Now you see me, now you don't: Children learn grammatical choices during online socially contingent video and audio interactions.

Leone Buckle[1*], Holly P. Branigan[2], Laura Lindsay[2], Katherine Messenger[1**]

[1]*Department of Psychology, University of Warwick, Coventry, UK*

[2]*Department of Psychology, University of Edinburgh, Edinburgh, UK*

Orcid IDs: LB: 0000-0002-8111-3685

HB: 0000-0002-7845-8850

KM: 0000-0002-9378-235X

The authors report there are no competing interests to declare.

*Corresponding author (and current affiliation): Dr Leone Buckle, Division of Psychology Communication & Human Neuroscience, University of Manchester, Manchester, UK; leone.buckle@manchester.ac.uk

** Current affiliation: Department of Psychology, Lancaster University, Lancaster, UK

Abstract

Previous research has established that children's experiences of language during in-person interactions (e.g., individual and cumulative experiences of structural choices) implicitly shape language learning. We investigated whether children also implicitly learn structural choices during online interactions, and whether this is affected by the visual co-presence of a partner. During an online conference call, three- and five-year-olds alternated describing pictures with an experimenter who produced active ('*a cat chased the dog*') and passive ('*the dog was chased by a cat*') prime descriptions; half the participants had video+audio calls, and half had audio-only. Children in both age groups produced more passives after passive than active primes, both immediately and with accumulating input across trials; neither effect was influenced by call format (video+audio vs audio-only). These results demonstrate that implicit grammar learning mechanisms, as evidenced by syntactic priming effects, operate during socially contingent online interactions. They also highlight the potential of online methodologies for developmental language production research.

155 words

Keywords: Syntactic priming; Structural priming; Online learning; Socially contingent interactions; Visual co-presence; Language development; Videoconferencing

Children's experiences of hearing conversational speech facilitate their language learning (e.g., Huttenlocher et al., 1991). Children often repeat words (Gershkoff-Stowe et al., 2006) and sentence structures (Branigan & Messenger, 2016) that they have heard their interlocutors use, and the diversity and complexity of children's vocabulary and syntactic structures is predicted by that of their caregivers (Huttenlocher et al., 2010). This vital linguistic input is assumed to come primarily from in-person interactions, but children are increasingly engaging in online interactions and using screen media for home learning and leisure, a trend that was particularly inflated by the recent COVID-19 global pandemic (Bergmann et al., 2022). For instance, 50% of UK-based 3- to 4-year-olds went online in 2021 to make video or voice calls or send messages, and 89% used video-sharing platforms (Ofcom, 2021-2022). Equally, 60% of US-based children under two years old made video calls several times a month (McClure, Chentsova-Dutton, Barr, Holochwost, & Parrott, 2015), while US-based parents of children aged 6 months to 6 years used screen media for their children's education almost as much as for their enjoyment (Nabi & Krcmar, 2016). A crucial question therefore is whether language experienced through live online interactions also supports children's language learning, and specifically grammatical learning in this case, and if so, under what circumstances.

Findings about the relationship between children's exposure to screen media, via television, computers, mobile devices or the internet, and their language abilities are conflicting (e.g., a negative relationship with productive grammar or vocabulary scores [Taylor et al., 2018; Kartushina et al., 2022] vs. no relationship with receptive or productive vocabulary scores [Selnow & Bettinghaus, 1982]). However, there is consistent evidence that under some but not all circumstances, screen-mediated language experience can support children's word learning. In particular, even though young children show reduced or non-existent word learning when they are exposed to recorded video+audio clips (the *video* *deficit;* Anderson & Pempek, 2005), they show significant learning when they experience the same input in "live" video+audio interactions (e.g., via platforms such as FaceTime, Zoom). Indeed, in some cases they show learning to the same extent from video+audio interactions as from in-person interactions. Thus, although children below 35 months show reduced or no learning of novel object or action labels from recorded video+audio clips compared to in-person demonstrations (Krcmar, 2010; Myers et al., 2017, 2018; Roseberry et al., 2014; Troseth et al., 2018), even 22-25 month olds are able to learn novel words from screen-mediated video+audio interactions (Myers et al., 2017), and 24-30 month olds learn novel verbs to the same extent from screen-mediated video+audio interactions as from in-person interactions (Roseberry et al., 2014).

Live video+audio conversations have been hypothesised to be beneficial for word learning because, like in-person interactions, they include social cues such as contingent responses to children's actions (in contrast to recorded video+audio utterances, which lack such cues and so are less likely to be treated as real-life interactions) (Troseth, 2010; see Glick, Saiyed, Kutlesa, Onisihi, and Nadig, 2022, for a review; see also Strouse and Samson, 2021). Socially-contingent interactions – in other words, interactions in which a partner responds in a timely and relevant way (Roseberry et al, 2014) – are argued to be effective for engaging children's attention and supporting joint attention to objects, including access to visual cues and gestures (Myers et al., 2017). Previous studies suggest that access to a partner's gaze may be particularly relevant for learning in these contexts, in part because eye contact can signal opportunities for learning relevant and generalizable information to young children (Gergely, Egyed, & Király, 2007). Accordingly, children who spent more time looking at the researcher's eyes during in-person and screen-mediated live video+audio conversations, and recorded video+audio demonstrations showed greater word learning (Roseberry et al., 2014).

Turning to children's grammatical development, there is strong evidence that language input experienced during in-person interactions supports children's learning of how syntactic structures map onto events. For example, children as young as 3 years old can learn to produce infrequent sentence constructions (e.g., passives) to describe depicted events immediately after hearing their partner use them within a turn-taking interaction task (Messenger, 2021; Branigan & Messenger, 2016). Like adults, children are thus susceptible to *syntactic priming* effects whereby they tend to re-use recently heard syntactic structures (e.g., passives vs actives: Messenger, 2021; double-object datives vs prepositional datives: Buckle et al., 2017; Rowland et al., 2012). Importantly, these effects are both cumulative, so that children more frequently attempt to produce these structures with increasing exposure to them (Branigan & McLean, 2016), and persistent, with children showing increased production of primed sentence structures a week, and in some circumstances a month, after initial exposure (Branigan & Messenger, 2016; Savage et al., 2006). These cumulative and lasting effects support the theory that children's individual experiences of hearing syntactic structures strengthen their nascent representations through implicit learning mechanisms. In particular, it has been suggested that these effects reflect error-based learning, whereby the syntactic processor makes predictions about the upcoming input and uses the error between the predicted and actual structure of a sentence to make long-lasting adjustments in the relative weights of the relevant representations that affect their subsequent likelihood of production (see Chang et al., 2006). In other words, syntactic priming effects are indicative of children's grammar learning experiences. However, all existing demonstrations of syntactic priming effects in young children come from in-person interactions; it is not known whether such experience-based grammar learning via syntactic priming also occurs during screen-mediated live video+audio interactions, as has been observed for word learning.

