COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 47 (2023) e13285

ISSN: 1551-6709 online
DOI: 10.1111/cogs.13285

Predictive Sentence Processing at Speed: Evidence from Online Mouse Cursor Tracking

Anuenue Kukona[a,b]

[a]*Division of Psychology, De Montfort University*
[b]*School of Human Sciences, University of Greenwich*

Received 6 October 2022; received in revised form 28 February 2023; accepted 27 March 2023

Abstract

Three online mouse cursor-tracking experiments investigated predictive sentence processing at speed. Participants viewed visual arrays with objects like a bike and kite while hearing predictive sentences like, "What the man will ride, which is shown on this page, is the bike," or non-predictive sentences like, "What the man will spot, which is shown on this page, is the bike." Based on the selectional restrictions of "ride" (i.e., vs. "spot"), participants made mouse cursor movements to the bike before hearing the noun "bike." Compellingly, this effect was observed at speech rates of ∼3 (Experiment 1), ∼6 (Experiment 2), and ∼9 (Experiment 3) syllables/s. While prior research suggests striking limits on prediction, these results highlight temporal dynamics that may impact comprehenders' ability to preactivate information when hearing impressively rapid speech. Implications for theories of sentence processing are discussed.

Keywords: Language comprehension; Mouse cursor tracking; Prediction; Semantics; Sentence processing; Speech rate

1. Introduction

Prediction is widely documented in the sentence processing literature (e.g., see reviews by Altmann & Mirković, 2009; Federmeier, 2007; Huettig, 2015; Kamide, 2008; Kuperberg & Jaeger, 2016; Pickering & Gambi, 2018; Van Petten & Luka, 2012). However, an essential

Correspondence should be sent to Anuenue Kukona, School of Human Sciences, University of Greenwich, Old Royal Naval College, Park Row, London, SE10 9LS, UK. E-mail: a.p.bakerkukona@greenwich.ac.uk

feature of language is poorly captured by prior studies: its capacity to transmit information at speed. The aim of the current research was to both assess comprehenders' ability to predict when hearing rapid speech and capture these processes using mouse cursor tracking.

Psycholinguistic research provides compelling evidence for the preactivation of information during sentence processing. For example, participants in Altmann and Kamide (1999) viewed visual scenes with objects like a cake while hearing sentences like, "The boy will eat the cake." Based on the selectional restrictions of the verb "eat" (e.g., vs. sentences replacing "eat" with "move"), participants made eye movements to the cake before hearing the noun "cake." Relatedly, participants in Wicha et al. (2004) read sentences in Spanish like, "El príncipe soñaba con tener el trono de su padre. El sabía que cuando su padre muriera podría al fin ponerse la corona por el resto de su vida" ("The prince dreamt about having the throne of his father. He knew that when his father died he would finally be able to wear the crown for the rest of his life"). Based on this discourse and the consistent grammatical gender of the determiner "la" and noun "corona" ("crown"; e.g., vs. sentences replacing feminine "la" with masculine "el"), participants showed an attenuated event-related brain potential (ERP) response to the "la" before "corona." Thus, prediction of both semantic (e.g., edible "cake") and syntactic (e.g., grammatical gender of "corona") information has been observed.

On the one hand, prediction has been theoretically linked to the speed of language processing. Kuperberg and Jaeger (2016) summarize the argument as follows: "given the noisiness, ambiguity and speed of our linguistic input, prediction is the most efficient solution for fast, efficient, and accurate comprehension" (p. 33). Psycholinguistic research also highlights the impressive speed with which language can be processed. As one classic example, Warren, Obusek, Farmer, and Warren (1969) compared the processing of a rapid sequence of spoken digits ("one," "three," "eight," and "two") to nonlinguistic sounds (high tone, hiss, low tone, and buzz). While all participants reported the correct order of the spoken digits, they were at chance for the nonlinguistic sounds (i.e., presented at a rate of five syllables/s or sounds/s). Classic studies such as Trueswell and Kim (1998) also reveal the rapid nature of sentence processing. Moreover, the precise (e.g., quantitative) speed of language comprehension has also been investigated. In the case of speech, Kuperman, Kyröläinen, Porretta, Brysbaert, and Yang (2021) estimated that the maximum speech rate yielding unhindered comprehension was 6.5 syllables/s. In the case of reading, Brysbaert (2019a) estimated that the average reading rate was 238 words/min. Kuperman et al. (2021) also emphasize the close parallel between their estimate for speech, which corresponded to 290 words/min, and Brysbaert's (2019a) estimate for reading.

On the other hand, the link between prediction and speed is poorly captured empirically. Rather, prior studies have typically presented language at the slower end of the rate continuum, and rate information is often not reported at all. For example, Altmann and Kamide (1999) do not report speech rate, but a rough estimate based on their figures (e.g., "eat the" spanning 700 ms) suggests an approximate speech rate of three syllables/s. Relatedly, Wicha et al. (2004) presented words one at a time for 300 ms with an interstimulus interval (ISI) of 200 ms, reflecting a reading rate of 120 words/min. Thus, these studies presented language at rates approximately one half of those estimated by Kuperman et al. (2021) and Brysbaert (2019a). Likewise, Brysbaert's (2019a) estimate only reflects an average, which is likely exceeded by participants' maximal abilities. Moreover, these studies are not unrepresentative of the literature; highly cited studies including Borovsky, Elman, and Fernald (2012; rough estimate of ~3 syllables/s based on "The pirate hides the treasure" spanning 2400 ms), Huettig and Janse (2016; rough estimate of ~2.5 syllables/s based on "de afgebeelde" spanning 2009 ms), DeLong, Urbach, and Kutas (2005; 120 words/min), and Otten and Van Berkum (2008; rough estimate of ~139 words/min based on an average word presentation of 326 ms with an ISI of 106 ms) presented language at similarly slow rates. Similarly, a sampling of 45 visual world studies by Fernandez, Engelhardt, Patarroyo, and Allen (2020) revealed that only three reported speech rate information, which spanned approximately one to three syllables/s. Thus, prior studies may not provide particularly clear insight into whether prediction is a solution for comprehending rapid speech.

