Anticipatory processing in a verb-initial Mayan language:
Eye-tracking evidence during sentence comprehension in Tseltal

Gabriela Garrido, Elisabeth Norcliffe, Falk Huettig, Penelope Brown, & Stephen Levinson

Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands

Corresponding author:

Gabriela Garrido

Email: Gabriela.Garrido@mpi.nl

Abstract

We present a visual world eye-tracking study on Tseltal (a Mayan language) and investigate whether verbal information can be used to anticipate an upcoming referent. Basic word order in transitive sentences in Tseltal is Verb-Object-Subject (VOS). The verb is usually encountered first, making argument structure and syntactic information available at the outset, which should facilitate anticipation of the post-verbal arguments. Tseltal speakers listened to verb-initial sentences with either an object-predictive verb (e.g., 'eat') or a general verb (e.g., 'look for') (e.g., "Ya slo'/sle ta stukel on te kereme", *Is eating/is looking (for) by himself the avocado the boy*/ "The boy is eating/is looking (for) an avocado by himself") while seeing a visual display showing one potential referent (e.g., avocado) and three distractors (e.g., bag, toy car, coffee grinder). We manipulated verb type (predictive *vs*. general) and recorded participants' eye-movements while they listened and inspected the visual scene. Participants' fixations to the target referent were analysed using multilevel logistic regression models. Shortly after hearing the predictive verb, participants fixated the target object before it was mentioned. In contrast, when the verb was general, fixations to the target only started to increase once the object was heard. Our results suggest that Tseltal hearers pre-activate semantic features of the grammatical object prior to its linguistic expression. This provides evidence from a verb-initial language for online incremental semantic interpretation and anticipatory processing during language comprehension. These processes are comparable to the ones identified in subject-initial languages, which is consistent with the notion that different languages follow similar universal processing principles.

Keywords: prediction, sentence comprehension, anticipatory eye-movements, verb-initial word order, Tseltal.

1. Introduction

To comprehend spoken sentences, listeners must extract words from the incoming speech signal, retrieve them from memory and integrate them into an interpretation. Despite the complexity of this process, listeners comprehend language with great speed and efficiency and with apparent ease. The results of a large and growing body of psycholinguistic research suggest that one reason why language processing may be so effortless, fast and efficient is because language users routinely make predictions about upcoming language input (Altmann & Mirković, 2009; DeLong, Urbach, & Kutas, 2005; Federmeier, 2007; Federmeier & Kutas, 1999; Huettig, Audring, & Jackendoff, in press; Kamide, 2008; Levy, 2008; Pickering & Garrod, 2013; Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005; Wicha, Moreno, & Kutas, 2004). Empirical evidence in support of anticipatory language processing has been demonstrated at many levels of linguistic structure (phonological, lexical, syntactic, conversational) using a variety of neurophysiological and behavioural research methods (e.g., event-related potentials, reaction times, eye-tracking). Yet in one way the empirical coverage of research on anticipatory processing remains small: it draws on a very restricted set of languages. Thus, while much psycholinguistic research emphasises the central, if not crucial role of anticipatory processing during language comprehension, there is currently very little data from speakers of typologically diverse languages that would allow us to test the generalisability of this claim.

The present study addresses this empirical gap by investigating anticipatory processing during sentence comprehension in Tseltal, a head-marking, verb-initial Mayan language. Using visual world eye tracking, our goal was to determine whether Tseltal listeners used information encoded in the sentence-initial verb to predict an upcoming direct object when listening to simple transitive sentences. As we discuss below, Tseltal verbs have certain semantic and morphological properties that might be expected to facilitate anticipation of an upcoming object. On the other hand, Tseltal is spoken largely in rural communities where there are lower levels of education and literacy compared to the kinds of populations typically studied by psycholinguists. Some research suggests that predictive processing may be attenuated by lack of experience with reading (Favier, Meyer, & Huettig, 2021; Huettig & Brouwer, 2015; Mani & Huettig, 2014; Mishra, Singh, Pandey, & Huettig, 2012; Ng, Payne, Stine-Morrow, & Federmeier, 2018). In our study, we assess whether such population-level factors modulate the extent of anticipatory processing in Tseltal.

In what follows, we first describe existing models of sentence comprehension that assume prediction as a key mechanism for language processing and then discuss eye tracking studies that support prediction during sentence comprehension. We then discuss the role of prediction in language processing and some of its limitations. Then we turn to Tseltal and introduce some of its linguistic properties. Finally, we present our experiment and discuss its results.

1.1 Anticipatory processing during sentence comprehension

Recent theoretical accounts suggest that prediction is a fundamental property of human information processing. In the domains of perception, action and learning, the brain is assumed to generate context-based predictions that guide and anticipate our processing goals (Bar, 2007, 2009; Clark, 2013; Friston, 2010). Many psycholinguists subscribe to the view that prediction, the pre-activation of linguistic input, is also central for language processing. Pickering and Garrod (2007, 2013) ascribe "a central role to prediction in language production, comprehension, and dialogue". Chang, Kidd, and Rowland (2013) argue that "prediction in processing is a by-product of language learning", though language acquisition constraints are critical for learning the syntactic and semantic representations that support prediction (Kidd, 2012; Rowland, Chang, Ambridge, Pine, & Lieven, 2012). Federmeier (2007; see also Federmeier and Kutas, 1999) concludes that the brain is continuously "thinking ahead" during sentence comprehension, using context to predict upcoming information while processing it at multiple levels. As a final example, Altmann and Mirković (2009) write that "most likely [...] prediction has a neural basis that pervades cortical function". In short, prediction is thought to be an essential component of language comprehension that allows people faster and more efficient processing of grammatical structure.

Predictive processing has been demonstrated at various levels of linguistic structure using a variety of experimental methods (e.g., reaction times, electrophysiology, eye-movements). Listeners use the information they have processed (linguistic and non-linguistic) to anticipate what will come next (Altmann & Mirković, 2009; DeLong, Groppe, Urbach, & Kutas, 2012; Federmeier & Kutas, 1999; Pickering & Garrod, 2013; Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005; Wicha, Moreno, & Kutas, 2004). Electrophysiological evidence shows that readers can form expectations about the phonological form of the word that will come next in sentences such as "The day was breezy, so the boy went outside to fly...*a kite vs. an airplane*" (e.g., DeLong, Urbach, & Kutas, 2005; but see Nieuwland et al., 2018 for a more nuanced account). People read

words faster when the word can be predicted from context, for example, predictable words are fixated less than unpredictable ones in highly-constraining contexts (e.g., Ehrlich & Rayner, 1981; Rayner & Well, 1996) or when the transitional probability between two words is greater than zero (McDonald & Shillcock, 2003). In addition, it has been demonstrated that syntactic structure can be predicted in coordinated sentences (Staub & Clifton, 2006) as well as in processing long-distance dependencies (Traxler, Bybee, & Pickering, 1997; Traxler & Pickering, 1996). Finally, there is evidence of prediction in every-day conversation. People predict upcoming turns in conversation from a variety of cues (e.g., pitch, the overall probability of turn durations of an interlocutor's current utterance, etc.) (Magyari, De Ruiter, & Levinson, 2017; Sacks, Schegloff, & Jefferson, 1974).

1.2 The Visual World Paradigm and sentence comprehension

Predictive processing during language comprehension has been studied using the Visual World Paradigm (VWP) (see Huettig, Rommers, & Meyer, 2011 for a review). This paradigm (Cooper, 1974; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995) provides a window into how visual perception, attention, working memory, and auditory input jointly determine the way speech is processed. Using the VWP, anticipatory eye-movements towards items within a visual display are taken as evidence for anticipatory processing.

In a seminal study, Altmann and Kamide (1999) demonstrated that semantic information extracted from verbs in English becomes rapidly available to listeners and guides their visual attention to an appropriate referent before it is mentioned. Participants' eye movements were recorded while they inspected a semi-realistic scene that depicted, for example, a cake, a toy car, a ball, or a toy train, and listened to sentences such as "The boy will eat the cake" or "The boy will move the cake". When participants heard the verb "eat", they directed their eye gaze anticipatorily towards the only edible item depicted (e.g., cake). In contrast, eye movements to the cake started much later when participants heard the verb "move". This pattern of looks was interpreted as reflecting anticipatory processing of an upcoming grammatical object that satisfied the semantic restrictions of the verb.

Using versions of this paradigm, subsequent work on English and a number of other languages has provided further evidence that prediction is important for sentence comprehension and that language users can rapidly integrate different kinds of information to generate predictions

about what they will hear next. Kamide, Scheepers and Altmann (2003) showed for German that case-marking information combined with verb semantic constraints can be used to predict upcoming post-verbal noun phrases (NPs). They found a similar effect for English postverbal NPs driven by the combination of verb semantics and voice information (active *vs*. passive marking). Kamide, Altmann and Haywood (2003) found that in both Japanese and English, information from multiple constituents can be rapidly integrated and used to predict upcoming elements. For English, listeners used information from the verb and its direct object to predict an upcoming goal argument (e.g., "bread", in "The woman will spread butter on the bread*"*). In Japanese verb-final ditransitive constructions (e.g., Waitress-NOM customer-DAT merrily hamburger-ACC bring, "The waitress will merrily bring the hamburger to the customer"), the combined information of the first and second NPs allowed the prediction of the third argument before hearing the verb and therefore, anticipation of the action itself. Anticipatory processing is not only driven by linguistic input, Knoeferle, Crocker, Scheepers, and Pickering (2005) demonstrated that information from visually presented events can be rapidly integrated with linguistic information to anticipate object referents in German verb-second and verb-final constructions.

The picture that emerges from these studies is that listeners exploit the particular 'affordances' of their language as they become available in order to generate predictions about upcoming language input. These findings, together with the wealth of converging evidence from electrophysiological and reading time studies point to an important, perhaps crucial role for prediction in language processing.

1.3 The role of prediction in language processing and its limitations

Experimental evidence of anticipatory processing in different fields such as language, action-perception and motor control have led to a widespread scientific consensus that prediction is a fundamental property of human language processing. Clark (2013) has claimed that "brains...are essentially prediction machines". There is, however, also evidence to suggest that prediction, while important, may not be *crucial* for processing language. These studies suggest that comprehenders do not always engage in predictive processing mechanisms and that some populations tend to rely less on predictive processing without compromising language comprehension. A good deal of research that derives from studies of language acquisition, second language learning, individual differences and cognitive resource limitations (e.g., working memory), and from different

populations such as healthy aged adults and illiterate speakers, has shown either no effects of predictive processing or only weak engagement of predictive mechanisms during language comprehension.

Several studies suggest that children's ability to predict upcoming linguistic input in spoken language processing is tied to a number of factors such as reading skills, oral language proficiency, and productive vocabulary size (Borovsky, Elmana, & Fernald, 2012; Kendeou, White, van den Broek, & Lynch, 2009; Mani & Huettig, 2012, 2014; Nation, Marshall, & Altmann, 2003). Studies on second language learning have shown that second language users are not able to predict (or appear to predict less well) upcoming expected words when reading in their second language compared to monolingual comprehenders (Hahne & Friederici, 2001; Kroll, Dussias, Bogulski, & Kroff, 2012; Martin et al., 2013; Zirnstein, van Hell, & Kroll, 2018). However, early exposed and highly proficient bilinguals appear to predict upcoming word referents to the same extent as monolingual speakers (Contemori & Dussias, 2020; Dijkgraaf, Hartsuiker, & Duyck, 2017; Zirnstein et al., 2018)

Research on healthy older adults has shown that older adults, as a group, are less able or slower to use contextual information and verbal knowledge to engage in predictive mechanisms during online language processing (e.g., N400 and P2 ERP effects are diminished) compared with younger adults (Federmeier & Kutas, 2005; Federmeier, Kutas, & Schul, 2010; Federmeier, McLennan, De Ochoa, & Kutas, 2002; Huang, Meyer, & Federmeier, 2012). In addition, there is evidence suggesting that working memory capacity and processing speed mediate anticipatory eye movements during language processing (Huettig & Janse, 2016; Huettig, Olivers, & Hartsuiker, 2011). Reductions in working memory capacity in older adults have been associated with delayed integration of contextual constraints in anticipating upcoming words (Federmeier & Kutas, 2005; Federmeier et al., 2010; Huang et al., 2012; Janse & Jesse, 2014; Wlotko & Federmeier, 2012). However, when controlling for age effects on memory and processing speed, advanced age participants were better at predicting words than younger adults (see Huettig & Janse, 2016) and older readers read slower, but showed larger predictability effects than their younger counterparts (Rayner, 2009; Rayner, Reichle, Stroud, Williams, & Pollatsek, 2006).

