Language prediction in multimodal contexts: The contribution of iconic gestures to anticipatory sentence comprehension

Florian Hintz[1*], Antje Strauß[2], Yung Khoe[3], Judith Holler[1,4]

[1]*Max Planck Institute for Psycholinguistics, Nijmegen, NL*

[2]*University of Konstanz, Konstanz, DE*

[3]*Center for Language Studies, Radboud University, Nijmegen, NL*

[4]*Donders Institute for Brain, Cognition & Behaviour, Radboud University, Nijmegen, NL*

Keywords: multimodal communication; language comprehension; prediction; gesture-speech integration; iconic co-speech gestures

***Corresponding author:**

Florian Hintz

florian.hintz@mpi.nl

Max Planck Institute for Psycholinguistics

P.O. Box 310

6500 AH Nijmegen

The Netherlands

Abstract word count: 298

Body word count: 11250

Abstract

There is a growing body of research demonstrating that during comprehension, language users predict upcoming information. Prediction has been argued to facilitate dialog in that listeners try to predict what the speaker will say next to be able to plan their own utterance early. Such behavior may enable smooth transitions between turns in conversation. In face-to-face dialog, speakers produce a multitude of visual signals, such as manual gestures, in addition to speech. Previous studies have shown that comprehenders integrate semantic information from speech and corresponding iconic gestures when these are presented simultaneously. However, in natural conversation, iconic gestures often temporally precede their corresponding speech units with substantial lags. Given the temporal lags in gesture-speech timing and the predictive nature of language comprehension, a recent theoretical framework proposed that listeners exploit iconic gestures in the service of predicting upcoming information. The proposed study aims to test this proposal. We will record electroencephalogram from 80 Dutch adults while they are watching videos of an actress producing discourses. The stimuli consist of an introductory and a target sentence; the latter contains a target noun. Depending on the preceding discourse, the target noun is either predictable or not. Each target noun is paired with an iconic gesture whose presentation in the video is timed such that the gesture stroke precedes the onset of the spoken target either by 520 ms (earlier condition) or by 130 ms (later condition). Analyses of event-related potentials preceding and following target onset will reveal whether and to what extent targets were pre-activated by iconic gestures. If the findings reveal support for the notion that iconic co-speech gestures contribute to predictive language comprehension, they lend support for the recent theoretical framework of face-to-face conversation and offer one possible explanation for the smooth transitions between turns in natural dialog.

Language prediction in multimodal contexts: The contribution of iconic gestures to anticipatory sentence comprehension

Introduction

In recent years, cognitive scientists have extensively studied prediction. This work has led to the notion of predictive processing being a fundamental principle of human cognition and prediction, "[offering] the best clue yet to the shape of a unified science of mind" (Clark, 2013). Therefore, prediction has become an integral part of a growing number of theories in the cognitive sciences. Theories of visual and auditory perception, for instance, proposed that "the human brain is continuously busy generating predictions that approximate the relevant future" (Bar, 2007: 280) and that viewers and listeners engage in prediction to prepare for upcoming visual and acoustic events (e.g., Bar, 2009; Bendixen, Schröger, & Winkler, 2009; De Lange et al., 2018; Friston, 2010; Schröger, 2007).

Language comprehension involves processing of auditory and visual information streams. Thus, it is not surprising that the prediction of upcoming information has been assigned an equally important role during language comprehension (e.g., Altmann & Mirković, 2009; Dell & Chang, 2014; Ferreira & Chantavarin, 2018; Kuperberg & Jaeger, 2016; Pickering & Gambi, 2018). Importantly, while the overwhelming majority of studies has focused on unimodal settings (e.g., spoken or written language processing), few studies have investigated prediction during comprehension in multimodal contexts, which – arguably – is the most natural and most frequent form of human language use (Levinson & Holler, 2014). In multimodal settings (e.g., face-to-face dialog), speakers convey information through auditory and visual signals, such as speech and manual gestures (e.g., Bavelas & Chovil, 2000; Kendon, 2004; McNeill, 1992). Critically, manual gestures often precede the elements in the speech stream they semantically correspond to most closely (e.g., ter Bekke et al., 2020; Chui, 2005; Church, Kelly & Holcombe, 2014; Ferré, 2010; Kok et al., 2016; Schegloff, 1984). Based on this observation, a recent framework proposed that recipients exploit the information conveyed by preceding gestures to aid the prediction of upcoming semantic information (Holler & Levinson, 2019). The proposed study aims to test this proposal. Specifically, we ask whether iconic gestures, a frequently used form of manual gestures conveying semantic information (e.g. about objects and actions) contribute to prediction during language comprehension.

Prediction during language comprehension

Prediction during language comprehension has frequently been studied by recording participants' electroencephalogram (EEG) and analysing modulations of the amplitude of the N400 component. The N400, seen as an indicator of semantic processing, is a negative-going, centro-parietally distributed event-related brain potential (ERP) that occurs approximately 400 ms after target word onset (e.g., Kutas & Federmeier, 2011; Kutas & Hillyard, 1984). The N400 is often interpreted as indexing the ease of processing a target word given the context in which it is presented. For example, when embedded in a predictive context, the same target word generates a reduced (i.e., less negative) N400 component, compared to being embedded in a non-predictive context. In such a case, reductions of the N400 amplitude are assumed to reflect the (partial) pre-activation of the target word by the predictive context (e.g., Federmeier & Kutas, 1999; Kutas, DeLong & Smith, 2011; Kutas & Federmeier, 2011). Importantly, there has been a methodological debate concerning the interpretation of N400 effects as reflecting either prediction (i.e., predictive pre-activation) of a given target word or the ease with which the target word can be integrated into the unfolding sentence context. The reason is that many previous EEG studies have measured the N400 starting from target onset, making it difficult to distinguish between the two accounts (e.g., Mantegna et al., 2019; Nieuwland et al., 2018, for discussion).

However, some EEG studies have addressed and overcome this methodological limitation. For example, DeLong et al. (2005) investigated the prediction of phonological form when reading predictable English sentences. The authors exploited the fact that the English indefinite article changes from 'a' to 'an' when the subsequent word, in their case a noun, starts with a vowel while maintaining the same meaning. In sentences such as "The day was breezy, so the boy went outside to fly a kite/an airplane", DeLong and colleagues recorded ERPs on the indefinite article preceding the noun and observed differences in N400 amplitude between the prediction-consistent form of the article ('a' in the example above) and the prediction-inconsistent form. Moreover, the authors found that the amplitude of the N400 component elicited by indefinite articles (and the target nouns) correlated negatively with the target words' predictability in the sentences, as assessed in an off-line cloze probability task (Taylor, 1956). DeLong et al. interpreted their results as demonstrating that readers used the predictive contexts to predict probabilistically the upcoming target word and that this prediction included the pre-activation of the words' phonological form. It is important to note that even though this study is often cited as providing unequivocal evidence for prediction during sentence comprehension, a recent large-scale multi-lab pre-registered study was unable to replicate the crucial condition difference on the articles (Nieuwland et al., 2018). Thus, the status of pre-activation of phonological form features during predictive language processing is currently unclear.

However, using a similar pre-nominal manipulation, Van Berkum et al. (2005; see Nieuwland et al., 2020 for a recent replication) observed evidence for prediction during sentence comprehension. The authors examined whether participants who listened to short Dutch stories use the given discourse information to predict upcoming nouns. Dutch nouns inflect for gender (common and neuter gender) and the gender marking is expressed on adjectives that modify and typically precede a given noun. Van Berkum and colleagues observed that the neural response on encountering an adjective gender-marker that mismatched the gender of the predicted noun differed from processing a gender marker that was congruent with that of the predicted target noun. The authors interpreted the more negative ERP in the mismatch condition as indicating that participants had used the discourse information to predict the target noun, and that this prediction involved the pre-activation of morpho-syntactic knowledge about the target word (i.e., gender). As the ERP was recorded on a word preceding the target, the finding supports the view that listeners predict upcoming words during comprehension. For further electrophysiological evidence for prediction during sentence comprehension see Wicha et al., (2003, 2004; Alemán Bañón & Martin, 2019; Nicenboim et al., 2020, for discussion).

In sum, the ERP literature offers support for the notion that language users often predict upcoming words during sentence comprehension. Motivated by this body of empirical work and in line with recent theoretical proposals, we here conceive of prediction as 'predictive pre-activation of linguistic structures', that is, activation of linguistic structures *before* bottom-up input has had a chance to activate them (Huettig et al., 2022; Huettig, 2015; see Kuperberg & Jaeger, 2016, for a related account). The proposal by Huettig and colleagues (2022) differentiates between within-item (e.g., hearing the beginning of a word pre-activates information about the remainder of that word at multiple levels of representation) and between-item (activation of an item at one or multiple levels of representation spreads to associated items) pre-activation. The authors stress that in their framework prediction is a natural by-product of the structure of the mental lexicon, where activation of connections between levels of word representations (within-item pre-activation) and activation of connections between associated items (between-item pre-activation) naturally result in pre-activation of interconnected information. Importantly, this proposal includes priming (sometimes termed 'expectation') from linguistic and non-linguistic information as contributing to linguistic pre-activation.

To date, the vast majority of empirical work has focused on unimodal language use. Whether, and to what extent, visual communicative signals that carry semantic information, such as iconic gestures, which accompany speech in face-to-face settings, contribute to predictive language processing merits further investigation.

The coupling of speech and co-speech gestures

During multimodal face-to-face interaction, interlocutors use a multitude of visual communicative signals in addition to speech, including manual gestures. Iconic gestures are one of the main carriers of gestural semantic information and can depict actions and objects as well as their attributes and relations (e.g., outlining the square shape of a window, or depicting the action of drinking by imitating to be holding a glass and leading it to the mouth). Iconic gestures are closely coupled with the speech that they accompany in terms of semantic content and timing during production (e.g., McNeill, 1992). Behavioural studies have demonstrated that co-speech gestures significantly contribute to the semantic information conveyed by speech during comprehension (e.g., Drijvers & Özyürek, 2017; Holler, Shovelton, & Beattie, 2009; Hostetter, 2011; Kelly et al., 1999, 2010; McNeill, Cassell & McCullough, 1994). Moreover, gesture comprehension has been studied using EEG mismatch paradigms (activity elicited by congruent speech-gesture pairings is compared with that elicited by conflicting speech-gesture pairings) as well as non-mismatch EEG paradigms. Both have demonstrated that our brain integrates (or attempts to integrate, in the case of conflict) the semantic information conveyed through speech and co-speech gestures (e.g., Drijvers & Özyürek, 2018; Holle & Gunter, 2007; Kelly, Kravitz & Hopkins, 2004; Özyürek et al., 2007; Sekine et al., 2020; Wu & Coulson, 2005, 2007), and that this integration happens early, at least when using naturalistically produced stimuli (already from 200 ms post-stimulus onset) and when the information conveyed is new (Wu & Coulson, 2010). Further evidence for the integration of iconic co-speech gestures during speech comprehension comes from fMRI studies (e.g., Dick et al., 2009; Green et al., 2009; He et al., 2018; Holle et al., 2008; Skipper et al., 2009; Willems, Özyürek & Hagoort, 2007, 2009).

A question that has not been addressed exhaustively by this body of research yet is whether gestures also play a role in prediction during spoken language comprehension. One prerequisite for this possibility is not only that the semantic information encoded in gestures is combined with the information in the speech (as shown by past studies where gestures co-occurred with target words), but also that gestures temporally precede related semantic information in the speech stream. Evidence that they often do so stems from qualitative, observational studies based on detailed examinations of individual gestures and their timing relative to speech (e.g., Kendon, 1980; Schegloff, 1984), as well as from quantitative evidence derived from multimodal language corpora and experimental studies (e.g., ter Bekke et al., 2020; Bergmann et al., 2011; Butterworth & Beattie, 1978; Ferré, 2010; de Kok et al., 2016; Morrel-Samuels & Krauss, 1992).

The recent study by ter Bekke et al. (2020) bears the most relevance to the proposed study since it, as the proposed study, focused on Dutch. This analysis revealed a varying extent to which co-speech gestures precede speech. When measured in their entirety, gestures were shown to precede the onset of the lexical affiliate (i.e., the linguistic component their depiction is most closely related to semantically) by an average of 680 ms. When measured from the onset of the stroke phase (i.e., the most meaning-bearing component of a gesture), the gestural semantic depiction preceded the lexical affiliate on average by 215 ms, with substantial variation ranging from 4300 ms to 6 ms.

In terms of multimodal utterance comprehension, researchers have previously addressed the variation in gesture-speech timing and its effect on the integration of both streams. Habets et al. (2011) manipulated the time interval between the onset of iconic gestures and the onset of corresponding spoken verbs using three different intervals: coinciding, gesture preceding by 160 ms, and by 320 ms. The authors used a mismatch paradigm with gestures and verbs in isolation either matching or not matching in meaning. The mismatch was indexed by a negative-going N400 (referenced to spoken verb onset) for incongruent compared to congruent gesture-verb pairs in the condition where the onsets of gesture and verb coincided, and when gestures slightly preceded the verbs by 160 ms. This effect was lost at asynchrony values of 320 ms, suggesting that early gesture and verb were not integrated.

Obermeier, Holle and Gunter (2011) used gesture fragments and embedded these in sentence contexts, either such that the gesture fragment preceded the noun (a homonym) it depicted, or occurred in synchrony with it. The gesture disambiguated a target word further downstream. The authors found that when gesture and speech occurred in synchrony, integration was more or less automatic. In the case of substantial asynchrony, however, more controlled, active memory processes were required for successful integration and disambiguation. Following-up on this result, Obermeier and Gunter (2015), used the same basic paradigm to investigate in more detail the specific time window during which semantic gesture-speech integration takes place in these sentence contexts. Their results suggested a window ranging from the gesture (or more specifically, the point at which the gesture was able to disambiguate the homonym) occurring 200 ms prior to 120 ms post-homonym within which semantic gesture-speech integration takes place. When the gesture preceded the homonym by 600 ms, they found no evidence of integration of gesture and speech at the homonym itself, but, intriguingly, they still found a disambiguation effect at the target word. On the basis of the latter finding and an exploratory post-hoc analysis, Obermeier and Gunter (2015) concluded that gestures occurring 600 ms prior to the homonym are perhaps integrated “in a different way, possibly in an earlier time window” (p. 303), thus suggesting the existence of multiple semantic integration points for gestures during an unfolding sentence. Note that the results of Obermeier and Gunter’s (2015) post-hoc analysis are in principle in line with the notion that gestures contribute to predictive processing in that they were integrated early, thereby potentially contributing to the pre-activation of the target. However, the conclusions that may be drawn in this regard are limited as in natural speech homonyms rarely disambiguate subsequent nouns. More importantly, their design did not tap into language prediction specifically as Obermeier and Gunter (2015) manipulated target predictability solely on the basis of the two-word homonym-target relationship. In a study tapping the potential of iconic gestures to prime semantic concepts, Wu and Coulson (2007) presented silent video clips of iconic gestures and measured the EEG response to subsequent target words, semantically related or unrelated to the gesture. The former elicited less negative ERPs during the N400 window. While this study supports the idea that iconic gestures can prime semantic concepts, it does not tell us whether such behaviour generalizes to stimuli where gestures are produced in discourses, affecting target word processing therein.