Helpfully, the link between prediction and speed is beginning to be explored empirically. For example, participants in Huettig and Guerra (2019) viewed visual arrays with objects like a bike while hearing sentences in Dutch like, "Kijk naar de afgebeelde fiets" ("Look at the pictured bike"), which were presented at either a slow or normal rate. They do not report speech rate, but a rough estimate based on their average sentence durations (e.g., "Kijk naar de afgebeelde fiets" spanning 4170.05 vs. 1815.58 ms) suggests an approximate speech rate of either two or 4.5 syllables/s. However, only at the slow but not normal rate, participants who previewed the visual arrays for 1 s made eye movements to the bike before hearing the noun "fiets" ("bike") based on the (i.e., consistent) grammatical gender of the determiner "de" (e.g., vs. objects consistent with the neuter "het" rather than common "de"). Relatedly, participants in Wlotko and Federmeier (2015) read sentences like, "They wanted to make the hotel look more like a tropical resort. So along the driveway they planted rows of palms," which were presented at a reading rate of either 120 or 240 words/min. However, only at the slower but not faster rate, participants showed an attenuated ERP response when "palms" was replaced with a word from the same (e.g., "pines") vs. a different (e.g., "tulips") category. Thus, rather than providing a solution for comprehending rapid speech, these studies suggest that preactivation (e.g., of the grammatical gender of "bike" or semantics of "palms") may be limited to processing rates well below (e.g., let alone equal to or faster than) those estimated by Kuperman et al. (2021) and Brysbaert (2019a). Moreover, Huettig and Guerra (2019) argue that their "findings are problematic for theoretical proposals that assume that prediction pervades cognition" (p. 196).

In summary, the sentence processing literature provides compelling evidence for prediction, which has been theoretically linked to the speed of language processing (e.g., see Kuperberg & Jaeger, 2016). In contrast, Huettig and Guerra (2019) and Wlotko and Federmeier (2015) reveal striking limits on preactivation at all but slower rates of processing. Thus, prediction may not provide a solution for comprehending rapid speech. We conjecture that such findings may have two diverging explanations, which prior studies have not distinguished. When hearing rapid speech, comprehenders may be unable to engage preactivation altogether (e.g., due to increased cognitive demands). Alternatively, the span of time that comprehenders have to preactivate information (i.e., before it simply arrives in the speech signal) also tends to be reduced. Thus, in Huettig and Guerra (2019) and Wlotko and Federmeier (2015),

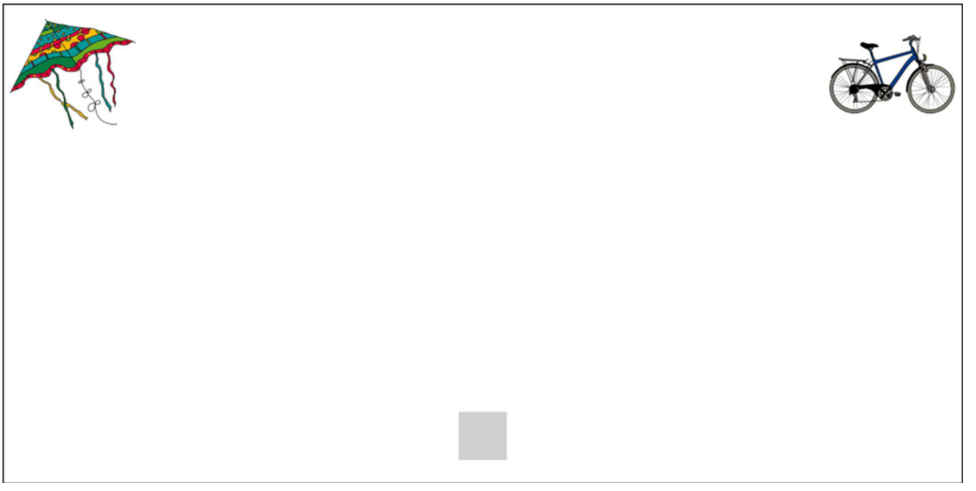

Fig. 1. Visual array for the example predictive sentence, "What the man will ride, which is shown on this page, is the bike."
The target bike is plotted on the right and distractor kite on the left.

participants may have failed to engage preactivation altogether, or they may have engaged preactivation, but its resolution may simply have lagged the speech signal.

The aim of the current research was two-fold: first, to investigate participants' ability to preactivate semantics when hearing rapid speech; and second, to capture these processes using mouse cursor tracking. Across three Internet-mediated experiments, participants viewed visual arrays like Fig. 1 with objects like a bike and kite while hearing sentences like, "What the man will ride, which is shown on this page, is the bike." Similar to Altmann and Kamide (1999), the selectional restrictions of the verb "ride" were predictive of the noun "bike" (e.g., vs. sentences replacing "ride" with "spot"). Thus, the current experiments focused on the prediction of semantic information, like Wlotko and Federmeier (2015), rather than syntactic information, like Huettig and Guerra (2019). Importantly, the current sentences always included filler words (i.e., "which is shown on this page") to provide a buffer between the verb and direct object that maximized the temporal expanse across which predictive effects could be detected. Thus, if engaged, preactivation was not expected to lag the speech signal. Sentences were presented at a natural speech rate in Experiment 1 ($M = 2.98$ syllables/s), doubled rate in Experiment 2 ($M = 5.96$ syllables/s), and tripled rate in Experiment 3 ($M = 8.94$ syllables/s). These rates were slower, approximately equal to and faster than those estimated by Kuperman et al. (2021). If participants are unable to engage preactivation altogether when hearing rapid speech, they were predicted to make mouse cursor movements to the bike before hearing the noun "bike" in Experiment 1 but not in Experiment 2 or 3.