Reading ability (Favier, Meyer, & Huettig, 2021; Huettig & Brouwer, 2015; James & Watson, 2013; Mani & Huettig, 2014) and vocabulary knowledge (Hintz, Meyer & Huettig, 2017; Kukona et al., 2016; Rommers, Meyer & Huettig, 2015) have been related to anticipatory

language-mediated eye movements in adults. It has been suggested that formal literacy enhances predictive language processing in a number of ways including by facilitating the processes that involve the retrieval of associated words and the pre-activation of the grammatical representations of upcoming words (see Huettig & Pickering, 2019). Eye tracking studies have shown that adult high literates predict object names following a sentential context while low literates do not (Mishra, Singh, Pandey, & Huettig, 2012). In addition, proficient readers can rapidly use phonological and semantic information to anticipate upcoming referents, while less proficient readers rely on these types of information to a lesser extent and do not engage in anticipatory language processing (Huettig, Singh, & Mishra, 2011). An important aspect of literacy is vocabulary knowledge. Recent studies suggest that anticipatory eye movements are associated with high vocabulary scores in adults. In three eye-tracking visual world studies, Hintz, Meyer & Huettig (2017) evaluated the effects of different potential predictors (e.g., functional and general associations, receptive vocabulary knowledge, production skills, and non-verbal intelligence) of verb-mediated anticipatory eye movements and found that vocabulary knowledge is a robust predictor of anticipatory eye movements during language comprehension.

Finally, another potential mediating factor of prediction in language processing is *language* itself. To date, predictive processing has only been studied for a small handful of the world's languages (predominantly English, Dutch, German—all closely related Germanic languages - with just a few studies on East Asian languages (e.g., Japanese, Korean, Chinese). Given how dramatically languages vary from one another at all levels of organization (Evans & Levinson, 2009), an important question is whether and to what extent different kinds of linguistic structures themselves might attenuate or facilitate prediction during language processing, and how linguistic context might interact with other factors that have been shown to moderate prediction, such as individual differences in working memory, bilingualism or literacy levels.

1.4 The present study

The goal of this study is to expand the empirical base of research on predictive language processing by investigating sentence comprehension in Tseltal, a Mayan language spoken in the state of Chiapas in southern Mexico. As we describe in detail below, Tseltal has a number of grammatical properties that may arguably facilitate the generation of predictions during sentence processing. On the other hand, Tseltal is largely spoken in rural communities where there are lower levels of

education and literacy compared to the kinds of populations typically studied by psycholinguists. Possibly, any facilitating effect of language structure on predictive processing in Tseltal might be modulated by the literacy levels of the population and lack of experience with experimental procedures.

In order to examine whether Tseltal listeners use information provided by the sentence-initial verb to anticipate an upcoming object, we carried out a visual-world eye tracking study in which participants heard simple verb-initial transitive sentences while looking at visual displays. The paradigm employed is adapted from Altmann and Kamide's (1999) study. We outline the properties of Tseltal most relevant for the present study in detail below (for a full grammatical description of the language, see Polian, 2013), and we then turn to the experiment and its results.

1.4.1   The Tseltal language

Tseltal is a Mayan language spoken in Mexico by over 400,000 people (INEGI, 2010). One of the most striking features of Tseltal grammar is its constituent order: VOS (verb-object-subject) for transitive sentences (Robinson, 2002). Only about 2% of the languages in the world follow this basic word order (Dryer, 2013). Tseltal is also a 'head-marking' language: verbs carry agreement markers which index the grammatical roles as well as the person and number of the verb's arguments. Head marking in Tseltal is sensitive to transitivity (i.e., it is ergatively aligned): the agent (grammatical subject) of active transitive verbs is cross-referenced on the verb distinctly from the single argument (grammatical subject) of intransitive verbs, which is instead marked the same way as the object argument of a transitive verb. The two sets of markers on verbs are customarily labelled "set A" (for the prefixes that encode information about the agent argument of transitive verbs) and "set B" (for the suffixes that index the object/intransitive subject) (Polian, 2013). Both set A and set B markers differentiate between 1st-2nd-3rd person singular and plural.

Tseltal verbs do not specify if the action takes place in the present, past or future. There is no grammatical tense in this language; instead, aspectual distinctions are marked in verb phrases with auxiliaries (e.g., completive '*la*'; incompletive '*ya*', perfective '*-oj*' and '*-ej*', and progressive '*yakal*') (Polian, 2013).

The structure of a basic Tseltal transitive sentence is given in [1][1].

[1]	Ya	s-nutz-Ø	ts'i'	te	kerem=e
	INC	A3-chase-[B3]	dog	DET	boy=DET

The boy is chasing a dog.

Many basic-level transitive verbs[2] in Tseltal are highly semantically specific, imposing strong restrictions on the types of objects that can be involved in the event described by the verb. There are, for example, many different verbs for describing transitive events of eating, carrying, breaking, washing and placing objects. These verbs encode properties of the objects that they select for, such as their shape, substance or position (*ti'* is an eat verb for meat foods; **top** is a break verb for pottery items, etc.). They are in frequent use, morphologically simple (monomorphemic), and are not restricted to any particular register (Brown, 2008).

1.4.2 Task and predictions

We used a visual world eye tracking task to test whether Tseltal listeners engage in anticipatory processing when comprehending simple Verb-Object-Subject transitive sentences. In this task, participants listened to pre-recorded spoken sentences while viewing displays featuring an array of simple objects (see Figure 1). Displays showed one target referent (for example, an avocado) and three distractors (e.g., bag, toy car, coffee grinder). We manipulated verb type (predictive: the verb can only refer to one object in the display, e.g. "eat (soft things)"; general (control): the verb can refer to any of the items, e.g. "look at") and we recorded participants' eye-movements while they listened and inspected the visual scene.

If Tseltal listeners use information at the sentence-initial verb to predict the upcoming object, then they should look anticipatorily at the target object (the avocado) more in the predictive condition than in the general condition. Previous studies on English have shown that presenting a sentence-initial agent in combination with a predictive verb (e.g., eat) is enough to yield anticipatory eye-movements to an upcoming object (e.g., cake) (Altmann & Kamide, 1999). Other

---

[1] The abbreviations used in this article are INC: incompletive aspect; A3: ergative 3rd person marker; B3: absolutive 3rd person marker; DET: determiner; PRON: pronoun; P: preposition.
[2] Positional verbs are also highly specific in Tseltal. Intransitive predicates, on the other hand, are generally less selectionally restrictive (Brown, 2008).

studies have shown that a specific agent in combination with a verb (e.g., man rides *vs*. girl rides) constrains anticipatory eye-movements to the most plausible direct object (e.g., The man will ride the motorbike *vs*. The girls will ride the carousel) (Kamide, Altmann, et al., 2003). Here, we wish to know whether, when listening to VOS sentences, there are sufficient linguistic cues from the sentence-initial verb *without* its subject/agent (which occurs *last* in the sentence) to enable Tseltal listeners to predict the upcoming object.

The only eye-tracking comprehension studies to so far be undertaken on a verb-initial language are studies of anticipatory processing in Tagalog (Austronesian). Sauppe (2016) found that verbal information in Tagalog drove anticipatory eye movements towards the agent of the sentence, not the patient, regardless of its syntactic function (subject or object) and its position in the sentence (immediately after the verb or sentence finally). In contrast, Garcia, Garrido Rodríguez and Kidd (2021) found that morphosyntactic markers are used by Tagalog adults to anticipate the upcoming agent but only in patient-voice sentences (compared to agent-voice utterances). These results might reflect an agent bias during comprehension (Bickel, Witzlack-Makarevich, Choudhary, Schlesewsky, & Bornkessel-Schlesewsky, 2015; Bornkessel-Schlesewsky & Schlesewsky, 2013a, 2013b; Bornkessel & Schlesewsky, 2006; Cohn & Paczynski, 2013; Kemmerer, 2012; Sauppe, 2016; Wang, Schlesewsky, Bickel, & Bornkessel-Schlesewsky, 2009). Our displays in the Tseltal experiment, unlike those of Sauppe and Garcia et al., do not depict agents, and so we do not measure shifts of attention to the agent in our study. Nevertheless, if such an agent bias is also operative during Tseltal comprehension, it is possible that this might attenuate predictive looks to the target object. On the other hand, as we outline below, transitive verbs in Tseltal have semantic and morphological properties that strongly orient towards object referents and that may therefore enhance listeners' ability to predict upcoming objects.

We hypothesised that three particular features of Tseltal verbs would facilitate the use of predictive processing in the predictive condition. First, *sentence-initial verb placement*: unlike most previously studied languages, Tseltal verbs come first in the sentence. This means that already by the first word of a simple sentence, listeners have information about the event structure of the unfolding utterance. Moreover, because the subject noun phrase (or any other verbal argument) has not yet been mentioned, listeners do not have to devote cognitive resources to

integrating the subject argument into the ongoing parse, potentially freeing up resources for predictive processing.

Second, *head-marking morphology*: Tseltal's head-markers indicate the argument structure of the clause (whether the verb selects for one or two core arguments) and the person and number of those arguments. This means that having parsed the initial head-marked verb, listeners already know whether the clause will contain an object referent, and what the person and number features of that referent will be. This information arguably serves to restrict the possibility space when generating predictions about upcoming arguments.

Third, *verb specificity*. As described above, many transitive verbs in Tseltal are highly specific; their semantics strongly restricts the range of objects they can select for. Brown (2008) draws an interesting connection between the high semantic specificity of many of Tseltal's transitive verbs and a salient property of spoken Tseltal: in natural speech, object arguments are very often not mentioned. She suggests that the reason object ellipsis is so frequent in Tseltal speech is because objects are often easily recoverable from the verb's semantics. Supporting this possibility, Brown found in a corpus study that in both adult and child speech, object ellipsis occurred more often with semantically specific transitive verbs than with light verbs such as 'give' and 'see', which do not restrict the referent of their object argument (see also Resnik, 1996 for a similar correlation in English). If listeners are able to easily recover the object referents of semantically specific verbs, they should find it easier to anticipate them, compared to the referents of semantically general verbs.

Together, these three features of Tseltal transitive verbs (initial placement, head-marking morphology and semantic specificity) may facilitate the prediction of object arguments during the comprehension of transitive clauses. There are, however, population-level factors that may influence the extent to which Tseltal listeners engage in predictive language processing. As discussed above, several studies have shown that language mediated prediction may be modulated by formal literacy (Huettig & Mani, 2016; Huettig, Singh, et al., 2011). In a study of Hindi-speaking low and high literates, for example, low literates used information from the unfolding spoken words to direct their eye gaze, but unlike high-literates, they did not use such information for prediction (Huettig, Singh, & Mishra, 2011). Tseltal is largely spoken in communities where there are low levels of education and literacy. If the attenuating effects of literacy are stronger than the facilitating effects of language structure, then we may not find robust evidence for predictive

processing. Furthermore, life in Tseltal communities revolves primarily around subsistence agriculture and there is limited access and exposure to technology such as computers and to experimental settings. It has been suggested that people's ability to predict upcoming events is influenced by their level of expertise in the task at hand (e.g., sports psychology research: Aglioti, Cesari, Romani, & Urgesi, 2008; Mori, Ohtani, & Imanaka, 2002; Starkes, Edwards, Dissanayake, & Dunn, 1995; Williams, Ward, Knowles, & Smeeton, 2002). The fact that our Tseltal population is less specialized than the usual populations under study may therefore also affect the extent to which listeners generate predictions in experimental settings.

In the present study we recruited Tseltal participants with a range of different educational levels, allowing us to explore indirectly the extent to which these population-level factors might interact with the effects of linguistic context (verb specificity/ highly predictive verbs) on object prediction. This is not a study of literacy and its effects on predictive language processing. However, given the literature mentioned above, it is important to consider our study population's characteristics. We additionally took into account participants' level of bilingualism (in Spanish), as this has also been shown in previous studies to affect comprehenders' reliance on predictive processing (Contemori & Dussias, 2020; Dijkgraaf, Hartsuiker, & Duyck, 2017; Zirnstein et al., 2018).

