-path effects were expected in the neutral intonation condition but not in the prosodic cue condition. As expected, neither group experienced garden-path effects in the prosodic cue condition. Interestingly, only the monolingual children experienced an increase in reaction times at the disambiguating segment for the neutral intonation condition. This was interpreted as evidence that the bilingual children did not experience garden-path effects, suggesting they may not have built a representation of the ambiguous sentence in real time.

Pontikas et al. (2023) used the visual-world paradigm to investigate how bilingual children processed local subject/object ambiguities in which-questions (e.g., "*Which donkey is carrying the zebra? / Which donkey is the zebra carrying?*"). Their eye-tracking data showed that both groups experienced initial garden-path effects with more looks towards images that reflected a subject-reading of the relevant sentences. Where the two groups showed differences was in the speed of processing; the bilingual children showed the same curve of looks to the target picture as the monolinguals but showed a relative shallower increase over time, in line with what previous studies have shown for reaction times in self-paced listening.

The studies available on bilingual child language processing are not directly comparable to Trueswell et al. as they used different types of ambiguous sentences (subject-object ambiguities vs. PP attachment ambiguities). The cues provided in the two studies are also different to those used in Trueswell et al. (1999) in that they were both linguistic in nature. No study has so far tested a non-linguistic cue, the impact of referential content in the visual stimuli, in bilingual children.

1.3 The present study

This study is the first of its kind to use eye-tracking to examine the same type of ambiguous sentences and the same type of facilitatory cue as in Trueswell et al. (1999) with bilingual children aged 8 to 11 years old, who are advanced simultaneous (2L1) or sequential (L2) learners of English with early naturalistic exposure. Overall, we address two gaps in the literature. Firstly, research into the processing of garden-path sentences has focused exclusively on monolingual children and adult L2 learners. Secondly, the limited research into locally ambiguous sentence processing in bilingual children has used either methods with limited ecological validity or types of sentence with a strong bias towards a specific interpretation. In the present study, we aim to examine not only the overall processing difficulty in garden-path sentences but also the timecourse of garden-path effects and the recovery from these. We do this by fitting non-linear models with time as a predictor variable to the data similarly to Pontikas et al. (2023) for the processing of which-questions. In the present study, participants saw the stimuli on a computer screen while listening to garden-path sentences and responded to a subsequent comprehension question by clicking on one of the images.

Our research questions were:

1) Do bilingual children differ from monolingual children in their comprehension of garden-path sentences?

Given prior work in bilingual sentence processing (Chondrogianni & Marinis, 2012, 2016; Chondrogianni et al., 2015; Vasić et al., 2012), it was expected that bilingual children would have similar response accuracy to monolinguals and that both groups would have lower response accuracy for garden-path sentences than unambiguous controls.

2) Do bilingual children differ from monolingual children in their ability to use contextual information from the visual scene to comprehend garden-path sentences?

Previous studies (Lemmerth & Hopp, 2019; Lew-Williams, 2017; Roesch & Chondrogianni, 2016) have pointed to a more nuanced use of cues by bilingual children in processing sentences. These studies only included morphosyntactic cues, the use of which is likely to be moderated by overall language proficiency (see Roesch & Chondrogianni). Studies examining bilingual children's ability to use referential context in the form of a preceding sentence do not show such a disadvantage (e.g., Chondrogianni et al., 2015; van Dijk et al., 2022a). It was unclear if this would be replicated for cues in the visual context. Previous studies with children aged 8-11 years suggest that this age group can utilize visual cues more robustly than 5-year-olds (Kidd & Bavin, 2005; Weighall, 2008). It was therefore conceivable that monolingual and bilingual children in this study would use referential context to aid comprehension of garden-path sentences.

3) Do bilingual children experience initial garden-path effects in the same way as monolingual children during processing?

Studies using self-paced listening have indicated overall slower processing for bilingual children compared to monolingual children. It is conceivable that slower processing could result in initial syntactic computations not being completed in real time and thus, garden-path effects not being experienced with only one representation of the sentence being formed upon hearing it complete. The findings from Papangeli & Marinis (2010), who did not show garden-path effects for ambiguous sentences in L2 Greek, may be a reflection of this. Alternatively, slower processing could be operationalised as the same parsing processes occurring in both monolingual and bilingual children but at different timepoints. In this study, this would have been reflected in effects of ambiguity being manifested at a later point in the critical sentence, or only in question accuracy or question response times, for the bilingual children.

4) Does the referential context in the visual stimuli aid monolingual and bilingual children in avoiding garden-path effects in real time processing?

If monolingual and bilingual children can make greater use of the referential context in real time, i.e., as they hear the ambiguous segment, an interaction between ambiguity and referential context was expected. The effects of ambiguity should be weaker, or absent, in the 2-referent context than in the 1-referent context.

Research questions 1 and 2 were relevant to children's ultimate comprehension of locally ambiguous sentences and were answered based on response accuracy and response (reaction) times. Research questions 3 and 4 address real time sentence processing and were answered based on the gaze data.

2. METHOD

2.1 Participants

We recruited 68 children for this study: 37 monolingual children (age range: 7;10-11;6, M=9;7, SD=1;1, 16 girls and 21 boys) and 31 bilingual/multilingual children (age-range: 7;4-11;5, M= 9;6, SD=1;2, 17 girls and 14 boys), all with a minimum exposure to English of two years.[1] The children in this study were older than in Trueswell et al. for two reasons. Firstly, the (monolingual) younger children in Trueswell et al. had greater difficulty recovering from garden-path effects and were not able to use the presence of a second referent in the visual stimuli to aid disambiguation. In Trueswell et al., it was the older children, around 8 years of age who showed some evidence for cue utilisation (subject to individual differences). As such, the age range selected for this study is perhaps the youngest age at which comparing monolingual and bilingual children can be done without the confound of the monolinguals not

[1] The children were the same participants as in Pontikas et al. (2023).

being able to process the tested linguistic feature. A second reason is that it is possible that bilingual children may experience an initial lag in the development of the second/additional language due to a later age of onset and less exposure. For this reason, younger bilingual children may not have been able to process the garden-path sentences used in this study.

Both monolingual and bilingual children were recruited from schools in the UK and none had a history of language impairment or learning difficulty. The participants undertook a series of baseline assessments including Clinical Evaluation of Language Fundamentals (Semel et al., 2006), Test for Reception of Grammar 2 (Bishop, 2003), Renfrew Test of Word Finding (Renfrew, 19995), The Children's Test of Nonword Repetition (Gathercole & Baddeley, 1996) and Raven's Coloured Progressive Matrices (Raven & Court, 1956). All scored within age-appropriate norms. At the group level, the bilingual children underperformed the monolingual children on some measures of language but not on other tasks (for details, please see the supplementary materials), and all bilingual children were within monolingual norms. The PABIQ questionnaire alongside semi-structured interviews were used to document the children's language profile and parental background.