The current research makes novel use of mouse cursor tracking to investigate prediction. Both visual world eye and mouse cursor tracking have been used to investigate phenomena such as spoken word recognition. For example, participants in Allopenna, Magnuson, and Tanenhaus (1998) made more eye movements to a cohort competitor beetle than unrelated carriage when hearing "beaker," supporting incremental lexical processing. Similarly, participants' mouse cursor movements in Spivey, Grosjean, and Knoblich (2005) were more attracted to a cohort competitor candy than an unrelated jacket when hearing (and using their mouse to click on a) "candle" (for closely related Internet-mediated evidence, see also Kukona & Jordan, 2023). In the sentence processing literature, both visual world eye and mouse cursor tracking have also been used to investigate the integration of visual and linguistic constraints (e.g., Farmer, Anderson, & Spivey, 2007; Farmer, Cargill, Hindy, Dale, & Spivey, 2007; Kukona, Gaziano, Bisson, & Jordan, 2022; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). Paralleling the use of predictive eye movements to measure preactivation (e.g., cake after hearing "eat"; Altmann & Kamide, 1999), the current research made use of predictive mouse movements to do the same. Nevertheless, the current design diverged from Altmann and Kamide (1999) in an important respect. As is typical of mouse cursor-tracking studies (e.g., Spivey et al., 2005), the current visual arrays included two objects in the upper corners of the visual display (e.g., vs. visual scenes), which maximized the spatial expanse across which predictive effects could be detected (i.e., with mouse cursor movements beginning at the bottom of the visual display).

2. Experiment 1

Experiment 1 investigated predictive sentence processing at a natural speech rate. Participants viewed visual arrays with objects like a bike and kite (see Fig. 1) while hearing sentences like, "What the man will ride, which is shown on this page, is the bike." All data and code are available at https://osf.io/vnst7.

2.1. Method

2.1.1. Participants

Fifty-two native English speakers from the United Kingdom with normal or corrected-to-normal vision and hearing (age $M = 39.26$, $SD = 10.18$, 6 unreported; 22 female, 30 male) were recruited through Prolific (https://www.prolific.co). The sample enabled detection of an average two-level within-participants psychological effect size ($d_z = 0.40$, power $= 0.80$, alpha $= 0.05$; Brysbaert, 2019b). The research received permission from the Faculty Research Ethics Committee, Faculty of Health and Life Sciences, De Montfort University.

2.1.2. Design and materials

Verb type (predictive and non-predictive) was manipulated within participants.

Thirty-six visual arrays with a target and distractor object (e.g., bike and kite, respectively) from Multipic (Duñabeitia et al., 2018) were created. Display properties (e.g., monitor sizes, screen resolutions, etc.) were not controlled across (i.e., Internet-mediated) participants, and visual arrays used normalized coordinates ranging from −1 to 1, such that (−1,−1) was the left-bottom, (0,0) was the center, and (1,1) was the right-top. Objects were 0.3×0.6 centered in the upper corners of the visual array (± 0.85, 0.70). Objects were square for a 2:1 aspect ratio and stretched for others.

For each visual array, a predictive sentence, in which the target (e.g., bike) but not distractor (e.g., kite) object satisfied the selectional restrictions of the verb (e.g., "ride"), and a non-predictive sentence, in which both objects did so (e.g., "spot"), were also created. Latent semantic analysis revealed that predictive verbs were more related to target ($M = 0.44$, $SD = 0.17$) than distractor ($M = 0.08$, $SD = 0.08$) objects, $t(35) = 11.74$, $p < .001$, while non-predictive verbs did not differ in their relatedness to target ($M = 0.12$, $SD = 0.08$) and distractor ($M = 0.14$, $SD = 0.08$) objects, $t(35) = -1.27$, $p = .21$. Sentences were recorded by a female native speaker of British English, who was instructed to do so at a natural speech rate (duration $M = 5.40$ s, $SD = 0.38$, min = 4.75, max = 6.54; rate $M = 2.98$ syllables/s, $SD = 0.24$, min = 2.47, max = 3.68). The full list of items is reported in the Appendix.

Two counterbalanced lists were created by rotating the visual arrays through the predictive and non-predictive conditions. Participants were presented with one list, which included all 36 visual arrays, one half presented with predictive sentences, and the other half presented with non-predictive sentences. Each non-predictive verb was also included only once on each list (e.g., participants presented "spot … the bike" were presented "play … the drum" rather than "spot … the drum").

2.1.3. Procedure

The experiment was created in PsychoPy (Peirce et al., 2019) and run on Pavlovia (https://www.pavlovia.org). Participants were instructed to use a computer mouse to click on the object referred to in each sentence. The experiment used static start and click response procedures without deadlines, such that participants clicked on an icon at the bottom of the visual array (0, −0.85) to begin each trial. Participants previewed the visual array for 0.50 s before hearing the sentence. There were no practice trials, feedback was not provided, and the order of trials and locations of objects (i.e., left vs. right) was randomized.