Finally, we tested whether participants learned to predict over the course of the experiment – one hypothesis is that as they receive regular input about the language structures used in the experiment, and as they get more exposure to the experimental set-up itself, they might get better at generating predictions about the object argument. Thus, given that the population is not "specialized" at the outset, they may gain enough specialization over the course of the experiment to alter their processing strategies.

2. Methods

2.1 Participants

Sixty-seven native speakers of Tseltal from the indigenous community of Majosik' were recruited and paid for their time to take part in this study. Majosik' is a small rural community of ~1100 inhabitants that is located in the municipality of Tenejapa (Chiapas, Mexico). The population is largely monolingual and lives along traditional lines. 16 participants had to be excluded due to bad calibration of the eye-tracker (n = 4), external distractions (n = 2, e.g., the participant could not

13

pay attention because she had to attend to her child, etc.), they fell asleep during the experiment (n = 8, e.g., before the experiment they had been working the land), or the participant did not want to continue with the experiment (n = 2). Our final sample consisted of 51 participants (30 women, *M*age = 25.2 years, *SD* = 6.8, range: 16-42 years).

Conducting experiments in the field comes with the risk of higher participant exclusion rates than might be expected under ideal lab-based conditions (e.g., due to failure to understand or complete the experiment, technical issues such as power outages; see Norcliffe, Konopka, Brown, & Levinson, 2015 and Whalen & McDonough, 2015 for discussion of the practical challenges of field-based experimental research). For these reasons it is typical to run a higher number of participants than in an equivalent lab-based experiment, on the expectation that exclusion rates will be higher. In our study, we tested 67 participants, of which we were able to retain 51. In this context, we note that this sample size is somewhat higher than previous lab-based visual world studies that have observed verb-mediated anticipatory effects, which have had between 20-40 participants, (see e.g. Altmann & Kamide, 1999; Boland, 2005; Kamide, Scheepers & Altmann, 2003; Kako & Trueswell, 2000) but comparable to other eye-tracking studies (production and comprehension) conducted in the field, (Sauppe, 2016: Tagalog comprehension N = 49; Norcliffe, Konopka, Brown, & Levinson, 2015: – Tseltal production N= 53; Nordlinger, Garrido Rodriguez, & Kidd, (in press): Murrinhpatha production N = 43; Nordlinger, Garrido Rodriguez, Wilmoth, & Kidd, 2020: Pitjantjatjara production N = 50; Tollan, Massam, & Heller, 2019: Niuean comprehension N = 46).

We administered a short oral questionnaire in Tseltal asking participants to state their level of education (formal schooling) and knowledge of Spanish: whether they spoke the language, if they understood it, and whether they read in Spanish or not. Indigenous students attending school in Majosik' are taught both in Tseltal (L1) and Spanish (L2) during primary education. After that, instruction is provided only in Spanish (SEP, 2017). Indigenous schools are multigrade: students in the same course are monolingual speakers (Tseltal) or have different levels of bilingualism (speaking another indigenous language or Spanish). The mean education level of participants was 7.8 years of schooling (*SD* = 3.3, range: 0-12 years). In Mexico, 7th grade is equivalent to the first year of middle school (e.g., the first year of Junior High School). Participants' level of education varied widely. Three were illiterate, 14 participants had completed primary school (6 years of formal schooling), 15 had completed secondary school (9 years of schooling), and 11 participants

had completed high school/ COBACH (12 years of formal schooling). The rest of the participants had received some level of formal education (n = 2: 3 years; n = 2: 4 years; n = 1: five years; n = 1: seven years; n = 2: 11 years of education). 33 participants reported knowing some Spanish (speaking, reading and writing) and the rest described themselves as monolingual Tseltal speakers. All participants gave written or oral consent before the experiment. The study was approved by the ethics board of the Faculty of Social Sciences of the Radboud University, Nijmegen.

2.2 Materials and Design

The stimuli consisted of 32 pairs of transitive active sentences recorded in two experimental conditions: one sentence contained a predictive verb (e.g., "eat-soft thing") and the other sentence a general verb (e.g., "look for"). Example sentences are provided in [2].

[2]   Predictive verb
         Ya   s-**lo'**   ta   s-tukel   **on**   te   kerem=e
         INC   A3-eat.soft[B3] P   A3-PRON   avocado   DET   boy=DET
         The boy is eating (soft) by himself an avocado.

         General verb
         Ya   s-**le**   ta   s-tukel   **on**   te   kerem=e
         INC   A3-look[B3]   P   A3-PRON   avocado   DET   boy=DET
         The boy is looking (for) by himself an avocado.

Each sentence pair was accompanied by a visual display that depicted four items. Only one of the items fulfilled the semantic restrictions of the predictive verb (i.e., the target direct object), while the three other items were distractors. The four depicted items were possible referents of the general verb condition (see Figure 1). In all sentences the aspect and the verb were mentioned first, followed by an adverbial phrase (*ta stukel*, "by himself/herself"), the direct object, and finally the subject. The adverbial phrase was included to give participants more time to process verbal semantic information and direct their gaze anticipatorily to the target object. Similar stimuli sentence structures have been used in previous research (Kamide, Scheepers & Altmann, 2003; Sauppe, 2016; Garcia, Garrido Rodriguez & Kidd, 2021). The sentences were recorded with neutral intonation by a male native speaker of Tseltal. The recordings were sampled at 44100 Hz. The onsets and offsets of all words were marked using Praat (Boersma, 2002).
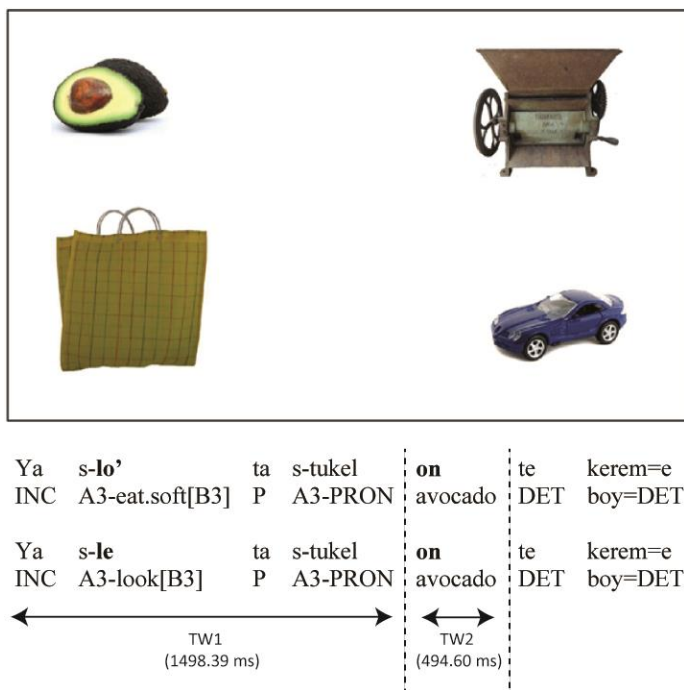
Ya	s-**lo'**		ta	s-tukel	**on**	te	kerem=e
INC	A3-eat.soft[B3]		P	A3-PRON	avocado	DET	boy=DET

Ya	s-**le**		ta	s-tukel	**on**	te	kerem=e
INC	A3-look[B3]		P	A3-PRON	avocado	DET	boy=DET

TW1
(1498.39 ms)

TW2
(494.60 ms)

Fig. 1 Stimulus example of visual display and sentential conditions. The auditory sentences and glosses are provided in Tseltal. The time windows relevant for analyses are indicated with arrows.

The visual displays were created using photographs of familiar items in daily use, which were taken at the field site by the researchers. These were supplemented with culturally appropriate free images that were available on the internet (Google Images). The images were processed using Photoshop so that they all had the same white background. 32 transitive sentences were used as fillers and each of them was paired to a visual display with four items. We followed Altmann and Kamide's (1999) within-subjects experimental design with respect to the filler sentences: these were divided into 4 sets of 8 sentences each. The direct object mentioned in the filler sentences was never depicted visually. However, the depicted items were chosen in a way that they could be possible referents of the verb. For example, in set 1, three of the items fulfilled the selectional requirements of the verb while one item was incompatible with the semantics of the verb. In set 2, two items depicted were possible referents of the verb and two items were not. In set 3, only one

16

item fulfilled the semantic requirements of the verb while the others were distractors. In set 4, none of the items depicted were possible targets of the verb.

Two lists of stimuli were created containing only one version of each sentential condition together with its accompanying visual display. The stimuli were arranged in a fixed-random order so that every experimental condition was followed by a filler sentence and the target items depicted would be in a different position on the screen from the previous trial. Participants were randomly assigned to the lists. The stimuli materials and a description of the objects in each scene are given in the Supplementary Materials (section A). All of our materials are available in an OSF repository (https://osf.io/gqv5c/?view_only=7fce527ed0a344888d2ea2d5dd80c5e9).

2.3 Apparatus and Procedure

Before the test session, participants read or were given overt instructions about the experiment in Tseltal. They were asked to complete a short questionnaire on their linguistic and educational background. After this, they were asked to sit in front of a 17" laptop computer with a resolution of 1024 x 768 pixels at a distance of approximately 58 cm. Their eye-movements were recorded with an SMI RED-M eye-tracker (SensoMotoric Instruments), which was attached to the base of the computer's screen, sampling at 120 Hz. The auditory stimuli were presented via headphones. Overt instructions were given for a second time by a native speaker of Tseltal. Participants were asked to listen to the sentences carefully and were told that they could look at whatever they wanted on the screen. They were not asked to perform any explicit task (i.e., this was a "look and listen" task, e.g., Huettig & Altmann, 2005; Huettig & McQueen, 2007; Huettig, Rommers, et al., 2011). It has been found that in the absence of any metalinguistic task, participants shift their visual attention around the scene as the acoustic stimulus unfolds (see Huettig, Rommers, et al., 2011 for discussion).

Each trial began with a centrally located fixation cross that appeared for 1000 ms. This was followed by a 1200 ms visual display preview that preceded the auditory stimulus. After this preview time, the sentence was played over the headphones while the display remained in view until the end of the trial. The experimental and filler visual displays were presented for 6000 ms in total. When the display disappeared, a fixation cross appeared in the center of the screen signalling the next trial (see Figure 2). There were four practice trials before the main experimental block. Before the practice and experiment session, the eye-tracker was calibrated using a 9-point

fixation dot. Calibration took about 20 seconds. The entire session lasted approximately 30 minutes.
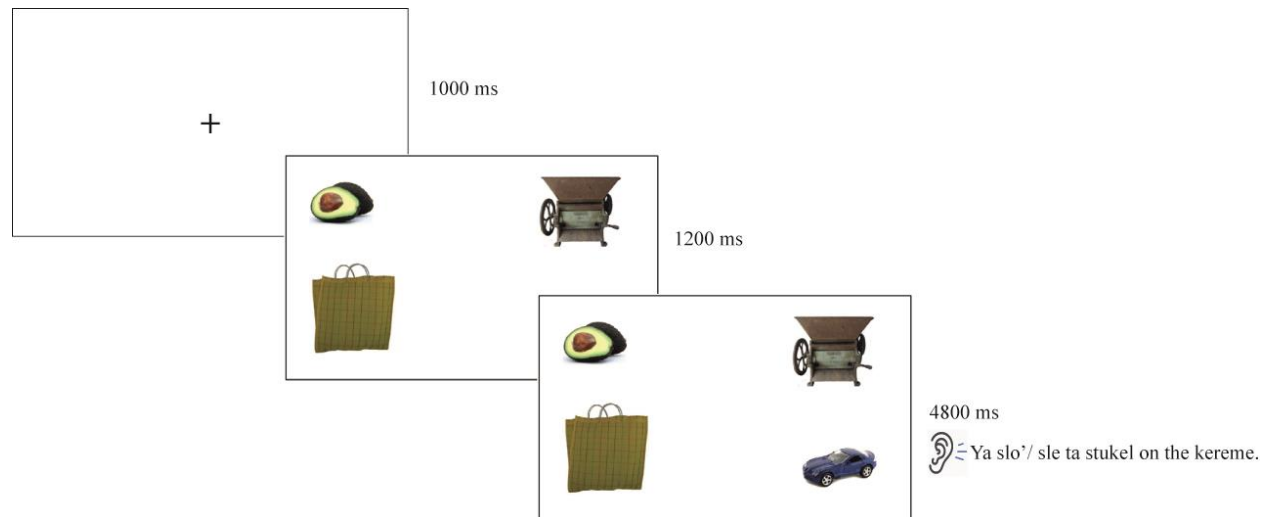

Fig. 2 Trial example. A fixation cross appeared for 1 s, this was followed by a 1200 ms visual preview of the visual display. After this, a sentence was played through the headphones while participants looked at the visual display.