Using syntactic priming paradigms to examine children's propensity to learn structural choices through screen-mediated interactions can address questions that have taken on particular urgency in view of children's increasing engagement in online interactions, especially since the COVID-19 pandemic (see Glick et al., 2022). Most fundamentally, it would provide valuable insight into whether the implicit learning mechanisms underlying syntactic priming effects are responsive to language experience beyond the in-person contexts hitherto studied – specifically, whether screen-mediated, live language experience supports grammar learning via priming. Furthermore, by manipulating whether such interactions involve video+audio or audio-only input, it would cast light on whether such language learning is supported by different kinds of socially contingent interactions, or only those that involve visual co-presence. This latter question is particularly relevant as the research on word learning suggests that visual co-presence and access to eye gaze is important for learning. Finally, investigating whether children learn grammar from screen-mediated interactions has timely methodological implications by establishing whether syntactic priming paradigms, and language production paradigms more generally, can be used to successfully collect robust developmental data in online environments.

We therefore conducted a syntactic priming study using an online picture-description and -matching game to investigate whether children learned to produce passives during different kinds of socially contingent (live video+audio vs live audio-only) screen-mediated interactions. Specifically, we investigated whether children learned to use these infrequent syntactic structures from their online partner's language input, as manifested in an increased tendency to produce passive sentences both immediately after hearing their partner use a passive, and across the experiment as exposure to passives increased. We also explored whether this learning was modulated by whether they experienced video+audio input (their partner was both audible and visible providing social contingency through visual co-presence) vs. audio-only input (their partner was audible but not visible, reducing the social contingency). If visual co-presence in particular is important for learning through screen-mediated interactions, children should be more susceptible to priming in the video+audio condition than the audio-only condition.

Furthermore, we tested children from two different age-groups, to explore learning through online interactions at different stages of acquiring the passive. We tested three-year-olds who are likely to be in the early stages of passive acquisition: three-year-olds can understand passives in some contexts (Ibbotson et al., 2011; Messenger & Fisher, 2018) but often make errors in comprehension and production and are unlikely to produce them without support (Bencini & Valian, 2008; Bever, 1970; Shimpi et al., 2007). We compared their performance to five-year-olds who are likely to be in the later stages of passive acquisition and more able to comprehend and produce passives accurately (Marchman et al., 1991). In addition to age, we measured sentence comprehension skills in all participants as an additional means to exploring how these learning effects vary with language ability. If there is error-based learning, then 3-year-olds, at earlier stages of learning, should be more susceptible to priming than 5-year-olds, if they have a sufficiently developed representation that can be facilitated through experience (and similarly, children with poorer comprehension skills should be more susceptible to priming than children with better comprehension skills). However, given that 3-year-olds might be less accurate in producing a passive response, we might only detect such a pattern when considering attempts to produce passives, rather than when considering accurate passive responses only (Messenger, Branigan, Buckle & Lindsay, 2022).

Open Practices statement

This study was not preregistered. The data, analysis code and materials are publicly accessible on the Open Science Framework (OSF) at https://osf.io/gdvs9/.

Method

Participants

Fifty-eight three-year-olds (28 females; *M*age: 3;6; range: 3;2–3;11 years) and sixty-two five-year-olds (31 females; *M*age: 5;7; range: 5;2–5;11 years) with no reported developmental or language delays took part in the experiment. Thirty-two three-year-olds and 32 five-year-olds were randomly allocated to the video+audio condition; the remaining participants were allocated to the audio-only condition. All participants were monolingual British English speakers except three, who were simultaneously acquiring another language but still heard English from their primary caregiver at least 80% of the time. We excluded 17 additional three-year-olds and 1 five-year-old who did not produce valid target responses for at least 50% of the experimental trials. Participants were recruited online as a convenience sample from all over the UK via lab databases and social media. We interpret our data using Bayes Factors which are considered more effective than traditional power analyses in determining the sensitivity of datasets for hypothesis testing (see Dienes, 2014). Bayes Factors indicate the extent to which datasets provide credible evidence to either accept or reject a hypothesis, while being influenced by sample size to a lesser degree than *p* values. The Bayes Factors reported in Table 3 confirm that our sample size was sufficient to obtain conclusive results. The research was approved by the Humanities and Social Sciences Research Ethics Committee at the University of Warwick (reference number 129_17/18).

Design

The experiment had a 2 (*Age:* three-year-olds vs. five-year-olds; between-participants, within-items) x 2 (*Prime Structure*: active vs. passive*;* within-participants and -items) x 2 (*Input Form:* video+audio vs. audio-only*;* between-participants, within-items) design.

Materials

The experiment contained 56 items, each consisting of the experimenter's picture, their picture description sentence (the *prime*) and the participant's *target* picture (see OSF for a full list). There were 48 unique prime-target pairs with no open class lexical overlap between primes and targets. Primes featured eight verbs (*carry*, *knock over*, *shake*, *poke*, *pinch*, *pat*, *pull*, *lift*) and target pictures depicted eight different actions (*kiss*, *bite*, *scratch*, *lick*, *chase*, *squash*, *push*, *tickle)*. Each prime verb was paired with three target actions (e.g., *carry – bite/tickle/lick*). Verbs were paired with nouns denoting human and animal characters, which could be the agent or patient of the event. Corresponding active and passive sentences (e.g., *a cow is carrying a robber*/*a robber is being carried by a cow*) were created for each prime picture (see Figure 1). These were allocated to two lists in a Latin-square design so that half the participants encountered the active version and half encountered the passive counterpart for each item and overall, all participants encountered equal numbers of active and passive primes; across participants, each item occurred an equal number of times in each version. An additional 8 plural-agent intransitive sentences (e.g., *the girls are dancing*) were created for (filler) *Snap* trials, in which the experimenter and participant had identical items. Scripts with randomised orders of items for each individual participant were generated in *R* (*R* Core Team, 2021); however, *Snap* items were evenly distributed throughout the experiment to keep children motivated and engaged.