2.2. Results and discussion

Four participants who responded before target word (e.g., "bike") onset on a majority of predictive trials were excluded from the analyses. An additional 25 trials (1.45%) in which the remaining participants responded before target word onset were also excluded from the analyses. Problematically, these trials did not yield trajectory data (i.e., horizontal x coordinates) at target word onset, which was the focus of the trajectory analyses (while this approach may thus underestimate preactivation, such trials typically reflected only a small minority). Accuracy was high across predictive ($M = 99.77\%$, $SD = 1.60$) and non-predictive ($M = 99.65\%$, $SD = 1.36$) conditions. Accuracy was submitted to a binomial mixed effects model with a deviation-coded fixed effect of verb type (predictive = −0.5; non-predictive = 0.5) and random uncorrelated intercepts and slopes by participants and random intercepts by items (the maximal model was simplified due to issues with fit). Models were run throughout in R using lme4 (Bates, Mächler, Bolker, & Walker, 2015) and lmerTest (Kuznetsova, Brockhoff, & Christensen, 2017). The analysis of accuracy revealed a non-significant effect of verb type, *Est.* = −0.87, $SE = 3.65$, $z = -0.24$, $p = .81$. Inaccurate trials, as well as trials with log response times (RTs) more than 2.5 standard deviations above the global mean (1.24%), were also excluded from the analysis of trajectories.

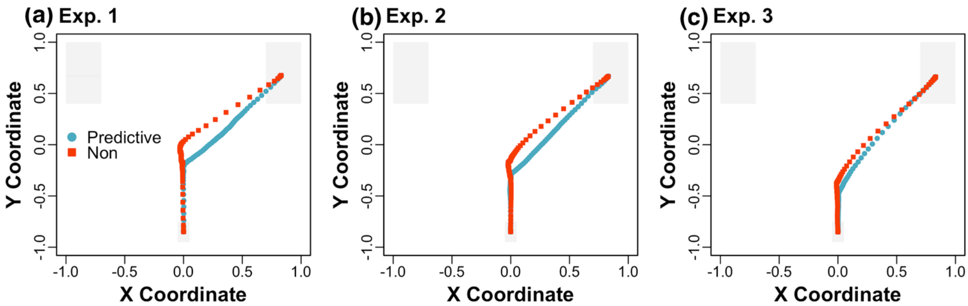

Fig. 2. Time-normalized mean mouse cursor trajectories across the visual array with predictive (e.g., “ride”) versus non-predictive (e.g., “spot”) sentences in Experiments 1–3 (a–c).
Speech rates were ∼3, ∼6, and ∼9 syllables/s in Experiments 1–3, respectively. Target objects (e.g., bike) are plotted on the right and distractor objects (e.g., kite) on the left.

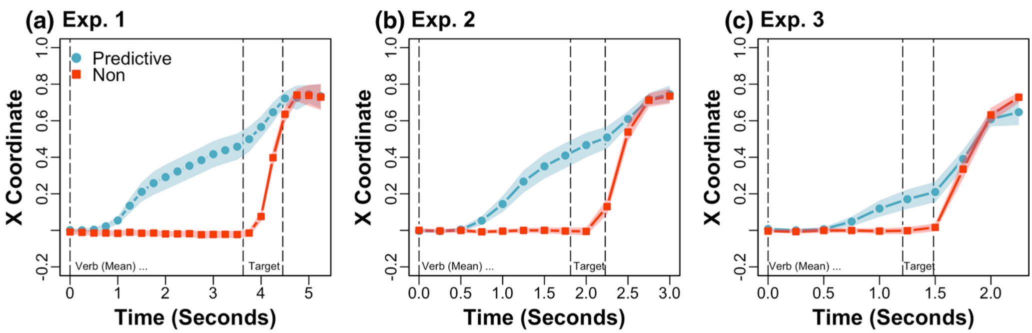

Fig. 3. Mean horizontal mouse cursor movements (i.e., x coordinates) from mean verb onset with predictive versus non-predictive sentences in Experiments 1–3 (a–c).
The shaded bands show 95% confidence intervals and the plots extend to mean sentence offset (+750 ms). Target objects are plotted toward an x coordinate of +1 and distractor objects toward −1.

Fig. 2a depicts mean trajectories across the visual array by verb type. Trajectories were aggregated by dividing each trial into 101 normalized time slices and inverting the horizontal axis for left target presentations (e.g., see Spivey et al., 2005). Fig. 3a depicts mean horizontal x coordinates by verb type in 250 ms time slices from mean verb onset. Horizontal x coordinates near 0 were biased toward neither the target nor distractor object, near 1 were biased toward the target object, and near −1 were biased toward the distractor object. Horizontal x coordinates at target word onset, reflecting predictive mouse cursor movements at the moment (i.e., time slice vs. temporal window) just prior to hearing the direct object noun, were submitted to a mixed effects model with a deviation-coded fixed effect of verb type and random intercepts and slopes by participants and items (intercepts and slopes were uncorrelated by participants; the maximal model was simplified due to issues with fit). The analysis of horizontal x coordinates revealed a significant effect of verb type, *Est.* = −0.49, *SE* = 0.04,

$t(54.01) = -12.00$, $p < .001$, such that trajectories were drawn toward the target object in the predictive ($M = 0.48$, $SD = 0.25$) as compared to non-predictive ($M = -0.02$, $SD = 0.06$) condition.

Participants hearing sentences like, "What the man will ride, which is shown on this page, is the bike" (e.g., vs. sentences replacing "ride" with "spot"), made mouse cursor movements to objects like a bike before hearing the noun "bike." Consistent with Altmann and Kamide (1999), these results reveal that participants can use verb selectional restrictions to preactivate semantics at speech rates of $\sim$3 syllables/s, which is typical of the (e.g., visual world) literature. Novelly, these results also extend prior research by revealing that online mouse cursor tracking is sensitive to preactivation. However, Kuperman et al. (2021) found that comprehension was unhindered at speech rates up to 6.5 syllables/s. In order to assess participants' ability to predict when hearing rapid speech, Experiment 2 presented sentences at speech rates comparable to those estimated by Kuperman et al. (2021).