In a recent study, Fritz and colleagues (2021) went beyond this limitation and manipulated discourse predictability to measure its effect on speech-gesture integration. They used stimuli in which iconic gestures were presented very early, (i.e., 5-7 syllables) before a target verb that was either predictable or not with respect to a preceding discourse. Similar to Obermeier and Gunter (2015), Fritz and colleagues (2021) observed EEG evidence that the presentation of early iconic gestures led to facilitated processing of the target verb, but only when the discourse was semantically related to the gesture (relative to an unrelated discourse). This effect was reflected in a late P600 ERP component, measured from spoken verb onset. However, in contrast to Obermeier and Gunter's (2015) post-hoc analysis, Fritz and colleagues found no evidence for immediate integration – that is, at the point where the early gestures occurred no integration took place (even when the speech context was predictable). It is worth highlighting that the gestures in Fritz et al.'s study preceded their affiliate by a substantial amount (on average 2 seconds, compared to 600 ms as in Obermeier & Gunter, 2015) and ended before its onset. Moreover, on their own, many of their gestures were low in interpretability. Finally, the discourse contexts used by Fritz et al. (2021) may not have sufficiently constrained the meaning of the early gestures for immediate integration to take place as acknowledged by the authors (p. 13). Fritz et al. therefore reasoned that the lack of an early integration effect might be due to gestures initially remaining ambiguous and that "gesture interpretation that has not been fully specified (disambiguated) can be re-analysed as more relevant information enters the discourse" (p. 15). Such an account can accommodate the downstream facilitatory effects of early gestures in predictable contexts after target word onset. However, such an account is not, at least not in a straightforward fashion, compatible with the notion that gestures contribute to predictive processing – in the sense of pre-activating a concept – as no integration effects were observed before target word onset.

In sum, the studies by Obermeier and Gunter (2015) and Fritz et al. (2021) have shown that gestures that occur substantially in advance of a related target word (even as early as 2 seconds) are still semantically integrated. This is an interesting finding as it shows that the temporal integration window for speech and gesture is considerably larger than previously assumed. The studies provide mixed evidence as to whether iconic gestures contribute to prediction during sentence comprehension. While the results by Obermeier and Gunter (2015) are in principle in line with that notion, the recent data by Fritz et al. (2021) suggest that early iconic gestures, embedded in constraining discourses, merely facilitate downstream integration of the target.

The potential benefit of early gestures in the context of conversational turn-taking

Whether or not iconic gestures contribute to predictive comprehension is an important topic for models of conversational turn-taking. For example, the situated framework for understanding human language processing by Holler and Levinson (2019) assumes a critical role of gestures in prediction during comprehension. The framework focuses on the various layers (from individual sounds and words to speech acts) and modalities of language. Specifically, the authors hypothesized that "the compositional and temporal architecture of multimodal utterances facilitates predictive coding on multiple levels, leading to a processing advantage over unimodal utterances" (p. 645). Indeed, this hypothesis is supported by behavioural studies showing that responses tend to be faster when speech is presented together with gestures than when speech is presented alone (e.g., Holle et al., 2008; Kelly, Özyürek & Maris, 2010; Nagels et al., 2015; Wu & Coulson, 2015). Moreover, Holler and Levinson assume that such a processing advantage benefits the fast pace of conversational turn-taking—especially when visual signals occur early on during the verbal utterance. Anticipating what the unfolding turn is going to be about is beneficial since response planning can begin early, facilitating the timely launch of the next conversational contribution (see also Levinson, 2006; Levinson & Torreira, 2015; Pickering & Garrod, 2015). The immediate integration of gestures that precede their lexical affiliate might thus benefit the tight temporal present conversational turns through early gestures pre-activating concepts (or contributing to their pre-activation) that are encoded verbally only further downstream. However, as previous studies provided inconclusive results, this hypothesis warrants further experimental investigation.

The proposed study

The proposed study will investigate whether iconic gestures contribute to predictive language comprehension. Albeit not employing an interactive paradigm requiring participants to respond, the proposed study sets out to test whether preceding iconic gestures pre-activate semantic concepts denoted by words occurring further downstream. Such pre-activation is a vital prerequisite for gestures potentially facilitating language comprehension in a turn-taking context. To address our main research question, we propose an EEG experiment similar to those conducted by Obermeier and Gunter (2015) and Fritz et al. (2021): Our participants will be watching videos of an actress producing two-sentence discourses each featuring a target word (i.e., a noun that can be gestured) at the end of the discourse. While remaining in a still pose when producing the preceding context, the actress will execute an iconic gesture shortly before producing the target word. Similar to Obermeier and Gunter (2015) and Fritz et al. (2021), the iconic gesture depicting the concept the target word will occur prior to the target word. In the present study, the gestures will be timed either such that its stroke starts 520 ms before target word onset (earlier condition) or 130 ms before target word onset (i.e., presented almost simultaneously with the target word, later condition; see Figure 1, for a schematic illustration of the trial structure). This temporal within-participants manipulation is substantially smaller than in the corresponding condition in Fritz et al. (2021) where gestures started on average a little more than 2 seconds before their lexical affiliate. The timing between gesture and target word we applied here was inspired by the natural timing variation established by recent corpus analyses of casual, unscripted face-to-face conversations (ter Bekke et al., 2020). Ter Bekke et al. (2020) found that for almost a quarter of representational (i.e., including iconic) gestures in the corpus (55 out of 258 gestures; 21.31%), the stroke onset preceded the lexical affiliate onset by 520 ms and sometimes more. We settled for 520 ms to apply a lag that is representative of the timing of gestures in conversation, and which still looked natural by intuitive judgement when examining stimuli created with a range of different lags (we started with the 130 ms lag applied in earlier studies and then doubled, tripled and quadrupled this lag for inspection). The timing we chose for our earlier-gesture condition was similar to the corresponding condition in Obermeier and Gunter (2015), where the temporal lag between gesture stroke onset and homonym was 600 ms. The lag of 130 ms in our later gesture condition was chosen such that gesture stroke onset still occurred before target word onset (as in the studies by Drijvers & Özyürek, 2018; Habets et al., 2011), but substantially later than in the earlier-gesture condition such that gesture and target word are to a large extent processed simultaneously.

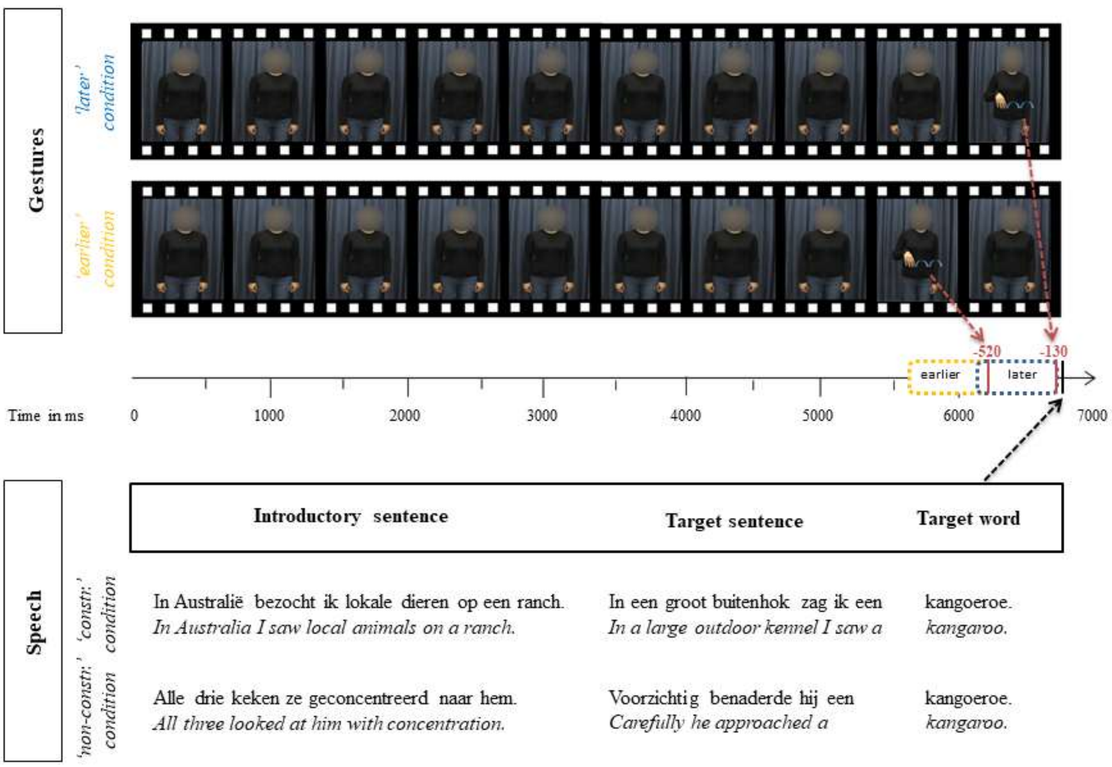

Figure 1: Schematic trial overview. Dotted boxes on the timeline indicate gesture onset (left-hand edge) and gesture stroke onset (right-hand edge, red vertical line). Gesture stroke onset was 520 ms (earlier condition) and 130 ms (later condition) before target word onset (black vertical line). Average gesture onset was 553 ms before gesture stroke onset.

In addition to gesture timing, the proposed study will manipulate target word predictability (within-participants) to gain insight into how it interacts with earlier and later gesture presentation. Thus, each target word will be embedded in a discourse, which is predictive of the word in question, as well as in a non-predictive discourse. In a rating study, we assessed the target word predictability using a cloze task (Taylor, 1956). Fritz et al. (2021) used discourses that were moderately constraining and did not observe effects of early gesture integration. For the proposed study, we therefore opted for highly constraining and non-constraining discourses, respectively, where the non-constraining condition will establish the baseline effects of iconic gesture processing in both earlier and later gesture presentation conditions.

A third aspect the proposed study will take into account is the interpretability of iconic gestures, which is naturally quite variable. While a lot of iconic gestures are not easily interpretable in the absence of speech (e.g. imagine someone drawing a square into the air to communicate 'TV' or 'window'), others are more pantomimic in nature and quite easily interpretable on their own (e.g., imagine someone depicting holding a glass and leading it to the mouth to communicate 'drinking'). For gestures produced without speech (silent gestures), it has been shown that there is a great deal of consistency in how different individuals iconically depict referents (Ortega & Özyürek, 2016, 2019; van Nispen et al., 2017). The basis for this convergence in gestural form and meaning across individuals is likely to be the combination of the manual modality's affordances (Perniss, 2015) and conceptual knowledge rooted in shared motor schemas (Ortega, Schiefner, & Özyürek, 2019). This may contribute to many iconic co-speech gestures also being quite interpretable in the absence of speech. Studies that have measured the communicativeness of iconic gestures have shown that, on the whole, they are most effective in conjunction with speech, but when shown without audio, they still communicated a third of the information they were able to convey together with the speech (Beattie & Shovelton, 1999).

In the present study, we are interested in how spoken discourse (predictable and non-predictable) and iconic gestures (early and late) conspire to influence language comprehension and whether iconic gestures contribute predictive pre-activation of linguistic structures. We will therefore ask an independent set of participants to interpret our iconic gesture stimuli, embedded in the predictable and non-predictable discourses, using three guesses (see Methods). Note that we conducted a previous rating study, involving 30 participants who interpreted the iconic gestures in isolation (without audio, so-called silent gestures), suggested good compatibility between the gestures and the concept they were meant to depict: In 38% of the cases, the target word was amongst those guesses; in 30% of the guesses, the target was mentioned as the first guess. This is quite remarkable, showing that at least in the present stimulus set, which is based on an actress spontaneously creating gestures with the target words, the iconic gestures were often relatively interpretable in the absence of speech. However, to capture the effect the gestures have on multimodal comprehension, in the proposed EEG experiment, we will use the variation in gesture interpretability, when presented in their respective discourse context, as a covariate in the planned analyses[1].

Hypotheses

Based on the literature discussed above (e.g., Fritz et al., 2021; Obermeier and Gunter, 2015), we adopt the notion that iconic gestures can facilitate language comprehension by providing an additional (visual) modality through which semantic representations can be accessed. Critically, our overarching hypothesis is that earlier iconic gestures can play a pivotal role in prediction during language comprehension as they may contribute to the pre-activation of the semantic concept of target words before these are encountered. Recall that in light of recent methodological debates (Mantegna et al., 2020; Nieuwland et al., 2019), recording ERP effects before target word onset is key for linking these effects to the notion of prediction. Such an account is in line with the within-item predictive pre-activation as recently proposed by Huettig et al. (2022), where semantic representations of the target concept may be pre-activated through non-verbal, gestural information denoting that concept.

In general, we expect facilitated processing to be reflected in reduced ERP amplitudes compared to scenarios where processing might be more difficult/demanding (Kutas & Federmeier, 2011, for discussion). In a similar vein as Federmeier and Kutas (2011) interpreted the results from studies on semantic feature pre-activation (Federmeier & Kutas, 1999) and Rommers et al. (2013) interpreted the results from studies on visual feature pre-activation[2], we view reduced ERP amplitudes as reflecting

[1] We thank an anonymous reviewer for this suggestion.
[2] We note that both of these studies analyzed N400 amplitude differences as measured from target word onset.

the degree to which the target has been pre-activated by iconic gestures and/or predictive discourses. It is likely that pre-activation will predominantly take place at the semantic level of representations. However, in cases of high certainty, pre-activation may also boost activation at other levels of representation (e.g., phonological, DeLong et al., 2005).

Our hypotheses concerning effects *after* target word onset will focus on the N400 component. Given the lack of a pre-target ERP equivalent to the N400, when describing hypothesized effects *before* target word onset, we will use a neutral term and refer to differences in 'ERP amplitude'. As for specific contrasts, we predict the following (see Table 1 for an overview; see also Statistical analysis):

(1) **Discourse predictability.** As has been demonstrated numerous times (e.g., DeLong et al., 2005; Van Berkum et al., 2005), we expect participants to exploit discourse information for generating predictions about upcoming targets (between-item pre-activation, as termed by Huettig et al., 2022). We have used the predictable and non-predictable discourses described below in a recent study ERP study and observed evidence for a main effect of discourse predictability (Hintz et al., submitted). Therefore, compared to the non-constraining conditions, target word processing should be facilitated in the constraining-discourse conditions. We believe this effect to be non-controversial.