The bilingual children came from varied backgrounds (summarised in Tables 1 and 2) but the majority would be classed as English dominant while one third had balanced language proficiency in both languages based on the PABIQ questionnaire. English was the language most used in the community and in education for almost all children but the bilingual children were equally divided in terms of language use in the home setting (English/L1 dominant or balanced). The total sample size meant that splitting the bilingual children into subgroups is not possible for the purpose of regression modelling. For this reason, measures of exposure to English, language proficiency and language dominance were included as continuous covariates in the accuracy and reaction time models for the bilingual children's data to control for individual variation. All the monolingual children were born in the UK apart from two monolingual children who were born in Australia and were raised in a monolingual English home environment. More a detailed breakdown of the bilingual children's background, see supplementary materials.

Ethical approval was granted from the School of Psychology and Clinical Language Sciences Research Ethics Committee (2015-115-IT).

2.2 Design

The visual world paradigm was used. Participants looked at pictures on a screen and listened to sentences while their looks to each picture were recorded in real time. They subsequently heard a comprehension question and responded by clicking the appropriate image on the screen.

The study used a 2x2 factorial design where syntactic ambiguity and the number of referents in the visual stimuli were manipulated. Participants heard either locally ambiguous garden-path sentences as in (2a) or unambiguous sentences with a full relative clause as in (2b), followed by a comprehension question, as in (3).

(2a) Peter put the apple on the plate in the bag before going to school.

(2b) Peter put the apple that is on the plate in the bag before going to school.

(3) Where did Peter put the apple?

The visual stimuli were manipulated to include one or two referents, in this case, apples, in the visual stimuli. For the 2-referent condition, there were two apples, one on the plate and one which was not. For the 1-referent condition, there was only one apple, which was on the plate and a distractor object in lieu of the second apple (e.g., a banana). Examples are illustrated in Figure 1. The positioning of the images was counterbalanced throughout the task. The ambiguous and unambiguous trials differed in terms of the sentences heard but used the same visual stimuli. The 1- and 2-referent trials differed in terms of visual stimuli but used the same sentences.

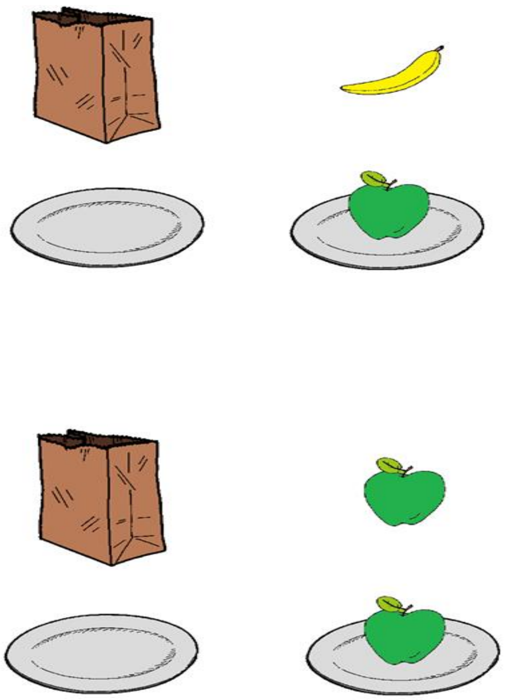

Figure 1 Sample stimuli for 1-referent (top) and 2-referent (bottom) conditions

2.3 Materials

Sixteen inanimate objects (e.g., apple) were paired with 32 potential locations to be moved to (e.g., table, desk, etc.). Sixteen pairs of sentences as in (2a) and (2b) were then created. For each pair four pictures were created to depict: a) the target i.e., the apple on a plate, b) an empty location which may initially be interpreted to be the destination of the 'putting' action (the Incorrect Destination, "ID", i.e., the empty plate), c) the destination of the object (Correct destination, "CD", i.e., a bag), d) a distractor image, either a different object (e.g., a banana) in the 1-referent condition or a second referent in the 2-referent condition (e.g. another apple). Sixteen comprehension questions as in (3) were created.

Twenty filler sentences were also included. Half the fillers used the same verb as in the experimental trials while the other half contained a different transitive verb. The fillers contained a mixture of sentence structures and were followed by a comprehension question using ‘what’, ‘where’ or ‘who’. Fillers had similar visual stimuli to the experimental trials. Two practice trials were constructed similarly to the fillers.

Four lists were created in a Latin-square design with each item appearing in each condition. The lists were then duplicated so that each trial appeared first as an ambiguous trial or as an unambiguous control. Each version of the task contained sixteen experimental trials, four from each condition with the order of the items and the fillers remaining constant. Participants were randomly allocated to one of the versions of the task.

The materials were recorded in a sound booth by a male adult native (Received Pronunciation) English speaker. The sentences were read out segmentally at a slow pace allowing for pauses between segments. Multiple iterations were recorded for each trial. The authors reviewed the auditory stimuli and selected recordings with the clearest quality of sound and where there the intonation was neutral/flat and did not bias the listener towards a specific interpretation. The same recording was used for the pair of garden-path sentences and the unambiguous controls.

2.4 Procedure

Participants sat in a quiet room wearing headphones at about 80cm from a laptop screen and a Tobii X-120 eye-tracker. A five-point calibration was performed before the beginning of the task. Adequacy of the calibration quality was judged by the first author done in accordance with the guidelines in the Tobii X-120 manual.

Participants needed to look at a fixation cross at the centre of the screen for 1,000ms in order for each trial to begin. Participants then heard a statement and saw four pictures of equal size in the four corners of the screen. Subsequently they heard the comprehension question and clicked one of the pictures to answer the question. The position of the correct picture varied throughout the task. The order of the stimuli was pseudorandomised ensuring that subsequent trials were of a different condition to previous ones. Experimental trials were always separated by fillers. Prior to commencing the task, participants undertook two practise trials. E-prime recorded participants' responses and response speed for each trial as well as their looks to each of the four images at a sampling rate of 120Hz.

2.5 Analyses

Mixed effects regression models were used to analyse all measures. Binary logistic models were fitted to response accuracy and linear models were fitted to response speed (henceforth reaction times) and spline regression models were fitted to the gaze data. For all measures, the fixed and random effects structure was largely the same. Fixed effects included a between-group comparison (monolingual vs. bilingual children), ambiguity (ambiguous sentences vs unambiguous controls) and referential context (1- vs 2-referents in visual stimuli). Deviation coding (-1, 1) was used for between subject variables (monolingual vs. bilingual) and fixed main effects of ambiguity (ambiguous vs. unambiguous sentences) and referential context (one or two referents in visual stimuli). Maximal random effects structures permitted by design were used following Barr, Ley, Scheepers & Tilly (2013). Random intercepts and slopes by subject and item were allowed to vary for all fixed effects. If the maximal model failed to converge, random slopes contributing minimally to the variance were removed iteratively until the model converged.

Response accuracy was coded correct or incorrect for each trial and was analysed with a binomial mixed effects model. Data about which picture was selected in the case of an inaccurate response was not recorded. Response speed was analysed only for trials with an accurate response. Unambiguous sentences were longer in duration than ambiguous ones because they contained the disambiguating words "that's" as part of the full relative clause. Therefore, reaction times were calculated by subtracting the duration of the sentence from times logged by E-prime. Computed reaction times were analysed with a linear mixed effects model.