3. Experiment 2

Experiment 2 investigated predictive sentence processing at twice the speech rate of Experiment 1. Experiment 2 was otherwise identical to Experiment 1.

3.1. Method

3.1.1. Participants

Fifty-two participants (age $M = 37.66$, $SD = 13.32$, 8 unreported; 35 female, 16 male, 1 other) were recruited through Prolific who satisfied the same criteria as Experiment 1 but did not participate in that experiment.

3.1.2. Design, materials, and procedure

The design, materials, and procedure were largely identical to Experiment 1. The only difference was that the durations of sentences in Experiment 1 were halved (i.e., their speech rates were doubled; $M = 5.96$ syllables/s) using the duration manipulation function in Praat (Boersma, 2001). Participants were also informed that sentences may sound distorted.

3.2. Results and discussion

One participant who responded before target word onset on a majority of predictive trials, and another participant with timing errors (i.e., non-monotonic time samples), were excluded from the analyses. An additional nine trials (1%) in which the remaining participants responded before target word onset were also excluded from the analyses. Accuracy was high across predictive ($M = 99.44\%$, $SD = 3.22$) and non-predictive ($M = 99.78\%$, $SD = 1.10$) conditions. The analysis of accuracy revealed a non-significant effect of verb type, $Est. = 1.04$, $SE = 0.88$, $z = 1.17$, $p = .24$. Inaccurate trials, as well as trials above the RT threshold (2.19%), were also excluded from the analysis of trajectories.

Fig. 2b depicts mean trajectories across the visual array, and Fig. 3b depicts mean horizontal x coordinates across time. The analysis of horizontal x coordinates at target word onset revealed a significant effect of verb type, *Est.* $= -0.43$, *SE* $= 0.04$, $t(63.34) = -10.24$, $p <$ .001, such that trajectories were drawn toward the target object in the predictive ($M = 0.42$, $SD = 0.24$) as compared to non-predictive ($M = 0.00$, $SD = 0.07$) condition.

At speech rates of $\sim$6 syllables/s, participants made mouse cursor movements to objects like a bike based on the selectional restrictions of verbs like "ride." Consistent with Experiment 1, these results reveal that online mouse cursor tracking is sensitive to preactivation. Novelly, these results also extend prior research by revealing that participants can preactivate information during sentence processing at faster speech rates than are typical of the literature but do not hinder comprehension (e.g., Kuperman et al., 2021). In contrast to prior research (e.g., Huettig & Guerra, 2019), the current sentences included a lengthy buffer (e.g., between "ride" and "bike") that maximized the temporal expanse across which predictive effects could be detected. In order to assess participants' ability to predict at even faster speech rates, Experiment 3 presented sentences at speech rates above those estimated by Kuperman et al. (2021).

4. Experiment 3

Experiment 3 investigated predictive sentence processing at thrice the speech rate of Experiment 1. Experiment 3 was otherwise identical to Experiments 1 and 2.

4.1. Method

4.1.1. Participants

Fifty-two participants (age $M = 35.70$, $SD = 13.40$, 1 unreported; 38 female, 14 male) were recruited through Prolific who satisfied the same criteria as Experiments 1 and 2 but did not participate in those experiments.

4.1.2. Design, materials, and procedure

The design, materials and procedure were largely identical to Experiments 1 and 2. The only difference was that the speech rates of sentences in Experiment 1 were tripled ($M = 8.94$ syllables/s).

4.2. Results and discussion

Accuracy was high across predictive ($M = 98.82\%$, $SD = 3.70$) and non-predictive ($M = 98.40\%$, $SD = 4.31$) conditions. The analysis of accuracy revealed a non-significant effect of verb type, *Est.* $= -1.12$, *SE* $= 1.19$, $z = -0.94$, $p = .35$. Inaccurate trials, as well as trials above the RT threshold (1.63%), were excluded from the analysis of trajectories.

Fig. 2c depicts mean trajectories across the visual array, and Fig. 3c depicts mean horizontal x coordinates across time. The analysis of horizontal x coordinates at target word onset revealed a significant effect of verb type, *Est.* $= -0.16$, *SE* $= 0.03$, $t(55.60) = -5.57$,

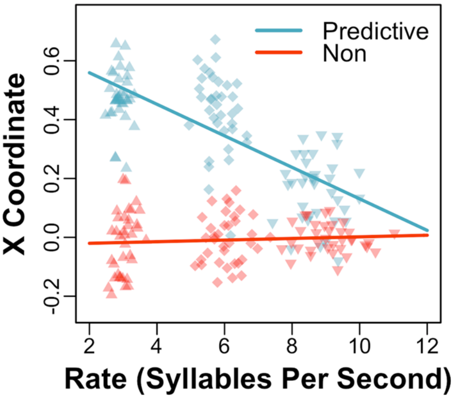

Fig. 4. Mean by-item x coordinates at target onset by speech rate with predictive versus non-predictive sentences across Experiments 1–3.
The upward triangles, diamonds, and downward triangles correspond to Experiments 1–3, respectively. Lines show model fits.

$p < .001$, such that trajectories were drawn toward the target object in the predictive ($M = 0.16$, $SD = 0.19$) as compared to non-predictive ($M = 0.00$, $SD = 0.07$) condition.

Finally, the effect of verb type was also assessed across experiments. Fig. 4 depicts by-item mean horizontal x coordinates at target word onset. These means were submitted to a by-items mixed effects model with a deviation-coded fixed effect of verb type, a fixed effect of rate (syllables per second), and their interaction and random intercepts and slopes by items. The analysis revealed a significant interaction of verb type and rate, $Est. = 0.06$, $SE = 0.01$, $t(147.33) = 11.12$, $p < .001$, such that the effect of verb type decreased as rates increased.