Hypothesized statistical effect 1: Main effect of *discourse predictability* **after** target word onset, with reduced N400 components in the constraining compared to the non-constraining discourse conditions.

(2) **Earlier and later gesture presentation in constraining and non-constraining discourses.** We expect that when iconic gestures are presented well in advance of the target words (earlier conditions), participants will extract lexico-semantic information from the gestures, which may – at least to some extent – pre-activate the target concepts (Obermeier & Gunter, 2015; within-item pre-activation, as termed by Huettig et al., 2022). We expect differential effects of earlier gesture presentation in the constraining and non-constraining discourse conditions. As explained in (1), we assume that listeners will exploit the constraining discourses for generating predictions about the

upcoming targets. Thus, the constraining discourses will pre-activate the targets to some degree before their onset. We assume that when iconic gestures are presented early and can be processed well before target onset, listeners will map the information extracted from the gestures onto their mental model of the unfolding discourse (cf. Hintz et al., 2020; Metusalem et al., 2012; Obermeier & Gunter, 2015), which includes the partially pre-activated target concept. The match between between-item pre-activation and within-item pre-activation may further contribute to the pre-activation of the target concept, resulting in facilitated processing before target onset. In contrast, when listening to non-constraining discourses, participants are not able to predict the upcoming targets via between-item pre-activation. When presented with early iconic gestures, they will still extract lexico-semantic information via the visual stream, which may partially pre-activate the upcoming target concept. However, unlike in the constraining discourse condition, listeners cannot straightforwardly map this information onto their discourse model before target onset as visually-derived and discourse-derived representations are not yet consistent with each other.

Hypothesized statistical effect 2a: Interaction between *gesture presentation* and *discourse predictability* **before** target word onset, with enhanced event-related potentials due to increased processing efforts during multimodal integration in the earlier-gesture-non-constraining-discourse condition than in the corresponding earlier-gesture-constraining discourse condition.

Relatedly, we predict the pre-target ERP response in the early-gesture-non-constraining condition, reflecting enhanced processing, to differ from that of the two corresponding pre-target time windows in the later-gesture conditions where no early audio-visual integration took place.

Hypothesized statistical effect 2b: Interaction between *gesture presentation and discourse predictability* **before** target word onset, with enhanced ERPs due to increased processing efforts during multimodal integration in the earlier-gesture-non-constraining-discourse condition compared to in the two later-gesture conditions.

Given the direct mapping of visually-derived information onto discourse-derived information, we further expect the pre-target ERP response of the earlier-gesture-constraining-discourse condition to differ from that of the later-gesture conditions as in the latter conditions gesture information and mapping will become available later. Specifically, compared to the later-gesture conditions, processing in the earlier-gesture-constraining-discourse-condition should be facilitated due to the availability of the gesture information.

Hypothesized statistical effect 2c: Interaction between *gesture presentation* and *discourse predictability* **before** target word onset, with reduced ERPs for the earlier-gesture-constraining-discourse than for the two corresponding time windows in the later-gesture conditions.

Note that given the hypothesized facilitatory pre-target effect in the earlier-gesture-constraining-discourse condition and the predicted enhanced processing effort before target onset in the earlier-gesture-non-constraining-discourse condition, we do not predict a main effect of gesture presentation. That is, the predicted pre-target effects in both conditions may cancel each other out, in which case we would most likely not observe a general statistical difference between earlier and later gesture presentation conditions. Given that the existence of both hypothesized interaction effects is yet to be determined and given that the size of both hypothetical effects is unknown, we will include the test for a main effect of gesture presentation as an exploratory analysis.

Exploratory effect 2d: Main effect of *gesture presentation* **before** target word onset.

Whereas the hypotheses for the pre-target main effect of gesture presentation are unclear, we predict a main effect for that condition after target word onset. That is, we assume that early gesture presentation provides a head start for gesture processing (before encountering the lexical target), which contributes to the activation of the target concept, and which will have downstream facilitatory consequences for target word processing (cf. Fritz et al., 2021; Obermeier & Gunter, 2015).

Hypothesized statistical effect 2e: Main effect of *gesture presentation* **after** target word onset, with reduced N400 components for earlier than for later gesture presentation conditions.

In addition to the post-target main effects of discourse predictability and gesture presentation, we predict an interaction between both factors. Whereas later gesture presentation is unlikely to yield post-target effects over and above those of discourse predictability, we hypothesize that early gesture presentation in the constraining-discourse condition will be most beneficial in that multimodal integration is likely to take place before target onset, in which case processing during the target word should be facilitated compared to the late-gesture-constraining-discourse condition.

Hypothesized statistical effect 2f: Interaction between *gesture presentation* and *discourse predictability* **after** target word onset, with a reduced N400 component for the earlier-gesture compared to the later-gesture constraining-discourse condition.

(3) **Effects of gesture interpretability.** We predict that the effects outlined above are moderated by gesture interpretability—that is, how well the gestures depict the target noun in the respective discourse they are embedded in. Specifically, we hypothesize that when a gesture is highly interpretable, it will attenuate the amplitude of the observed ERP components. Multimodal integration costs for earlier gestures are expected to be lower for gestures with high compared to low interpretability. Similarly, amplitudes post-target word onset should be reduced for the later-gesture condition, as the integration for gestures presented almost simultaneously with the targets should be easier for gestures with high rather than low interpretability.

Hypothesized statistical effect 3: Interaction between *gesture interpretability* and *gesture presentation* **before** and **after** target word onset, with reduced ERP amplitudes for gestures with high rather than low interpretability.

Table 1: Overview of hypothesized and exploratory EEG effects.

Effect	**Pre-/post-target**	**Effect type**	**Comparison**
1	post	main effect	*constr.* vs. non-constr.
2a*	pre	interaction	*early-constr.* vs. early-non-constr.
2b	pre	interaction	early-non-constr. vs. *late-constr.* & *late-non-constr.*
2c	pre	interaction	*early-constr.* vs. late-constr. & late-non-constr.
2d	pre	main effect	early vs. late (exploratory)
2e	post	main effect	*early* vs. late
2f*	post	interaction	*early-constr.* vs. late-constr.
3	pre & post	interaction	covariate ‘interpretability’ (*high* = *more positive*)

Note. Of the listed conditions in the ‘comparison’ column, the one in italic font is hypothesized to elicit the reduced ERP in the statistical comparison. Effects with * are minimally required for providing evidence for the notion that early iconic gestures contribute to predictive language comprehension.

It is important to highlight that statistical evidence in favour of the hypothesized effects 2a, 2c, 2e, and 2f would constitute evidence for the notion that iconic gestures, presented well before target word onset, contribute to predictive language comprehension. Note, however, that not all effects are necessary for confirming that notion. Minimally required are effects 2a and 2f. Effect 2a would demonstrate that earlier gestures presented in constraining discourses yield facilitatory processing before target word onset compared to earlier gestures presented in non-constraining discourses. Similarly, effect 2f would demonstrate the positive downstream consequences (i.e., after target onset) of early compared to late gesture presentation, over and above the effects of discourse predictability.

Methods

Pilot data

Given the sparse previous EEG research on language prediction in multimodal contexts, we decided to assess the feasibility of the proposed experiment in a pilot study. To that end, we recruited 13 pilot participants before the first Covid-19-related lockdown in the Netherlands in March 2020. The participants were selected based on the same criteria as described for the proposed study (see below). The data of two pilot participants could not be analysed: One had to be excluded due to a technical error during testing; the other one turned out to be an early bilingual speaker of Farsi and Dutch. The remaining eleven pilot participants were on average 24 years old (*SD* = 2.82, range = 21-31); eight of which were female. We had initially planned to run the experiment using a mixed-design with gesture presentation as within-participants and with discourse predictability as between-participants manipulation. Five of the pilot participants therefore heard the target words embedded in constraining-discourse contexts and six participants heard the target words in non-constraining-discourse contexts. The data[3] were pre-processed using the same pipeline as described for the proposed study. By-participant averages for each condition were created, which were then used to create grand-averages. For inspecting the pilot data, we focused on the Cz and Pz electrodes, which have previously been described to be sensitive to context effects on semantic processing (e.g., the N400, Kutas & Federmeier, 2011; Lau et al., 2008). Note that due to this being a pilot dataset consisting of a small number of participants, we did not conduct any inferential statistics on the data. For the main analysis of the experiment, we will consider all electrodes.

[3] As for the main study data, the data of the pilot participants will be made available at the archive of the Max Planck Institute for Psycholinguistics.

Panel A in Figure 2 plots grand-average ERPs for the four conditions, relative to spoken target onset (time zero), one second before and one second after target word onset. Visual inspection suggests that the early-gesture-constraining-context condition (red solid line) started to diverge (i.e., became more positive) from all other conditions around 400 ms prior to target word onset and continued to be more positive than all other conditions until target word onset. During this period, there was no major difference between the other three conditions. At target word onset, we observed a difference between the two early-gesture conditions (solid red and solid black lines), with constraining trials eliciting more positive ERPs than non-constraining trials. Around 400 ms after target word onset, the two constraining conditions (solid and dashed red lines) differed from the two non-constraining conditions (solid and dashed black lines). Moreover, while the early-non-constraining condition appeared to be more negative than the late-non-constraining condition after target word onset, there was no major difference between the two constraining conditions.

In sum, visual inspection of the data of the eleven pilot participants suggested that the design for the proposed experiment is suited for observing meaningful differences between conditions. That is, the large difference between constraining and non-constraining trials 400 ms after target word onset most likely reflected facilitatory effects of context predictability on target word processing (Kutas & Federmeier, 2011). However, more relevant for the proposed study, there appeared to be differential effects associated with earlier and later gesture presentation such as the pre-target positive deflection in the earlier-gesture-constraining-discourse condition (shortly after gesture stroke onset) and the post-target negative deflection in the earlier-gesture-non-constraining-discourse condition[4].

[4] We recently also used the stimuli for the proposed study in a similar ERP experiment (n = 60, Hintz et al., submitted). In that experiment, we manipulated discourse predictability (predictbale vs. non-predictbale, between-participants) and gesture type (iconic vs. control, within-participants). Gesture stroke onset was in all conditions 130 ms before spoken target word onset. We observed main effects of both manipulations (facilitated processing for predictable and iconic gestures, respectively) as well as an interaction between both factors: Compared to all other conditions, processing was most facilitaed in the predictable-discourse-iconic-gesture condition.

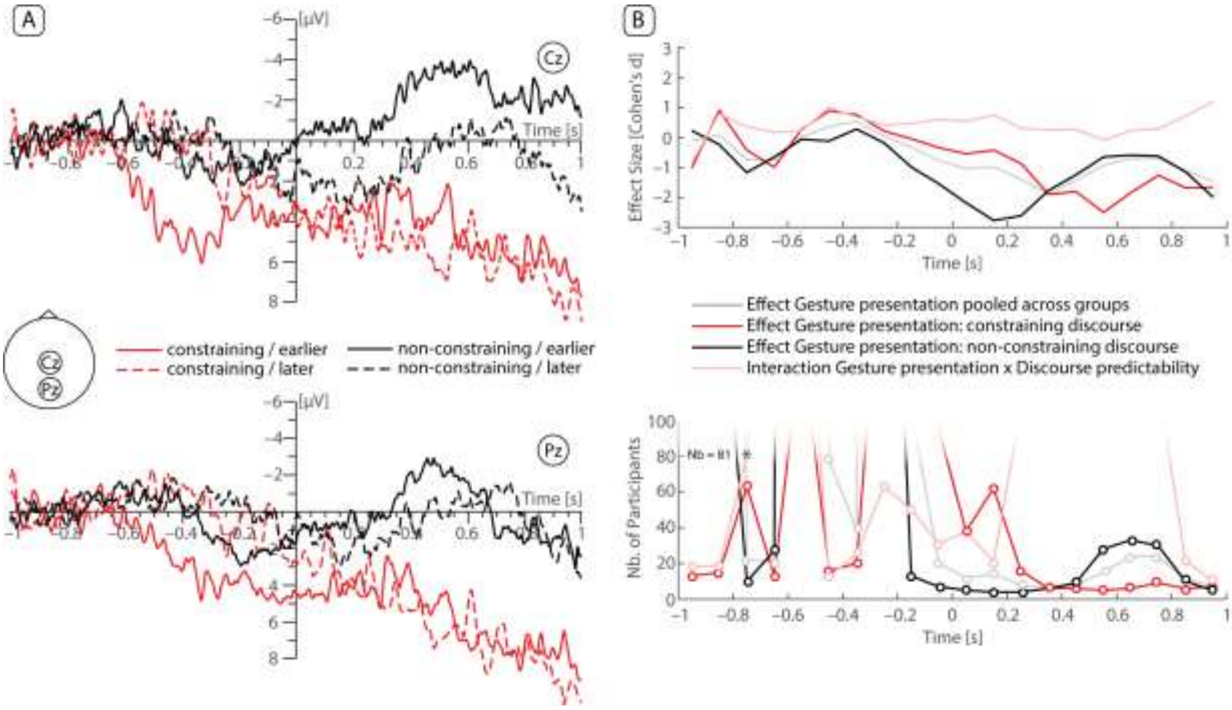

Figure 2: **Summary of EEG pilot data.** *Panel A* Grand-average ERPs plotted for the four experimental conditions based on eleven pilot participants at Cz and Pz electrodes. Time zero marks the onset of the spoken target word. Gesture stroke onset was 520 ms before target word onset (earlier, solid lines) and 130 ms before target word onset (later, dashed lines), respectively. *Panel B* Outcome of power analysis/simulation based on mixed design ('discourse predictability' as between-participants manipulation): Upper plot shows fluctuations in achieved effect sizes using the pooled voltages of Cz and Pz electrodes, based on the eleven pilot participants, for main effects and interactions over time (in 100 ms increments). Lower plot shows the required number of participants for the same main effects and interactions over time (i.e., 100 ms increments) to achieve a power of 0.9.

Participants

To determine the number of required participants for the proposed study, we first looked at similar previous studies and their effect sizes. Obermeier et al. (2011, Experiment 1) reported medium effect sizes ($f^2 = 0.38$ and $f^2 = 0.32$) in a within-participants design. Given their sample size of 32 participants and a significance level of $p < 0.05$, this means their effects achieved a power of 0.85 and 0.71, respectively. None of the other EEG studies relevant to the proposed experiment (e.g., Fritz et al., 2021; Habets et al., 2011; Obermeier & Gunter, 2015; Özyürek et al., 2007) reported effect sizes. The number of participants tested in these studies varied between 14 and 41.