Participants also completed the Sentence Structure subtest of the Clinical Evaluation of Language Fundamentals (CELF) Preschool-2 UK assessment (Wiig et al., 2006), which is standardized for children aged 3 to 6 years. The Sentence Structure subtest asked children to pick the picture, from a choice of four, which best matched the sentence they heard; sentences increased in difficulty across trials, for example, from short declaratives such as intransitives or prepositional phrases to longer sentences such as passives, conjoined clauses, and adverbial clauses.

<Insert Figure 1 about here>

Procedure

We used a ‘Snap’ game task (Branigan et al., 2005), which is based on a popular and engaging UK children’s game and has been used in a number of in-person syntactic priming studies (e.g. Branigan & McLean, 2016; Branigan & Messenger, 2016; Messenger et al., 2011, 2012, 2021). In the task, the experimenter and participant alternated in describing pictures that were displayed side-by-side within a *PsychoPy* experiment (Peirce, et al., 2019) and shown via screen-share on a Microsoft Teams video call. The experimenter began by describing the first picture using the scripted prime sentence and then pressed a button to reveal the participant’s target picture, which the participant then described. The experimenter’s picture remained visible during the participant’s turn. Likewise, the participant’s picture remained while the experimenter revealed and described their next trial picture. Players shouted “*Snap*” whenever the two pictures onscreen matched. The experimenter and participant took turns describing the pictures until all 56 trials had been completed; whoever correctly shouted “*Snap*” first on the greatest number of trials won the game, however the experimenter always ensured that it was the child who won the Snap trials to maintain motivation. The experimenter kept their web camera on for participants in the video+audio condition but turned it off for participants in the audio-only condition (see Figure 2). All participants kept their camera on (except two three-year-olds, who were subsequently excluded for producing too few valid responses) and were therefore visible to themselves and the experimenter during the task. Each session was recorded and participants' target utterances were later transcribed and coded for syntactic form. At the end of the session, 47 of the three-year-olds and 58 of the five-year-old participants agreed to complete the CELF Sentence Structure test. There were 22 trials in total, but the test was stopped early if a participant made 5 consecutive incorrect responses, per the discontinuation rule of the assessment. Each participant received a score out of 22, having been awarded a score of 1 for each correct response, and 0 for each incorrect response.

<Insert Figure 2 about here>

Coding

For each experimental trial, we coded participants' first attempt at producing a target picture description, as well as the cumulative number of passive primes that the participant had heard the experimenter produce on all trials up to and including the current trial. We excluded instances where the child did not hear the experimenter (e.g., due to temporary loss of online connection) or the experimenter failed to produce the prime correctly. Two coders independently coded the responses in two ways, to reflect the fact that learning via syntactic priming can manifest both as increased successful production of structures (strict coding), and as increased attempted production of structures (lax coding); see Messenger et al. (2022).

We first applied a strict coding scheme to investigate how language experience affected children's ability to produce adult-like passive structures (see OSF for the full coding scheme and examples). Under this scheme, participant responses were coded as:

(1) *Active* responses if they had an agent subject, transitive verb and patient object and could otherwise be paraphrased as a passive;

(2) *Passive* responses if they had a patient subject, auxiliary and transitive past participle, by-phrase and agent object and could otherwise be paraphrased as an active;

(3) *Other structures* if they were incomplete, non-paraphrasable or reversed active/passives, truncated passives, passives with prepositions other than 'by', passives with present progressive verb morphology, or intransitives/ditransitives;

(4) *No response* if they did not contain a matrix verb or were produced immediately after either direct prompts from caregivers or an incorrect prime (e.g., wrong structure used).

Second, we applied a lax coding scheme to explore how language experience affected children's propensity to attempt to produce passive-like structures. Under this scheme, incomplete, non-paraphrasable or reversed utterances (i.e. those coded as *Other structures* in the strict scheme) were coded as either *Active* or *Passive* as appropriate. Truncated passives and passives with either incorrect prepositions or verb morphology were also coded as *Passive*. The frequency of participant responses under each coding scheme are shown in Tables 1 and 2. Cohen's Kappa (k) was calculated to determine the rate of inter-coder agreement between the two coders for 10% of the data obtained from each age group. Agreement was high for both the strict coding scheme (3-year-olds' data: 96.59% agreement, *Cohen's kappa,* $k = 0.948$; 95% CI [0.92 - 0.98]; $SE = 0.02$; 5-year-olds' data: 96.82% agreement, $k = 0.952$; 95% CI, [0.93-0.98]; $SE = 0.01$) and lax coding scheme (3-year-olds' data: 93.21% agreement, *k* = 0.882; 95% CI, [0.84 - 0.93]; *SE* = 0.024; 5-year-olds' data: 97.11% agreement, *k* = 0.948; 95% CI, [0.92-0.98]; *SE* = 0.016).

<Insert Tables 1 and 2 about here>

Results

We performed Bayesian statistical analysis in *R* (version 4.1.2; R Core Team, 2021) using sum contrast coding to examine the effects of Age (three-year-olds, coded 1, vs five-year-olds, coded -1), Prime Structure (passive=1 vs active=-1), and Input Form (audio-only=1 vs video+audio =-1) on the likelihood of participants producing passive target responses (coded 1, active responses were coded as 0, other responses were excluded; see Figure 3 for the proportion of strict-coded passive targets in each condition). We fitted the data to two Bayesian generalised regression models using the package *brms* (version 2.16.3, Bürkner, 2017); we ran one model for strict-coded target responses and another, identical model for lax-coded responses.

The models included a three-way interaction term with Age, Prime Structure, and Input Form as fixed effects (which were centred to reduce multicollinearity; Neter, 1985) plus the cumulative number of passive primes heard as a co-variate. We ran an additional model for each dataset with Sentence Structure score added and age removed. We added subject and item random effects. (Barr, 2013). We initially fitted maximal models with random slopes for all parameters but removed all random slopes for the models to converge (Barr, et al., 2013). For each model we set the initial parameters to zero and ran four chains for 6000 iterations using a warm-up period of 3,000 iterations. The estimate, estimated error and 95% credible interval values for each parameter were obtained from the model summary outputs and are reported in Table 3. Based on the observed data, there is a 95% probability that the true effect for each predictor lies within the reported credible intervals. In this instance, positive effect estimates and credible intervals with lower and upper values above 0 indicate that the predictor increases the likelihood of participants producing passive targets. Negative effect estimates and credible intervals with both values below 0 suggest that the predictor decreases the likelihood of passives being produced. Since 0 indicates a null effect, credible intervals that contain 0 indicate that the predictor does not reliably impact production of passives (see Hespanhol et al., 2019).