At speech rates of ~9 syllables/s, participants made mouse cursor movements to objects like a bike based on the selectional restrictions of verbs like "ride." Consistent with Experiment 2, these results reveal that participants can preactivate information during sentence processing at faster speech rates than are typical of the literature. Novelly, these results reveal that participants can even do so at speech rates that can hinder comprehension (e.g., Kuperman et al., 2021). Contrasting with Huettig and Guerra (2019) and Wlotko and Federmeier (2015), these results suggest that the fast pace of natural conversation may be supported by predictive sentence processing, at least when the (e.g., temporal) dynamics of an unfolding utterance allow preactivation to reach its resolution before information arrives in the speech signal (e.g., see the buffer between "ride" and "bike").

5. General discussion

Three experiments investigated predictive sentence processing at speed. Based on the selectional restrictions of verbs like "ride" in sentences like, "What the man will ride, which is shown on this page, is the bike" (e.g., vs. sentences replacing "ride" with "spot"), participants made mouse cursor movements to objects like a bike before hearing the noun "bike." Compellingly, these mouse cursor movements were observed at impressively rapid speech rates, which both exceeded prior studies (e.g., Huettig & Guerra, 2019) and have been observed to hinder comprehension (e.g., Kuperman et al., 2021). These results advance the understanding of predictive sentence processing in two novel respects: first, they reveal that participants are able to preactivate semantics when hearing rapid speech; and second, they suggest that motor movements of the hand are sensitive to these processes.

The current results contrast with the striking limits on prediction observed by Huettig and Guerra (2019) and Wlotko and Federmeier (2015). Again, Huettig and Guerra (2019) observed preactivation (e.g., of "fiets" following "de") when participants previewed visual arrays for 1 s and heard sentences at (i.e., roughly estimated) speech rates of ~2 but not ~4.5 syllables/s. Relatedly, Wlotko and Federmeier (2015) observed an attenuated ERP response when predictable words (e.g., "palms" following "planted…") were replaced with words from the same (e.g., "pines") versus a different (e.g., "tulips") category at a reading rate of 120 but not 240 words/min. In contrast, the current experiments revealed preactivation when participants previewed visual arrays for half a second and heard sentences at average speech rates of ~3 (Experiment 1), ~6 (Experiment 2), and even ~9 (Experiment 3) syllables/s. We conjecture that the diverging patterns across this set of studies depend at least partly on the span of time (i.e., buffer) available to preactivate information. In Huettig and Guerra (2019), the predictive "de" was separated from "fiets" by only a single word. Relatedly, in Wlotko and Federmeier (2015), the sentences were more complex (e.g., see "tropical resort"), but the predictive "planted" was separated from "palms" by only two words. In contrast, in the current experiments, the predictive "ride" was separated from "bike" by seven words (i.e., "which is shown on this page," alongside "the"). Thus, even at the most rapid speech rate (i.e., Experiment 3), there was over 1 s on average between the onsets of the verb and noun (e.g., see Fig. 3c). This lengthy temporal buffer likely allowed preactivation to reach its resolution before information arrived in the speech signal, contrasting with Huettig and Guerra (2019) and Wlotko and Federmeier (2015).

The current results complement the recent findings of Fernandez et al. (2020). Their participants viewed visual arrays with objects like a wolf and deer while hearing sentences like, "the wolf attacked the deer… Who did the wolf…?" Based on this discourse and the unfolding wh-question, younger adult participants made eye movements to the deer (e.g., even before hearing the verb in the wh-question) at speech rates up to 5.5 syllables/s. Fernandez et al. (2020) findings address a different aspect of comprehension (e.g., responding to questions), but like the current results, they suggest that participants are able to integrate visual and linguistic constraints at faster speech rates than are typical of the (e.g., visual world) literature. Relatedly, a closer inspection of Wlotko and Federmeier's (2015) within-participants blocked design (i.e., their participants were presented with both reading rates across separate blocks) revealed evidence for preactivation in the faster block when it followed the slower block. The presentation of sentences in ERP studies like Wlotko and Federmeier (2015) diverges considerably from natural reading, and thus one possibility is that their participants had to adapt to the task before they could preactivate semantics effectively, including effectively integrating information from earlier in their utterances (e.g., "tropical resort"). On balance, the findings of Fernandez et al. (2020) and Wlotko and Federmeier (2015) thus provide evidence for prediction at rates comparable to those estimated by Kuperman et al. (2021) and Brysbaert (2019a), complementing the current results.

Taken together, this evidence suggests that prediction may be fundamental to cognition (e.g., see Clark, 2013). If participants can only consistently preactivate information at slower speech rates, then preactivation may be unable to support the fast pace of natural conversation. In contrast, the current results revealed preactivation at speech rates of up to 9 syllables/s (e.g., which is fast enough to hinder comprehension; Kuperman et al., 2021), suggesting that preactivation may provide an important solution for fast comprehension given the speed of linguistic processing (e.g., see Kuperberg & Jaeger, 2016). However, like Huettig and Guerra (2019), the current experiments also suggest that predictive sentence processing is not without limits. The analysis of all three experiments revealed a negative relationship between preactivation and speech rate, predicting that these effects would disappear at a speech rate above ~12 syllables/s or four times Experiment 1 (e.g., see Fig. 4; however, this number exceeds the data range and should be interpreted with caution). Rather than reflecting a hard rate limit, we again conjecture that there is an important trade-off between preactivation and the span of time (i.e., buffer) available to preactivate information. Again, at more rapid speech rates, comprehenders will tend to have less time to preactivate information before it simply arrives in the speech signal. Alternatively, it is conceivable that the current results would be detectable at even more rapid speech rates if the buffer between "ride" and "bike" were increased, providing additional time for processing (e.g., see the 19 words separating "príncipe" and "la corona" in Wicha et al., 2004).