We complemented our literature survey with a power analysis (according to Cohen, 1988) and a power simulation on the basis of our eleven pilot participants. In line with our proposed statistical approach, we used the pre-processed EEG time series data and fitted, for each of the eleven participants, at each channel and time sampling point a linear model using the Fieldtrip toolbox (Oostenveld et al., 2011). This model contained the baseline EEG activity as continuous predictor (Alday, 2019), gesture presentation as a fixed factor, and item (i.e., target word) as random effect (with a random intercept and a random slope by gesture presentation (syntax: EEGofInterest ~ BaselineEEG * GesturePresentation + (1 + GesturePresentation | TargetWord)). We obtained a time-resolved t-map of the contrast of earlier versus later gesture presentation (i.e., the within-participant manipulation).

In Figure 2B (upper panel), we plot effect sizes of the four effects of interest obtained by averaging t-values over 100 ms time bins and – for the purpose of the power calculation – over electrodes Cz and Pz (Kutas & Federmeier, 2011; Lau et al., 2008): (1) Main effect of gesture presentation (pooling t-values from constraining and non-constraining discourses, grey line), (2) Effect of gesture presentation in constraining discourses (red line), (3) Effect of gesture presentation in non-constraining discourses (black line), and (4) Interaction between gesture presentation and discourse predictability (pink line), obtained by subtracting the mean t-values of the non-constraining-discourse group from the mean t-values of the constraining-discourse group. Note that we selected Cz and Pz electrodes since these electrodes have previously been shown to respond to discourse manipulations in the context of predictive processing and to provide a 'proof-of-principle'. As described below, our main analysis will involve cluster-based permutation testing on all electrodes, using a more complex model fitted at the individual participant's level.

Based on these effect sizes, we calculated the number of participants required to detect the four effects with a statistical power of 0.9 for each 100-ms bin, using the package 'pwr' (Champely, 2020) in R (R Core Team). Taking financial and time resources into account (Lakens, 2022), we visualized numbers up to 100 participants in the lower panel of Figure 2B. The numbers varied considerably depending on the type of effect and time bin. Recall that critical for providing evidence for our main research question, whether or not iconic gestures contribute to predictive language comprehension, were interaction effects 2a and 2f. In line with our hypotheses, we therefore focused on the size of the interaction effect over time and the number of required participants to detect this effect with sufficient power. The numbers ranged between 13 and 81 for the pre-target epoch and between 10 and 38 for the post-target epoch. To ensure sufficient statistical power pre- *and* post-target, we will recruit 81 participants. This number of participants is sufficient – based on the pilot data – to detect well-powered interaction effects in twelve of the twenty time bins. In nine bins, 81 participants are twice as many as actually required on the basis the between-participants design.

As noted above, we revised our initial plans and decided to run the proposed experiment using a 2x2 *within*-participants design, with both discourse predictability and gesture presentation manipulated within-participants. In contrast to the power calculation reported above, the within-participants design will enable us to fit a linear model that incorporates the interaction between discourse and gesture timing at the individual participant's level. More generally, compared to a mixed design, a within-participants design increases statistical power by decreasing inter-individual variability (Brysbaert, 2019). Thus, 81 participants will be sufficient to yield well-powered interaction effects in the proposed study.

All participants will be native speakers of Dutch (raised in a monolingual household), right-handed, and recruited through the participant database 'SONA' of Radboud University. The participants will range in age between 18 and 35 years. They will be screened for neurologic impairments, traumata and neuroleptics, as well as for known hearing and language impairments. Only healthy participants without a history of neurological problems will be included. They will have normal or corrected to normal vision, and they will not be colour-blind. Participants must not have taken part in any of the pre-tests (see below). Participants will give written informed consent to take part in the experiment. The study has been approved by the Ethics Board of the Faculty of Social Sciences at Radboud University in compliance with the Declaration of Helsinki. Participants will consent to share their anonymized experimental data for research purposes, and participation will be compensated with 15 Euros or 1.5 ECTS.

Materials

The stimuli of the proposed study were based on the materials developed by Hintz et al. (submitted). The stimulus set consists of 80 concrete target nouns (mean Zipfian frequency = 3.92, SD = 0.90, range = 2.06 – 6.47, Keuleers & Brysbaert, 2010; mean prevalence = 0.99, SD = 0.02, range = 0.91 - 1, Keuleers et al., 2015), which were embedded in contexts. The contexts comprise short Dutch discourses consisting of two sentences, ending in the target noun. In 80 discourses, the target word can be predicted from the preceding context; in the remaining 80, the target word cannot be predicted. Each target word is paired with an iconic gesture, which is timed to have its stroke onset either 520 ms (‘earlier’ condition) or 130 ms (‘later’ condition) prior to target word onset, yielding a total of 320 unique stimuli (Figure 1, for an example).

Video recordings of the stimuli were made in the video recording laboratory of the Max Planck Institute for Psycholinguistics. A female native speaker of Dutch was videotaped while producing the spoken discourses using normal intonation and a regular speaking rate. Next to producing a discourse, she executed an iconic gesture that depicted the target noun. The speaker wore clothes in a neutral dark colour and stood in front of a unicolor curtain. She was positioned to be in the centre of the screen. At sentence onset, her arms were hanging casually by her sides. She was instructed to memorize and subsequently reproduce the stimulus. She was instructed to produce the gesture at a point in time that felt natural to her, always close to the target word, but no specific instructions on the timing were given (i.e., the actress was blind to the goal of the present study). At least three versions of each stimulus were recorded. From these three versions, the best recording was selected based on the naturalness of speech and gesture, consistency of speech and gesture across different conditions, and quality of the recording (e.g., absence of background noise, video recording artefacts, etc.). We used ELAN (Version 4.1.2, Wittenburg et al., 2006) to annotate the onset and offset of several events in the video: the target word, the gesture phrase (i.e., from the first to the last frame in which manual movement could be observed that belonged to a gesture), and the gesture stroke phase (the most meaning-bearing part of the gesture, Kita et al., 1998). The video recordings were further edited using Adobe After Effects© to add a mask blurring the speaker's face, such that facial movements and expressions were not visible. Finally, we used ffmpeg© to shift the video track of the stimuli recordings relative to the audio track such that the onset of the gesture stroke preceded the onset of the spoken target by either 520 ms or 130 ms in every stimulus video. The spoken target words are on average 647 ms long (SD = 172, range = 289 - 1052); gesture strokes are on average 668 ms long (SD = 237, range = 240 – 1320). In both gesture conditions, gesture strokes and target words overlapped. In the earlier-gesture condition, the overlap was on average 148 ms, ranging from -280 ms (no overlap) to 800 ms. In the later-gesture condition, gesture stroke and target word overlap on average 538 ms, ranging from 110 ms to 1190 ms.

The 320 videos were on average 8302 ms long (SD = 1169 ms, range = 6120 - 11952 ms). Target word onset occurred on average after 6703 ms (SD = 1155 ms, range = 4374 - 10336 ms) and was comparable across the predictable (M = 6715 ms, SD = 1168 ms, range = 4423 - 10336 ms) and the non-predictable (M = 6693 ms, SD = 1146 ms, range = 4374 - 10010 ms) conditions.

Rating studies

We conducted two web-based sentence completion studies to assess the cloze probability of the target words in the predictive and non-predictive discourses (Taylor, 1956). Moreover, a lab-based rating study was run to assess how unambiguously the iconic gestures represented the target nouns in the absence of speech. Additionally, we will measure the interpretability of all gestures in the two spoken discourses in which they appeared (predictable/non-predictable), since this may capture most closely how participants perceive and process them in the context of the multimodal stimulus presentations used in the experiment.

Both sentence completion studies were implemented in LimeSurvey (LimeSurvey GmbH). The participants read the discourses up until and including the determiner preceding the target word, and were instructed to fill in the word they thought would be the most likely continuation of the running sentence. Thirty participants took part in the first rating study involving 80 predictive contexts (22 female, M = 26.2 years, SD = 3.5 years, range = 21–33 years); another thirty-one participants took part in the second rating study involving 80 non-predictive contexts (18 female, M = 24.0 years, SD = 4.0 years, range = 18–33 years). Participants' responses were coded as 'match' in case the word in question was provided. In the case of a non-target response, the pairwise semantic distance to the target word was calculated using the Dutch version of Snaut[5] (Mandera, Keuleers & Brysbaert, 2017). The semantic

[5]Snaut was accessed through http://meshugga.ugent.be/snaut-dutch/.

distance values were then converted to similarity values by subtracting them from 1. Finally, the cloze probability for each target word was calculated by summing up ‘matches’ (value of 1) and similarity values for non-target responses (value between 0 and 1) and by dividing this sum by the number of participants who responded. For the predictable contexts, the average cloze probability was 0.85 (SD = 0.13, range = 0.51–1). For the non-predictable contexts, the average cloze probability was 0.

Thirty-two participants (23 female, M = 23 years, SD = 2.9 years, range = 19–31 years) took part in the laboratory-based iconic gesture interpretability rating study, which was implemented in Presentation (version 20.0; Neurobehavioral Systems Inc.). On each trial, the participants first saw a video recording of one of the 80 iconic gestures *without audio*. They were then asked to provide a maximum of three guesses (nouns) of what the gesture might denote (free entry format, interpretability measure). Finally, they were shown the target word and asked to rate the compatibility of the just-seen gesture and the target word it depicted, using a scale ranging from 1 (incompatible) to 7 (fully compatible, compatibility measure). On average, the probability of the target word being among the three words provided by participants was 0.38 (SD = 0.32, range = 0 – 1); the mean probability of the target word being the first guess was 0.30 (SD = 0.30, range = 0 – 1). The average compatibility rating was 5.16 (SD = 1.24, range = 1.75 – 7), indicating good compatibility. As it is a more unbiased measure than the compatibility rating (independent of any cues), we calculated the adjusted probability of the target word being the participants’ first guess, while using similarity values for non-target responses in the same way as was done for the two sentence completion tasks. The average adjusted probability was 0.49 (SD = 0.25, range = 0.05–1).

In sum, the three rating studies confirmed the suitability of the stimuli for the purposes of the present study: The cloze probability studies demonstrated that predictable and non-predictable items were classified appropriately. The gesture rating study demonstrated that the iconic gestures we selected to embody the target words are – with some variability – well interpretable when presented on their own. However, since in the proposed study the gestures will be presented in context, we will additionally assess the interpretability of the iconic gestures within the spoken discourses in which they appear. To that end, we will conduct an additional rating study in parallel to running the main EEG experiment. As for the previous gesture rating study, we will invite 30 Dutch participants (none of which will have participated in the main experiment or any of the rating studies). The rating study will be set up such that each participant will be presented with predictable and non-predictable versions (in that order) of a target word-iconic gesture pairing. We will rotate across gesture presentation conditions (early vs. late) using two experimental lists. Participants will be presented with predictable and non-predictable discourses up until the frame preceding the onset of the gesture retraction (which does not contain any semantic information). In both conditions, participants will be asked to provide up to three words (nouns) that they feel best denote the meaning depicted by the gesture. The data from that rating study will be analysed in the same way as those from the other rating study (i.e., taking into account semantic distance between the provided responses and the target words).

Experimental design and lists

To re-cap, the proposed experiment will use a 2 (gesture presentation: earlier vs. later) x 2 (discourse predictability: constraining vs. non-constraining) within-participants design. We will present each participant with all predictable items, consisting of 40 earlier gesture and 40 later gesture trials, first. Following the predictable block, the same target word-iconic gesture pairings will be presented in non-predictable discourses. The reason that motivated presenting the predictable block first was that a recent encounter of target word-gesture pairings may influence processing of a subsequent predictable discourse, but may not or to a much lesser extent influence processing of a subsequent non-predictable discourse. That is, on comprehending a predictable discourse, lingering activation of the target concept from the recent encounter may be boosted to full activation much earlier than would have been the case without a repetition. Non-predictable discourses, on the other hand, are very unlikely (according to our cloze probability norms) to benefit from a previous encounter of the target concept since the two sentences preceding the target word do not contain information that contributes to (further) activation of the target concept. We may thus assume that comprehending non-predictable discourses following an item repetition is comparable to comprehending the same non-predictable discourse without a repetition. Half of the trials on each block will belong to the earlier-gesture condition; the other half will be later gesture trials.

Within a participant, we will present earlier and later gesture presentation versions of an item, each paired with one of the discourse conditions. Furthermore, we will set up two experimental lists for rotating items across the gesture presentation conditions. On the basis of the two resulting experimental lists, we will create 80 pseudo-randomized versions (one for each participant; 40 for each experimental list) prior to testing using the program 'Mix' (van Casteren & Davis, 2006). The pseudo-randomization will be constrained by allowing a maximum of three repetitions of the same gesture stroke onset (i.e., earlier or later gesture trials). The 160 experimental trials on each list will be preceded by two practice trials.

Procedure

Following the general informed consent procedure, participants will be fitted with an EEG cap. During EEG recording, participants will be seated in front of a computer monitor, with speakers placed on either side. Participants will be seated in a sound-attenuating and electrically shielded booth. They will receive the following instructions (translated from Dutch): "In the present experiment, we investigate how people comprehend discourse. It is therefore important that you are concentrated during the experiment and always pay attention to what is being communicated. On each trial, a small story is told, consisting of two sentences. Try not to blink your eyes or move your head too much while the speaker communicates (if you need to blink please do so in between trials). Sometimes we ask you to indicate if you saw a red * in the middle of the screen at the end of the story. Press one of the two buttons (Yes or No) to answer the question. The next trial starts again in the same way."

The stimuli will be presented full screen on a 23-inch monitor operating at a 1920 x 1080 pixels native resolution, using Presentation software (version 20.0; Neurobehavioral Systems, Inc.). Forty of the 160 experimental trials (20 per block, appearing in a pseudo-random order with variable intervals in between) will be followed by a yes/no question to ensure that participants are looking at the computer screen during the experiment. On such a catch trial, a red asterisk will be presented in the centre of the screen, on top of the video, 200 ms after the offset of the spoken target. The asterisk will be presented for 500 ms. Participants will be instructed to indicate whether they saw the red asterisk or not by pressing the ‘Z’ key on the keyboard to provide a no-response or by pressing the ‘M’ key to provide a yes-response. The asterisk will be presented on half of the catch trials. We opted for a visual rather than a language comprehension task to ensure that participants are looking at the screen. In principle, a language comprehension task can be carried out without looking at the screen. After every 20 trials, participants will be able to take a short, self-timed break before continuing the experiment.