For the gaze data, four areas of interest (AOI) were defined a priori capturing one of the four quarters of the screen. For each trial, specific AOIs were defined as "target" (the apple on the plate), "distractor" (standalone apple or other object), "incorrect destination (ID)" (the plate), "correct destination (CD)" (the bag) or "out" (when no look was recorded). Looks were automatically coded in E-prime. Looks were time-locked to the onset of particular phrases of interest (henceforth regions) in the sentence; the incorrect destination ("on the plate"), the correct destination ("in the bag"), and the segment following disambiguation until the end of the sentence ("before going to school"), defined as the final region. Looks to the AOIs were aggregated and proportions were calculated relative to the total number of looks excluding "outs". These were transformed to compute the empirical logit (Barr, 2008); weights were added to each observation based on the reciprocal of the variance. Gaze data only from trials with an accurate response to the comprehension question were included. For the purposes of this study, we analysed looks to the incorrect destination (ID) as these best reflect garden-path or misinterpretation effects. In contrast to the response speed analysis, for the gaze data, we did not remove trials with an incorrect response as these trials may have stronger garden-path effects.

As it takes around 200ms to program a saccadic eye-movement (Matin, Bao & Boff, 1993), the time window earlier than 200ms in each phrase was excluded from any analyses. This means that the time window analysed was 200-1800ms post phrase onset. Trials where looks to "out" exceeded 50% were removed from further analyses (ca 25-30% of trials depending on the region analysed and 28% of all trials). As a practice, this is consistent with other studies, for example Atkinson, Wagers, Lidz, Phillips & Omaki, (2018).

For the gaze data models, time as bin number was scaled and coded as a restricted cubic spline, as it is a continuous variable. For a similar analysis and an outlining of the underlying rationale, see Contemori et al. (2018) and Pontikas et al. (2023). This enables us to capture non-linear changes as a function of time. Interactions of within/between subject variables with time are taken to reflect a different change in looks over time for the different levels of the given variable. We fitted models with three, four and five knots (Harrell, 2001) but report the results from the models with the simplest splines as these provided the best fit. For significant effects of variables on the intercept term, i.e., overall differences irrespective of time, we computed the mean looks to the ID in the relevant conditions. To better interpret main effects or interactions over time, we rely on visual inspection of the timecourse of looks to the ID in Figure 2. Where there were significant interactions, these were qualified by fitting separate models with the data split by the relevant variables. Due to concerns about statistical power and the risk of false positives, as well as for reasons of brevity and conciseness, we further explored only up to two-way interactions for the gaze data and did not explore 3-way interactions further. For the same reasons, we do not report trends for any of the analyses.

For the accuracy and reaction time analyses, age and Core Language Scores (henceforth, CLS) from the CELF-4 were added as covariates to account for individual variability and to ensure that potential group differences were not confound of differences in language proficiency. Moreover, for the same reason, separate models were fitted by default to the data from bilingual children including CLS score and age but also composite scores of length of exposure and language dominance as established from the PABIQ questionnaire. The inclusion of these background measures as well as the aforementioned analyses for the bilingual children were not undertaken for the gaze data analyses. The motivation for this is two-fold; firstly, the complexity of the model relative to the data and second the interpretation of the parameters. Furthermore, as the gaze data models are non-linear regressions (time has been coded as a spline), the parameters cannot be interpreted in the same way as they would in a linear regression.[2]

3. RESULTS

3.1 Accuracy and reaction times

The results for accuracy and reaction times are summarised in Tables 3-5 below. For accuracy, there was a main effect of ambiguity with a large effect size in terms of Odds Ratios (ORs) and Relative Risk Ratios (RRs) (*OR* = 4.84, 95% *CI* = 4.56 – 5.12, *RR* = 3.83, 95% *CI* = 3.58 - 4.08), with ambiguous garden-path sentences having lower comprehension accuracy than the unambiguous control sentences. Participants were approximately three times as likely to respond inaccurately to a garden-path sentence than an unambiguous control. There were no other significant main effects or interactions. The bilingual children were practically indistinguishable from the monolinguals in terms of accuracy (for the harder ambiguous sentences accuracy was 72.9%, bootstrapped 95% *CI* = 66.8-78.5 vs 73.6% bootstrapped 05% *CI* = 68.4-78.5 for the bilinguals and monolinguals, respectively) and neither group benefitted from the presence of the second referent in the visual stimuli. A separate model fitted to the data from the bilingual children only including the individual differences measures showed the same main effect while none of the measures quantifying bilingual experience were significant (for the full model see online supplement).

For reaction times, there was a main effect of age with reaction times becoming faster for older children. No other main effects or interactions were significant. Again, reaction times were

[2] In other words, that a change of unit in the predictor variable corresponds incrementally to a change of beta in the dependent variable.

almost indistinguishable between the two groups of children (Hedge's g= 0.07 for ambiguous sentences, g = 0.06 for unambiguous controls)[3]. Differences in reaction times between ambiguous and unambiguous sentences were also extremely small (Hedge's g = 0.09 for monolinguals; g = 0.03 for bilinguals). A separate model fitted to the bilingual children only which included measures of bilingualism replicated the effect of age but other measures were not significant predictors.

Table 1 Accuracy scores (bootstrapped 95% *CI*s) by group and condition

	Monolinguals	Bilinguals
Ambiguous, One referent	72.9% (65.3-79.9)	75% (66.4-82.8)
Ambiguous, Two referents	74.3% (67.4-81.3)	70.7% (62.1-78.5)
Unambiguous, One referent	93.1% (88.9-97.2)	88.8% (82.8-93.9)
Unambiguous, Two referents	93.1% (88.9-96.6)	87.9% (81.9-93.9)

Table 2 Mean RTs (SD) by group and condition [bootstrapped 95% CIs]

	Monolinguals	Bilinguals
Ambiguous, One referent	5488 (1928) [5145-5871]	5221 (1084) [5004-5456]
Ambiguous, Two referents	5634 (1946) [5303-6030]	5610 (2138) [5189-6099]

[3] Hedge's g has been used as there is an imbalance in the size of the data set as it used the pooled SD in contrast to Cohen's d. As with Cohen's d, an effect size of 0.2/0.5/0.8 would be classed as small, medium, large, respectively.

