

SHORT REPORTS

Neural signatures of syntactic variation in speech planning

Sebastian Sauppe^{1,2*}, Kamal K. Choudhary³, Nathalie Giroud⁴, Damián E. Blasi^{5,6,7,8}, Elisabeth Norcliffe⁹, Shikha Bhattamishra³, Mahima Gulati³, Aitor Egurtzegi^{1,2}, Ina Bornkessel-Schlesewsky^{10,11}, Martin Meyer^{2,12,13}, Balthasar Bickel^{1,2}

1 Department of Comparative Language Science, University of Zurich, Zurich, Switzerland, **2** Center for the Interdisciplinary Study of Language Evolution (ISLE), University of Zurich, Zurich, Switzerland, **3** Department of Humanities and Social Sciences, Indian Institute of Technology Ropar, Rupnagar, India, **4** Department of Computational Linguistics, University of Zurich, Zurich, Switzerland, **5** Human Evolutionary Biology Department, Harvard University, Cambridge, Massachusetts, United States of America, **6** Department of Linguistic and Cultural Evolution, Max Planck Institute for the Science of Human History, Jena, Germany, **7** Linguistic Convergence Laboratory, National Research University Higher School of Economics, Moscow, Russia, **8** Human Relations Area Files, Yale University, New Haven, Connecticut, United States of America, **9** Department of Psychology, University of York, York, United Kingdom of Great Britain and Northern Ireland, **10** School of Psychology, Social Work and Social Policy, University of South Australia, Adelaide, Australia, **11** Cognitive and Systems Neuroscience Research Hub, University of South Australia, **12** Division of Neuropsychology, Department of Psychology, University of Zurich, Zurich, Switzerland, **13** Cognitive Psychology Unit, Psychological Institute, University of Klagenfurt, Klagenfurt, Austria



OPEN ACCESS

Citation: Sauppe S, Choudhary KK, Giroud N, Blasi DE, Norcliffe E, Bhattamishra S, et al. (2021) Neural signatures of syntactic variation in speech planning. *PLoS Biol* 19(1): e3001038. <https://doi.org/10.1371/journal.pbio.3001038>

Academic Editor: David Poeppel, New York University, UNITED STATES

Received: July 23, 2020

Accepted: December 31, 2020

Published: January 26, 2021

Copyright: © 2021 Sauppe et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Raw data, models and analysis scripts are available from <https://osf.io/uhtcn/>.

Funding: This research was funded by Swiss National Science Foundation Grant No. 100015_160011 awarded to BB and MM (www.snf.ch). IBS acknowledges the support of an Australian Research Council Future Fellowship (FT160100437, www.arc.gov.au). DEB acknowledges that this article was prepared within the framework of the HSE University Basic Research Program and funded by the Russian

* sebastian.sauppe@uzh.ch

Abstract

Planning to speak is a challenge for the brain, and the challenge varies between and within languages. Yet, little is known about how neural processes react to these variable challenges beyond the planning of individual words. Here, we examine how fundamental differences in syntax shape the time course of sentence planning. Most languages treat alike (i.e., align with each other) the 2 uses of a word like “gardener” in “the gardener crouched” and in “the gardener planted trees.” A minority keeps these formally distinct by adding special marking in 1 case, and some languages display both aligned and nonaligned expressions. Exploiting such a contrast in Hindi, we used electroencephalography (EEG) and eye tracking to suggest that this difference is associated with distinct patterns of neural processing and gaze behavior during early planning stages, preceding phonological word form preparation. Planning sentences with aligned expressions induces larger synchronization in the theta frequency band, suggesting higher working memory engagement, and more visual attention to agents than planning nonaligned sentences, suggesting delayed commitment to the relational details of the event. Furthermore, plain, unmarked expressions are associated with larger desynchronization in the alpha band than expressions with special markers, suggesting more engagement in information processing to keep overlapping structures distinct during planning. Our findings contrast with the observation that the form of aligned expressions is simpler, and they suggest that the global preference for alignment is driven not by its neurophysiological effect on sentence planning but by other sources, possibly by aspects of production flexibility and fluency or by sentence comprehension. This challenges current theories on how production and comprehension may affect the evolution and distribution of syntactic variants in the world’s languages.

Academic Excellence Project 5-100. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: EEG, electroencephalography; ERD, event-related desynchronization; ERS, event-related synchronization; IAF, individual alpha frequency; ROI, region of interest.