EEG data recording

Participants’ EEG will be recorded throughout the whole test session. The EEG signal will be recorded from 27 active scalp electrodes (Fz, FCz, Cz, Pz, Oz, F3/4, F7/8, FC1/2, FC5/6, C3/4, CP1/2, CP5/6, T7/8, P3/4, P7/8, O1/2), which will be mounted in an elastic cap (ActiCAP), and will be placed according to the 10-20 convention. The EEG signal will be recorded with an online reference to the left mastoid. Additionally, activity will be recorded at the right mastoid and at four bipolar electrooculogram (EOG) channels (two horizontal and 2 vertical). The ground electrode will be located on the forehead. The impedance for all active electrodes will be kept below 5kΩ, and triggers will be time-locked to both gesture stroke onset and target word onset. EEG signals will be recorded using BrainVision Recorder software (version 1.20.0401; Brain Products GmbH), at a sampling rate of 500 Hz, using a time constant of 8 s (0.02 Hz) and high cut-off of 100 Hz in the hardware filter.

Data pre-processing

We will include participants whose accuracy on the catch trials is 80% (32 out of 40 correct responses) or higher. We will replace excluded participants to ensure sufficient statistical power.

The pre-processing of the EEG data will be performed using BrainVision Analyzer (version 2.2) and will involve five main steps (cf. Paul et al., 2021): re-referencing, filtering, segmentation, ocular correction and artefact rejection. First, the raw data will be inspected to identify insufficient EEG signals (e.g., electrodes that show poor signal due to large-amplitude artefacts or deficient connectivity for at least half of the experiment). Deficient channels with a voltage > 2 SD of the EEG voltage

(Duma et al., 2019) will be interpolated through spherical splines. A maximum of four EEG channels will be interpolated and a summary of the number of interpolated channels will be reported (cf. Nieuwland & Arkhipova, 2020). Should the number of deficient channels in a participant exceed four, the data from that participant will be excluded from further analysis. Next, the data will be re-referenced to the average of both mastoid channels. Then, the data will be filtered using a Butterworth IIR filter with a low cut-off of .01 Hz and a high cut-off of 30 Hz. In the next step, the continuous data will be segmented into epochs. We will create one epoch ranging from - 1000 to 0 ms relative to the target word onset and one epoch ranging from 0 to 1000 ms relative to target word onset. Both will be screened for eye movements, large muscle artefacts, electrode drifts, and amplifier blocking. Ocular correction using the Gratton and Coles (1983) method will be used to detect and straighten out artefacts in the EOG channels, such as blinks and other eye movements. Then, a semiautomatic artefact rejection will be performed. To that end, BrainVision Analyzer will highlight trials with channels whose values exceed ±100 μV. These trials will be inspected and rejected on an individual base. Following artefact rejection, participants must retain ¾ of trials in each condition (60 out of 80, at least 120 trials overall), otherwise their data will be excluded. The average number of removed trials per condition (and SD) will be reported. In the penultimate step, mastoid and EOG channels will be removed.

Following these pre-processing steps, we will extract baseline EEG activity on each electrode during a two-hundred millisecond window (-1200 to -1000 ms relative to target word onset)—a point in time where in earlier and later gesture conditions, participants have not seen the gesture stroke yet. This baseline time window does overlap with the gesture preparation phase in most cases. While the stroke phase is deemed to be the meaning bearing part of the gesture (McNeill, 1992), the preparation phase could of course already contain some semantic information. However, if this was the case, then the discrepancy between baseline and effect measurement window in the gesture condition would become smaller. Thus, if anything, overlap of the baseline window with the gesture preparation phase would reduce the likelihood that we find the predicted effect rather than facilitate hypothesis-conform results. Since choosing an even earlier time window as baseline would have led to differences in the speech stimuli, we opted for the -1200 to -1000 ms time window prior to target word onset.

Statistical Analysis

Due to the lack of relevant prior research, the spatio-temporal locus of our effects – especially for the period preceding target word onset – is unknown. Therefore, we need a statistical technique that deals with our prior ignorance. Cluster-based permutation (CBP) analysis is a non-parametric randomization technique that identifies clusters of significant differences between conditions in time and space while minimizing the multiple-comparisons problem (Maris & Oostenveld, 2007). This approach allows for analysing the data without selecting a priori time windows and/or sets of electrodes. Since its inception in 2007, the technique has been considerably improved. One improvement is that CBP can now incorporate random-effects structures (e.g., to account for idiosyncratic by-participant and by-item variance), continuous predictors, interaction effects between conditions and between conditions and continuous predictors. One implementation of such an approach is referred to as 'massive univariate general linear analyses' (Pernet et al., 2011). In such analyses, the data are analysed using a hierarchical general linear model (GLM) where parameters of the GLM are estimated for each participant at each time point and each electrode independently (referred to as '1st level analyses'). The estimated parameters from the first level analyses are then integrated into CBP analyses across all participants (referred to as '2nd level analyses'). We recently applied variants of this approach in two ERP projects (Hintz et al., submitted; Strauß et al., 2022).

Here, we will fit linear models at each channel and time sampling point for each individual (1st level analysis), estimating the parameters for the within-participants conditions (discourse and gesture timing, henceforth referred to as 'fixed effects') and baseline EEG activity. On the second level, we will apply cluster-based permutation testing across all participants. The two-stage approach will be applied separately to the two extracted epochs (pre-target: -1000 ms to 0 ms relative to target word onset; post-target: 0 ms to 1000 ms relative to target word onset). The pre-target period includes the earlier and later gesture presentation; the post-target epoch includes the unfolding of the spoken target.

To be specific, on the first level, we will fit a linear model comprising discourse predictability (predictable = 0.5 vs. non-predictable = - 0.5) and gesture presentation (early = 0.5 vs. late = -0.5) as difference-coded fixed effects and baseline EEG activity (Alday, 2019) as a continuous predictor (scaled and centred). Interaction effects between all predictors will be added. The model will further contain item as random effect (i.e., target)—with a random intercept and a random slope for the interaction of both fixed factors (cf. Barr et al., 2013). Below we list the model formula to be used:

EEGofInterest ~ DiscoursePredictability * GesturePresentation * BaselineEEG + (1 + Discourse * GesturePresentation | Target)

On the second level, we will submit z-scored t-maps, that have been estimated for the main effects and interactions of interest, to a cluster search. By using Monte-Carlo non-parametrical permutation testing (1000 randomizations) as implemented in Fieldtrip (Oostenveld et al. 2011), type I-error controlled cluster significance probabilities ($\alpha = 0.025$) will be estimated. The minimum number of neighbours will be set to 1 to reduce the risk that two spatially distinct electrodes are considered one cluster by the algorithm. Note that we are using a relatively sparse montage of 32 electrodes, which – in our experience – already substantially reduces the risk of merging spatially distinct electrodes. The minimum size of a cluster will remain unconstrained. To quantify their stability, we will – for each significant effect (including post-hoc comparisons of interaction effects) – report the effect size and the Bayes Factor (Jeffreys, 1961).

The permutation method will be used to test for main effects of discourse predictability after target word onset (hypothesised effect 1) by comparing the respective z-transformed t-values of all participants against zero in a two-tailed dependent samples t-test.

To test for main effects of gesture presentation pre- and post-target onset (hypothesised effects 2d and 2e), z-transformed t-values of the respective main effects of all participants will be compared against zero in a two-tailed dependent samples t-test.

To test for interactions between gesture presentation and discourse predictability, a two-tailed dependent samples t-test comparing z-transformed t-values of the interaction against zero will be used. We will then extract the raw data from significant clusters (i.e., time and sets of electrodes) and submit these to planned post-hoc t-tests (hypothesised effects 2a, 2b, 2c and 2f) to test for differences between the conditions.

Exploring the moderating influence of gesture interpretability (hypothesised effect 3)

To explore whether potential main effects of discourse predictability and gesture presentation, and their interaction are influenced by how interpretable the iconic gestures are perceived in the respective sentence contexts, we will repeat the first-level and the second-level analyses incorporating the to-be-collected gesture interpretability rating. Specifically, on the first level, the new rating measure will be added as a continuous predictor (scaled and centred), interacting with all other predictor terms. The model formula will be:

EEGofInterest ~ DiscoursePredictability * GesturePresentation * GestureInterpretability * BaselineEEG + (1 + Discourse * GesturePresentation | Target)

On the second level, we will run cluster searches for (1) the interaction between discourse and gesture interpretability, for (2) the interaction between gesture presentation and gesture interpretability, and (3) for the three-way interaction between discourse predictability, gesture presentation and gesture interpretability. We will discuss the results of this planned exploratory analysis in relation to the two-stage analysis described above, without the gesture interpretability rating.

References

Alday, P. M. (2019). How much baseline correction do we need in ERP research? Extended GLM model can replace baseline correction while lifting its limits. *Psychophysiology, 56*(12), e13451.

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

Bañón, J. A., & Martin, C. (2019). Anticipating information structure: An event-related potentials study of focus assignment via the it-cleft. *Neuropsychologia*, *134*, 107203.

Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends in Cognitive Sciences, 11*(7), 280-289.

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

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.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1e48.

Bavelas, J. B., & Chovil, N. (2000). *Visible acts of meaning: An integrated message model of language in face-to-face dialogue. 19*(2), 163–194.

Beattie, G., & Shovelton, H. (1999). Do iconic hand gestures really contribute anything to the semantic information conveyed by speech? An experimental investigation. *Semiotica 123*(1-2), 1-30.

Bendixen, A., Schröger, E., & Winkler, I. (2009). I heard that coming: event-related potential evidence for stimulus-driven prediction in the auditory system. *Journal of Neuroscience, 29*(26), 8447-8451.

Bergmann, K., Aksu, V., & Kopp, S. (2011). The relation of speech and gestures: Temporal synchrony follows semantic synchrony. In *Proceedings of the 2nd Workshop on Gesture and Speech in Interaction (GeSpIn 2011).*

Butterworth, B., & Beattie, G. (1978). Gesture and silence as indicators of planning in speech. *Recent advances in the psychology of language* (pp. 347-360). Springer, Boston, MA.

Brysbaert, M. (2019). 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.

Champely, S. (2020). pwr: Basic Functions for Power Analysis. R package version 1.3-0. 2020.

Chui, K. (2005). Temporal patterning of speech and iconic gestures in conversational discourse. *Journal of Pragmatics, 37*(6), 871-887.

Church, R. B., Kelly, S., & Holcombe, D. (2014). Temporal synchrony between speech, action and gesture during language production. *Language, Cognition and Neuroscience, 29*(3), 345-354.

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

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Academic press.

De Kok, I., Hough, J., Schlangen, D., & Kopp, S. (2016). Deictic gestures in coaching interactions. In *Proceedings of the Workshop on Multimodal Analyses Enabling Artificial Agents in Human-Machine Interaction - MA3HMI '16*, 10–14.

De Lange, F. P., Heilbron, M., & Kok, P. (2018). How do expectations shape perception?. *Trends in Cognitive Sciences*, 22(9), 764-779.

Dell, G. S., & Chang, F. (2014). The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1634), 20120394.

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.

Dick, A. S., Goldin-Meadow, S., Hasson, U., Skipper, J. I., & Small, S. L. (2009). Co-speech gestures influence neural activity in brain regions associated with processing semantic information. *Human Brain Mapping, 30*(11), 3509-3526.

Drijvers, L., & Özyürek, A. (2018). Native language status of the listener modulates the neural integration of speech and iconic gestures in clear and adverse listening conditions. *Brain and Language 177-178*, 7-17.

Duma, G. M., Mento, G., Semenzato, L., & Tressoldi, P. E. (2019). EEG anticipation of random high and low arousal faces and sounds. *F1000Research, 8*(1508), 1508.

Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language, 41*(4), 469-495.

Ferré, G. (2010). *Timing relationships between speech and co-verbal gestures in spontaneous French*. In *Language Resources and Evaluation, Workshop on Multimodal Corpora* (Vol. 6, pp. 86-91).

Ferreira, F., & Chantavarin, S. (2018). Integration and prediction in language processing: A synthesis of old and new. *Current Directions in Psychological Science, 27*(6), 443-448.

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

Fritz, I., Kita, S., Littlemore, J., & Krott, A. (2021). Multimodal language processing: How preceding discourse constrains gesture interpretation and affects gesture integration when gestures do not synchronise with semantic affiliates. *Journal of Memory and Language, 117*, 104191.

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography & Clinical Neurophysiology, 55*(4), 468– 484.

Green, A., Straube, B., Weis, S., Jansen, A., Willmes, K., Konrad, K., & Kircher, T. (2009). Neural integration of iconic and unrelated coverbal gestures: A functional MRI study. *Human Brain Mapping, 30(10)*, 3309-3324.

Habets, B., Kita, S., Shao, Z., Özyurek, A., & Hagoort, P. (2011). The Role of Synchrony and Ambiguity in Speech–Gesture Integration during Comprehension. *Journal of Cognitive Neuroscience*, *23*(8), 1845–1854.

He, Y., Steines, M., Sommer, J., Gebhardt, H., Nagels, A., Sammer, G., Kircher, T., & Straube, B. (2018). Spatial–temporal dynamics of gesture–speech integration: a simultaneous EEG-fMRI study. *Brain Structure and Function, 223*(7), 3073-3089.

Hintz, F., Khoe, Y.H., Strauß, A., Psomakas, A.J.A., & Holler (submitted). Electrophysiological evidence for the enhancement of gesture-speech integration by linguistic predictability during multimodal discourse comprehension.

Hintz, F., Meyer, A. S., & Huettig, F. (2020). Activating words beyond the unfolding sentence: Contributions of event simulation and word associations to discourse reading. *Neuropsychologia*, 141: 107409.

Holle, H., & Gunter, T. C. (2007). The Role of Iconic Gestures in Speech Disambiguation: ERP Evidence. *Journal of Cognitive Neuroscience, 19*(7), 1175–1192.

Holle, H., Gunter, T. C., Rüschemeyer, S. A., Hennenlotter, A., & Iacoboni, M. (2008). Neural correlates of the processing of co-speech gestures. *Neuroimage, 39*(4), 2010- 2024.

Holler, J., & Levinson, S. C. (2019). Multimodal Language Processing in Human Communication. *Trends in Cognitive Sciences*, *23*(8), 639–652.

Holler, J., Shovelton, H., & Beattie, G. (2009). Do iconic hand gestures really contribute to the communication of semantic information in a face-to-face context?. *Journal of Nonverbal Behavior*, *33*(2), 73-88.

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin*, *137*(2), 297.

Huettig, F., Audring, J., & Jackendoff, R. (2022). A parallel architecture perspective on pre-activation and prediction in language processing. *Cognition*, *224*: 105050.

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

Kelly, S.D., Barr, D., Church R.B. & Lynch K. (1999). Offering a hand to pragmatic understanding: the role of speech and gesture in comprehension and memory. *Journal of Memory and. Language. 40*, 577– 592.

Kelly, S. D., Kravitz, C., & Hopkins, M. (2004). Neural correlates of bimodal speech and gesture comprehension. *Brain and Language*, *89*(1), 253–260.

Kelly, S. D., Özyürek, A., & Maris, E. (2010). Two Sides of the Same Coin: Speech and Gesture Mutually Interact to Enhance Comprehension. *Psychological Science*, *21*(2), 260–267.