Unambiguous, One referent	5399 (1619) [5154-5700]	5334 (1340) [5088-5602]
Unambiguous, Two referents	5363 (1533) [5119-5628]	5618 (2098) [5253-6056]

Table 3 Models fitted for accuracy and reaction times

	Accuracy				Reaction times			
	β	*SE*	*z*	*p*	*β*	*SE*	*t*	*p*
Intercept	2.33	0.27	8.81	<.001	5468	119	45.8	<.001
Group	-0.06	0.23	-0.24	.813	-50	113	-0.44	.662
Ambiguity	0.80	0.12	6.53	<.001	-34	55	-0.62	.536
Context	-0.03	0.11	-0.25	.806	76	64	1.2	.237
Age	0.35	0.22	1.58	.115	-348	106	-3.29	.002
CLS score	0.39	0.21	1.81	.070	-198	106	-1.86	.067
Group*Ambiguity	-0.10	0.13	-0.77	.439	49	70	0.69	.497
Group*Context	-0.07	0.11	-0.61	.543	61	63	0.96	.344
Ambiguity*Context	0.03	0.11	0.32	.747	-28	55	-0.51	.609
Group*Ambiguity*Context	0.06	0.11	0.57	.569	14.81	55	0.27	.787
For bilinguals only								
Ambiguity	0.74	0.21	3.55	<.001	20	82	0.24	.881
Age	0.69	0.34	0.98	.325	-537	137	-3.92	.001
CLS score	0.69	0.38	1.83	.067	-4	145	-0.03	.977
Length of exposure	-0.61	0.38	-1.61	.108	109	160	0.68	.503
Dominance	-0.45	0.44	-1.04	.299	262	189	1.39	.177
Exposure	0.142	0.44	0.32	.747	-78	198	-0.39	.698

3.2 Gaze data

The gaze data is presented in Figure 2 (referential contexts have been collapsed due to absence of effects for timecourse models). Gaze proportions here have been calculated over 100ms bins for a more fine-grained presentation, even though for the timecourse analyses bins of 200ms were used for aggregation to reduce autocorrelation (see Choo, Brown-Schmidt & Yee, 2018; Mirman 2014; Stone, Lago & Schad, 2021). LOESS smoothing has been applied in the visualisation in order to facilitate the presentation of the gaze data trajectories and to generate the best fitting curve to the data. The shaded area shows bootstrapped 95% confidence intervals used to estimate the uncertainty about the trajectory of the curves.

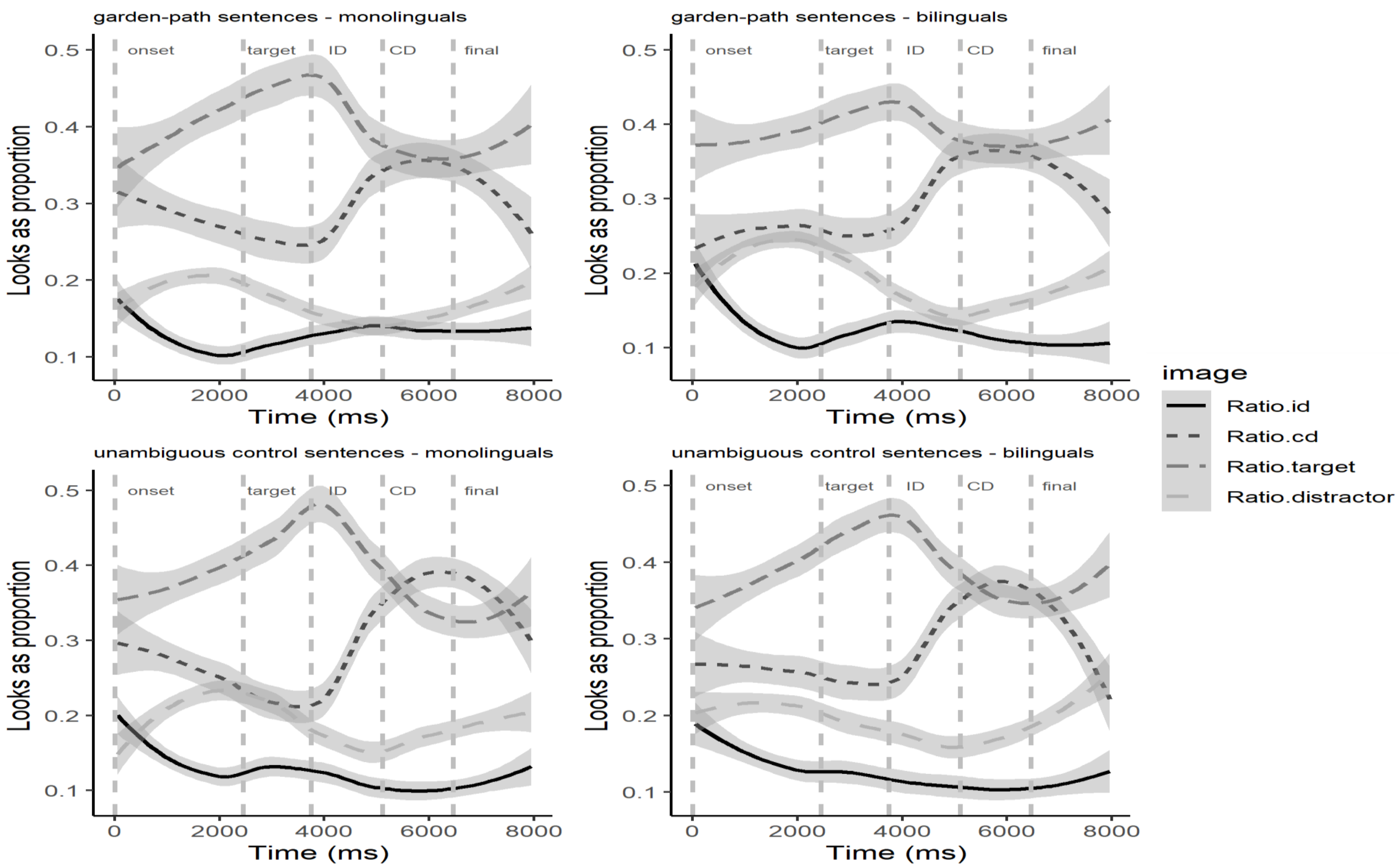
garden-path sentences - monolinguals
garden-path sentences - bilinguals
unambiguous control sentences - monolinguals
unambiguous control sentences - bilinguals
onset
target
ID
CD
final
Looks as proportion
Time (ms)
image
Ratio.id
Ratio.cd
Ratio.target
Ratio.distractor

Figure 2 Looks to all images by group and ambiguity over time for example sentence "Peter put (onset) the apple (target) on the plate (incorrect destination: ID) in the bag (correct destination: CD) before going to school (final).

Looks to the ID during the ID region (Peter put the apple ON THE PLATE in the box before going to school)**:**

In this region there was only a significant group by ambiguity interaction on the intercept term and a group by ambiguity interaction on the first spline term. These interactions were qualified by separate models to the data split by group and ambiguity. A model fitted to the data from the monolinguals showed a significant effect of ambiguity (β = 0.06, SE = 0.03, t = 1.96, p = .050) and an interaction of ambiguity with time (β = -0.28, SE = 0.07, t = -4.18, p < .001), with cumulatively more looks to the ID for the ambiguous sentences than the unambiguous controls (mean = 16.3%, 95% *CI* = 14.5-18.2% vs 11.8%, 95% *CI* = 10.2-13.5%, respectively). This effect was not found for the model fitted to the bilingual data (all *p*s > 0.05; mean looks = 13.6%, 95% *CI* = 11.8-15.3% vs. mean = 14.5%, 95% *CI* = 12.7-16.3%). These results suggest more looks overall towards the incorrect destination for monolingual children in the ambiguous condition. In terms of the effects of group and ambiguity on the first spline, visual inspection of the best fitting curve suggests an increase in looks to the ID at the start of the ID region for monolingual children when they heard ambiguous sentences. This increase in the region of interest was not found for bilingual children. This may be interpreted as significantly weaker garden-path effects for the bilinguals immediately after hearing the ambiguous segment. For the unambiguous sentences, the monolingual children show a decrease in looks to the ID in the ID region but for the bilingual children, the decline is less steep.