Relatedly, Huettig and Guerra's (2019) findings also suggest important situational constraints on predictive sentence processing. They observed an effect of preview length: While participants did not show evidence of preactivation at a normal speech rate with a 1 s preview, they did with a 4 s preview. Minimally, the current experiments suggest that an extended preview is not a (e.g., situational) prerequisite for preactivation when hearing rapid speech. Rather, the span of time that is required to preactivate information may simply be reduced by an extended preview. Nevertheless, the current results may be partly "scaffolded" by a related situational constraint: as is typical of mouse cursor-tracking studies (e.g., Spivey et al., 2005), the current visual arrays included only two (e.g., vs. four) objects. On the one hand, it is conceivable that weaker effects may have been observed with more complex visual arrays, although preactivation has previously been observed with much more complex visual stimuli (e.g., Coco, Keller, & Malcolm, 2016; Heyselaar, Peeters, & Hagoort, 2021; Reuter, Dalawella, & Lew-Williams, 2021; Staub, Abbott, & Bogartz, 2012). On the other hand, while the current visual arrays were less complex, they were also previewed for less time than in Huettig and Guerra (2019). Thus, the interplay among situational constraints like these remains an interesting direction for future research.

In addition, the current experiments emphasize non-natural speech and semantics, which reflect important limitations. Similar to Fernandez et al. (2020), the auditory stimuli in Experiments 2 and 3 were manipulated in Praat, which enabled the precise manipulation of duration with minimal effect on pitch. However, these stimuli sounded modified (e.g., distorted, if minorly; note that accuracy remained high across experiments) rather than natural. Thus, it is conceivable that even stronger effects may have been observed with natural speech. Relatedly, the speaker who recorded the stimuli in Experiment 1 was instructed to do so at a natural speech rate, but this yielded a rate at the slower end of the continuum, which is not atypical of the (e.g., visual world) literature. Thus, focusing on speech rate when recording stimuli (e.g., alongside clarity) reflects an important consideration for future research. In addition, while the current manipulations focused on semantics, they did not address the preactivation of other (e.g., syntactic, phonological, etc.) information. Underpinning the preactivation of varying informational sources, prediction has been hypothesized to depend on multiple theoretical mechanisms (e.g., Huettig, 2015; Pickering & Gambi, 2018). In fact, the diverging patterns observed in Huettig and Guerra (2019) versus the current experiments may also be connected to the differing influences of syntactic versus semantic information. For example, "de" is likely predictive of far more words syntactically than "ride" is semantically. In addition, given the impressive speed with which associated representations interact (e.g., see examples such as the masked priming literature; Van den Bussche, Van den Noortgate, & Reynvoet, 2009), we conjecture that prediction based on associations (e.g., see also Kukona, Fang, Aicher, Chen, & Magnuson, 2011) may be particularly resilient to rapid speech rates. However, while the preactivation of "bike" when hearing "ride" may depend on associations, it remains unclear whether comprehenders also activate associated but unpredictable representations when hearing rapid speech (e.g., see findings such as Kukona, Cho, Magnuson, & Tabor, 2014, 2016). Thus, exploring multiple mechanisms at speed reflects another important direction for future research.

Finally, the current experiments demonstrate that online mouse cursor tracking is a powerful tool for investigating prediction. The literature is dominated by eye tracking and ERP, which reflect lab-based methods that rely on specialized tools. In the current experiments, mouse cursor tracking provided a continuous (i.e., online) measure of behavior, similar to eye tracking and ERP, but data collection was Internet-mediated (i.e., online). Thus, online mouse cursor tracking provides an exciting new opportunity to move research beyond the lab, reaching wider and more diverse (e.g., vs. WEIRD; Henrich, Heine, & Norenzayan, 2010a, 2010b) populations. Of course, Internet-mediated research is not without limitations. For example, Internet-mediated participants' hardware and software are likely to vary widely, which may introduce considerable noise. Moreover, it may be difficult to comprehensively and accurately capture this variability. Focusing on display properties, participants' screen resolutions were recorded automatically in the current experiments and did vary widely (e.g., approximately $\frac{1}{4}$, the most frequent, used 1920×1080; however, 26 different screen resolutions were recorded). In contrast, recording monitor sizes (e.g., or physical sizes of visual stimuli and distances between them) was not possible automatically. It is conceivable that display properties such as these may interact with speech rate, thus reflecting a potential direction for future (e.g., lab-based) research. Nevertheless, the current results suggest that online mouse cursor tracking is strikingly sensitive to the moment-by-moment dynamics of prediction.

Acknowledgments

This work was supported by the Institute for Psychological Science, De Montfort University. Preliminary results from this study were presented at the 35th Annual Conference on Human Sentence Processing and the Architectures and Mechanisms for Language Processing 28 conference.

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, *38*(4), 419–439.

Altmann, G. T., & Mirković, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science*, *33*(4), 583–609.

Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73*(3), 247–264.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International*, *5*(9), 341–345.

Borovsky, A., Elman, 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.

Brysbaert, M. (2019a). How many words do we read per minute? A review and meta-analysis of reading rate. *Journal of Memory and Language*, *109*, 104047.

Brysbaert, M. (2019b). How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *Journal of Cognition*, *2*(1), 16.

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

Coco, M. I., Keller, F., & Malcolm, G. L. (2016). Anticipation in real-world scenes: The role of visual context and visual memory. *Cognitive Science*, *40*(8), 1995–2024.

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.

Duñabeitia, J. A., Crepaldi, D., Meyer, A. S., New, B., Pliatsikas, C., Smolka, E., & Brysbaert, M. (2018). MultiPic: A standardized set of 750 drawings with norms for six European languages. *Quarterly Journal of Experimental Psychology*, *71*(4), 808–816.