Kendon, A. (1980). Gesticulation and speech: Two aspects of the Process of Utterance. , (25), 207. In M. Richie Key (Ed.), *The relationship of verbal and nonverbal communication* (pp. 207–227). Mouton.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge University Press.

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, *42*(3), 627-633.

Keuleers, E., Stevens, M., Mandera, P., & Brysbaert, M. (2015). Word knowledge in the crowd: Measuring vocabulary size and word prevalence in a massive online experiment. *Quarterly Journal of Experimental Psychology*, *68*(8), 1665-1692.

Jeffreys, H. (1961). *Theory of Probability* (3rd ed.). Oxford: Oxford University Press.

Kita, S., van Gijn, I., & van der Hulst, H. (1998). Movement phases in signs and co-speech gestures, and their transcription by human coders. In I. Wachsmuth & M. Fröhlich (Eds.), *Gesture and Sign Language in Human-Computer Interaction* (Vol. 1371, pp. 23– 35). Springer Berlin Heidelberg.

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

Kutas, M., DeLong, K. A., & Smith, N. J. (2011). A look around at what lies ahead: Prediction and predictability in language processing. In M. Bar (Ed.),*Predictions in the brain: Using our past to predict the future* (pp. 190–207). New York, NY: Oxford University Press.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62,* 621-647.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*(12), 463-470.

Kutas, M., & Hillyard, S. A. (1984). Event-Related Brain Potentials (ERPs) Elicited by Novel Stimuli during Sentence Processing. *Annals of the New York Academy of Sciences*, *425*(1), 236-241.

Lakens, D. (2022). Sample Size Justification. *Collabra: Psychology*, *8*(1). https://doi.org/10.1525/collabra.33267

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience*, *9*(12), 920-933.

Levinson, S. C. (2006). On the human "interaction engine". In N. J. Enfield, & S. C. Levinson (Eds.), *Roots of human sociality: Culture, cognition and interaction* (pp. 39-69). Oxford: Berg.

Levinson, S. C., & Holler, J. (2014). The origin of human multimodal communication. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1651), 20130302.

Levinson, S. C., & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. *Frontiers in psychology*, *6*, 731.

Mandera, P., Keuleers, E., & Brysbaert, M. (2017). Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language*, *92*, 57– 78.

Mantegna, F., Hintz, F., Ostarek, M., Alday, P. M., & Huettig, F. (2019). Distinguishing integration and prediction accounts of ERP N400 modulations in language processing through experimental design. *Neuropsychologia*, *134*, 107199.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG- data. *Journal of Neuroscience Methods, 164*(1), 177-190.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago press.

McNeill, D., Cassell, J., & McCullough, K. E. (1994). Communicative effects of speech-mismatched gestures. *Research on language and social interaction*, *27*(3), 223-237.

Metusalem, R., Kutas, M., Urbach, T. P., Hare, M., McRae, K., & Elman, J. L. (2012). Generalized event knowledge activation during online sentence comprehension. *Journal of Memory and Language*, *66*(4), 545-567.

Morrel-Samuels, P., & Krauss, R. M. (1992). Word familiarity predicts temporal asynchrony of hand gestures and speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(3), 615.

Nagels, A., Kircher, T., Steines, M., & Straube, B. (2015). Feeling addressed! The role of body orientation and co-speech gesture in social communication. *Human Brain Mapping, 36(5),* 1925-1936.

Nicenboim, B., Vasishth, S., & Rösler, F. (2020). Are words pre-activated probabilistically during sentence comprehension? Evidence from new data and a Bayesian random-effects meta-analysis using publicly available data. *Neuropsychologia*, *142*, 107427.

Nieuwland, M. S., Arkhipova, Y., & Rodríguez-Gómez, P. (2020). Anticipating words during spoken discourse comprehension: A large-scale, pre-registered replication study using brain potentials. *Cortex*, *133*, 1-36.

Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., Von Grebmer Zu Wolfsthurn, S., Bartolozzi, F., Kogan, V., Ito, A., Mézière, D., Barr, D. J., Rousselet, G. A., Ferguson, H. J., Busch-Moreno, S., Fu, X., Tuomainen, J., Kulakova, E., Husband, E. M., … Huettig, F. (2018). Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. *ELife*, *7*, e33468.

Nieuwland, M. S., Barr, D. J., Bartolozzi, F., Busch-Moreno, S., Darley, E., Donaldson, D. I., ... & Von Grebmer Zu Wolfsthurn, S. (2020). Dissociable effects of prediction and integration during language comprehension: Evidence from a large-scale study using brain potentials. *Philosophical Transactions of the Royal Society B*, *375*(1791), 20180522.

Obermeier, C., Holle, H., & Gunter, T. C. (2011). What Iconic Gesture Fragments Reveal about Gesture–Speech Integration: When Synchrony Is Lost, Memory Can Help. *Journal of Cognitive Neuroscience*, *23*(7), 1648–1663.

Obermeier, C., & Gunter, T. C. (2015). Multisensory Integration: The Case of a Time Window of Gesture–Speech Integration. *Journal of Cognitive Neuroscience*, *27*(2), 292– 307.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.*, 156869.

Ortega, G., & Özyürek, A. (2016). Generalisable patterns of gesture distinguish semantic categories in communication without language: Evidence from pantomime. Talk presented at the 7th Conference of the International Society for Gesture Studies (ISGS7). Paris, France.

Ortega, G., Özyürek, A. Systematic mappings between semantic categories and types of iconic representations in the manual modality: A normed database of silent gesture. *Behavior Research Methods 52*, 51–67 (2020).

Ortega, G., Schiefner, A., & Özyürek, A. (2019). Hearing non-signers use their gestures to predict iconic form-meaning mappings at first exposure to signs. *Cognition*, *191*, 103996.

Özyürek, A. (2014). Hearing and seeing meaning in speech and gesture: Insights from brain and behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1651), 20130296.

Özyürek, A., Willems, R. M., Kita, S., & Hagoort, P. (2007). On-line Integration of Semantic Information from Speech and Gesture: Insights from Event-related Brain Potentials. *Journal of Cognitive Neuroscience, 19*(4), 605–616.

Paul, M., Govaart, G. H., & Schettino, A. (2021). Making ERP research more transparent: Guidelines for preregistration. *International Journal of Psychophysiology, 164*, 52–63.

Pernet, C. R., Chauveau, N., Gaspar, C. & Rousselet, G. A. (2011). LIMO EEG: a toolbox for hierarchical linear modeling of electroencephalographic data. *Comput. Intell. Neurosci.*, 831409.

Perniss, P. (2015). Collecting and analyzing sign language data: Video requirements and use of annotation software. *Research methods in sign language studies: A practical guide*, 55-74.

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

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36*(04), 329-347.

Rommers, J., Meyer, A. S., Praamstra, P., & Huettig, F. (2013). The contents of predictions in sentence comprehension: Activation of the shape of objects before they are referred to. *Neuropsychologia, 51*(3), 437-447.

Schegloff, E. A. (1984). *On some gesture's relation to talk*. Structures of social action: Studies in conversation analysis, 266-296.

Schröger, E. (2007). Mismatch negativity: A microphone into auditory memory. *Journal of Psychophysiology*, *21*(3-4), 138.

Sekine, K., Schoechl, C., Mulder, K., Holler, J., Kelly, S., Furman, R., & Özyürek, A. (2020). Evidence for children's online integration of simultaneous information from speech and iconic gestures: An ERP study. *Language, Cognition and Neuroscience*. Advance online publication.

Skipper, J. I., Goldin-Meadow, S., Nusbaum, H. C., & Small, S. L. (2009). Gestures orchestrate brain networks for language understanding. *Current Biology, 19*(8), 661-667.

Strauß, A., Wu, T., McQueen, J. M., Scharenborg, O., & Hintz, F. (2022). The differential roles of lexical and sublexical processing during spoken-word recognition in clear and in noise. *Cortex, 151*, 70-88.

Taylor, W. L. (1956). "Cloze procedure": A new tool for measuring readability. *Journalism Quarterly, 30*(4), 415-433.

Ter Bekke, M., Drijvers, L., & Holler, J. (2020). The predictive potential of hand gestures during conversation: An investigation of the timing of gestures in relation to speech [Preprint]. *PsyArXiv*. https://doi.org/10.31234/osf.io/b5zq7.

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.

Van Casteren, M. & Davis, M. H. (2006). Mix, a program for pseudorandomization. *Behavior Research Methods*, *38*(4), 584–589.

Van Nispen, K., van de Sandt-Koenderman, W. M., & Krahmer, E. (2017). Production and comprehension of pantomimes used to depict objects. *Frontiers in Psychology*, *8*, 1095.

Wicha, N. Y., Moreno, E. M., & Kutas, M. (2003). Expecting gender: An event related brain potential study on the role of grammatical gender in comprehending a line drawing within a written sentence in Spanish. *Cortex, 39*(3), 483-508.

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.

Willems, R. M., Özyürek, A., & Hagoort, P. (2007). When Language Meets Action: The Neural Integration of Gesture and Speech. *Cerebral Cortex*, *17*(10), 2322–2333.

Willems, R. M., Özyürek, A., & Hagoort, P. (2009). Differential roles for left inferior frontal and superior temporal cortex in multimodal integration of action and language. *Neuroimage 47*(4), 1992-2004.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: a professional framework for multimodality research. *5th International Conference on Language Resources and Evaluation (LREC 2006)* (pp. 1556-1559).

Wu, Y. C., & Coulson, S. (2005). Meaningful gestures: Electrophysiological indices of iconic gesture comprehension. *Psychophysiology*, *42*(6), 654–667.

Wu, Y. C., & Coulson, S. (2007). Iconic gestures prime related concepts: An ERP study. *Psychonomic Bulletin & Review, 14*(1), 57–63.

Wu, Y. C., & Coulson, S. (2010). Gestures modulate speech processing early in utterances. *Neuroreport, 21*(7), 522-526.

Wu, Y. C., & Coulson, S. (2015). Iconic gestures facilitate discourse comprehension in individuals with superior immediate memory for body configurations. *Psychological Science, 26*(11), 1717-1727.

Appendix: Materials

Predictive sentence context	Non-predictive sentence context	Target word	Cloze probability	Silent gesture interpretability
In de dierentuin is het oude verblijf van Kiki omgetoverd tot een waar klimparadijs. Iedereen is namelijk dol op de *Kiki's old residence in the zoo has been transformed into a true climbing paradise. Everybody loves the*	Wendy is het populairste meisje van de klas. Ze heeft de coolste stickers van de *Wendy is the most popular girl in the class. She has the coolest stickers of the*	aap *(monkey)*	0.75	0.60
De dochters van de kermisexploitant vinden spannende attracties erg leuk. Het liefst spenderen ze de hele dag in de *The daughters of the fairground operator love exciting attractions. They prefer to spend the whole day in the*	Gerda is helderziend en heeft geregeld een visioen. Toen ze in de bus stapte zag ze ineens de *Gerda is a psychic and regularly has a vision. When she got on the bus she suddenly saw the*	achtbaan *(rollercoaster)*	0.62	0.58
Jos had het helemaal gehad om met de trein te reizen. Hij bedacht zich geen moment en kocht een *Jos had enough of travelling by train. He didn't hesitate for another moment and bought a*	Samen keken ze voor de 10e keer naar deze film. Hun favoriete scene was die met een *They watched this movie together for the 10th time. Their favorite scene was the one with a*	auto *(car)*	0.94	0.75
Pieter is op kraamvisite bij zijn beste vriendin. In zijn armen houdt hij de *Pieter is on a maternity visit with his best friend. In his arms he holds the*	Hij kon zijn aandacht niet bij de les houden. Al snel was hij aan het dagdromen over de *He couldn't focus his attention on the lesson. Soon he was daydreaming about the*	baby *(baby)*	0.98	0.96
Het rommelhok was door een stylist omgetoverd tot een logeerkamer. In de hoek van de kamer stond nu een	Iedereen weet dat uitstel leidt tot afstel. Patricia had dan ook eerder moeten beginnen aan haar verslag over een	bed	0.88	0.88

The messy room had been turned into a guest room by a stylist. In the corner of the room was now a	*Everyone knows that you shouldn't put off until tomorrow what you can do today. Patricia should have started earlier on her report about a*	*(bed)*		
Als het warm weer is gaat Maarten graag naar het terras met zijn vrienden. Hij geniet dan nog meer van een *When the weather is warm Maarten likes to go to the café with his friends. There he enjoys a*	Het dagje uit van de familie Jansen ging niet helemaal zoals gedacht. Vader had alleen maar aandacht voor een *The day out of the Jansen family didn't go quite as planned. Father was only paying attention to a*	biertje *(beer)*	0.9	0.51
Het was een koude winter en er moest nieuw hout worden klaargemaakt voor de open haard. Vader ging naar de schuur en pakte de *It was a cold winter and new wood had to be prepared for the fireplace. Father went to the barn and grabbed the*	Maarten staat bij de kassa in de bouwmarkt. De mensen achter hem blijven het maar hebben over de *Maarten is at the checkout in the DIY store. The people behind him keep talking about the*	bijl *(axe)*	0.85	0.38
Het werd buiten steeds donkerder en het begon hard te waaien. Niet lang daarna zag Tom een *It was getting darker and darker outside and it started blowing hard. Not long after, Tom saw a*	Het werd helemaal stil in de volle kamer. Iedereen keek naar een *It got completely quiet in the crowded room. Everybody was looking at a*	bliksemschicht *(lightning bolt)*	0.63	0.36
Nienke ziet in het park een grote struik in bloei staan. Glimlachend ruikt ze aan de *Nienke sees a large blooming bush in the park. Smiling she smells the*	Gelukkig had Arnold zijn camera meegenomen. Hij maakte meer dan tien foto's van de *Luckily Arnold had brought his camera. He took more than ten pictures of the*	bloem *(flower)*	0.87	0.21
Johan loopt 's ochtends de keuken in. Voor zijn ontbijt eet hij graag een *Johan walks into the kitchen in the morning. For his breakfast he likes to eat a*	Rudolf is een heel belangrijk examen aan het maken. Hij raakte in de war bij de vraag over een *Rudolf is taking a very important exam. He got confused when he was asked about a*	boterham *(sandwich)*	0.66	0.60