A reviewer queried whether an increase in looks to the ID in the segment preceding the ambiguous region could undermine our claims about garden-path effects in monolingual but not bilingual children in the ID region. To address these concerns, we fitted spline regression models to the looks to the ID in the preceding region (i.e., when participants heard the NP "the frog"). This analysis did not show any differences between monolingual and bilingual children in terms of ambiguity either on the intercept term or on the spline term in this earlier region. For this region, we maintain our interpretation of results at the ID region. For further details of this analysis, please see our data repository on OSF.

Looks to the ID during the CD region (Peter put the apple on the plate IN THE BOX before going to school):

Analysis revealed a significant main effect of time in the second part of the spline, a significant interaction of group and ambiguity overall, an interaction between time and ambiguity (second component of the spline), a group by ambiguity by time interaction (second component of the spline) and a group by context by time interaction (second component of spline). To qualify these interactions, separate effects were fitted by group and ambiguity. The model fitted to the monolinguals' data showed a main effect of ambiguity ($\beta = 0.08$, $SE = 0.03$, $t = 2.03$, $p = .024$) with more looks to the ID for the ambiguous sentences than the unambiguous ones (mean = 13.6%, 95% *CI* = 11.9-15.4% vs. mean = 9.5%, 95% *CI* = 7.9-10.9%), an interaction of ambiguity with referential context ($\beta = 0.08$, $SE = 0.03$, $t = 2.18$, $p = .028$) and an ambiguity by time interaction (second component of spline, $\beta = 0.21$, $SE = 0.09$, $t = 2.53$, $p = .011$).

To further qualify the ambiguity by context interaction, separate models were fitted to the monolingual children's data by ambiguity type. For the ambiguous sentences, the monolingual children showed an effect of referential context with more looks to the ID in the two-referent context than the one-referent context (mean = 11%, 95% *CI* = 8.7-13.3% vs. mean = 16.1%, 95% *CI* = 13.5-18.9%; $\beta = 0.12$, $SE = 0.05$, $t = 2.57$, $p = .012$) and an effect of time in the first component of the spline ($\beta = -0.18$, $SE = 0.09$, $t = -1.99$, $p = .047$). This effect was in the opposite direction of what would be expected as the presence of the second referent is assumed to help avoid rather than increase garden-path effects. The model fitted to the monolingual data for the unambiguous sentences found no main effects or interactions (all *p*s > .05).

The model fitted to the bilingual data found no effect of ambiguity nor an interaction between ambiguity and referential context but only an interaction between ambiguity and time (first component of spline; β = 0.12, *SE* = 0.05, t = 2.00, *p* = .046).

The model fitted to the ambiguous sentences only showed a main effect of group (β = 0.08, *SE* = 0.04, t = 2.05, *p* = .045) mirroring the effect of ambiguity in the monolinguals-only model. Looks to the ID were more for the monolinguals (means = 13.5%, 95% *CI* = 11.8-15.3%) than the bilinguals (mean = 10%, 95% *CI* = 8.4-11.7%). There were no other main effects or interactions. This was not found for the model fitted to the unambiguous data (relevant *p*s > 0.05).

The analysis for this region suggests a similar picture to the ID region. There are overall more looks to the ID for ambiguous sentences relative to unambiguous controls for the monolinguals only. This indicates that garden-path effects may still be lingering in the monolinguals but not in the bilinguals.

Looks to the ID during the FINAL region (Peter put the apple on the plate in the box BEFORE GOING TO SCHOOL):

Analysis revealed only a significant main effect of time (β = 0.10, *SE* = 0.05, *t* = 2.12, *p* = .034). There was also a group by ambiguity by time interaction (β = 0.09, *SE* = 0.05, *t* = 2.00, *p* = .045) and a group by ambiguity by context by time interaction (β = 0.10, *SE* = 0.04, *t* = 2.36, *p* = .019). However, as these are three- and four-way interactions, they are not explored further.

4. DISCUSSION

The present study is the first of its kind to examine the processing of garden-path sentences in bilingual children. Our main findings are the following.

(1) Bilingual children experienced garden-path effects evidenced by lower comprehension accuracy for garden-path sentences relative to unambiguous controls.

(2) Bilingual children showed comparable comprehension accuracy to monolinguals for garden-path sentences and unambiguous controls.

(3) Only the monolingual children experienced garden-path effects in real time, evidenced by cumulatively more looks to the incorrect destination for ambiguous sentences relative to unambiguous controls.

(4) Referential context did not aid ambiguity resolution for either monolingual or bilingual children in any measure.

Below, we discuss these main findings in relation to our four research questions.

RQ1: Do bilingual children differ from monolingual children in their comprehension of garden-path sentences?

The findings suggest similar comprehension of garden-path sentences for the two groups. For both bilingual and monolingual children, garden-path sentences were more difficult relative to unambiguous controls. In our study, the question 'Where did Peter put the apple?' requires participants to understand the correct destination, and the only other location in the visual display is the incorrect destination. Although we acknowledge we did not record the image that participants did ultimately choose, we maintain that inaccurate responses to the comprehension question meant that the participant was likely to have failed to interpret the correct destination.

There was no group by ambiguity interaction in accuracy rates, suggesting there was no additional difficulty for the bilingual children, and that reanalysis was equally difficult for both groups. The overlapping non-parametrically bootstrapped confidence intervals in both groups further suggests that the absence of group differences here was not merely a case of absence of evidence but evidence for similar or equivalent processing abilities at an end-state of comprehension.

The results were consistent with previous studies which have shown difficulties in offline accuracy for ambiguous sentences albeit for act-out rather than responding to a comprehension question (e.g., Choi & Trueswell, 2010; Hurewitz et al., 2000; Kidd & Bevan, 2005; Kidd et al., 2011;; Trueswell et al., 1999). Accuracy for ambiguous sentences was, however, considerably higher for children in this study than in some other studies (e.g., Hurewitz et al., 2000; Trueswell et al. but not Kidd et al., 2011). For example, in Trueswell et al., accuracy was considerably lower than we found, with approximately 60% of ambiguous trials resulting in an incorrect act out. The participants in the present study were older than in Trueswell et al. however, and there is evidence of improvement with age in the comprehension of garden-path sentences (Weighall, 2008). This difference in age we believe can explain the different results between this study and Trueswell et al. (1999).

RQ2: Do bilingual children differ from monolingual children in their ability to use contextual information from the visual scene to comprehend garden-path sentences?