Farmer, T. A., Anderson, S. E., & Spivey, M. J. (2007). Gradiency and visual context in syntactic garden-paths. *Journal of Memory and Language*, *57*(4), 570–595.

Farmer, T. A., Cargill, S. A., Hindy, N. C., Dale, R., & Spivey, M. J. (2007). Tracking the continuity of language comprehension: Computer mouse trajectories suggest parallel syntactic processing. *Cognitive Science*, *31*(5), 889–909.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, *44*(4), 491–505.

Fernandez, L. B., Engelhardt, P. E., Patarroyo, A. G., & Allen, S. E. (2020). Effects of speech rate on anticipatory eye movements in the visual world paradigm: Evidence from aging, native, and non-native language processing. *Quarterly Journal of Experimental Psychology*, *73*(12), 2348–2361.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010a). Most people are not WEIRD. *Nature*, *466*(7302), 29.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010b). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2-3), 61–83.

Heyselaar, E., Peeters, D., & Hagoort, P. (2021). Do we predict upcoming speech content in naturalistic environments? *Language, Cognition and Neuroscience*, *36*(4), 440–461.

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, *1626*, 118–135.

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.

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.

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

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

Kukona, A., Cho, P. W., Magnuson, J. S., & Tabor, W. (2014). Lexical interference effects in sentence processing: Evidence from the visual world paradigm and self-organizing models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(2), 326–347.

Kukona, A., Fang, S. Y., Aicher, K. A., Chen, H., & Magnuson, J. S. (2011). The time course of anticipatory constraint integration. *Cognition*, *119*(1), 23–42.

Kukona, A., Gaziano, O., Bisson, M. J., & Jordan, A. (2022). Vocabulary knowledge predicts individual differences in the integration of visual and linguistic constraints. *Language*, *Cognition and Neuroscience*, *37*(6), 750–765.

Kukona, A., & Jordan, A. (2023). Online mouse cursor trajectories distinguish phonological activation by linguistic and nonlinguistic sounds. *Psychonomic Bulletin & Review*, *30*, 362–372.

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, *31*(1), 32–59.

Kuperman, V., Kyröläinen, A. -J., Porretta, V., Brysbaert, M., & Yang, S. (2021). A lingering question addressed: Reading rate and most efficient listening rate are highly similar. *Journal of Experimental Psychology: Human Perception and Performance*, *47*(8), 1103–1112.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26.

Otten, M., & Van Berkum, J. J. (2008). Discourse-based word anticipation during language processing: Prediction or priming? *Discourse Processes*, *45*(6), 464–496.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203.

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin*, *144*(10), 1002–1044.

Reuter, T., Dalawella, K., & Lew-Williams, C. (2021). Adults and children predict in complex and variable referential contexts. *Language*, *Cognition and Neuroscience*, *36*(4), 474–490.

Spivey, M. J., Grosjean, M., & Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proceedings of the National Academy of Sciences*, *102*(29), 10393–10398.

Staub, A., Abbott, M., & Bogartz, R. S. (2012). Linguistically guided anticipatory eye movements in scene viewing. *Visual Cognition*, *20*(8), 922–946.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632–1634.

Trueswell, J. C., & Kim, A. E. (1998). How to prune a garden path by nipping it in the bud: Fast priming of verb argument structure. *Journal of Memory and Language*, *39*(1), 102–123.

Van den Bussche, E., Van den Noortgate, W., & Reynvoet, B. (2009). Mechanisms of masked priming: A meta-analysis. *Psychological Bulletin*, *135*(3), 452–477.

Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology*, *83*(2), 176–190.

Warren, R. M., Obusek, C. J., Farmer, R. M., & Warren, R. P. (1969). Auditory sequence: Confusion of patterns other than speech or music. *Science*, *164*(3879), 586–587.
Wicha, N. 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.
Wlotko, E. W., & Federmeier, K. D. (2015). Time for prediction? The effect of presentation rate on predictive sentence comprehension during word-by-word reading. *Cortex*, *68*, 20–32.

Appendix

Predictive and non-predictive verbs, as well as target and distractor objects, in Experiments 1–3 are reported in Table A1. Participants heard sentences with a predictive or non-predictive verb that ended with the target (e.g., "What the boy will ride/spot, which is shown on this page, is the bike"). Target and distractor objects were from Multipic (Duñabeitia et al., 2018).

Table A1
Experimental items

Predictive	Non-predictive	Target	Distractor
ring	glance at	bell	shower
win	look at	medal	airport
climb on	notice	roof	knife
melt	see	chocolate	helicopter
play	spot	drum	pencil
dive into	stare at	pool	curtain
slice	view	pizza	keyboard
salute	watch	flag	net
recharge	ask about	battery	kitchen
drive	chat about	car	tomato
bake	discuss	potato	scissors
prune	enquire about	tree	bible
cradle	hear about	baby	doughnut
peel	learn about	banana	sock
swim with	speak about	dolphin	basket
poach	talk about	egg	beard
inflate	think about	balloon	queen
sit on	wonder about	chair	belt
tie	glance at	rope	hair
freeze	look at	ice cream	wave
hike up	notice	mountain	bomb
navigate with	see	compass	sink
ride	spot	bike	kite
shoot	stare at	gun	bed
strum	view	guitar	cow
focus	watch	camera	door

(*Continued*)

Table A1
(Continued)

Predictive	Non-predictive	Target	Distractor
dethrone	ask about	king	microphone
hunt	chat about	lion	spoon
tune	discuss	piano	candle
eat with	enquire about	fork	ruler
lock	hear about	cage	bottle
ripen	learn about	fruit	gym
publish	speak about	book	dragon
pave	talk about	road	scarf
roast	think about	carrot	dress
wear	wonder about	helmet	clown