De laatste tijd heeft opa wat moeite met het lezen van kleine lettertjes. Gelukkig heeft hij nu een *Lately grandpa has been having some trouble reading small print. Luckily he now has*	Toen hij het tijdschrift weg wilde gooien, hield zijn vriendin hem tegen. Ze wilde nog het stuk lezen over een *When he wanted to throw away the magazine, his girlfriend stopped him. She still wanted to read the article about*	bril *(glasses)*	0.68	0.81
De familie Kuiper kwam verschillende planten tegen tijdens hun tocht door de woestijn. Toen ze even niet opletten, bezeerde hun dochter zich aan een *The Kuiper family came across several plants during their journey through the desert. When they didn't pay attention for a moment, their daughter hurt herself on a*	De buurman van Nadia is op excursie geweest. Ze luistert aandachtig als hij vertelt over een *Nadia's neighbor has been on an excursion. She listens attentively when he tells about a*	cactus *(cactus)*	0.97	0.09
Op 1 januari vindt altijd een groot klassiek concert plaats in het centrum. Het orkest dat daar speelt wordt geleid door de *On the 1st of January there is always a big classical concert in the city centre. The orchestra that plays there is conducted by the*	Julia heeft nu langs alle zenders gezapt maar er is niets wat haar leuk lijkt. Uit verveling zet ze toch maar het programma op over de *Julia's been zapping all the stations now, but there's nothing she likes. Out of boredom she puts on the program about the*	dirigent *(conductor)*	0.97	0.58
Patricia wil de grote groep jongeren passeren. Om ze aan de kant te laten gaan gebruikt ze de *Patricia wants to pass the large group of young people. In order to make them move aside, she uses the*	Melanie werd 6 jaar en dat werd uitgebreid gevierd. Het eerste cadeau dat ze kreeg was de *Melanie turned 6 and that was extensively celebrated. The first gift she received was the*	fietsbel *(bike bell)*	0.51	0.44
Oma heeft de laatste tijd steeds meer moeite om de televisie te verstaan. De audicien raadde haar aan om gebruik te gaan maken van een	Over drie dagen is Huub jarig. Voor zijn verjaardag wil hij graag een	gehoorapparaat	0.91	0.40

Grandma's been having more and more trouble understanding television lately. The Hearing Care Professional advised her to use a	*It's Huub's birthday in three days. For his birthday he would like a*	*(hearing aid)*		
De postcodeloterij was afgelopen maand in onze straat gevallen. We besloten een mooie roadtrip door Amerika te gaan maken van het *The postcode lottery had fallen on our street last month. We decided to make a nice roadtrip through America from the*	Andre kan de slaap al uren niet vatten. Als hij zijn ogen sluit ziet hij steeds het *Andre hasn't been able to sleep for hours. When he closes his eyes, he always sees the*	geld *(money)*	0.91	0.95
De man is achter het stuur gekropen, ook al had hij teveel gedronken. Hij bracht de nacht door in de *The man got behind the wheel, even though he'd had too much to drink. He spent the night in the*	Ondanks dat hij in de file stond, moest Richard toch lachen. Op de radio hoorde hij een grap over de *Even though he was stuck in traffic, Richard still had to laugh. He heard a joke on the radio about the*	gevangenis *(prison)*	0.71	0.15
Toen de jager omkeek stond er een hert achter hem. Snel pakte hij het *When the hunter turned around, there was a deer behind him. Quickly he grabbed the*	Na drie koppen koffie kon Vera zich eindelijk op haar werk concentreren. Niet langer bleef ze verwonderd kijken naar het *After three cups of coffee Vera could finally concentrate on her work. No longer did she keep looking in amazement at the*	geweer *(weapon)*	0.94	0.88
Jenny zag dat de plantjes water nodig hadden. Ze ging naar de kraan en vulde de *Jenny saw that the plants needed water. She went to the tap and filled up the*	Karel was zijn schuur aan het opruimen. In de hoek onder alle spullen vond hij de *Charles was cleaning out his barn. In the corner under all the stuff he found the*	gieter *(watering can)*	0.94	0.21
Op zaterdagavond ga ik graag naar een rockconcert. Ik ben dol op het geluid van de	Thomas at snel zijn bord leeg zodat hij niet bij zijn ouders aan tafel hoefde te blijven zitten. Hij wilde geen woord meer horen over de	gitaar	0.66	0.98

I like to go to a rock concert on Saturday night. I love the sound of the	*Thomas finished his meal quickly so that he didn't have to sit at his parents' table. He didn't want to hear another word about the*	*(guitar)*		
Rapunzel liet haar blonde lokken uit het raam van de toren vallen. De knappe prins ging vervolgens langzaam omhoog via het *Rapunzel dropped her blonde locks from the window of the tower. The handsome prince then slowly climbed up her*	In sommige gevallen slaat een hobby om naar een obsessie. Het liep dan ook uit de hand met Wilma's liefde voor het *In some cases, a hobby turns into an obsession. Things got out of hand with Wilma's love for the*	haar *(hair)*	0.91	0.35
Sam gleed bij het snowboarden zo de steile afgrond in. Gelukkig werd hij snel gered met een *Sam slipped into the steep abyss while snowboarding. Fortunately, he was quickly rescued with a*	Leo luistert naar een spannend audioboek. Hij is net bij het stuk over een *Leo is listening to an exciting audiobook. He's just at the piece about the*	helicopter *(helicopter)*	0.86	0.49
Hendrik ging vissen aan de waterkant. Toen zijn dobber onderging, pakte hij meteen de *Henry went fishing on the waterfront. When his float went down, he immediately grabbed the*	Bas heeft veel spullen die hij niet meer gebruikt. Naar de kringloopwinkel brengt hij de *Bas has a lot of stuff he doesn't use anymore. To the thrift store he brings the*	hengel *(fishing rod)*	0.91	0.44
Valerie heeft gisteren teveel gedronken en heeft nu een enorme kater. Het meeste last heeft ze van de *Valerie had too much to drink yesterday and now she's got a huge hangover. She’s suffering most from a*	Grietje kwam helemaal chagrijnig thuis. Haar hele dag was verpest door de *Gretel was grumpy when she came home. Her whole day was ruined by a*	hoofdpijn *(headache)*	0.79	0.31
Lucas wil altijd weten hoe laat het is. Voor zijn verjaardag kreeg hij van zijn vrienden een *Lucas always wants to know what time it is. For his birthday, his friends gave him a*	Harry heeft wel 30 verschillende t- shirts. Zijn nieuwste shirt heeft een plaatje van een *Harry's got 30 different t-shirts. His latest shirt has a picture of a*	horloge *(watch)*	0.98	0.89

Ze verveelden zich al de hele middag. Ineens sprong Karin op, ze had eindelijk een *They've been bored all afternoon. Suddenly Karin jumped up, she finally had an*	Moeder zegt de hele dag al geen woord tegen vader. Zij was nog steeds boos op hem vanwege zijn kritiek op een *Mother hasn't said a word to Father all day. She was still angry with him because he criticized her*	idee *(idea)*	1	0.69
In Australië bezocht ik lokale dieren op een ranch. In een groot buitenhok zag ik een *In Australia I visited the local animals on a ranch. In a large outdoor kennel I saw a*	Alle drie keken ze geconcentreerd naar hem. Voorzichtig benaderde hij een *All three of them looked at him with concentration. Carefully he approached the*	kangoeroe *(kangaroo)*	0.85	0.22
Juffrouw Jannie liet in de les over de herfst een paar stekelige groene bolsters zien. Voorzichtig opende ze er eentje en de kleuters keken verwonderd naar de *Miss Jannie showed some prickly green husks in the lesson about autumn. Carefully she opened one and the toddlers looked in amazement at the*	Nancy verzamelt graag dingen uit de natuur. Het mooist vindt ze de *Nancy enjoys collecting things from nature. What she likes best is the*	kastanje *(chestnut)*	0.76	0.16
Garfield, het huisdier van de familie Jansen, was weggelopen. Loesje hing overal in de buurt posters op met een foto van de *Garfield, the pet of the Jansen family, had run away. Loesje hung up posters all over the neighborhood with a picture of their*	Carolien bladert alle tijdschriften op tafel door. Ze is op zoek naar het artikel over de *Carolien flipped through all the magazines on the table. She is looking for the article about the*	kat *(cat)*	0.98	0.21
Nicole had maar een uur voor de moeilijke toets. Tijdens het invullen van de vragen keek ze daarom regelmatig op de *Nicole only had an hour before the difficult test. Whilst answering the questions she therefore regularly looked at the*	Beatrice weet op de housewarming niet zo goed waar ze over moet praten. Twijfelend begint ze een gesprek over de *Beatrice doesn't know what to talk about on the housewarming. Doubtfully she starts a conversation about the*	klok *(clock)*	1	0.25

Ik maakte een heerlijk kopje thee voor mezelf. Uit de trommel pakte ik een *I made myself a lovely cup of tea. From the tin I took a*	De kinderen waren buiten verstoppertje aan het spelen. Susan kon de anderen niet vinden maar ze vond wel een *The kids were playing hide and seek outside. Susan couldn't find the others but she did find a*	koekje *(cookie)*	0.81	0.24
Justin was een uitstekende journalist. Om op de hoogte te blijven van de actualiteiten keek hij elke dag in de *Justin was an excellent journalist. To keep abreast of current affairs, every day he reads the*	Rianne speelde een spelletje op haar mobiel. Ineens verscheen er een reclame over de *Rianne played a game on her mobile. Suddenly an advertisement appeared about the*	krant *(newspaper)*	1	0.42
De klaar-overs stonden elke ochtend klaar om de kinderen te helpen. Zo kwamen zij zonder problemen over het *The lollipop men and women were ready to help the children every morning. Without problems they acrossed the*	De familie was bijeengekomen om te spreken over de situatie. Ze waren het erover eens dat er iets gedaan moest worden aan het *The family had gathered to discuss the situation. They agreed that something had to be done at the*	kruispunt *(intersection)*	0.57	0.50
Roel wilde een fles wijn open maken voor bij het diner. Uit de keukenlade haalde hij de *Roel wanted to open a bottle of wine for dinner. From the kitchen drawer he took the*	Quinten was een cryptogram puzzel aan het maken. Hij wist niet wat er bedoeld werd met de cryptische beschrijving van de *Quinten was making a cryptogram puzzle. He didn't know what was meant by the cryptic description of the*	kurkentrekker *(corkscrew)*	0.78	0.47
De meeste kleuters doe je een groot plezier met een bezoekje aan de dierentuin. Maar soms schrikken ze van het gebrul van de *Most toddlers give you lots of happiness during a visit to the zoo. But sometimes they are frightened by the roar of the*	Nerveus stond de kleine jongen voor de groep. Hij haalde diep adem en begon zijn spreekbeurt over de *The nervous little boy stood up in front of the group. He took a deep breath and started his talk about the*	leeuw *(lion)*	0.87	0.55

Ivo houdt van wat extra room in zijn koffie. Hij deed er melk in en gebruikte een *Ivo likes some extra cream in his coffee. He put milk in it and used a*	Norbert kreeg een onvoldoende tijdens zijn beoordelingsgesprek. Zijn baas vond dat hij teveel bezig was met een *Norbert got a fail at his appraisal interview. His boss thought he was too preoccupied with an*	lepeltje *(spoon)*	0.86	0.56
Anna moest op de twintigste verdieping zijn voor een vergadering. Ze ging meteen naar de *Anna had to be on the 20th floor for a meeting. She went straight to the*	Marjolein is de hele dag druk bezig geweest op haar werk. Ze is vergeten de reparateur te bellen voor de *Marjolein has been busy at work all day. She forgot to call the repairman for the*	lift *(elevator)*	1	0.27
Op bijna elke middelbare school wordt er een eindejaarsgala georganiseerd. Perfect opgedofte leerlingen worden graag naar de feestlocatie gebracht met een *At almost every secondary school there is an end-of-year gala. Perfectly dressed up pupils are brought to the party location with a*	Paula ging naar de bibliotheek in het dorp. Voor de ingang stond een *Paula went to the library in the village. In front of the entrance there was a*	limousine *(limousine)*	0.92	0.07
Mieke wilde een paar sfeervolle kaarsjes aansteken. Daarvoor gebruikte ze een *Mieke wanted to light some decorative candles. To do so she used a*	Alle kinderen uit de straat waren aan het stoepkrijten. De grappigste tekening was die van een *All the kids in the street were drawing chalk on the pavement. The funniest drawing was that of a*	lucifer *(match)*	0.64	0.30
De kok ging de verse zalm snijden. Hij nam de vis uit de koelkast en pakte een *The cook was going to cut the fresh salmon. He took the fish out of the fridge and grabbed a*	In een sloppenwijk is het meestal een grote smeerboel. Tussen alle rommel vond Pablo een *A slum is usually a big mess. Among all the rubbish Pablo found a*	mes *(knife)*	0.91	0.73

De zangeres betrad het podium van de concertzaal. Zij zwaaide naar het publiek en ging staan bij de *The singer entered the stage of the concert hall. She waved to the audience and stood up behind the*	Ijverig maakte Dave de spellingstoets. Hij twijfelt over de zin met de *Dave took the spelling test diligently. He had doubts about the sentence with the*	microfoon *(microphone)*	0.79	0.60
Voor jongeren is het tegenwoordig heel belangrijk dat ze altijd hun vrienden kunnen bereiken. Elke puber zeurt daarom bij zijn ouders om een *Nowadays, it is very important for young people to always be able to reach their friends. That is why every adolescent is nagging their parents for a*	Nu Theo even zonder werk zit, kan hij niet meer zomaar alles kopen. Hij heeft dan ook geen geld voor een *Now that Theo is out of work for a while, he can't just buy everything anymore. He doesn't have the money for a*	mobieltje *(mobile phone)*	0.81	0.79
Met mooi weer gaat Jos in het weekend graag touren met de vriendengroep. Ook doordeweekst gebruikt hij graag de *With nice weather Jos likes to ride during the weekend with his group of friends. During the week he also likes to use his*	Julius kijkt al de hele week uit naar het weekend. Zaterdag krijgt hij eindelijk de *Julius has been looking forward to the weekend all week. On Saturday, he will finally get the*	motor *(motorbike)*	0.70	0.77
Het water in de gootsteen kan niet weglopen door alle etensresten in de afvoerbuis. Gelukkig kwamen deze los toen Willem gebruik maakte van de *The water in the sink cannot drain all the food residues in the drainpipe. Luckily, it came loose when Willem used the*	Rita zette de boodschappen op het aanrecht. Het enige wat ze is vergeten te halen is de *Rita put the groceries on the sink. The only thing she forgot to pick up is the*	ontstopper *(plunger)*	0.83	0.36
Paul is heel muzikaal. Op zondag gaat hij altijd naar de kerkdienst en speelt hij op het *Paul is very musical. On Sunday, he always goes to church and plays the*	Hugo werkt zes dagen in de week. Zijn vrije dag brengt hij het liefste door achter het *Hugo works six days a week. He prefers to spend his day off behind the*	orgel *(organ)*	0.91	0.30
Er dreigden grote grijze wolken net toen ik naar de markt wilde lopen. Ik durfde het huis niet uit te gaan zonder een	De kinderen waren verwikkeld in een hevige discussie. Ze maakten elkaar goed duidelijk wat ze nu vonden van een	paraplu	1	0.38