The hypothesis that bilingual children would benefit from the second referent in offline comprehension was not confirmed. This study did not find any evidence for a benefit from the referential context for either monolingual or bilingual children in the end of sentence measures, namely response accuracy and reaction times. These results are in line with Trueswell et al. (1999) who found a benefit of the second referent in the visual stimuli only for adults but not children. However, the children in Trueswell et al. were 5 years old, younger than those in the current study. Weighall (2008) tested both 5- and 8-year-olds and neither age group utilised context to aid ambiguity resolution. This suggests the ability to utilise context continues to develop with age. From this perspective, we maintain that the lack of context effects in our sample, aged up to 11, is not inconsistent with previous research.

Note that this was the first study to examine the utilisation of referential context as a visual cue to facilitate processing in bilingual children and thus, direct comparisons with other studies should be cautious. Van Dijk et al. (2022a) showed that bilingual children were equally able as monolingual children to use gender and referential context during pronoun ambiguity resolution offline. Chondrogianni et al. (2015) also showed that bilingual children could utilise referential context from preceding sentences. One logical explanation for this discrepancy between studies is that our study, which required participants to use information in the visual stimuli to interpret the sentence they heard, was more demanding as it involved utilising information across domains (even if there is nothing inherently more demanding with visual stimuli per se). It may be that the ability to integrate non-linguistic, visual cues during sentence processing has not yet adequately developed. This may take longer for adult-like real-time competence to be attained as it involves two sources or domains of information (linguistic and non-linguistic) being utilised in tandem. On the other hand, reanalysis of initially misinterpreted sentences requires solely syntactic information.

RQ3: Do bilingual children experience initial garden-path effects in the same way as monolingual children during processing?

The gaze data suggest a different time-course of garden-path effects in bilingual children compared to monolinguals. In the earliest time window, when hearing the ambiguous PP, there were significantly more looks to the incorrect destination for the ambiguous sentences for the monolingual children as well as statistically significant changes to the trajectory of looks over time. Subsequently, looks to the ID appeared to remain stable. The latter is what would be expected if the participants had misinterpreted the phrase as the destination argument of the verb instead of a modifier of the noun. This was not found for the bilingual children's gaze data.

The absence of effects of ambiguity at the earliest point suggests an absence of garden-pathing in real time for bilingual children. In this sense, the study replicates what was found by Papangeli & Marinis (2010) for subject-object ambiguities in Greek with the self-paced listening paradigm. The limitation of the methodology in that study is that, although one can pinpoint segments in the sentence which are difficult to process by means of an increase in reaction times, the exact interpretation of the sentence cannot be inferred. This is not the case for the visual world study, where looks towards specific images are broadly associated with specific interpretations of the sentence. In this case, the observed change in looks towards the incorrect destination for the monolingual children suggests an early misinterpretation of the PP as the destination of the 'putting' action, which did not appear to occur in the bilingual children.

On the surface, this contrasts to Pontikas et al. (2023), who found that both monolingual and bilingual children showed an initial subject-bias and misinterpreted locally ambiguous object which-questions as subject-questions. Upon closer consideration of the results, we believe it was not the case that bilingual children failed to experience garden-path effects as a by-product of incremental processing. The absence of a group by ambiguity interaction for the response accuracy indicates that bilingual children had similar difficulty with garden-path sentences in comparison to the unambiguous controls. This is in line with previous studies on garden-path sentences with monolingual children (e.g., Trueswell et al., 1999), which is typically taken to reflect the fact that children have indeed been garden-pathed during incremental processing. We argue that the effect of ambiguity on response accuracy was due to the fact that the bilingual children misinterpreted the ambiguous sentences and that they had a similar pattern of recovery to the monolingual children.

This begets the question as to what accounts for the lack of ambiguity effects at the point of the ambiguous PP for bilingual children. Our speculative answer would be processing speed. Previous studies using the self-paced listening paradigm have consistently shown overall longer reaction times for bilingual children relative to monolinguals (e.g., Chondrogianni et al. 2015; Chondrogianni & Marinis, 2012, 2016; Vasić & Blom, 2011), which has been interpreted as evidence for slower processing. We argue that our findings related to experiencing garden-path effects can be explained on the basis of differences in slower processing speed. We contend that bilingual children were slower to process the ambiguous sentences and thus, there were no immediate garden-path effects. However, the bilingual children were able to form an erroneous interpretation by the time they had heard the entire sentence and had similar difficulty revising this representation as the monolingual children, as evidenced by the effects of ambiguity on response accuracy. This entails that ultimately the bilingual children understood the ambiguous sentences but needed more time. This interpretation is consistent with the operationalisation of slower processing speed as the same parsing process occurring at later timepoints, rather than failing to be completed at all.

It is plausible that these differences in processing speed may be due to the bilingual children's two linguistic systems being active during processing. The research design of the present study does not allow us to disentangle this interpretation of our results from others, such as potential effects of L1 influence. However, we contend this hypothesis is not inconsistent with the available literature. Previous studies have shown slower lexical processing for bilinguals (e.g., Blumenfeld & Marian, 2007; Costa, Caramazza & Sebastian-Galles, 2000; Colomé, 2001; de Bruin, Della Sala & Bak, 2016). Other studies have shown evidence for cross-linguistic influence in bilingual children in syntactic priming (Unsworth, 2023; Vasilyeva et al., 2010) and more recently during real-time sentence processing (van Dijk et al., 2022b). These results provide evidence of two linguistic systems being active during bilingual processing, and we maintain this could provide an explanation for the slow processing observed in our bilingual group.

Note that the garden-path effects were relatively small in terms of the gaze data and that the effect of ambiguity was substantially larger for response accuracy. Overall, the proportion of looks towards the incorrect destination was low, with the majority of looks being towards the target image. This may reflect lexical processing, such that the participants were more likely to look towards the word they had just heard. Note that this is not inconsistent with previous studies. For example, looks to target also make up more than half the looks in Trueswell et al. (1999).

RQ4: Does the referential context in the visual stimuli aid monolingual and bilingual children in avoiding garden-path effects in real time processing?

In this study, the gaze data suggested that neither group of children seemed to utilise visual context to aid syntactic ambiguity resolution during incremental processing. This suggests that the children could not integrate contextual information rapidly. In Trueswell et al. (1999), young monolingual children aged around 5, could not use referential context but adults did. Our results suggest that older children do not yet utilise referential context robustly during processing, either. The age range of children tested in this study is similar to that in Weighall (2008) who also found children had difficulty using referential context. Weighall (2008) manipulated the number of referents across an act-out task in three groups of monolingual children, aged 5, 8 and 11 years. Overall accuracy was comparably high to the children in this study ranging between 72-91% (with quite large SDs for each condition). This suggests ability to use context continues to develop until a later age than tested here.

It is possible that individual differences in the processing of garden-path sentences can be attributed to differences in cognitive abilities, including executive functions such as switching and inhibition. This study did not include such background measures and as such is not informative in this respect. This would be a potential avenue for future research but would require a substantially larger number of participants and establishing a suitable age range for testing this feature.