Introduction

Language is not a disparate and haphazard collection of idiosyncratic templates for how to formulate sentences. Instead, sentence templates form intricate systems of partial overlaps and alignments, an observation that has fueled inquiry since the Indian scholar Pāṇini wrote the first explicit grammar of a language over 2,500 years ago. For example, sentences like “The gardener planted a tree,” “The gardener crouched,” “The gardener worked hard,” or “A tree was planted by the gardener” align with each other in an abstract way by employing templates that begin with the same structure, here, with a noun phrase (the subject: “the gardener” and “a tree”) followed by a verb (“plants,” “crouches,” “works,” and “was”), while they differ in the remainder. Such alignments differ between languages in striking ways.

Some languages align noun phrases like “the gardener” in a sentence with 1 noun phrase (intransitives, e.g., “the gardener crouched”) with the agent noun phrase in sentences with 2 or more noun phrases (transitives, e.g., “the gardener planted trees,” “the gardener poured water into the trough,” etc.). Other languages do not align the 2 sentence types and instead add a special marker in sentences with more than 1 noun phrase, keeping them formally distinct from sentences with only 1 noun phrase. These languages contrast a plain, unmarked noun phrase in intransitives (“the gardener crouched”) with a marked agent noun phrase in transitives (“the gardener+*x* planted trees,” with an additional marker represented here by “+*x*”; cf. Fig 1). This phenomenon, known as ergativity [1,2], is found in about one-third of the world’s languages [3].

Alignments and other kinds of partial overlaps are recognized as central properties of the human language faculty in all theories of grammar, albeit using widely different representational formats, e.g., derivations or inheritance schemata [4]. Despite their prominence in linguistic theory and the striking nature of variation, however, alignments have received remarkably little attention in neuroscience and psychology. The difference between aligned and nonaligned expressions poses 2 unresolved questions for neural processing and its relation to the variation between languages:

First, are there differences in the neural planning processes for aligned and nonaligned expressions? More concretely, is the way speakers plan sentences shaped by whether initial noun phrases are unmarked (“the gardener. . .,” aligned expression) or whether there is an opposition between unmarked and marked noun phrases (“the gardener” versus “the gardener +*x*,” nonaligned expression; Fig 1)? Does this difference affect sentence planning already in its early phases or only later, when speakers encode the phonological form of and articulatory motor plans for words with and without an additional agent noun phrase marker “+*x*” [5,6]? Specifically, sentence structures with aligned expressions might allow speakers to delay commitment to the choice between an intransitive and a transitive sentence plan because their beginnings overlap (“the gardener. . . crouches/plants trees”) [7,8]. At the same time, this intermittent compatibility with multiple sentence plans might also require that 2 possible options need to be kept distinct while speakers construct a syntactic plan [9,10].

Second, if neural information processing is sensitive to how structures are aligned, does the neural processing mirror the observation that aligned structures are structurally simpler (with no additional marking) and more common in the world’s languages [11]? Such a mirroring has been motivated by evidence from event-related potentials in sentence comprehension: Across languages, both with and without nonaligned expressions, comprehenders initially interpret unmarked initial noun phrases as referring to agents, i.e., following an aligned pattern [3,12–16]. A bias for aligned structures is also found in phylogenetic models of linguistic evolution [3]: After controlling for contingencies of history (such as language contact and language shifts), languages are universally more likely to develop and maintain aligned