There were big grey clouds just as I was about to walk to the market. I didn't dare leave the house without a	*The children were engaged in a fierce discussion. They made it very clear to each other what they now thought of the*	*(umbrella)*		
Max was de hele morgen bezig geweest met het versieren van de kerstboom. Nu miste op de top van de boom alleen nog de *Max had been decorating the Christmas tree all morning. Now all that was missing on top of the tree was the*	Trudy kon haar aandacht niet bij het gesprek houden. Ze werd afgeleid door de *Trudy couldn't keep her attention on the conversation. She was distracted by the*	piek *(peak)*	0.97	0.11
Op de zuidpool leven grote groepen dieren. Mijn favoriet is de *The South Pole is home to large groups of animals. My favorite is the*	Ze verslikte zich toen ze het nieuws op de televisie zag. Vol verbazing keek ze naar het bericht over de *She choked when she saw the news on television. In amazement, she watched the news about the*	pinguïn *(penguin)*	0.87	0.58
Het jonge stel ging op huwelijksreis naar Egypte. Daar bezochten ze als eerste de *The young couple went on their honeymoon to Egypt. There they visited the*	De kinderen in de klas hielden een kringgesprek. Erik vertelde enthousiast over de *The children in the class had a discussion in a circle. Erik told enthusiastically about the*	piramiden *(pyramids)*	0.92	0.24
Voordat ik de bioscoopzaal in ga haal ik altijd iets lekkers. Er is niets zo leuk als een film kijken terwijl je geniet van de *Before I go into the cinema, I always get something nice. There's nothing like watching a movie while you're enjoying the*	Ze stonden al tien minuten op haar te wachten. Uiteindelijk arriveerde ze met de *They had been waiting for her for ten minutes. Eventually she arrived with the*	popcorn *(popcorn)*	0.82	0.26
Annette vond het muf ruiken in haar kamer. Om frisse lucht te krijgen opende ze het *Annette thought her room smelled stale. To get some fresh air, she opened the*	Het was boekenweek op de basisschool. Sarah nam een griezelboek mee over het *It was book week in elementary school. Sarah brought a horror book about the*	raam *(window)*	0.94	0.58

De verstrooide leerling was al zijn sportspullen vergeten. Hij leende bij zijn tennisleraar een *The absent-minded student had forgotten all of his sports gear. From his tennis coach he borrowed a*	Mensen met autisme hebben vaak een obsessie voor een bepaald voorwerp of onderwerp. Daniel heeft dat met een *People with autism often have an obsession with a certain object or subject. Daniel has that with a*	racket *(racket)*	0.73	0.39
De deelnemer aan het televisiespelprogramma keek gespannen naar de presentatrice. Op haar teken mocht hij draaien aan het *The participant in the television game show looked tensely at the presenter. On her signal he was allowed to spin the*	Olivier dreigt een burn-out te krijgen als hij op deze manier doorgaat. Hij draagt teveel verantwoordelijkheid met zich mee, waaronder die over het *Olivier is in danger of getting burned out if he goes on like this. He takes on too much responsibility, including for the*	rad *(wheel)*	0.91	0.18
In de zomermaanden is het een natte boel in de tropen. Je kan dan nergens schuilen voor de *In the summer months the tropics are a wet place. There's nowhere to hide from the*	Peuters kunnen de hele dag door hetzelfde zeggen. Kleine Lucas heeft het nu al een uur over de *Toddlers can say the same thing all day long. For an hour now, little Lucas has been talking about the*	regen *(rain)*	0.85	0.87
De zon scheen fel aan de hemel toen het plotseling in de verte begon te regenen. Toen we in verwondering omhoog keken zagen we een *The sun was shining brightly in the sky when it suddenly started raining in the distance. When we looked up in amazement we saw a*	Hij werd verrast toen hij de bladzijde omsloeg. Daar stond op twee pagina's groot een *He was surprised when he turned the page. Spread out over two pages there was a large*	regenboog *(rainbow)*	0.79	0.71
Nora moest een aantal moeilijke rekensommen maken. Daarom pakte ze meteen de *Nora had to make some difficult calculations. That's why she immediately grabbed the*	Annebel vindt het leuk om te tekenen. Ze is bijna klaar met het tekenen van de *Annebel likes to draw. She's almost done drawing the*	rekenmachine *(calculator)*	0.98	0.33

Mariska vertelde enthousiast over het huwelijksaanzoek van haar verloofde. Iedereen keek vol verwondering naar de *Mariska talked enthusiastically about her fiancé's marriage proposal. Everyone looked amazed at the*	Fiona heeft altijd veel geluk. In de garage vond ze vanochtend de *Fiona is always very lucky. This morning in the garage she found the*	ring *(ring)*	0.90	0.95
Het lint hing klaar voor de opening van de nieuwe school. De burgemeester pakte hiervoor een *The ribbon hung ready for the opening of the new school. The mayor grabbed*	Tina komt thuis van werk en hangt haar jas op. Op de tafel ziet ze een *Tina comes home from work and hangs up her coat. On the table she sees*	schaar *(scissors)*	1	1.00
De kinderen vermaakten zich kostelijk op het strand. Stijn maakte een diepe gracht met de *The children had a great time on the beach. Stijn made a deep moat with the*	Hij zat al twee uur op internet, op zoek naar het perfecte cadeau. Uiteindelijk koos hij voor de *He had been on the internet for two hours, looking for the perfect gift. In the end he chose the*	schep *(shovel)*	0.97	0.50
De inbreker had geen moeite om de geheime familiekluis te vinden. Deze bevond zich namelijk achter het *The burglar had no trouble finding the secret family safe. It was behind the*	Muziek kan sterke herinneringen oproepen. Bij dit nummer denkt Karel meteen aan het *Music can evoke strong memories. With this song Karel immediately thinks of the*	schilderij *(painting)*	0.88	0.21
In het bezoekerscentrum is een grote vijver met verschillende dieren aangelegd. Op de steen onder de warmtelamp ziet Hans een *A large pond with various animals has been dug in the visitor centre. On the stone under the heat lamp Hans sees a*	Richard begon zich zorgen te maken om Simone. Ze staarde al 10 minuten strak naar een *Richard started to worry about Simone. For 10 minutes, she had been staring at a*	schildpad *(turtle)*	0.64	0.38
Caroline heeft al meerdere pogingen gedaan om te stoppen, maar ze vindt het moeilijk haar slechte gewoonte af te leren. Uiteindelijk pakte ze toch een	Toen ik wakker werd kon ik me mijn droom nog levendig herinneren. Het was niet de eerste keer dat ik droomde over een	sigaret	0.91	0.61

Caroline has tried to stop several times, but she finds it difficult to get rid of her bad habit. Eventually, she picked up a	*When I woke up, I could still vividly remember my dream. It wasn't the first time I dreamt about a*	*(cigarette)*		
Kleine Tony kon nog niet goed met mes en vork eten en knoeide. Hij kreeg daarom van zijn moeder een *Little Tony couldn't eat with knife and fork yet and messed around. That's why his mother gave him a*	Voor het buurtfeest mocht de vierjarige Michael een muurschildering maken. Ze hadden niet verwacht dat hij een tekening zou maken van een *Four-year-old Michael was allowed to make a mural for the neighbourhood party. They didn't expect him to draw a*	slabbetje *(bib)*	0.52	0.23
Simone is altijd alles kwijt. Toen ze vanochtend naar haar werk wilde rijden, moest ze een uur zoeken naar de *Simone always loses everything. When she wanted to drive to work this morning, she had to spend an hour looking for the*	Marta werd verrast toen ze door het oude boek bladerde. Er zat namelijk een foto tussen van de *Marta was surprised when she flipped through the old book. There was a photo of the*	sleutel *(key)*	0.72	0.85
Moeder maande de tienjarige Tobias de computer nu echt uit te zetten. Hij leek wel verslaafd aan het *Mother urged ten-year-old Tobias to turn off the computer for real. He seemed addicted to the*	Christa kon haar ogen niet geloven toen Chiel in een t-shirt naar buiten kwam. Op zijn onderarm had hij een tatoeage van het *Christa couldn't believe her eyes when Chiel came out in a t-shirt. On his forearm he had a tattoo of the*	spelletje *(game)*	0.57	0.57
Het vuilnis in het zomerse Napels was al een week niet opgehaald. Alle inwoners van de stad klaagden over de *The garbage in summery Naples hadn't been collected for a week. All of the city's inhabitants complained about the*	Het was al drie uur 's nachts maar hij kon nog steeds niet slapen. Hij bleef maar piekeren over de *It was three o'clock in the morning but he still couldn't sleep. He kept worrying about the*	stank *(smell)*	0.96	0.33

Mijn vrienden en ik hebben dit weekend een wedstrijdje hockey gespeeld. Jan had nog nooit eerder gespeeld en had veel moeite met het vasthouden van de *My friends and I played a game of hockey this weekend. Jan had never played before and had a lot of trouble holding the*	Midden in de nacht werd Susan wakker omdat ze naar de wc moest. Onderweg naar de badkamer struikelde ze over de *In the middle of the night Susan woke up because she had to go to the toilet. On the way to the bathroom she tripped on the*	stick *(stick)*	0.73	0.38
Het tapijt in de huiskamer lag na het kinderfeestje vol met kruimels en confetti. Martijn besloot om snel de rommel op te ruimen met de *The carpet in the living room was full of crumbs and confetti after the children's party. Martijn decided to quickly clean up the mess with the*	De gegijzelde man had zo lang niets gegeten of gedronken dat hij begon te hallucineren. Toen hij knipperde met zijn ogen zag hij de *The hostage hadn't eaten or drunk for so long that he started to hallucinate. When he blinked he saw a*	stofzuiger *(vacuum cleaner)*	0.91	0.35
Maria kreeg het niet voor elkaar om de computer aan te zetten. Blijkbaar zat de stekker niet in het *Maria didn't manage to turn on the computer. Apparently the plug wasn't in the*	Ze wist niet wat ze kon verwachten in de andere kamer. Toen ze de deur opendeed was het eerste was ze zag het *She didn't know what to expect in the other room. When she opened the door, the first thing she saw was the*	stopcontact *(socket)*	1	0.16
Mijn blouses kwamen helemaal gekreukt uit de was. Om ze weer glad te krijgen gebruikte ik het *My blouses came out of the wash all wrinkled. To get them smooth again, I used the*	Katy stond met haar vrienden op de jaarlijkse rommelmarkt. Het eerste wat ze verkocht was het *Katy was at the annual flea market with her friends. The first thing she sold was*	strijkijzer *(iron)*	1	0.46
Het hele dorp werkte samen om het kleine meisje uit de waterput te redden. Met z'n allen trokken ze aan het *The whole village worked together to save the little girl from the well. Together they pulled the*	Guus liep wel 200 treden omhoog. Boven in de toren pakte hij het *Guus walked up 200 steps. At the top of the tower he grabbed the*	touw *(rope)*	0.93	0.78

De oude kerk heeft een hoge toren vanwaar je een mooi uitzicht hebt. Om boven te komen moest Stefan met de *The old church has a high tower from where you have a nice view. To get upstairs Stefan had to take the*	Toen ze naar het restaurant liep, wist ze niet wat ze zag. Op het raam hing een poster van de *When she walked to the restaurant, she couldn't believe what she saw. On the window hung a poster of*	trap *(stairs)*	0.93	0.86
Hans en Lydia hebben vorige week de Kilimanjaro beklommen. Toen ze de top bereikten waren ze onder de indruk van het *Hans and Lydia climbed Kilimanjaro last week. When they reached the top they were impressed by the*	Vader besloot oude foto's te laten digitaliseren. Helaas miste de foto van het *Dad decided to have old photos digitized. Unfortunately, one was missing: the photo of the*	uitzicht *(view)*	1	0.23
Hanneke trakteert zichzelf op een manicure. De schoonheidsspecialiste haalt de scherpe punten weg met een *Hanneke treats herself to a manicure. The beautician removes the sharp edges with a*	De muur is bezaaid met foto's en plaatjes. Franks favoriet is die van een *The wall is littered with photos and pictures. Frank's favorite is one of a*	vijl *(file)*	0.93	0.55
Jos moest naar Engeland voor een conferentie. Omdat hij snel zeeziek wordt, kocht hij een ticket voor het *Jos had to go to England for a conference. Because he gets seasick quickly, he bought a ticket for the*	Een duur merk zegt niet altijd iets over de kwaliteit. Dat merkten ze duidelijk op bij het *An expensive brand doesn't always say something about quality. They clearly noticed that at the*	vliegtuig *(airplane)*	1	0.74
Jan zag dat hij de planken korter moest maken. Hij liep naar de gereedschapskist en pakte de *Jan saw that he had to shorten the boards. He walked to the toolbox and grabbed the*	Ankie begon wat vergeetachtig te worden. Ze wist niet meer wat ze had gedaan met de *Ankie started to get a little forgetful. She didn't know what she'd done with the*	zaag *(saw)*	0.98	0.67
Voor ieders veiligheid is het belangrijk dat de verkeersregels worden nageleefd. Zo moeten mijn kinderen altijd oversteken bij een	Elke week hebben ze in de kleuterklas een ander thema. Deze week gaat het thema over een	zebrapad	0.93	0.21

It is important for everyone's safety that the traffic rules are observed. For example, my children must always cross at a	*Every week in kindergarten they have a different theme. This week the theme is about a*	*(zebra crossing)*		
Na een uitgebreide restauratie was het schip klaar om de zee op te gaan. De matroos maakte de trossen los en hees het *After an extensive restoration the ship was ready to go to sea. The sailor loosened the mooring lines and hoovered up the*	Iedereen is al op het feestje, behalve Gerda. Zij is nog bezig met het repareren van het *Everyone's already at the party, except Gerda. She's still repairing the*	zeil *(sail)*	0.80	0.05
De dappere ridder zag dat de draak de tovenaar bedreigde. Uit zijn schede pakte hij snel het *The brave knight saw that the dragon threatened the sorcerer. From his scabbard he quickly grabbed the*	De drie broertjes keken erg schuldig toen moeder thuiskwam. Ze wisten heel goed dat ze niet hadden mogen komen aan het *The three brothers looked very guilty when Mother came home. They knew very well that they shouldn't have touched the*	zwaard *(sword)*	1	0.57
Kelly ging voor het eerst kajakken in Frankrijk. Toen haar kajak omsloeg op de wilde rivier was ze blij met het *Kelly went kayaking in France for the first time. When her kayak tipped over on the wild river she was happy with the*	Toen ze elkaar tegenkwamen op de gang bleven ze even staan om te kletsen. Het gesprek ging al snel over het *When they bumped into each other in the hallway, they stopped to chat. Soon, the conversation was about the*	zwemvest *(life jacket)*	0.60	0.49