Our findings in respect to the use of cues are not directly comparable with previous studies on ambiguous sentence processing in bilingual children, as this is the first to examine the use of a non-linguistic cue (number of referents in visual stimuli). Previous studies have pointed to a more nuanced or reduced use of linguistic cues in bilingual children (e.g. Lemmert & Hopp, 2019 and Lew-Williams, 2017; Papangeli & Marinis, 2010; Roesch & Chondrogianni, 2016). These studies however examined phenomena where monolingual children could utilise the relevant cues, while both the monolingual and bilingual children did not utilise the non-linguistic cue in this study. Therefore, this study cannot be informative about the developmental trajectory and/or the time course of qualitative characteristics of cue utilisation in bilingual children.

Some final comments on the methodology and analyses

This study alongside the Pontikas et al. (2023) study differs from other studies in terms of the populations studied. The bilingual children were of mixed linguistic backgrounds in contrast to those in the previous studies (heritage Romanian in Bentea & Marinis, 2022; L1 Russian in Papangeli & Marinis, 2010; Hebrew-Russian in Meir et al., 2020; German/Turkish-Dutch in van Dijk, 2022a, English/German-Dutch in van Dijk, 2022b). To address this limitation, we fitted models with age and English proficiency as covariates for the accuracy and reaction time data. These did not alter the findings. For the bilingual data, we also fitted separate models with length and balance of exposure to English as well as dominance in English as covariates. None of these were significant predictors or alter the results. We did not attempt this with the gaze data, as the model would have been too complex for the number of observations. We cannot ascertain the impact of the participants' first language as there was no available information in the World Atlas of Language Structures (Dryer, 2013). However, the use of random slopes by subject for all fixed effects reduced the probability of our findings being driven by a subset of participants who may or may not speak certain language(s). Further studies could opt to include speakers from specific L1s and examine the impact of cross-linguistic influences on sentence processing.

It should be noted that the majority of bilingual children were English-dominant. This is to be expected given that the children had entered formal mainstream English-medium education and consequently had more English input. This is likely to be the case in any context of successful social integration. In this sense, that the bilingual children show similar processing patterns should come as no surprise even if more fine grained eye-tracking measures suggest slower processing. What is still unclear is what happens with complex sentences when processing in the heritage language where input is reduced and outcomes are likely to be more variable (for a review, see Hoff, 2017). Further research addressing this question is needed.

Finally, a reviewer commented that the critical sentences in the current study were spoken at a slow rate (see Figure 2), especially when compared to previous research using similar designs (e.g. Trueswell et al., 1999). Whilst we acknowledge this difference to previous studies, we are unsure how this will have influenced our results. If anything, it makes the absence of garden-path effects during online processing in the bilingual children more striking. How speech rate may influence incremental processing in both monolingual and bilingual children is a question for future research.

5. CONCLUSION

We examined real time processing of locally ambiguous (garden-path) sentences in English in bilingual children aged 8-11 years. The results show that the bilingual children did not underperform relative to their monolingual peers. Both groups had lower accuracy responding to comprehension questions following garden-path sentences compared to unambiguous controls. This suggests both groups experienced effects of initial misinterpretation from which they had equal difficulty recovering. The gaze data showed effects of ambiguity for the monolinguals but this was not observed for the bilingual children. Taken together, this indicates that the bilingual children experienced garden-path effects due to misinterpretation but at a later point relative to the monolinguals. The latter is consistent with slower processing observed in other studies. The presence of a second referent in the visual stimuli did not appear to have any effect for the processing of the ambiguous sentences in either group.

Data availability

All materials, data and model code are available on OSF through the following link: https://osf.io/fbt6y/.

Acknowledgments

We are immensely thankful all our participants for their willingness and time, their parents who consented to their participation and to the schools which helped us recruit participants and test them on their premises. Recruitment of participants was supported financially by the School of

Psychology and Clinical Language Sciences, University of Reading (SPCLS). Ethics approval was granted by the SPCLS Ethics Committee (2015-115-IT).

REFERENCES

Atkinson, E., Wagers, M. W., Lidz, J., Phillips, C., & Omaki, A. (2018). Developing incrementality in filler-gap dependency processing. *Cognition*, 179, 132-149.

Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, *59*(4), 457-474.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255-278.

Bentea, A., & Marinis, T. (2021). Not all wh-dependencies are created equal: processing of multiple wh-questions in Romanian children and adults. *Applied Psycholinguistics*, 42(4), 825-864.

Blumenfeld, H. K., & Marian, V. (2007). Constraints on parallel activation in bilingual spoken language processing: Examining proficiency and lexical status using eye-tracking. *Language and Cognitive Processes*, 22(5), 633-660.

Bosch, J. E., & Foppolo, F. (2023). Prediction during spoken language processing in monolingual and multilingual children: Investigating the role of literacy. *Linguistic Approaches to Bilingualism*, doi.org/10.1075/lab.22099.bos.

Cho, S. J., Brown-Schmidt, S., & Lee, W. Y. (2018). Autoregressive generalized linear mixed effect models with crossed random effects: An application to intensive binary time series eye-tracking data. *Psychometrika*, 83, 751-771.

Choi, Y., & Trueswell, J. C. (2010). Children's (in) ability to recover from garden paths in a verb-final language: Evidence for developing control in sentence processing. *Journal of Experimental Child Psychology*, 106(1), 41-61.

Chondrogianni, V., & Marinis, T. (2012). Production and processing asymmetries in the acquisition of tense morphology by sequential bilingual children. *Bilingualism: Language and Cognition*, *15*(1), 5-21.

Chondrogianni, V. & Marinis, T. (2016). L2 children do not fluctuate: Production and on-line processing of indefinite articles in Turkish-speaking child learners of English. In: Haznedar, B. & Ketrez, F. N. (Eds.). *The Acquisition of Turkish in Childhood*. John Benjamins. [*Trends in Language Acquisition Research 20*], 361-388.

Chondrogianni, V., Vasić, N., Marinis, T., & Blom, E. (2015). Production and on-line comprehension of definiteness in English and Dutch by monolingual and sequential bilingual children. *Second Language Research*, *31*(3), 309-341.

Colomé, À. (2001). Lexical activation in bilinguals' speech production: Language-specific or language-independent? *Journal of Memory and Language*, 45(4), 721-736.

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: implications for models of lexical access. J*ournal of Experimental Psychology: Learning, Memory, and Cognition*, 26(5), 1283.

de Bruin, A., Della Sala, S., & Bak, T. H. (2016). The effects of language use on lexical processing in bilinguals. *Language, Cognition and Neuroscience*, 31(8), 967-974.

Dryer, S.M. (2013). Position of Interrogative Phrases in Content Questions. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Dussias, P. E., & Scaltz, T. R. C. (2008). Spanish–English L2 speakers' use of subcategorization bias information in the resolution of temporary ambiguity during second language reading. *Acta psychologica*, 128(3), 501-513.