2.4 Data analyses

The auditory sentential stimuli were segmented in Praat (Boersma, 2002) to measure the mean duration of verbs, adverbial phrases, direct objects, and subjects across the two sentential conditions in order to account for differing word lengths across our stimuli (see Table 1). We calculated the proportion of fixations to the target (e.g., avocado) and to the averaged distractor objects (e.g., market bag, toy car, coffee mill) and white space (empty areas on the screen) and their corresponding 95% confidence intervals (calculated by-participant and by-item for each sampling step) for both the predictive verb condition and the general verb condition. Figure 3 illustrates the time-course of fixations to the target and averaged distractors in the two sentential conditions.

We used multilevel logistic regression (cf. Barr, 2008, 2013; Jaeger, 2008) with random intercepts and slopes to analyse the dependent variable (i.e., fixations to the target: categorical dependent variable coded as 1 = yes, 0 = no) as a function of verb type (coded as 1 = predictive verb, 0 = general verb) and time (continuous variable measured in milliseconds). We estimated the contribution of verb type to the anticipatory eye-movements towards the target object referent and how these eye-fixations changed over time in two different time windows (TWs).

The first time window (TW 1), reflecting the *Verb + Adverb* region, included the aspect marker, the verb and the adverbial phrase, and was defined as from 200 ms after aspect onset until 200 ms after adverb offset (duration: $M = 1498.39$ ms, $SD = 172.99$ ms). Time window 1 is our predictive window. We expect verb-mediated anticipatory eye movements towards the appropriate object in this time window. The second time window (TW 2) covered the auditory presentation of the first noun phrase, the target *Object*. TW 2 started 200 ms after the onset of the object and ended 200 ms after the object was mentioned (duration: $M = 494.35$, $SD = 113.79$). This time window serves a control purpose: we expect fixations only to the target object during this region because participants are hearing the target noun in both sentential conditions. All of our analysis time windows were segmented according to the different regions of interest (i.e., word durations: *Verb + Adverb* and *Object* regions). For each analysis time window, we added 200 ms to *Verb + Adverb* and *Object* word onsets to adjust for the time it takes to program and launch a saccadic eye movement (Matin et al., 1993; Saslow, 1967; Duchowsky, 2007). The SMI Red-m eye tracker (sampling rate 120 Hz) sampled the eye position every 8.3 ms (e.g., fixations, saccades and blinks). Therefore, to account for variations in the duration of regions across stimuli due to differing word durations and the fixation pattern of each participant, the duration of each time window was standardized. That is, each time window was centered around the grand mean and divided by the standard deviation of the specific time window. We considered only fixations that occurred within the specific time window to compute such values. Thus, the fixed effect of time within each model has its own time scale and a value of 0 in this variable represents the scaled grand mean of the time window of interest. More details about the operationalization of the time variable can be found in an OSF repository (https://osf.io/gqv5c/?view_only=7fce527ed0a344888d2ea2d5dd80c5e9).

We conducted multilevel logistic regression models that included our main predictors of interest, verb type and time and their interaction (Agresti, 2019). The random effects structure included intercepts and slopes for verb type and time calculated for participants and items. The maximal random effects structure justified by design and that allowed the models to converge was used (Barr, 2013; Barr, Levy, Scheepers, & Tily, 2013). In addition, we conducted an exploratory analysis in which participants' level of education (continuous measure that ranged from 0 years of schooling to 12 years of formal education) and knowledge of Spanish (categorical coded as yes = 1, no = 0) were included as fixed effects to explore the extent to which these population

characteristics might interact with the effects of verb specificity on object prediction. The effect of experiment (i.e., number of trials) was also included as a predictor (categorical coded as first part of the experiment *vs*. second part) in this model, to test whether participants had learnt to anticipate an object target over the course of the experiment.

To evaluate the contribution of each predictor and see if it improved model fit, we generated models from the more reduced version (only one predictor) to the most complex one with the relevant effects and interactions (evaluated via forward model comparison, using the likelihood ratio test criterion) (Pinheiro & Bates, 2000). All models were fit with the Laplace Approximation for Maximum Likelihood using lme4 (version 1.1-21) function (Bates, Machler, Bolker, & Walker, 2015). Confidence intervals (95%) are provided for the regression coefficients (Snijders & Bosker, 2012). All statistical analyses were conducted in R (version 3.6.1). The model that best described the data had an interaction of verb type and time and random intercepts and slopes of our main predictors for participants and items. The summaries of model fit of the interaction of verb type and time during TW 1 (*Verb + Adverb* region) and TW 2 (*Object* region) are given in Table 4. We also report the estimates of the variance of the distribution of the random effects. Additional model summaries that included level of education, knowledge of Spanish, and part of the experiment can be found in the Supplementary Materials (section B). All analyses are available in an OSF repository (https://osf.io/gqv5c/?view_only=7fce527ed0a344888d2ea2d5dd80c5e9).

3. Results

The mean duration of verbs, adverbial phrases, direct objects, and subjects across the predictive and general sentential conditions can be found in Table 1. There were no significant differences in word duration across the two sentential conditions.

Table 1
t-test Results comparing mean word duration across sentential conditions (duration in ms).

Sentence	Condition				
	Predictive	General	*t*	*df*	*p*
Aspect + Verb	717.73	718.98	-0.04	62	0.966
Adverbial phrase	788.02	772.04	0.53	62	0.599
Object	498.27	490.43	0.27	62	0.785

Determiner + Subject	987.26	978.75	0.29	62	0.775

Figure 3 shows the time-course graph of the proportion of fixations to the target object (i.e., avocado-like items) when participants heard the predictive verb "eat-soft things" (solid line), compared to fixations to the same object when hearing the general verb "look" (dotted lines). Visual inspection of the graph shows that fixations towards the appropriate object started to increase during the *Verb + Adverb* region in the predictive verb condition and continued to increase during the *Object* region and all the way until the subject was heard. In contrast, in the general verb condition, participants directed their attention towards the target only once it was mentioned (i.e., during the *Object* region). Then, fixations to the target increased once the object was heard and continued until the end of the sentence. This pattern of fixations suggests that Tseltal speakers anticipated the upcoming direct object before it was encountered in the sentence. In addition, participants fixated the target object in both sentential conditions once the linguistic expression referring to that object was encountered in the utterance, and they continued fixating it until the end of the sentence.
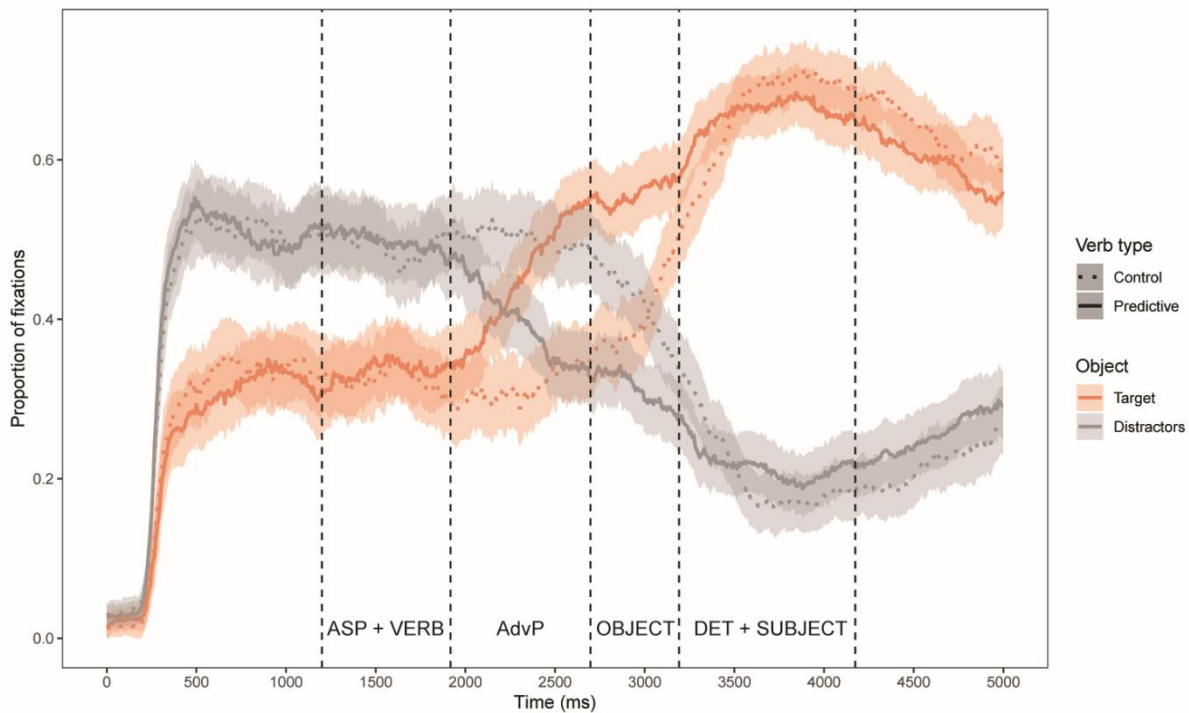

Fig. 3 Time course of fixation proportions for target and averaged distractor objects in the predictive and general conditions. Ribbons indicate confidence intervals (95%), calculated for each sampling step. Dotted lines indicate the mean onset and offset of word durations in the sentential conditions.

Table 2 presents the proportion of fixations to the target as a function of verb type and knowledge of Spanish (during TW 1). It shows that the proportion of fixations to the target was very similar across verb types and between participants with and without knowledge of Spanish. Descriptive statistics of fixations to the target as a function of verb type and education (TW 1) can be found in Table 3. The fixation distributions were very similar across participants' years of formal schooling and verb type, with subtle differences in some groups. The proportion of fixations to the target was higher for the illiterate participants (note that there are only three illiterate participants in the sample), for those with 4 years of schooling and for participants with 9-12 years of education, in the predictive verb condition. Finally, Table 4 summarises the estimated regression coefficients and variance components of the multilevel logistic regression model consisting of the interaction between verb type and time during TW 1 (*Verb +Adverb* region) and TW 2 (*Object* region). The model that best described the data had an interaction of verb type and time and random intercepts and slopes of our predictors for participants and items. We discuss each time window next.

TW 1 *Verb + Adverb* region: There was no significant effect of time ($\beta = 0.06$, *SE*= 0.09, 95% CI [-0.12, 0.23]), but there was a significant effect of verb type ($\beta = 0.54$, *SE* = 0.15, 95% CI [0.25, 0.83]), suggesting that Tseltal speakers are more likely to fixate the target when hearing a predictive verb compared to a general one. In addition, there was a significant interaction between verb and time ($\beta = 0.41$, *SE*= 0.12, 95% CI [0.18, 0.64]). Thus, speakers are more likely to fixate the target when hearing a predictive verb and as more verbal information becomes available over time. The results show that Tseltal participants anticipated the target object before it was mentioned, during this TW.

TW 2 *Object* region: There were significant effects of verb type ($\beta = 0.88$, *SE*= 0.25, 95% CI [0.40, 1.36]) and time ($\beta = 0.53$, *SE*= 0.13, 95% CI [0.28, 0.78]), suggesting that participants were more likely to fixate the target, both when hearing a predictive verb compared to a general one and as time increased. The interaction between verb type and time was not significant ($\beta = -0.26$, *SE*= 0.16, 95% CI [-0.58, 0.06]). As expected, participants directed their visual attention to the object that was being mentioned during this time window.