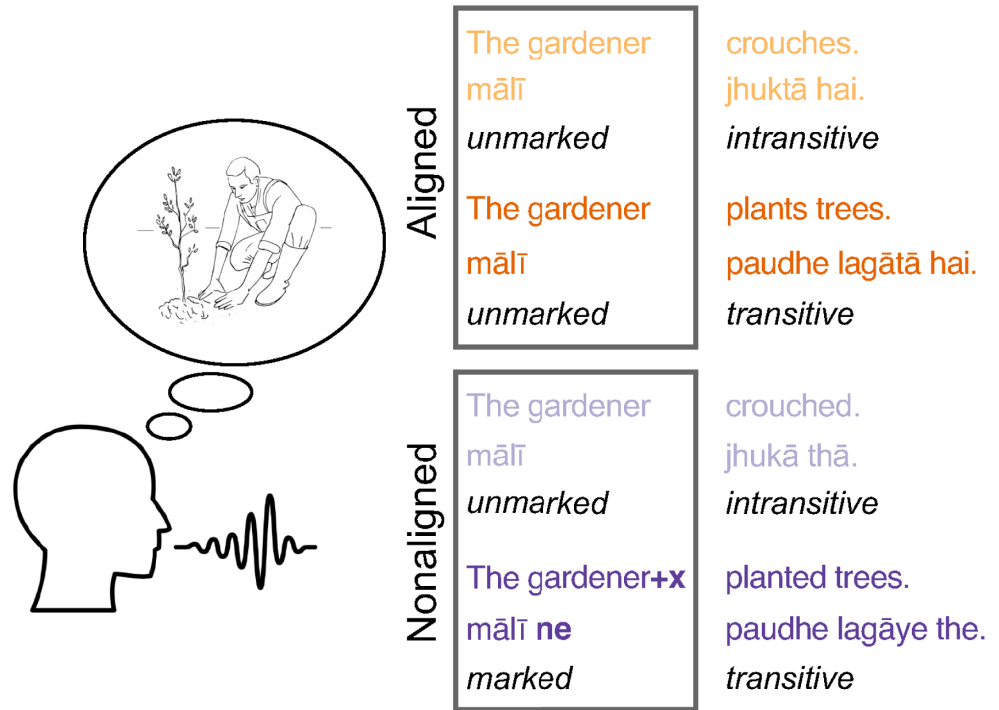


Fig 1. Illustration of the contrast between aligned and nonaligned agent expressions in Hindi, which allows both types of expressions (for detailed examples see S1 Table). In aligned expressions (used with imperfective aspect), unmarked noun phrases are compatible with both intransitive (“crouches”) and transitive sentences (“plants trees”). In nonaligned expressions (used with perfective aspect), unmarked noun phrases are only compatible with intransitive sentences (“crouched”) because in transitive sentences (“planted trees”) noun phrases are marked by an additional element “+x” (“ne” in Hindi).

<https://doi.org/10.1371/journal.pbio.3001038.g001>

expressions than the opposite. Such a correlation of neural processing and evolutionary developments might stem from a tight interlacing of comprehension and production processes [17] or from shared neural underpinnings [18,19], and would be consistent with other findings on how the brain’s processing constraints shape the form of languages [20–24].

Here, we seek to resolve these questions by exploiting a contrast between aligned and nonaligned syntax in Hindi for an experiment on sentence planning. Hindi aligns noun phrases in what is known as the imperfective aspect system (“the gardener crouches/was crouching” and “the gardener plants/was planting trees”) and keeps them nonaligned in the perfective aspect system (“the gardener crouched/has crouched” versus “the gardener+x planted/has planted trees,” Fig 1).

Speakers of Hindi described pictures of 1-participant (intransitive) and 2-participant (transitive) events by using sentences with either aligned or nonaligned syntax, with alignment condition split between groups of participants. While they prepared their responses, i.e., planned their sentences, we measured neural oscillatory processing using electroencephalography (EEG) simultaneously with overt visual attention as reflected in eye gaze. Visual attention allocation during planning is related to the syntactic structures being planned [25–28] and thus indicates how contrasts in structures shape the preparation of speakers’ sentence plans. With respect to neural processes, we focus on changes in total oscillatory power (event-related desynchronization [ERD] and event-related synchronization [ERS], [29]) in the theta and alpha frequency bands, typically ranging from 3 to 7 Hz and 8 to 12 Hz, respectively. While

subserving a wide range of processes [30], these frequency bands are implicated in the processing of syntactic (and possibly semantic) dependencies during the comprehension of sentences (theta: [31–39]; alpha: [34,35,40–42]). ERD and ERS are currently among the most suitable means to study the neural processes underlying early stages of sentence planning because oscillatory power can capture both evoked (phase-locked) and induced (not phase-locked) responses [43,44]. This allows the effective study of neural events involved in sentence planning that either begin at the moment when the to-be-described picture is presented or emerge during the process of sentence planning (and are therefore not phase-locked).

We target relational and structural planning processes in the first 800 ms after the picture onset, in line with crosslinguistic results on eye tracking during sentence planning [45,46]. Under the hypothesis that different processes underlie the preparation of aligned and non-aligned structures, we expected to find behavioral and neural dissociations in these early stages of planning. Specifically, we expected differences between sentence types in gazes toward agents as well as in neural oscillatory activity in the theta and alpha bands. These effects are taken to reflect speakers' differential engagement of working memory and attentional processes [36,47–51] when committing to a specific structure earlier or later during the planning of aligned and nonaligned sentences.