We calculated the Bayes Factor for each predictor to determine the probability of the alternative hypothesis (H_A), that the predictor has an effect on the frequency of passive target utterances, in relation to the null hypothesis (H_0), that the predictor has no effect. To do this, we compared a model with the predictor to a null model in which the predictor had been omitted. Our results are interpreted using Jeffery's (1961) evidence categories for Bayes Factors, as cited by Wetzels et al. (2011). Bayes Factors above 1 indicate evidence for the H_A; the evidence is considered strong when Bayes Factors are greater than 10 and decisive when they exceed 100. On the other hand, Bayes Factors below 1/3 are regarded as evidence for the H_0; the evidence is said to be strong when Bayes Factors are less than 1/10 and decisive when they are below 1/100. Bayes Factors between 1 and 1/3 indicate that there is insufficient evidence to accept/reject either the alternative or null hypotheses. The Bayes Factor for each of our predictors is also shown in Table 3.

<Insert Figure 3 about here>

<Insert Table 3 about here>

We obtained no Bayes Factors between 1 and 1/3 for any of our predictors, indicating credible evidence to either accept or reject each of our alternative hypotheses. Participants' age influenced their production of passives: Three-year-olds generally produced fewer passives than five-year-olds (Strict: M=0.13, SE=0.01 vs. M=0.23, SE=0.01; Lax: M=0.16, SE=0.01 vs. M=0.25, SE=0.01). Crucially, we found decisive evidence that syntactic priming occurs during online interaction: Overall, participants produced more passive targets immediately after hearing passive primes than active primes (Strict: M=0.28, SE=0.01 vs. M=0.11, SE=0.01; Lax: M=0.30, SE=0.01 vs. M=0.12, SE=0.01).

The lax-coded target responses showed evidence of cumulative priming effects: Participants were more likely to attempt to produce a passive response as the number of passive primes that they had heard increased across the priming task. However, this effect was not present for the strict-coded target responses: Participants were not more likely to produce a grammatically correct passive with increased exposure to passive primes. Importantly, both the strict and lax models strongly indicate that language input form (video+audio vs. audio-only) did not interact with prime structure to influence children's production of passives. We found no evidence that the effect of prime structure on children's production of passives differed between age groups, or that input form had a different effect between age groups.

Our results demonstrate children's increasing developmental competence in producing passives, indicating that syntactic representations are reinforced with increased prior experience producing and comprehending language. In comparison to three-year-olds, five-year-olds attempted to produce passive structures more frequently (with and without an immediate prime) and did so with fewer grammatical errors (grammatically correct passives: 3-year-olds: 54.79%; 5-year-olds: 70.56%). We ran a second analysis, replacing Age with the Sentence Structure measure of syntactic comprehension (SC Score; M_{score} out of 22 = 15.78, range = 3-22, SE= 0.41). This showed that higher sentence comprehension scores increased the likelihood of passive responses independently of prime structure (Bayes Factors: strict model = 5.26; lax model =1.35; see Table 4); the remaining pattern of results was the same with the exception that there were no cumulative priming effects evident in lax coded responses. Children's overall propensity to attempt to produce passives and their ability to do so using correct grammar were therefore closely tied to their ability to comprehend complex syntax.

<Insert Table 4 about here>

Discussion

Our study presents credible novel evidence (Jeffrey, 1969, as cited by Wetzels et al., 2011) that children at early and later stages of acquiring the passive are susceptible to immediate and cumulative effects of syntactic experience, manifested as syntactic priming effects, during socially contingent screen-mediated interactions. During a conference call, children were more likely to attempt (lax coding) and successfully use (strict coding) passive structures immediately after hearing an experimenter use a passive structure than after an active structure; though five-year-olds produced more passives overall, this priming effect did not differ across age groups. Children were also more likely to attempt to produce passives with increasing cumulative exposure to them (lax coding). Importantly, we demonstrate that strong priming effects can be elicited via screen-mediated online interactions, as in in-person interactions, and moreover, children showed these priming effects to the same extent irrespective of the form of screen-mediated interaction (audio+video vs audio-only).

Our results indicate that the experience-based implicit learning mechanisms that have been proposed to support syntactic priming function in a wider range of contexts than the in-person interactions previously studied, including contexts that do not involve visual co-presence with a conversational partner. These findings provide novel evidence that children can learn not only single words (e.g. Gaudreau et al., 2020) but also how to use infrequent syntactic structures via socially contingent interactions online, showing immediate priming effects (M=17% priming) that are remarkably consistent with those reported in in-person syntactic priming studies using a similar method and materials (e.g., Branigan & Messenger, 2016 – 17% priming; Messenger, 2021 – 18% priming) and cumulative priming effects across the session, which – as in in-person studies – emerged only on the lax-coded data (Branigan & McLean, 2016). It is difficult to disentangle whether cumulative effects were observed only in the lax coding due to differences in knowledge representations captured by the different coding schemes or simply due to the greater number of datapoints in the lax-coded data.

We did not find any evidence for age- or comprehension skill-related modulations of these effects, as might have been expected under an error-based learning account. Our findings in this respect are in keeping with previous studies, which have largely found no reliable effects of age or language skill on priming (e.g., Branigan & McLean, 2016; Kidd, 2012 [age]; Messenger, Branigan, & McLean, 2011; Rowland et al., 2012; though cf. Kidd, 2012 [language skill]). We suggest that the failure to find an age-related effect – here and in previous studies – may in part reflect variability among the sample in terms of both competence and performance. Priming is contingent on children having a sufficiently developed representation that can be facilitated through experience, and the detection of priming effects is in turn contingent on children being able to produce the primed structure (Messenger et al., 2022). It may therefore be difficult to detect age-related effects, and specifically the predicted decrease in priming with increasing age, because, while some younger children may have both the relevant representation and the ability to produce it accurately (and thereby demonstrate the predicted larger effects), some younger children may have such a representation but be unable to produce it accurately (making it difficult to detect priming effects reliably), and yet others may not have even a nascent representation that can be primed. This possibility is supported by recent findings of a decrease in priming with increasing age, but only in a subset of participants with more advanced language skills (Kumarage, Donnelly, & Kidd, 2022). This account cannot of course explain the failure to find comprehension-skill-related effects, however, we note that a number of participants (disproportionately in the three-year-old group) did not complete the CELF, reducing our power to detect this effect.