We conducted complementary analyses to test the effects that other factors, such as level of education, knowledge of Spanish, and part of the experiment might have on object prediction. These models compared fixations to the target as a function of verb type, time, education, knowledge of Spanish, and section of the experiment. The interactions between these predictors and verb type were included as well to address effects specific to predictive (vs. general) verbs. The models were estimated only during TW 1 (*Verb + Adverb* region)

because this is the predictive region of interest, where we could test whether these factors have an influence on verb-mediated anticipatory processing. The model summary can be found in the Supplementary Materials (section B, Table SB1). We found that none of the additional predictors evaluated had an effect on fixating the target. In addition, there were no significant interactions between education and verb type (β = -0.01, SE= 0.04, 95% CI [-0.08, 0.06]), knowledge of Spanish and verb type (β = 0.04, SE= 0.25, 95% CI [-0.44, 0.53]) and section of the experiment and verb type (β = 0.19, SE= 0.25, 95% CI [-0.30, 0.68]). Nevertheless, the interaction between verb type and time remained significant (β = 0.40, SE= 0.01, 95% CI [0.38, 0.42]). Overall, the results suggest that education, knowledge of Spanish and section of the experiment do not play a role in verb-mediated anticipatory eye fixations in Tseltal.

Table 2
Proportions and standard deviations of fixations to the target as a function of both verb type and knowledge of Spanish. (N = 251,717 obs.)

| | Knowledge of Spanish | | | |
| | Yes | | No | |
Verb type	Mean	*SD*	Mean	*SD*
Predictive	.29	.45	.30	.45
General	.20	.40	.21	.41

Note. Fixations to the target (vs. distractors and white space) during TW 1.

Table 3
Proportions and standard deviations of fixations to the target as a function of verb type and education (N = 251,717 obs.)

| | Verb type | | | |
| | Predictive | | General | |
Education (years of formal schooling)	Mean	*SD*	Mean	*SD*
0 (illiterates- no formal schooling)	.32	.47	.21	.41
3 years	.33	.47	.26	.44
4 years	.39	.49	.20	.40
5 years	.26	.44	.20	.40
6 years	.27	.45	.20	.40
7 years	.40	.49	.39	.49
9 years	.29	.46	.20	.40
11 years	.32	.47	.20	.40
12 years	.30	.46	.21	.40

Note. Fixations to the target (vs. distractors and white space) during TW 1.

Table 4

Regression coefficients and variance components for the multilevel logistic regression models. Fixations to the target were modelled as a function of verb type and time during two time windows: TW 1 Verb + Adverb region and TW 2 Object region.

Variable	TW1 Verb + Adverb			TW2 Object		
	Regression coefficient (95% CI)	SE	z	Regression coefficient (95% CI)	SE	z
Fixed effects						
Intercept	-1.66 (-2.03, -1.29)	0.19	-8.72***	-1.23 (-1.75, -0.71)	0.27	-4.64***
Verb (predictive)	0.54 (0.25, 0.83)	0.15	3.65***	0.88 (0.40, 1.36)	0.25	3.59***
Time	0.06 (-0.12, 0.23)	0.09	0.64	0.53 (0.28, 0.78)	0.13	4.17***
Verb (predictive): Time	0.41 (0.18, 0.64)	0.12	3.49***	-0.26 (-0.58, 0.06)	0.16	-1.61
Random effects	**Variance**	**SD**		**Variance**	**SD**	
Participant						
Intercept	0.55	0.74		1.70	1.30	
Verb (predictive)	0.35	0.59		1.15	1.07	
Time	0.16	0.40		0.23	0.48	
Verb (predictive): Time	0.25	0.50		0.53	0.73	
Item						
Intercept	0.81	0.90		1.14	1.07	
Verb (predictive)	0.47	0.69		1.16	1.08	
Time	0.16	0.40		0.34	0.58	
Verb (predictive): Time	0.27	0.52		0.45	0.67	

Note *** $p < 0.001$.

4. General discussion

While predictive processing is often assumed to be a central, if not crucial aspect of language comprehension, very little evidence from typologically diverse languages has been marshalled to support this view. Our study makes a small contribution towards filling this empirical gap by investigating real-time sentence comprehension in Tseltal, a verb-initial Mayan language. Using the visual world paradigm, our goal was to test whether Tseltal speakers use verbal information, which is provided upfront in the sentence, to anticipate the upcoming grammatical object. Our study followed Altmann and Kamide's (1999) study design, adapted to a smaller community of non-Western speakers to assess whether there are similar effects to those found in the literature for subject-initial languages.

In our experiment, Tseltal speakers listened to verb-initial transitive sentences while seeing a visual display showing one potential referent and three distractors. We manipulated verb type (predictive *vs*. general) and recorded participants' eye-movements while they listened and inspected the visual scene. We estimated the contribution of verb type to anticipatory eye-movements towards the target object referent in two different time windows. In the first time window, which covered the initial verb (and its aspect marker) together with the adverbial phrase, we found a significant effect of verb type, indicating that participants fixated the target more during this window in the predictive condition compared to the general condition. There was also a significant interaction between verb type and time, showing that fixations to the target increased over time when hearing a predictive verb. Fixations to the target remained until the object was heard (the second time-window) in the predictive condition. In contrast, in the general verb condition, participants directed their attention to the relevant object in the visual display only when the word for it was encountered in the sentence. By the time the object was heard in both sentential conditions, participants were fixating the only object depicted visually that matched that referring expression.

This pattern of language-mediated eye-movements shows two things: 1) there are anticipatory looks to the most plausible referent depicted that will follow a verbal expression; 2) Tseltal participants use verbal information to direct their visual attention towards the external world (i.e., the visual display). These results replicate what has been previously found in the literature for subject-initial languages (Altmann & Kamide, 1999; Arai & Keller, 2013; Boland, 2005; Hintz, Meyer, & Huettig, 2017; Kako & Trueswell, 2000; Kamide, Altmann, & Haywood, 2003; Kamide, Scheepers, & Altmann, 2003; Knoeferle, Crocker, Scheepers, & Pickering, 2005; Mani & Huettig, 2012; etc.) that verbal information is extracted very quickly and guides anticipatory looks to whichever object in the visual display satisfies the selectional restrictions of the verb.

In line with previous studies, the results suggest that conceptual overlap between objects in the visual display and the unfolding language is one way that mediates the direction of visual attention (Altmann & Kamide, 2007; Altmann & Mirković, 2009; Kamide, Altmann, et al., 2003; Tanenhaus, Magnuson, & Chambers, 2000). The items depicted in the visual display activate their own conceptual features (see Huettig and McQueen, 2007; McQueen and Huettig, 2014) possibly creating an "episodic trace" (Altmann & Kamide, 2007, p.512). When linguistic input is presented, the verb's own specific semantic features are activated. Because the semantic features of the predictive verbs match only one of the objects depicted visually, the matching activations between the object and its referring linguistic expression cause a shift of attention towards the target object (Altmann and Kamide, 2007; Altmann and Mirković, 2009).

What is interesting about the Tseltal results is that semantic information encoded in the verb can rapidly guide eye-movements, despite the fact that Tseltal verbs come first in the sentence. It is possible that the absence of an initial subject noun phrase (or any other verbal argument) may in fact facilitate anticipatory processing: because the subject noun phrase has not yet been mentioned, listeners do not have to devote cognitive resources to integrating the subject argument into the ongoing parse, which might free up resources for anticipating upcoming arguments. Future work could capitalise on Tseltal's word order flexibility (the grammar permits fronted sentence-initial subjects, see Polian, 2013) to investigate whether anticipatory looks to the object are modulated by the position of the grammatical subject of the sentence.

Predictive processing in Tseltal may also be facilitiated by the particular semantic properties of transitive verbs in the language. A property of Mayan languages is that many verb roots (i.e., transitive and positional verb roots) incorporate into their semantics physical properties (e.g., shape, substance, or position) of the object that they select for (cf. Brown, 2008). In Altmann and Kamide's study, when English speakers hear *The boy will eat*, the verb *eat* restricts interpretation possibilities to edible items a boy could eat. In Tseltal, by contrast, when hearing *Ya slo'* ('He/ She/ It is eating-soft things'), the possible referents are restricted even more to only those edible items that are soft. The referents of Tseltal verbs are more concrete and the range of contexts where they might occur is more easily determined.

An important question for future research concerns the nature of the representations that are activated upon hearing such semantically rich verbs: Are the perceptual properties of upcoming objects (e.g. their shape, their texture) activated in such cases? (see Huettig & Altmann, 2005, 2007; Huettig & McQueen, 2007). Rommers, Meyer and Huettig (2013) showed that Dutch listeners can predict perceptual attributes that will be referred to in an utterance, even in the absence of a visual depiction of the target word. In their study, participants eye-movements were recorded while they listened to sentences that were predictive of a specific word, for example, "moon" in "In 1969 Neil Armstrong was the first man to set foot on the moon". Participants inspected

a visual display with three unrelated distractors and either the target object (e.g., moon), a shape competitor of the target object (e.g., tomato) or an unrelated control object (e.g., rice). Participants looked significantly more at the target object as expected, but interestingly, they also directed their attention towards the shape competitor of the target word more than to unrelated objects before "moon" was heard. Rommers et al's data provide evidence that the pre-activation of a predicted concept can also activate the visual representation of that object. In Rommers et al's study, listeners' predictions of perceptual attributes of noun referents were generated from the entire multi-word sentential context before the target noun. In Tseltal, by contrast, verbal information alone may be sufficient to activate visual representations of upcoming objects, given the relative 'nouniness' of transitive verbs in the language.

The only other visual world eye tracking studies to investigate anticipatory processing in a verb-initial language, to the best of our knowledge, are Sauppe's (2016) and Garcia et al.'s (2021) studies of Tagalog (Austronesian). In Sauppe's study, participants heard sentences of the type 'Eat frog fly' or 'Eat fly frog' (both with the meaning 'The frog will eat the fly') while inspecting a visual display that depicted an agent, a patient and a distractor. The verb carried morphological marking that allowed listeners to infer the order and the syntactic status of the agent and patient. After hearing the verb, Tagalog listeners directed their gaze more to the agent, regardless of its syntactic function and position in the sentence. In a different study, Garcia and colleagues found that Tagalog adult comprehenders (the study was also conducted with children) did not always anticipate the agent when hearing verb-initial sentences. In their study, participants heard sentences such as "Bite last Tuesday diligent cow monkey" or "Bite last Tuesday diligent monkey cow" (both meaning 'A (diligent) cow was biting a monkey last Tuesday') while viewing a picture depicting a transitive event between two animals (e.g., cow biting a monkey). Voice marking (agent *vs*. patient) and noun phrase argument order (VAP: agent-initial sentences *vs*. VPA: patient-initial sentences) were manipulated. Participants anticipated the agent in the patient-voice condition (compared to agent-voice) in agent-initial sentences (compared to patient-initial sentences), thus showing that Tagalog speakers use morphosyntactic markers in the verb to anticipate an upcoming argument. These results might reflect a bias towards agent identification in the early stages of sentence processing (cf. Bornkessel and Schlesewsky 2006). In the present study, we did not include depictions of agents in the visual display so we cannot determine whether such a bias is also operative in Tseltal. We can conclude, however, that at least *in the absence* of a competing visual representation of an agent, verbal information rapidly guides listeners' eye movements to a depicted object. In other words, an agent bias does not fully attenuate the rapid anticipation of a grammatical object, at least when a visual agent competitor is not present. We have suggested that the object-oriented nature of Tseltal's transitive verbs may encourage early attention to object referents; this may be a point

of contrast with Tagalog. Future work should directly test this by examining whether Tseltal listeners preferentially anticipate agents or patients when both are visually depicted.