Results

In the first 800 ms after picture onset, speakers fixated on the agent characters in the pictures more when planning aligned than when planning nonaligned sentences (Fig 2A). This is reflected in an interaction between the cubic time term and sentence type in a logistic mixed-effects growth curve regression [52] ($\hat{\beta} = 1.13$, SE = 0.26, 95% CI = [0.63, 1.63], $p < 0.001$, S4 Table). EEG oscillatory power changes were analyzed with mixed-effects regression trees [53], which partition the data set by sentence type and regions of interest (ROIs; averaging across electrode positions within these regions) with respect to the time course of power changes, analyzed as polynomial growth curves (see Materials and methods). Our model identifies broadly distributed ERS in the theta band (Fig 2 B–D, S5 Table). Both intransitive and transitive aligned sentences exhibited stronger theta ERS than nonaligned sentences, starting from around 200 ms, especially over mid-frontal electrode sites, as well as subsisting after 600 ms in right-central and posterior electrode sites (Fig 2C and 2D).

In the alpha frequency band, we found a related but different effect (Fig 3, S6 Table). Here, the main distinction is between all marked and all unmarked noun phrases (Fig 3A), rather than between aligned and nonaligned syntax (Fig 1). Specifically, we found broadly distributed ERD starting after around 400 ms and peaking around 700 ms (Fig 3B), which was larger in sentences with unmarked noun phrases (“The gardener. . .”) than in sentences with marked noun phrases (“The gardener+x. . .”). The alpha ERD effect was most pronounced over central and posterior regions (Fig 3C and 3D).

All models controlled for a series of potential confounds and group-level variables (see Materials and methods). We specifically also controlled for (cumulative) syntactic priming, the phenomenon that the repetition of sentence structures alters and eases their subsequent processing [54,55]. We included the position of each utterance in the experiment as predictor so that effects persisted above and beyond any priming effects that may accumulate over an experimental session.

There were no differences in reaction times, i.e., speech onset latencies, between the production of aligned and nonaligned sentences (see S2 Fig, S3 Table). This suggests that the task demands were indistinguishable between describing the same picture as representing an ongoing (imperfective) or a recently completed (perfective) situation. Even though we cannot rule