Our results suggest that previously observed deficits in children's word learning from pre-recorded videos (e.g. Roseberry et al., 2009) may have a basis in a lack of social contingency rather than screen-mediated linguistic input per se (see Myers et al., 2017). It is worth noting, however, that our evidence indirectly suggests that different forms of social contingency may impact learning at different developmental stages. The subset of children who *were* able to complete the task showed similar priming effects (and produced the same overall rate of passives) irrespective of whether they interacted with the experimenter in video+audio or audio-only contexts. However, in the wider sample of children tested, we found broad deficits in target response quality during audio-only interactions that diminished with age. Recall that seventeen three-year-olds were excluded because they did not produce sufficient valid responses; sixteen were in the audio-only condition (versus one in the video+audio condition); many of these were highly disengaged and some said that they did not like being unable to see the experimenter's face. In contrast, all but one five-year-old completed the experiment, and five-year-olds generally produced fewer invalid responses (N=87; cf. three-year-olds: N=315), with noticeably more task engagement during audio-only interactions. We cannot deduce from our data which aspect of the missing visual context was important for the three-year-olds in the audio-only condition, nor can we rule out the possibility that audio-only interactions might induce greater engagement if the audio were altered to make it more engaging when experienced alone. Nonetheless, this discrepancy in three- and five-year-olds' behaviour suggests that, in this experiment at least, visual access to a conversational partner may facilitate interactions that subsequently lead to language learning in three-year-olds more broadly (see Roseberry et al, 2014, for more pronounced effects in younger children), but may be less impactful for five-year-olds (see also Strouse, Troseth, O'Doherty, and Saylor, 2018).

Methodologically, our study is the first demonstration that this particular language production paradigm is effective in an online, screen-mediated setting for robust developmental data collection. Online testing provides flexible and efficient participant testing and may reach participants on larger geographical scales than often possible with in-person studies. Our study also suggests that such paradigms may be most effective with younger children when they include visual contact with participants. Moreover, our results suggest that children make use of input via screen-mediated conversations much as they do from in-person interactions. As such, children's increasing use of online contexts to communicate and interact with others (Bergmann et al., 2022; McClure et al., 2015; Ofcom 2021-2022) may in fact provide beneficial implicit linguistic input that supports their learning about how to use structures within their language. Although our participants were typically developing children, a substantial body of research has suggested that syntactic priming methodologies can be an effective way of supporting language learning in atypically developing populations (e.g., children with developmental language disorder; see Leonard, Krok & Wisman Weil, 2022, for a review). Our findings further suggest that there may be potential to carry out such interventions effectively online, thereby increasing their accessibility. Follow-up studies on whether learning effects are equally persistent after online, screen-mediated interactions as in-person interactions (i.e., lasting for at least a week;

Branigan & Messenger, 2016) would provide valuable insight into the extent to which the implicit learning mechanisms that support syntactic priming also function to strengthen linguistic representations and support longer-term language learning in different contexts.

In conclusion, we show that linguistic input from different forms of socially contingent screen-mediated interactions (video+audio and audio-only) can facilitate children's production of dispreferred, complex sentence structures. These results are consistent with implicit learning accounts of syntactic priming effects and extend developmental findings to show that such effects are robust to online contexts, through which young children may increasingly experience linguistic input. Importantly, these findings indicate that even short, task-based online interactions with children can provide a rich source of input that children are able to exploit for grammar development, in much the same way as they do when interacting in-person.

Word count: 5147 (excluding abstract and references)

Acknowledgements

This research was funded in whole by the Economic and Social Research Council, grant number ES/R007721/1. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence (where permitted by UKRI, 'Open Government Licence' or 'Creative Commons Attribution No-derivatives (CC BY-ND) licence' may be stated instead) to any Author Accepted Manuscript version arising from this submission.

We thank Abi Taylor and Abbey Parker for their assistance with data collection.

References

Anderson, D. R., & Pempek, T. A. (2005). Television and very young children. *American Behavioral Scientist*, *48*(5), 505-522.

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in psychology*, *4*, 328.

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.

Bencini, G., & Valian, V. (2008). Abstract sentence representations in 3-year-olds: Evidence from language production and comprehension. Journal of Memory and Language, 59, 97–113.

Bever, T. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.). Cognition and the development of language (pp. 279–362). New York: John Wiley.

Branigan, H. P., & McLean, J. F. (2016). What children learn from adults' utterances: An ephemeral lexical boost and persistent syntactic priming in adult–child dialogue. *Journal of Memory and Language*, *91*, 141-157.

Branigan, H. P., & Messenger, K. (2016). Consistent and cumulative effects of syntactic experience in children's sentence production: Evidence for error-based implicit learning. *Cognition*, *157*, 250-256.

Branigan, H., McLean, J., & Jones, M. (2005). A blue cat or a cat that is blue? Evidence for abstract syntax in young children's noun phrases. In A. Brugos, M. Clark-Cotton, & S. Ha (Eds.), *BUCLD 29: The Proceedings of the Twenty-Ninth Boston University Conference on Language Development* (pp. 109–121). Somerville, MA: Cascadilla Press.

Buckle, L., Lieven, E., & Theakston, A. L. (2017). The effects of animacy and syntax on priming: A developmental study. *Frontiers in Psychology, 8*, 2246.

Bürkner, P., (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software, 80*(1), 1-28.

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, *113*(2), 234.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in psychology*, *5*, 781.

Dimitrova, Nevena, Bergmann, Christina, Alaslani, Khadeejah, Almohammadi, Alaa, Alroqi, Haifa, Aussems, Suzanne, Barokova, Mihaela, Davies, Catherine, Gonzalez-Gomez, Nayeli, Gibson, Shannon P., Havron, Naomi, Horowitz-Kraus, Tzipi, Kanero, Junko, Kartushina, Natalia, Keller, Christina, Mayor, Julien, Mundry, Roger, Shinskey, Jeanne and Mani, Nivedita (2022) *Young children's screen time during the first COVID-19 lockdown in 12 countries.* Scientific Reports, 12 . 2015.