It has been suggested that literacy, vocabulary knowledge and second language learning might mediate prediction during language processing. In our study, literacy levels were not manipulated systematically (i.e., we did not include measures from standardized tests on reading abilities, vocabulary size and level of bilingualism in our experimental design). We instead made the assumption that years of formal (rural) schooling and knowledge of Spanish are somewhat related to literacy skills (and vocabulary size). Bearing in mind that we could only examine these factors in an exploratory way, it is nevertheless notable that we did not find any effects of level of Education or knowledge of Spanish as factors mediating anticipatory processing and incremental interpretation of objects in Tseltal. These results are different from those reported in the literature, where participants with low literacy levels showed either reduced or no anticipation of upcoming language input compared to high literates (see Huettig & Brouwer, 2015; Huettig, Singh, et al., 2011; Mishra et al., 2012). There are two possible (mutually compatible) explanations for these differences. First, the Tseltal speakers had more years of formal schooling (on average 7-8 years) than the low literate Hindi speakers in Mishra et al.'s (2012) study (2 mean years of formal education). This might suggest that some exposure to formal education already has an effect in facilitating prediction in spoken language (Araújo, Fernándes, & Huettig, 2019; Huettig & Pickering, 2019; Favier, Meyer, & Huettig, 2021). An interesting question in this regard is the extent to which it is education (and therefore literacy) *per se* that is at issue here, as opposed to education and literacy specifically obtained in the target language. This is relevant because it is only during the first four years of formal schooling that Tseltal children are taught to read and write in Tseltal. After that, formal schooling is conducted in Spanish. Further work is required to determine whether the capacity for predictive processing in Tseltal can be tied specifically to those four years of schooling in Tseltal and/or whether Spanish literacy levels and vocabulary size are also factors mediating anticipatory processing in Tseltal.

Second, it is possible that Tseltal verbs may have provided a richer set of predictive cues than those used in the Hindi sentences. In this study we pointed to a number of linguistic features of Tseltal verbs that sets them apart from those of previously studied languages (their rich semantics and morphology) and which may serve to facilitate predictive processing. An important area of future work will be to untangle these various potentially contributing factors to predictive processing in Tseltal.

Finally, we did not find significant effects of the role of experiment associated with anticipatory eye-movements. That is, Tseltal participants did not acquire any predictive strategy during the experiment that might have driven anticipatory eye-movements towards the target object. Taken together, the data suggest that anticipatory processing of upcoming linguistic information in Tseltal relies only on the verb's meaning,

independent of education, knowledge of Spanish or experiment effects. The visual display together with the semantics of the verb contributed to the speed with which Tseltal listeners activated upcoming constituents.

5. Conclusions

This study shows that it is possible to create a laboratory setting in the field to investigate an aspect of sentence processing (incrementality and prediction during sentence comprehension) that has received little typological coverage. Our results provide evidence that in Tseltal, a verb-initial Mayan language, predictive processing is an integral part of real-time language processing. This lends support the view that predictive processing during language comprehension might be a universal processing principle.

References

Aglioti, S. M., Cesari, P., Romani, M., & Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nature Neuroscience, 11*(9), 1109-1116. doi:10.1038/nn.2182

Agresti, A. (2019). *An introduction to categorical data analysis* (3rd ed.). Hoboken, NJ: John Wiley & Sons.

Altmann, G., & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition, 73*(3), 247-264. doi:10.1016/s0010-0277(99)00059-1

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition, 30*(3), 191-238. doi:10.1016/0010-0277(88)90020-0

Altmann, G. T. M., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language, 57*(4), 502-518. doi:10.1016/j.jm1.2006.12.004

Altmann, G. T. M., & Mirković, J. (2009). Incrementality and Prediction in Human Sentence Processing. *Cognitive Science, 33*(4), 583-609. doi:10.1111/j.1551-6709.2009.01022.x

Arai, M., & Keller, F. (2013). The use of verb-specific information for prediction in sentence processing. *Language and Cognitive Processes, 28*(4), 525-560. doi:10.1080/01690965.2012.658072

Araújo, S., Fernándes, T., & Huettig, F. (2019). Learning to read facilitates the retrieval of phonological representations in rapid automatized naming: Evidence from unschooled illiterate, ex-illiterate, and schooled literate adults. *Developmental Science, 22*(4). doi:10.1111/desc.12783

Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions (vol 11, pg 280, 2007). *Trends in Cognitive Sciences, 11*(9), 372-372. doi:10.1016/j.tics.2007.08.004

Bar, M. (2009). The proactive brain: memory for predictions. *Philosophical Transactions of the Royal Society B-Biological Sciences, 364*(1521), 1235-1243. doi:10.1098/rstb.2008.0310

Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language, 59*(4), 457-474. doi:10.1016/j.jml.2007.09.002

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology, 4*. doi:10.3389/fpsyg.2013.00328

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. doi:10.1016/j.jml.2012.11.001

Bates, D., Machler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Bickel, B., Witzlack-Makarevich, A., Choudhary, K. K., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2015). The Neurophysiology of Language Processing Shapes the Evolution of Grammar: Evidence from Case Marking. *Plos One, 10*(8). doi:10.1371/journal.pone.0132819

Boersma, P. (2002). Praat, a system for doing phonetics by computer. *Glot International, 5*(9/10), 341-345.

Boland, J. E. (2005). Visual arguments. *Cognition, 95*(3), 237-274. doi:10.1016/j.cognition.2004.01.008

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2013a). Neurotypology: modelling crosslinguistic similarities and differences in the neurocognition of language comprehension. In M. Sanz, I. Laka, & M. K. Tanenhaus (Eds.), *Language down the garden path: The cognitive and biological basis for linguistic structures* (pp. 241-252). Oxford: Oxford University Press.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2013b). Reconciling time, space and function: A new dorsal-ventral stream model of sentence comprehension. *Brain and Language, 125*(1), 60-76. doi:10.1016/j.bandl.2013.01.010

Bornkessel, I., & Schlesewsky, M. (2006). The extended argument dependency model: A neurocognitive approach to sentence comprehension across languages. *Psychological Review, 113*(4), 787-821. doi:10.1037/0033-295x.113.4.787

Borovsky, A., Elmana, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology, 112*(4), 417-436. doi:10.1016/j.jecp.2012.01.005

Brouwer, S., Mitterer, H., & Huettig, F. (2013). Discourse context and the recognition of reduced and canonical spoken words. *Applied Psycholinguistics, 34*(3), 519-539. doi:10.1017/s0142716411000853

Brown, P. (2008). Verb specificitry and argument realization in Tzeltal child language. In M. Bowerman & P. Brown (Eds.), *Cross-linguistic perspectives on argument structure: Implications for learnability* (pp. 167-189). New York: Lawrence Erlbaum.

Chang, F., Kidd, E., & Rowland, C. F. (2013). Prediction in processing is a by-product of language learning. *Behavioral and Brain Sciences, 36*, 350-351.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(3), 181-204. doi:10.1017/s0140525x12000477

Cohn, N., & Paczynski, M. (2013). Prediction, events, and the advantage of Agents: The processing of semantic roles in visual narrative. *Cognitive Psychology, 67*(3), 73-97. doi:10.1016/j.cogpsych.2013.07.002

Contemori, C., & Dussias, P. E. (2020). The processing of subject pronouns in highly proficient L2 speakers of English. *Glossa-a Journal of General Linguistics, 5*(1). doi:10.5334/gjgl.972

Cooper, R. (1974). The control of eye fixation by the meaning of spoken language. *Cognitive Psychology, 6*, 84-107.

Cutler, A., & Clifton, C. (1999). Comprehending spoken language: A blueprint of the listener. In C. M. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 123-165). Great Britan: Oxford University Press.

DeLong, K. A., Groppe, D. M., Urbach, T. P., & Kutas, M. (2012). Thinking ahead or not? Natural aging and anticipation during reading. *Brain and Language, 121*(3), 226-239. doi:10.1016/j.bandl.2012.02.006

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*(8), 1117-1121.

Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2017). Predicting upcoming information in native-language and non-native-language auditory word recognition. *Bilingualism-Language and Cognition, 20*(5), 917-930. doi:10.1017/s1366728916000547

Dryer, M. S. (2013). Order of Subject, Object and Verb. In M. S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online* (Vol. (Available online at http://wals.info, Accessed on 2021-01-21)). Leipzig: Max Planck Institute for Evolutionary Anthropology.

Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior, 20*, 641-655.

Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences, 32*(5), 429-+. doi:10.1017/s0140525x0999094x

Favier, S., Meyer, A. S., & Huettig, F. (2021). Literacy can enhance syntactic prediction in spoken language processing. *Journal of Experimental Psychology: General*.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology, 44*(4), 491-505. doi:10.1111/j.1469-8986.2007.00531.x

Federmeier, K. D., & Kutas, M. (1999). Right words and left words: electrophysiological evidence for hemispheric differences in meaning processing. *Cognitive Brain Research, 8*(3), 373-392. doi:10.1016/s0926-6410(99)00036-1

Federmeier, K. D., & Kutas, M. (2005). Aging in context: Age-related changes in context use during language comprehension. *Psychophysiology, 42*(2), 133-141. doi:10.1111/j.1469-8986.2005.00274.x

Federmeier, K. D., Kutas, M., & Schul, R. (2010). Age-related and individual differences in the use of prediction during language comprehension. *Brain and Language, 115*(3), 149-161. doi:10.1016/j.bandl.2010.07.006

Federmeier, K. D., McLennan, D. B., De Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. (vol 39, pg 133, 2002). *Psychophysiology, 39*(5), 684-687.

Friston, K. J. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience, 11*(2), 127-138. doi:10.1038/nrn2787

Garcia, R., Garrido Rodriguez, G., & Kidd, E. (2021). Developmental effects in the online use of morphosyntactic cues in sentence processing: Evidence from Tagalog. *Cognition, 216*, 104859.

Hahne, A., & Friederici, A. D. (2001). Processing a second language: Late learners' comprehension mechanisms as revealed by event-related brain potential. *Bilingualism: Language and Cognition, 4*(2), 123-141.

Hawkins, J. A. (2007). Processing typology and why psychologists neet to know about it. *New Ideas in Psychology, 25*, 87-107.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2-3), 1-75. doi:10.1017/s0140525x0999152x

Hintz, F., Meyer, A. S., & Huettig, F. (2014). *The influence of verb-specific featural restrictions, word associations and production-based mechanisms on language-mediated anticipatory eye-movements*. Paper presented at the 27th CUNY Conference on Human Sentence Processing, Ohio State University, Columbus/Ohio (US).

Hintz, F., Meyer, A. S., & Huettig, F. (2017). Predictors of Verb-Mediated Anticipatory Eye Movements in the Visual World. *Journal of Experimental Psychology-Learning Memory and Cognition, 43*(9), 1352-1374. doi:10.1037/xlm0000388

Huang, H. W., Meyer, A. M., & Federmeier, K. D. (2012). A "concrete view" of aging: Event related potentials reveal age-related changes in basic integrative processes in language. *Neuropsychologia, 50*(1), 26-35. doi:10.1016/j.neuropsychologia.2011.10.018

Huettig, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. *Cognition, 96*(1), B23-B32. doi:10.1016/j.cognition.2004.10.003

Huettig, F., & Altmann, G. T. M. (2007). Visual-shape competition during language-mediated attention is based on lexical input and not modulated by contextual appropriateness. *Visual Cognition, 15*(8), 985-1018. doi:10.1080/13506280601130875

Huettig, F., Audring, J., & Jackendoff, R. (in press). A parallel architecture perspective on pre-activation and prediction in language processing. *Cognition*.

Huettig, F., & Brouwer, S. (2015). Delayed Anticipatory Spoken Language Processing in Adults with Dyslexia-Evidence from Eye-tracking. *Dyslexia, 21*(2), 97-122. doi:10.1002/dys.1497

Huettig, F., & Guerra, E. (2019). Effects of speech rate, preview time of visual context, and participant instructions reveal strong limits on prediction in language processing. *Brain Research, 1706*, 196-208. doi:10.1016/j.brainres.2018.11.013

Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language Cognition and Neuroscience, 31*(1), 80-93. doi:10.1080/23273798.2015.1047459

Huettig, F., & Mani, N. (2016). Is prediction necessary to understand language? Probably not. *Language Cognition and Neuroscience, 31*(1), 19-31. doi:10.1080/23273798.2015.1072223

Huettig, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language, 57*(4), 460-482. doi:10.1016/j.jml.2007.02.001

Huettig, F., Olivers, C. N. L., & Hartsuiker, R. J. (2011). Looking, language, and memory: Bridging research from the visual world and visual search paradigms. *Acta Psychologica, 137*(2), 138-150. doi:10.1016/j.actpsy.2010.07.013

Huettig, F., & Pickering, M. J. (2019). Literacy Advantages Beyond Reading: Prediction of Spoken Language. *Trends in Cognitive Sciences, 23*(6), 464-475. doi:10.1016/j.tics.2019.03.008

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica, 137*(2), 151-171. doi:10.1016/j.actpsy.2010.11.003

Huettig, F., Singh, N., & Mishra, R. K. (2011). Language-mediated visual orienting behavior in low and high literates. *Frontiers in Psychology, 2*. doi:10.3389/fpsyg.2011.00285

INEGI. (2010). Censo de Poblacion y Vivienda 2010.   Retrieved from www.inegi.gob.mx

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*(4), 434-446. doi:10.1016/j.jml.2007.11.007

Jaeger, T. F., & Norcliffe, E. (2009). The cross-linguistic study of sentence production. *Language and Linguistic Compass, 3*.