Friedmann, N., Belletti, A. & Rizzi, L. (2009). Relativized relatives: types of intervention in the acquisition of A-bar dependencies, *Lingua* 119(1), 67–88.

Hoff, E. (2018). Bilingual development in children of immigrant families. *Child Development Perspectives*, 12(2), 80-86.

Hopp, H. (2015). Individual differences in the second language processing of object–subject ambiguities. *Applied Psycholinguistics*, 36(2), 129-173.

Hurewitz, F., Brown-Schmidt, S., Thorpe, K., Gleitman, L. R., & Trueswell, J. C. (2000). One frog, two frog, red frog, blue frog: Factors affecting children's syntactic choices in production and comprehension. *Journal of Psycholinguistic Research*, 29(6), 597-626.

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, 8(4), 441-480.

Kidd, E., & Bavin, E. L. (2005). Lexical and referential cues to sentence interpretation: An investigation of children's interpretations of ambiguous sentences. *Journal of Child Language*, 32(4), 855-876.

Kidd, E., Stewart, A. J., & Serratrice, L. (2011). Children do not overcome lexical biases where adults do: The role of the referential scene in garden-path recovery. *Journal of Child Language*, 38(1), 222-234.

Lemmerth, N., & Hopp, H. (2019). Gender processing in simultaneous and successive bilingual children: cross-linguistic lexical and syntactic influences. *Language Acquisition*, 26(1), 21-45.

Lew-Williams, C. (2017). Specific Referential Contexts Shape Efficiency in Second Language Processing: Three Eye-Tracking Experiments With 6-and 10-Year-Old Children in Spanish Immersion Schools. *Annual Review of Applied Linguistics*, 37, 128-147.

Matin, E., Shao, K. & Boff, K. (1993). Saccadic overhead: information processing time with and without saccades. *Perception & Psychophysics* 53, 372–80.

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, 55(3), 381-401.

Meir, N., Parshina, O., & Sekerina, I. A. (2020). The interaction of morphological cues in bilingual sentence processing: An eye-tracking study. In *Proceedings of the 44th Boston University Conference on Language Development* (pp. 367-389). Somerville, MA: Cascadilla.

Meroni, L., & Crain, S. (2003). On not being led down the kindergarten path. In *Proceedings of the 27th Boston University Conference on language development* (pp. 531-544). Somerville, MA: Cascadilla Press.

Mirman, D. (2014). Growth curve analysis: A hands-on tutorial on using multilevel regression to analyze time course data. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 36, No. 36).

Pan, H. Y., & Felser, C. (2011). Referential context effects in L2 ambiguity resolution: Evidence from self-paced reading. *Lingua*, 121(2), 221-236.

Pan, H. Y., Schimke, S., & Felser, C. (2015). Referential context effects in non-native relative clause ambiguity resolution. *International Journal of Bilingualism*, 19(3), 298-313.

Papangeli, A. & Marinis, T. (2010). Επεξεργασία δομικά αμφίσημων προτάσεων στην Ελληνική ως Γ1 και ως Γ2. [Processing of structurally ambiguous sentences in Greek as L1 and L2]. In: *Proceedings of the Annual Meeting of the Department of Linguistics, School of Philology, Faculty of Philosophy, Aristotle University of Thessaloniki*, May 2–3, 2009. Thessaloniki: Institute of Modern Greek Studies [Institute Manoli Triantafillidi], 477-486.

Paradis, J. (2005). Grammatical morphology in children learning English as a second language. *Language, Speech, and Hearing Services in Schools*, 36, 172-187.

Pontikas, G., Cunnings, I., & Marinis, T. (2023). Online processing of which-questions in bilingual children: Evidence from eye-tracking. *Journal of Child Language*, *50*(5), 1082-1118.

Pozzan, L., & Trueswell, J. C. (2016). Second language processing and revision of garden-path sentences: a visual word study. *Bilingualism: Language and Cognition*, 19(3), 636-643.

Roberts, L. (2012). Individual differences in second language sentence processing. *Language Learning*, 62, 172-188.

Roberts, L., & Felser, C. (2011). Plausibility and recovery from garden paths in second language sentence processing. *Applied Psycholinguistics*, 32(2), 299-331.

Roesch, A. D., & Chondrogianni, V. (2016). "Which mouse kissed the frog?" Effects of age of onset, length of exposure, and knowledge of case marking on the comprehension of wh-questions in German-speaking simultaneous and early sequential bilingual children. *Journal of Child Language*, *43*(3), 635-661.

Schoonbaert, S., Hartsuiker, R. J., & Pickering, M. J. (2007). The representation of lexical and syntactic information in bilinguals: Evidence from syntactic priming. *Journal of Memory and Language*, 56(2), 153-171.

Stone, K., Lago, S., & Schad, D. J. (2021). Divergence point analyses of visual world data: Applications to bilingual research. *Bilingualism: Language and Cognition*, 24(5), 833-841.

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.

Unsworth, S. (2007). Child L2, adult L2, child L1: Differences and similarities. A study on the acquisition of direct object scrambling in Dutch. *Language Acquisition*, 14(2), 215-217.

Unsworth, S. (2023). Shared syntax and cross-linguistic influence in bilingual children: Evidence from between-and within-language priming. *Linguistic approaches to bilingualism*. doi.org/10.1075/lab.22093.uns.

van Dijk, C., Aumeistere, A., Brouwer, S., Dijkstra, T., & Unsworth, S. (2022a). Cross-linguistic Influence Online: An Eye-Tracking Study on Pronoun Resolution in Simultaneous

Bilingual Turkish-Dutch Children. In Y. Gong & F. Kpogo (Eds.), *Proceedings of the 46th annual Boston University Conference on Language Development*, 153-166.

van Dijk, C., Dijkstra, T., & Unsworth, S. (2022b). Cross-linguistic influence during online sentence processing in bilingual children. *Bilingualism: Language and Cognition*, 1-14.

Vasić, N., & Blom, W. B. T. (2011). Production and processing of determiners in Turkish-Dutch child L2 learners. In *BUCLD35: proceedings of the 35th annual Boston University Conference on Language Development* (pp. 616-627). Cascadilla Press.

Vasić, N., Chondrogianni, V., Marinis, T., & Blom, W. B. T. (2012). Processing of gender in Turkish-Dutch and Turkish-Greek child L2 learners. In *BUCLD36: proceedings of the 36th annual Boston University Conference on Language Development* (pp. 646-659). Cascadilla Press.

Vasilyeva, M., Waterfall, H., Gámez, P. B., Gómez, L. E., Bowers, E., & Shimpi, P. (2010). Cross-linguistic syntactic priming in bilingual children. *Journal of Child Language*, 37(5), 1047-1064.

Weighall, A. R. (2008). The kindergarten path effect revisited: Children's use of context in processing structural ambiguities. *Journal of Experimental Child Psychology*, 99(2), 75-95.

Williams, J. N., Möbius, P., & Kim, C. (2001). Native and non-native processing of English wh-questions: Parsing strategies and plausibility constraints. *Applied Psycholinguistics*, 22(4), 509-540.