Gaudreau, C., King, Y. A., Dore, R. A., Puttre, H., Nichols, D., Hirsh-Pasek, K., & Golinkoff, R. M. (2020). Preschoolers benefit equally from video chat, pseudo-contingent video, and live book reading: implications for storytime during the coronavirus pandemic and beyond. *Frontiers in Psychology*, *11*, 2158.

Gergely, G., Egyed, K., & Király, I. (2007). On pedagogy. Developmental Science, 10, 139–146. https://doi.org/10.1111/j.1467-7687.2007.00576.x

Gershkoff-Stowe, L., Connell, B., & Smith, L. (2006). Priming overgeneralizations in two- and four-year-old children. *Journal of Child Language*, *33*(3), 461-486.

Glick, A. R., Saiyed, F. S., Kutlesa, K., Onishi, K. H., & Nadig, A. S. (2022). Implications of video chat use for young children's learning and social–emotional development: Learning words, taking turns, and fostering familial relationships. In *Wiley*

Interdisciplinary Reviews: Cognitive Science (Vol. 13, Issue 5). John Wiley and Sons Inc. https://doi.org/10.1002/wcs.1599

Hespanhol, L., Vallio, C. S., Costa, L. M., & Saragiotto, B. T. (2019). Understanding and interpreting confidence and credible intervals around effect estimates. *Brazilian Journal of Physical Therapy, 23*(4), 290-301.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*(2), 236.

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, *61*(4), 343-365.

Ibbotson, P., Theakston, A., Lieven, E., & Tomasello, M. (2011). The role of pronoun frames in early comprehension of transitive constructions in English. Language Learning and Development, 7, 24–39.

Kidd, E. (2012). Individual differences in syntactic priming in language acquisition. *Applied Psycholinguistics*, *33*(2), 393–418. https://doi.org/10.1017/S0142716411000415

Krcmar, M. (2010). Can social meaningfulness and repeat exposure help infants and toddlers overcome the video deficit?. *Media Psychology*, *13*(1), 31-53.

Kuhl, P. K., Tsao, F. M., & Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, *100*(15), 9096-9101.

Kumarage, S., Donnelly, S., & Kidd, E. (2022). Implicit learning of structure across time: A longitudinal investigation of syntactic priming in young English-acquiring children. *Journal of Memory and Language*, *127*, 104374. https://doi.org/10.1016/J.JML.2022.104374

Leonard, L. B., Krok, W. & Wisman Weil, L. (2022). Syntactic priming and language intervention for children with grammatical deficits In. K. Messenger (Ed.) *Trends in Language Acquisition Research, 31: Syntactic Priming in Language Acquisition: Representations, Mechanisms and Applications* (pp. 203–224). John Benjamins

McClure, E. R., Chentsova-Dutton, Y. E., Barr, R. F., Holochwost, S. J., & Parrott, W. G. (2015). "Facetime doesn't count": Video chat as an exception to media restrictions for infants and toddlers. *International Journal of Child-Computer Interaction*, *6*, 1–6. https://doi.org/10.1016/j.ijcci.2016.02.002

Marchman, V. A., Bates, E., Burkardt, A., & Good, A. B. (1991). Functional constraints of the acquisition of the passive: Toward a model of the competence to perform. First Language, 11, 65–92.

Messenger, K. (2021). The persistence of priming: Exploring long-lasting syntactic priming effects in children and adults. *Cognitive Science*, *45*(6), e13005.

Messenger, K., Branigan, H. P., & McLean, J. F. (2011). Evidence for (shared) abstract structure underlying children's short and full passives. *Cognition*, *121*(2). https://doi.org/10.1016/j.cognition.2011.07.003

Messenger, K., Branigan, H. P., McLean, J. F., & Sorace, A. (2012). Is young children's passive syntax semantically constrained? Evidence from syntactic priming. *Journal of Memory and Language*, *66*(4), 568–587. https://doi.org/10.1016/j.jml.2012.03.008

Messenger, K., Branigan, H., Buckle, L., & Lindsay, L. (2022). How does syntactic priming experience support language development? In. K. Messenger (Ed.) *Trends in Language Acquisition Research, 31: Syntactic Priming in Language Acquisition: Representations, Mechanisms and Applications* (pp. 57–82). John Benjamins

Messenger, K., & Fisher, C. (2018). Mistakes weren't made: Three-year-olds' comprehension of novel-verb passives provides evidence for early abstract syntax. *Cognition*, *178*, 118–132. https://doi.org/10.1016/J.COGNITION.2018.05.002

Myers, L. J., Keyser, H., & Cors, M. (2019). Co-viewers support participation in video chat interactions, but live experiences promote richer word learning for 24-to 36-month-olds in the USA. *Journal of Children and Media*, *13*(4), 415-432.

Myers, L. J., LeWitt, R. B., Gallo, R. E., & Maselli, N. M. (2017). Baby FaceTime: Can toddlers learn from online video chat? *Developmental Science*, *20*(4), e12430.

Nabi, R. L., & Krcmar, M. (2016). It takes two: the effect of child characteristics on US parents' motivations for allowing electronic media use. *Journal of Children and Media*, *10*(3), 285-303.

Neter, J., Wasserman, W., & Kutner, M. H. (1985). Multicollinearity, influential observations, and other topics in regression analysis—II. *Applied statistical linear models*, *101*, 390-393.

Ofcom https://www.ofcom.org.uk/data/assets/pdf_file/0025/217825/children-and-parents-media-use-and-attitudes-report-2020-21.pdf

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. Behavior Research Methods.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Roseberry, S., Hirsh-Pasek, K., & Golinkoff, R. (2014). Skype me! Socially contingent interactions help toddlers learn language. *Child Development, 85 3*, 956-970

Roseberry, S., Hirsh-Pasek, K., Parish-Morris, J., & Golinkoff, R. M. (2009). Live action: Can young children learn verbs from video? *Child Development*, *80*(5), 1360-1375.

Rowland, C. F., Chang, F., Ambridge, B., Pine, J. M., & Lieven, E. V. (2012). The development of abstract syntax: Evidence from structural priming and the lexical boost. *Cognition*, *125*(1), 49-63.