James, A., & Watson, D. (2013). *Language experience and prediction in auditory sentence comprehension.* Paper presented at the Proceedings of the 19th Architectures and Mechanisms for Language Processing Conference, Marseille, France.

Janse, E., & Jesse, A. (2014). Working memory affects older adults' use of context in spoken-word recognition. *Quarterly Journal of Experimental Psychology, 67*(9), 1842-1862. doi:10.1080/17470218.2013.879391

Kako, E., & Trueswell, J. C. (2000, August 13-15). *Verb meanings, object affordances, and the incremental restriction of reference.* Paper presented at the Proceedings of the Twenty-Secod Annual Conference of the Cognitive Science Society, Philadelphia, PA.

Kamide, Y. (2008). Anticipatory processes in sentence processing. *Language and Linguistics Compass, 2*(4), 647-670.

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language, 49*(1), 133-156. doi:10.1016/s0749-596x(03)00023-8

Kamide, Y., Scheepers, C., & Altmann, G. T. M. (2003). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research, 32*(1), 37-55. doi:10.1023/a:1021933015362

Kemmerer, D. (2012). The cross-linguistic prevalence of SOV and SVO word orders reflects the sequential and hierarchical representation of action in Broca's area. *Language and Linguistics Compass, 6*, 50-66.

Kendeou, P., White, M. J., van den Broek, P., & Lynch, J. S. (2009). Predicting Reading Comprehension in Early Elementary School: The Independent Contributions of Oral Language and Decoding Skills. *Journal of Educational Psychology, 101*(4), 765-778. doi:10.1037/a0015956

Kidd, E. (2012). Implicit Statistical Learning Is Directly Associated With the Acquisition of Syntax. *Developmental Psychology, 48*(1), 171-184. doi:10.1037/a0025405

Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: evidence from eye-movements in depicted events. *Cognition, 95*(1), 95-127. doi:10.1016/j.cognition.2004.03.002

Kroll, J. F., Dussias, P. E., Bogulski, C. A., & Kroff, J. R. V. (2012). Juggling two languages in one mind: What bilinguals tell us about language processing and its consequences for cognition. In B. H. Ross (Ed.), *Psychology of Learning and Motivation, Vol 56* (Vol. 56, pp. 229-262).

Kukona, A., Braze, D., Johns, C. L., Mencl, W. E., Van Dyke, J. A., Magnuson, J. S., . . . Tabor, W. (2016). The real-time prediction and inhibition of linguistic outcomes: Effects of language and literacy skill. *Acta Psychologica, 171*, 72-84.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition, 106*(3), 1126-1177. doi:10.1016/j.cognition.2007.05.006

Magyari, L., De Ruiter, J. P., & Levinson, S. C. (2017). Temporal Preparation for Speaking in Question-Answer Sequences. *Frontiers in Psychology, 8*. doi:10.3389/fpsyg.2017.00211

Mani, N., & Huettig, F. (2012). Prediction During Language Processing is a Piece of Cake - But Only for Skilled Producers. *Journal of Experimental Psychology-Human Perception and Performance, 38*(4), 843-847. doi:10.1037/a0029284

Mani, N., & Huettig, F. (2014). Word reading skill predicts anticipation of upcoming spoken language input: A study of children developing proficiency in reading. *Journal of Experimental Child Psychology, 126*, 264-279. doi:10.1016/j.jecp.2014.05.004

Marslen-Wilson, W. D. (1975). Sentence perception as an interactive parallel process. *Science, 189*(4198), 226-228. doi:10.1126/science.189.4198.226

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken-word recognition. *Cognition, 25*(1-2), 71-102. doi:10.1016/0010-0277(87)90005-9

Martin, C. D., Thierry, G., Kuipers, J. R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language, 69*(4), 574-588. doi:10.1016/j.jml.2013.08.001

McDonald, S. A., & Shillcock, R. C. (2003). Eye movements reveal the on-line computation of lexical probabilities during reading. *Psychological Science, 14*(6), 648-652. doi:10.1046/j.0956-7976.2003.psci_1480.x

McQueen, J. M., & Huettig, F. (2014). Interference of spoken word recognition through phonological priming from visual objects and printed words (vol 76, pg 190, 2014). *Attention Perception & Psychophysics, 76*(5), 1503-1503. doi:10.3758/s13414-014-0698-z

Mishra, R. K., Singh, N., Pandey, A., & Huettig, F. (2012). Spoken language mediated anticipatory eye movements are modulated by reading ability: Evidence from Indian low and high literates. *Journal of Eye Movement Research, 5*(1), 1-10.

Mori, S., Ohtani, Y., & Imanaka, K. (2002). Reaction times and anticipatory skills of karate athletes. *Human Movement Science, 21*(2), 213-230. doi:10.1016/s0167-9457(02)00103-3

Nation, K., Marshall, C. M., & Altmann, G. T. M. (2003). Investigating individual differences in children's real time sentence comprehension using language-mediated eye movements. *Journal of Experimental Child Psychology, 86*, 314-329.

Ng, S., Payne, B. R., Stine-Morrow, E. A. L., & Federmeier, K. D. (2018). How struggling adult readers use contextual information when comprehending speech: Evidence from event-related potentials. *International Journal of Psychophysiology, 125*, 1-9. doi:10.1016/j.ijpsycho.2018.01.013

Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., . . . Huettig, F. (2018). Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. *Elife, 7*. doi:10.7554/eLife.33468

Norcliffe, E., Harris, A. C., & Jaeger, T. F. (2015). Cross-linguistic psycholinguistics and its critical role in theory development: early beginnings and recent advances. *Language Cognition and Neuroscience, 30*(9), 1009-1032. doi:10.1080/23273798.2015.1080373

Norcliffe, E., Konopka, A. E., Brown, P., & Levinson, S. C. (2015). Word order affects the time course of sentence formulation in Tzeltal. *Language Cognition and Neuroscience, 30*(9), 1187-1208. doi:10.1080/23273798.2015.1006238

Nordlinger, R., Garrido Rodriguez, G., & Kidd, E. (in press). Sentence planning and production in Murrinhpatha, an Australian 'free word order' language. *Cognition*.

Nordlinger, R., Garrido Rodriguez, G., Wilmoth, S., & Kidd, E. (2020). *What drives word order flexibility? Evidence from sentence production experiments in two Australian Indigenous languages*. Paper presented at the 53rd Annual Meeting of the Societas Linguistica Europaea.

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences, 11*(3), 105-110.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36*(4), 329-347. doi:10.1017/s0140525x12001495

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed effects models in S and S-PLUS*. New York: Springer.

Polian, G. (2013). *Gramatica del tseltal de Oxchuc*. Chiapas, Mexico: CIESAS.

SEP (2017). *Aprendizajes clave para la educacion integral. Educacion primaria indigena*. Ciudad de Mexico: Secretaria de Educacion Publica-DGEI.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62*(8), 1457-1506. doi:10.1080/17470210902816461

Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. *Psychology and Aging, 21*(3), 448-465. doi:10.1037/0882-7974.21.3.448

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review, 3*(4), 504-509. doi:10.3758/bf03214555

Resnik, P. (1996). Selectional constraints: An information-theoretic model and its computational realization. *Cognition, 61*(1-2), 127-159. doi:10.1016/s0010-0277(96)00722-6

Robinson, S. (2002). Constituent order in Tenejapa Tzeltal. *International Journal of American Linguistics, 68*(1), 51-80.

Rommers, J., Meyer, A. S., & Huettig, F. (2013). Object Shape and Orientation Do Not Routinely Influence Performance During Language Processing. *Psychological Science, 24*(11), 2218-2225. doi:10.1177/0956797613490746

Rowland, C. F., Chang, F., Ambridge, B., Pine, J. M., & Lieven, E. V. M. (2012). The development of abstract syntax: Evidence from structural priming and the lexical boost. *Cognition, 125*(1), 49-63. doi:10.1016/j.cognition.2012.06.008

Sacks, H., Schegloff, E., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50*(4), 696-735.

Sauppe, S. (2016). Verbal Semantics Drives Early Anticipatory Eye Movements during the Comprehension of Verb-Initial Sentences. *Frontiers in Psychology, 7*. doi:10.3389/fpsyg.2016.00095

Snijders, T., & Bosker, R. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). London: Sage Publishers.

Starkes, J. L., Edwards, P., Dissanayake, P., & Dunn, T. (1995). A new technology and field-test of advance cue usage in volleyball. *Research Quarterly for Exercise and Sport, 66*(2), 162-167. doi:10.1080/02701367.1995.10762223

Staub, A., & Clifton, C. (2006). Syntactic prediction in language comprehension: Evidence from either ... or. *Journal of Experimental Psychology-Learning Memory and Cognition, 32*(2), 425-436. doi:10.1037/0278-7393.32.2.425

Tanenhaus, M. K., Magnuson, J. S., & Chambers, C. (2000). Eye movements and lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research, 29*(6), 557-580.

Tanenhaus, M. K., Spiveyknowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*(5217), 1632-1634. doi:10.1126/science.7777863

Tollan, R., Massam, D., & Heller, D. (2019). Effects of case and transitivity on processing dependencies: Evidence from Niuean. *Cognitive Science, e12736*, 1-14. doi:10.1111/cogs.12736

Traxler, M. J., Bybee, M. D., & Pickering, M. J. (1997). Influence of connectives on language comprehension: Eye-tracking evidence for incremental interpretation. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 50*(3), 481-497. doi:10.1080/027249897391982

Traxler, M. J., & Pickering, M. J. (1996). Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language, 35*(3), 454-475. doi:10.1006/jmla.1996.0025

Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology-Learning Memory and Cognition, 31*(3), 443-467. doi:10.1037/0278-7393.31.3.443

Wang, L. M., Schlesewsky, M., Bickel, B., & Bornkessel-Schlesewsky, I. (2009). Exploring the nature of the 'subject'-preference: Evidence from the online comprehension of simple sentences in Mandarin Chinese. *Language and Cognitive Processes, 24*(7-8), 1180-1226. doi:10.1080/01690960802159937

Whalen, D. H., & McDonough, J. (2015). Taking the laboratory into the field. *Annual Review of Linguistics, 1*, 395-415.

Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience, 16*(7), 1272-1288. doi:10.1162/0898929041920487

Williams, A. M., Ward, P., Knowles, J. M., & Smeeton, N. J. (2002). Anticipation skill in a real-world task: Measurement, training, and transfer in tennis. *Journal of Experimental Psychology-Applied, 8*(4), 259-270. doi:10.1037/1076-898x.8.4.259

Wlotko, E. W., & Federmeier, K. D. (2012). Age-related changes in the impact of contextual strength on multiple aspects of sentence comprehension. *Psychophysiology, 49*(6), 770-785. doi:10.1111/j.1469-8986.2012.01366.x

Zirnstein, M., van Hell, J. G., & Kroll, J. F. (2018). Cognitive control ability mediates prediction costs in monolinguals and bilinguals. *Cognition, 176*, 87-106. doi:10.1016/j.cognition.2018.03.001

Anticipatory processing in a verb-initial Mayan language: Eye-tracking evidence during sentence comprehension in Tseltal: Supplementary Materials

Supplementary materials section A

List of 32 sentence pairs used in the experiment. For presentation purposes, each sentence (VOS word order) contains a predictive verb followed by a general verb and then the rest of the sentence, which is the same for both verbs. Verb forms and target objects are in bold. The items in parenthesis refer to the target and distractor objects included in the visual displays. The first sentence corresponds to the original Tseltal auditory stimuli. Abbreviations: A3 (ergative 3rd person marker), B3 (absolutive 3rd person marker), DET (determiner), INC (incompletive aspect), P (preposition), PRON (pronoun).

1. Ya   s-**lo'**   / s-**le**   ta   s-tukel   **on**   te   kerem=e
   INC   A3-eat.soft[B3]   / A3-look[B3]   P   A3-PRON   avocado   DET   boy=DET
   The boy **is eating** (soft)/ **is looking** for an **avocado** by himself.
   (avocado, coffee mill, traditional market bag, toy car)

2. Ya   s-**ti'**   / y-**il**   ta   s-tukel   **ti'bal**   te   ants=e
   INC   A3-eat.meat[B3]   / A3-see[B3]   P   A3-PRON   meat   DET   woman=DET
   The woman **is eating** (meat)/ **is seeing meat** by herself.
   (meat, necklace, coins, flowers)

3. Ya   s-**set'**   / s-**tzaj**   ta   s-tukel   **laso**   te   ants=e
   INC   A3-cut (scissors)[B3]   / A3-choose[B3]   P   A3-PRON   rope   DET   woman=DET
   The woman **is cutting**/ **is choosing** the **rope** by herself.
   (rope, earrings, cookies, clay pot)

4. Ya   x-**chojtan**   / y-**elk'an**   ta   s-tukel   **xila**   te   winik=e
   INC   A3-set.down.four.legs[B3]   / A3-steal[B3]   P   A3-PRON   chair   DET   man=DET
   The man **is setting** (down on four legs)/ **is stealing** a chair by himself.
   (chair, wood, traditional hat, whistle)

5. Ya   s-**top'**   / s-**muk**   ta   s-tukel   **baso**   te   ach'ix=e
   INC   A3-break. pottery[B3]   / A3-covering[B3]   P   A3-PRON   glass   DET   girl=DET
   The girl **is breaking**/ **is covering** a **glass** by herself.
   (glass, mushrooms, paper mask, piece of leather)

6. Ya   s-**lik**   / x-**chon**   ta   s-tukel   **balti**   te   winik=e
   INC   A3-carry.by.handle[B3]   / A3-sell[B3]   P   A3-PRON   bucket   DET   man=DET
   The man **is carrying**/ **is selling** a bucket by himself.
   (bucket, comb, traditional blouse, metal tube)

7. Ya   s-**k'as**   / s-**kus**   ta   s-tukel   m**exa**   te   kerem=e
   INC   A3-break.non-flexible[B3]   / A3-wipe[B3]   P   A3-PRON   table   DET   boy=DET
   The boy **is breaking** (non-flexible object)/ **is wiping** a **table** by himself.
   (table, cable, watch, ball)

8. Ya   x-**ch'il**   / s-**tzak**   ta   s-tukel   **tux ak**   te   ants=e
   INC   A3-fry.without.grease[B3]   / A3-grab[B3]   P   A3-PRON   onion   DET   woman=DET
   The woman **is frying**/ **is grabbing** an **onion** by herself.
   (onion, toilet paper, mirror, bottle gourd)

9. Ya   s-**pak**   / s-**man**   ta s-tukel   **pak'**   te   ants=e
   INC A3-fold[B3]   / A3-buy[B3]   P   A3-PRON   tablecloth   DET   woman=DET
   The woman **is folding**/ **is buying** a **tablecloth** by herself.
   (tablecloth, batteries, scrub brush, melon)

10. Ya   s-**jal**   / s-**pas**   ta s-tukel   **chujkilal**   te   ants=e
    INC A3-weave[B3]   / A3-make[B3]   P   A3-PRON   traditional belt   DET   woman=DET
    The woman **is weaving**/ **is making** a **belt** by herself.
    (traditional belt, molcajete, violin, bracelet)

11. Ya   s-**sil**   / x-**chol**   ta s-tukel   **elemonex**   te   winik=e
    INC A3-cut (slices)[B3] / A3-line up objects[B3]   P   A3-PRON   lemon   DET   man=DET
    The man is cutting (slices)/ **is lining** up a **lemon** by himself.
    (lemon, button, glass bottle, thread)

12. Ya   s-**jojk'an**   / s-**lok'ta**   ta s-tukel   **moral**   te   kerem=e
    INC A3-hang[B3]   / A3-photograph[B3]   P   A3-PRON   bag   DET   boy=DET
    The boy **is hanging**/ **is photographing** a **bag** by himself.
    (bag, tab, roof tiles, potatoes)

13. Ya   s-**pet**   / s-**majan**   ta s-tukel   **prensa**   te   ants=e
    INC A3-carry (arms)[B3] / A3-borrow[B3]   P   A3-PRON   tortilla press   DET   woman=DET
    The woman **is carrying**/ **is borrowing** a **tortilla press** by herself.
    (tortilla press, money, yo-yo, ribbon)

14. Ya   s-**k'ut**   / s-**k'an**   ta s-tukel   **ixim**   te   ach'ix=e
    INC A3-shell.corn[B3]   / A3-want[B3]   P   A3-PRON   corn   DET   girl=DET
    The girl **is shelling**/ **is wanting** (desiring) **corn** by herself.
    (corn, light bulb, traditional clothes, bicycle)

15. Ya   s-**wuy**   / s-**bon**   ta s-tukel   **ton**   te   kerem=e
    INC A3-smash[B3]   / A3-paint[B3]   P   A3-PRON   stone   DET   boy=DET
    The boy **is smashing**/ **is painting** a **stone** by himself.
    (stone, balloon, paper, shoe)

16. Ya   s-**bul**   / s-**pik**   ta s-tukel   **alwanex**   te   ach'ix=e
    INC A3-uproot[B3]   / A3-touch[B3]   P   A3-PRON   radish   DET   girl=DET
    The girl **is uprooting**/ **is touching** a **radish** by herself.
    (radish, hair clip, lighter, soap)

17. Ya   s-**we'**   / s-**ta**   ta s-tukel   **waj**   te   ants=e
    INC A3-eat.tortilla.like[B3] / A3-find[B3]   P   A3-PRON   tortilla   DET   woman=DET
    The woman **is eating** (tortilla-like)/ **is finding** a **tortilla** by herself.
    (tortilla, camera, leaf, plastic bag)

18. Ya   s-**ts'is**   / s-**k'ej**   ta s-tukel   **sk'u'**   te   ach'ix=e
    INC A3-sow[B3]   / A3-put.away[B3]   P   A3-PRON   shirt   DET   girl=DET
    The girl **is sowing**/ **is putting** (away) a **shirt** by herself.
    (shirt, aluminum sheet, padlock, coffee huller)

19. Ya   y-**uch'**   / s-**jip**   ta s-tukel   **pepsi**   te   kerem=e

INC A3-drink[B3] / A3-throw[B3] P A3-PRON pepsi-cola DET boy=DET
The boy **is drinking**/ **is throwing** a **pepsi** by himself.
(pepsi, suitcase, drum, sandal)

20. Ya s-**tz'u'** / y-**ajtaj** ta s-tukel **wale'** te kerem=e
INC A3-suck.juice[B3] / A3-count[B3] P A3-PRON sugar cane DET boy=DET
The boy **is sucking**/ **is counting** a **sugar cane** by himself.
(sugar cane, cell phone, socks, amphora)

21. Ya s-**t'uman** / s-**tzob** ta s-tukel **chenek'** te ach'ix=e
INC A3-soak[B3] / A3-gather[B3] P A3-PRON beans DET girl=DET
The girl **is soaking**/ **is gathering beans** by herself.
(beans, books, candies, clothe pegs)

22. Ya s-**kojkon** / s-**busan** ta s-tukel **lew** te ants=e
INC A3-pour[B3] / A3-pile.up[B3] P A3-PRON oil DET woman=DET
The woman **is pouring**/ **is piling** up **oil** (plastic bottles) by herself.
(oil, chicken bones, shovel, clothes)

23. Ya s-**jos** / x-**ut'** ta s-tukel **manko** te winik=e
INC A3-peel.w/knife[B3] / A3-pinch[B3] P A3-PRON mango DET man=DET
The man **is peeling** (with knife)/ **is pinching** a **mango** by himself.
(mango, toy puppy, tortilla dough, eraser)

24. Ya s-**bulantes** / s-**k'abu** ta s-tukel **ja'** te ants=e
INC A3-boil[B3] / A3-look[B3] P A3-PRON water DET woman=DET
The woman **is boiling**/ **is looking** at the **water** by herself.
(water, flag, pillow, scale)

25. Ya x-**jujch'i** / s-**tzum** ta s-tukel **kantela** te kerem=e
INC A3-blow[B3] / A3-ignite.turn.on[B3] P A3-PRON candle DET boy=DET
The boy **is blowing**/ is **(lightning) turning** on a **candle** by himself.
(candle, television, chain saw, mixer)

26. Ya s-**joy** / s-**nak'** ta s-tukel **ch'ajan tak'in** te ach'ix=e
INC A3-bend.encircle[B3] / A3-hide[B3] P A3-PRON wire DET girl=DET
The woman **is bending**/ **is hiding** a **wire** by herself.
(wire, boots, can opener, traditional shawl)

27. Ya s-**kets** / s-**pix** ta s-tukel **ak'** te ants=e
INC A3-gather.bunches[B3] / A3-wrap[B3] P A3-PRON hay DET woman=DET
The woman **is gathering** (in bunches)/ **is wrapping hay** by herself.
(hay, cartons, bottle, sand)

28. Ya s-**jul** / y-**ich'** ta s-tukel **akuxa** te winik=e
INC A3-insert[B3] / A3-take[B3] P A3-PRON stickpin DET man=DET
The man **is inserting**/ **is taking** a **stickpin** by himself.
(stickpin, detergent, hammer, bottle opener)

29. Ya s-**tij** / x-**woch'** ta s-tukel **amay** te kerem=e
INC A3-play.instrument[B3] / A3-crush[B3] P A3-PRON flute DET boy=DET
The boy **is playing**/ **is crushing** the **flute** by himself.

(flute, glasses, reed, gourd)

30. Ya   x-**luch**   /   s-**tsoy**   ta   s-tukel   **tsekil**   te   ants=e
    INC   A3-embroider[B3]   /   A3-place[B3]   P   A3-PRON   skirt   DET   woman=DET
    The woman i**s embroidering**/ **is placing** a **skirt** (traditional) by herself.
    (traditional skirt, cart, traditional wooden hook, corn mill)

31. Ya   x-**yoch**   /   s-**meltzan**   ta   s-tukel   **torniyo**   te   winik=e
    INC   A3-loosen[B3]   /   A3-fix[B3]   P   A3-PRON   screw   DET   man=DET
    The man **is loosening**/ **is fixing** a **screw** by himself.
    (screw, toilet, iron, bookcase)

32. Ya   y-**ub**   /   s-**kil**   ta   s-tukel   **chojak'**   te   winik=e
    INC   A3-close (net bag)[B3]   /   A3-drag B3]   P   A3-PRON   net bag   DET   man=DET
    The man **is closing**/ **is dragging** a **net bag** by himself.
    (net bag, spoke, gas cylinder, wood)

Supplementary materials section B

Table SB1
Estimated regression coefficients and variance components for multilevel
logistic regression models predicting fixations to the target as a function
of verb type, time, education, knowledge of Spanish and experiment effects
during time window 1- *Verb + Adverb* region.

Variable	TW1 Verb + Adverb		
	Regression coefficient (95% CI)	*SE*	*z*
Fixed effects			
Intercept	-1.49 (-2.13, -0.84)	0.33	-4.54***
Verb (predictive)	0.50 (-0.01, 1.02)	0.26	1.92
Time	0.004 (-0.07, 0.08)	0.04	0.09
Experiment: second part	-0.14 (-0.67, 0.38)	0.27	-0.53
Education (years)	-0.01 (-0.09, 0.07)	0.04	0.21
Spanish: yes	-0.004 (-0.60, 0.61)	0.31	0.01
Verb (predictive)*Time	0.40 (0.38, 0.42)	0.01	36.89***
Verb (predictive)*Experiment	0.19 (-0.30, 0.68)	0.25	0.77
Verb (predictive)*Education	-0.01 (-0.08, 0.06)	0.04	-0.34
Verb (predictive)*Spanish	0.04 (-0.44, 0.53)	0.25	0.17
Random effects	**Variance**	*SD*	
Participant			
Intercept	0.45	0.67	
Verb (predictive)	0.28	0.53	
Time	0.08	0.28	
Item			
Intercept	1.29	1.13	
Verb (predictive)	0.50	0.71	
Spanish	0.39	0.62	

Note ***$p < .001$