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2006). Structural priming as implicit learning in language acquisition: The persistence of lexical and structural priming in 4-year-olds. *Language Learning and Development*, *2*(1), 27-49.

Selnow, G. W., & Bettinghaus, E. P. (1982) Television exposure and language development. *Journal of Broadcasting & Electronic Media*, 26:1, 469-479.

Shimpi, P. M., Gámez, P. B., Huttenlocher, J., & Vasilyeva, M. (2007). Syntactic priming in 3-and 4-year-old children: Evidence for abstract representations of transitive and dative forms. Developmental Psychology, 43, 1334–1346.

Strouse, G. A., & Samson, J. E. (2021). Learning From Video: A Meta-Analysis of the Video Deficit in Children Ages 0 to 6 Years. *Child Development*, *92*(1), e20–e38. https://doi.org/10.1111/cdev.13429

Strouse, G. A., Troseth, G. L., O'Doherty, K. D., & Saylor, M. M. (2018). Co-viewing supports toddlers' word learning from contingent and noncontingent video. *Journal of Experimental Child Psychology*, *166*, 310–326. https://doi.org/10.1016/j.jecp.2017.09.005

Taylor, G., Monaghan, P., & Westermann, G. (2018) Investigating the association between children's screen media exposure and vocabulary size in the UK. *Journal of Children and Media,* 12:1, 51-65.

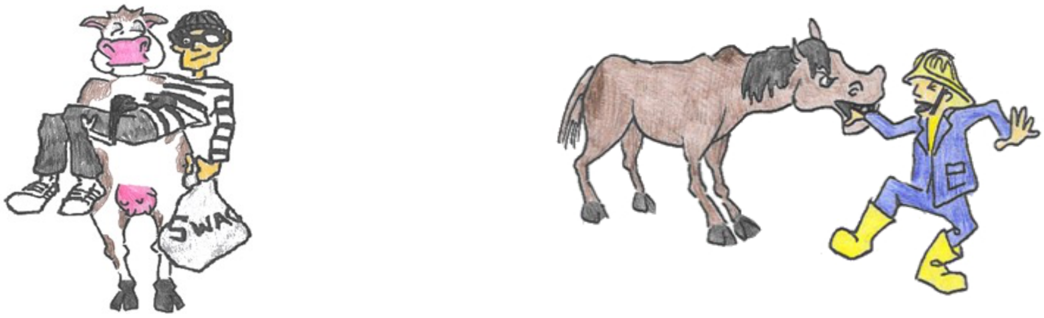
Troseth, G. L. (2010). Is it life or is it Memorex? Video as a representation of reality. *Developmental Review*, *30*(2), 155-175.

Troseth, G. L., Strouse, G. A., Verdine, B. N., & Saylor, M. M. (2018). Let's chat: On-screen social responsiveness is not sufficient to support toddlers' word learning from video. *Frontiers in Psychology*, *9*, 2195.

Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E. J. (2011). Statistical evidence in experimental psychology: An empirical comparison using 855 t tests. *Perspectives on Psychological Science*, *6*(3), 291-298.

Wiig, H., Secord, W. A., & Semel, E. (2006). *CELF preschool 2 UK: Clinical evaluation of language fundamentals, preschool*. Harcourt Assessment.

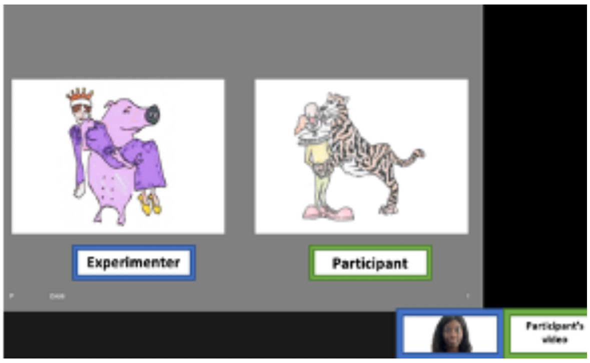
Figure 1. Example experimental item



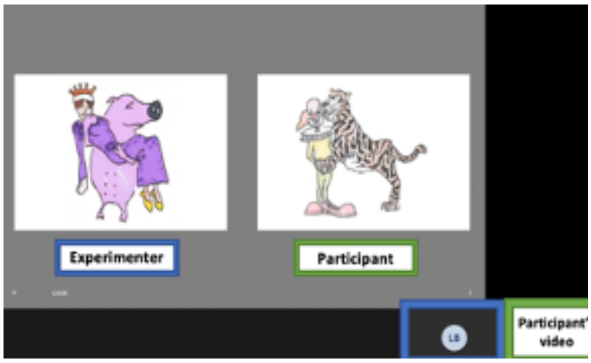
Active prime: *A cow is carrying a robber* Target: *horse biting fireman*

Passive prime: *A robber is being carried by a cow*

Figure 2. Experiment display in video+audio and audio-only conditions



Video+audio condition



Audio-only condition

Figure 3. Mean proportion of strict-coded passive targets produced by three- and five-year-olds after passive and active primes in each input form condition (SE in error bars)

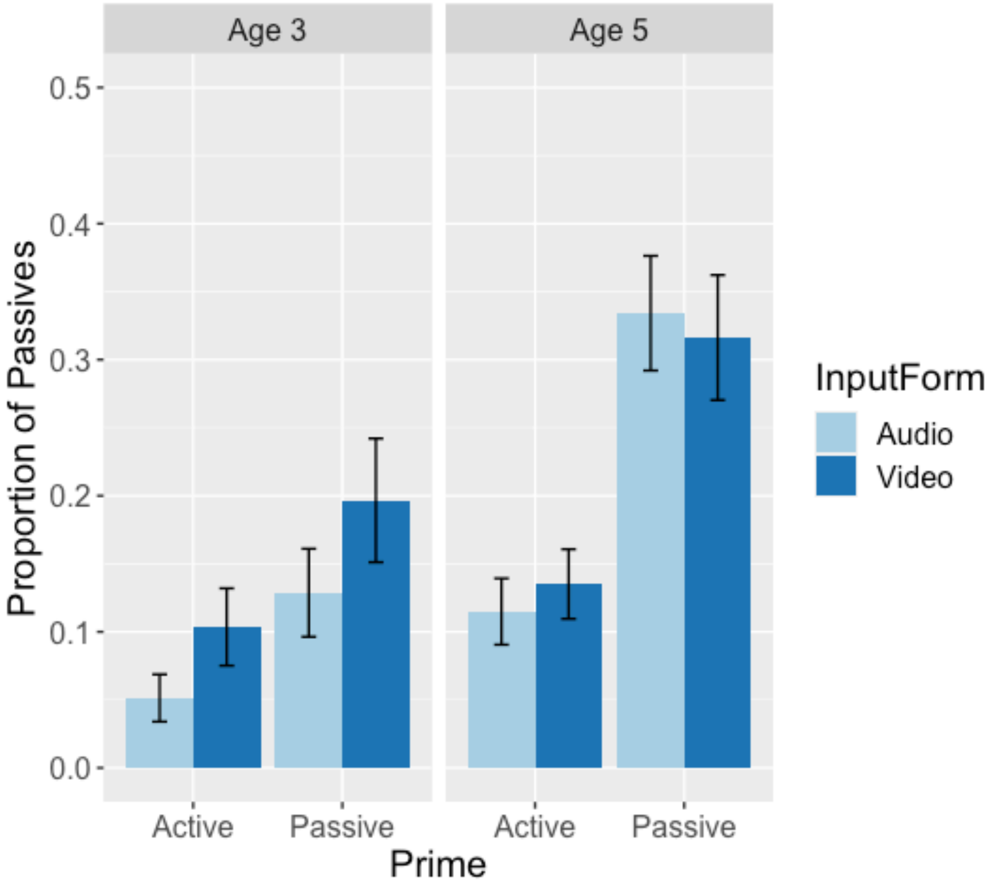


Table 1. Frequency of participant response types by Age, Input Form and Prime Structure according to strict coding scheme

		3-year-olds		5-year-olds	
Prime structure	**Response Type**	**Video+audio**	**Audio-only**	**Video+audio**	**Audio-only**
Active	Active	292	294	449	429
	Passive	31	20	73	61
	Other structures	359	246	231	213
	No response	89	67	15	15
Passive	Active	235	237	314	313
	Passive	67	42	153	166
	Other structures	369	277	266	221
	No response	94	65	35	22

Table 2. Frequency of participant response types by Age, Input Form and Prime Structure according to lax coding scheme

		3-year-olds		5-year-olds	
Prime structure	**Response Type**	**Video+audio**	**Audio-only**	**Video+audio**	**Audio-only**
Active	Active	486	382	562	526
	Passive	54	28	98	88
	Other structures	142	150	93	89
	No response	89	67	15	15
Passive	Active	390	321	427	395
	Passive	113	97	229	227
	Other structures	168	138	77	78
	No response	94	65	35	22

Table 3. Summary for Bayesian regression models (strict and lax) and Bayes Factors for predictors

Coding	Fixed effect	Estimate	Est. Error	2.5% CI	97.5 % CI	Bayes Factor
Strict	Intercept	-2.22	0.17	-2.58	-1.89	-
	Age	-0.45	0.14	-0.73	-0.18	23.49‡
	Prime Structure	0.66	0.06	0.55	0.77	>1000,000‡
	Input Form	-0.12	0.14	-0.39	0.15	0.21
	Cumulative number of passives	0.02	0.01	0.00	0.05	0.09
	Age x Prime Structure	-0.08	0.06	-0.20	0.04	0.16
	Prime Structure x Input Form	0.03	0.06	-0.08	0.14	0.06
	Age x Input Form	-0.11	0.14	-0.39	0.16	0.20
	Age x Prime Structure x Input Form	-0.02	0.06	-0.14	0.10	0.06
Lax	Intercept	-2.15	0.16	-2.47	-1.83	-
	Age	-0.46	0.14	-0.74	-0.19	39.20‡
	Prime Structure	0.71	0.05	0.62	0.80	>1000,000‡
	Input Form	-0.01	0.14	-0.28	0.27	0.14
	Cumulative number of passives	0.02	0.01	0.01	0.04	2.47†
	Age x Prime Structure	0.03	0.05	-0.07	0.12	0.06
	Prime Structure x Input Form	0.06	0.05	-0.04	0.15	0.10
	Age x Input Form	-0.05	0.14	-0.32	0.23	0.14
	Age x Prime Structure x Input Form	0.04	0.05	-0.06	0.14	0.07

†*Bayes Factor >1 indicates evidence for an effect.*

‡*Bayes Factor indicates strong (>10) or decisive (>100) evidence for an effect*

Table 4. Summary for Bayesian regression models (strict and lax) and Bayes Factors for predictors including sentence comprehension (SC) scores

Coding	**Fixed effect**	**Estimate**	**Est. Error**	**2.5% CI**	**97.5 % CI**	**Bayes Factor**
Strict	Intercept	-4.16	0.65	-5.47	-2.93	-
	SC Score	0.12	0.04	0.05	0.20	7.22†
	Prime Structure	0.68	0.27	0.16	1.23	7.31†
	Input Form	-0.27	0.52	-1.31	0.73	0.63
	Cumulative Number of Passives	0.02	0.01	-0.00	0.04	0.04
	SC Score x Prime Structure	0.00	0.02	-0.03	0.03	0.02
	Prime Structure x Input Form	-0.02	0.27	-0.55	0.51	0.27
	SC Score x Input Form	0.01	0.03	-0.05	0.07	0.03
	SC Score x Prime Structure x Input Form	0.00	0.02	-0.03	0.03	0.02
Lax	Intercept	-3.77	0.62	-5.00	-2.59	-
	SC Score	-0.10	0.04	0.03	0.17	1.52†
	Prime Structure	1.00	0.22	0.57	1.45	11697.06‡
	Input Form	0.41	0.51	-0.62	1.42	0.61
	Cumulative Number of Passives	-0.02	0.01	0.01	0.04	0.39
	SC Score x Prime Structure	-0.02	0.01	-0.04	0.01	0.03
	Prime Structure x Input Form	0.34	0.22	-0.09	0.77	0.77
	SC Score x Input Form	-0.02	0.03	-0.09	0.04	0.04
	SC Score x Prime Structure x Input Form	-0.02	0.01	-0.04	0.01	0.03

†*Bayes Factor >1 indicates evidence for an effect.*

‡*Bayes Factor indicates strong (>10) or decisive (>100) evidence for an effect.*