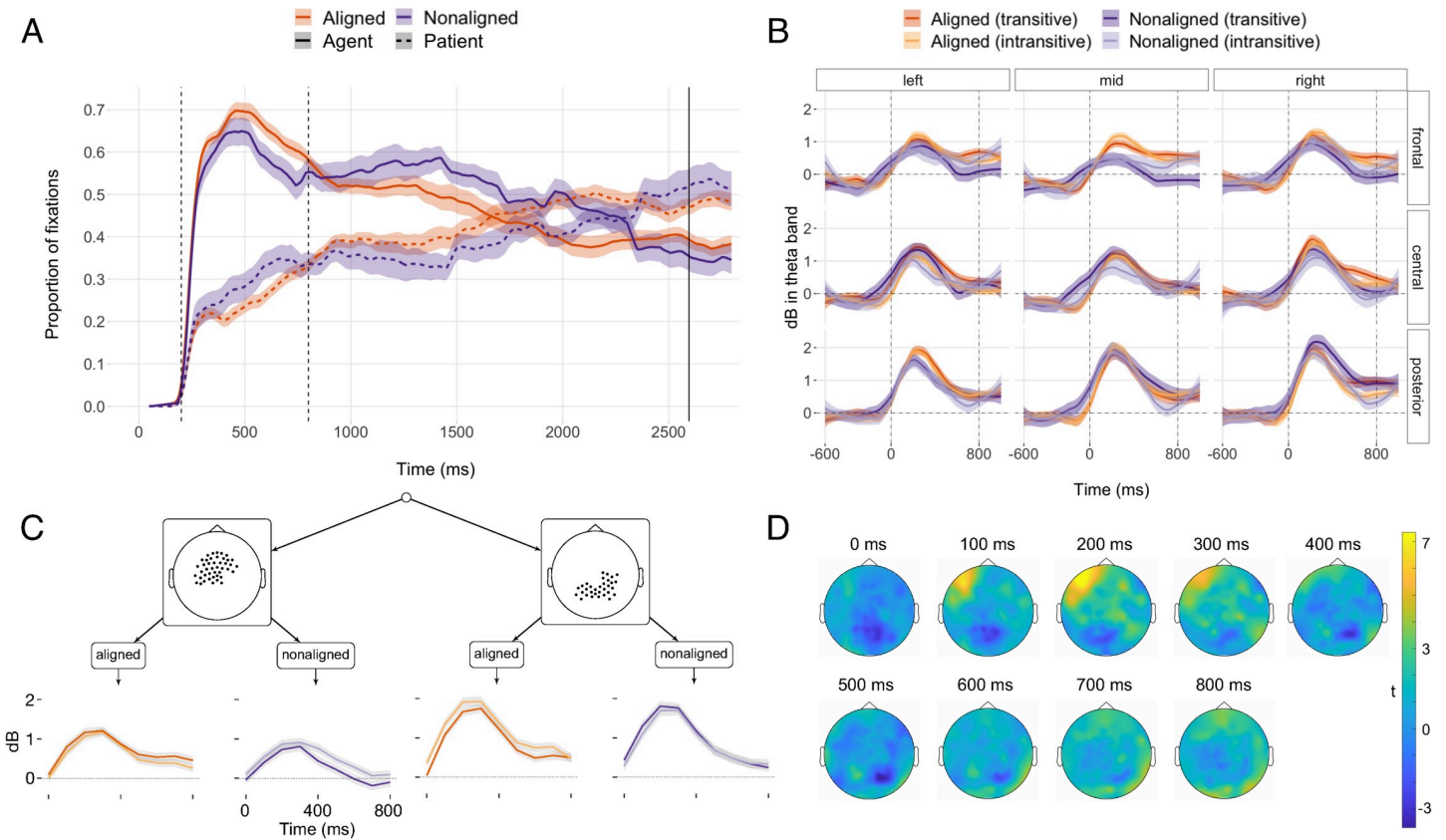


Fig 2. (A) Smoothed grand average of proportions of eye fixations to agent and patient characters in the picture stimuli during planning of transitive sentences (“The gardener plants trees” and “The gardener+x planted trees”). The solid vertical line represents speech onset (grand mean of transitive sentence responses). Overall, the planning of aligned sentences reveals more visual attention to agents ($p < 0.001$ in linear mixed-effects regression). (B) Smoothed grand average of event-related power changes (dB relative to a baseline period of -600 to -200 ms) in individually defined theta frequency bands. Grid cells in this panel represent ROIs (see S7 Table for details). Dashed vertical lines (in A and B) indicate analysis time windows, ribbons indicate standard errors, and $t = 0$ is the onset of stimulus picture presentation. (C) Regression model tree for power changes in individually defined theta frequency bands between 0 and 800 ms. Scalp maps show electrode positions of ROIs included in the respective grouping. All splits are statistically significant at $p < 0.001$. Tree tips show model fits for the respective grouping; ribbons represent 95% CIs, and colors code the same distinctions as in A and B. The model also identified lower-order splits which are subsidiary to the main difference between aligned and nonaligned sentences (see S4 Fig, S5 Table). (D) Topographic maps of power differences of aligned minus nonaligned sentences in t -values for a fixed theta frequency band (3–7 Hz). (Underlying data and scripts are available from <https://osf.io/uhtcn/> and in the Supporting information file S1 Data.) Overall, the planning of aligned sentences reveals ERS (more positive dB) in the theta band. ERS, event-related synchronization; ROI, region of interest.

<https://doi.org/10.1371/journal.pbio.3001038.g002>

it out categorically, the eye tracking and EEG results are therefore unlikely to stem from the fact that the planning of aligned sentences necessitated a completed conceptualization of the depicted events, while the planning of nonaligned sentences necessitated an ongoing conceptualization of the same event.

Our choice of the analysis time window makes it unlikely that we measured the encoding of phonological word forms rather than differences in syntax. It is generally assumed that the generation of grammatical structure precedes lexical encoding and the planning of phonological word forms [26,56,57], although the exact timing of the switch from grammar to phonology is not well understood [58,59]. In line with this, phonological encoding is unlikely to have influenced grammatical encoding in our measurements because variation in phonological lengths of noun phrases (in syllables) has only small effects in our models, magnitudes smaller than the effects of the sentence type differences (S4–S6 Tables). Furthermore, our analysis time window and frequency bands largely guard against possible confounds from the articulatory encoding of phonological forms in terms of speech-related muscle movement plans since these

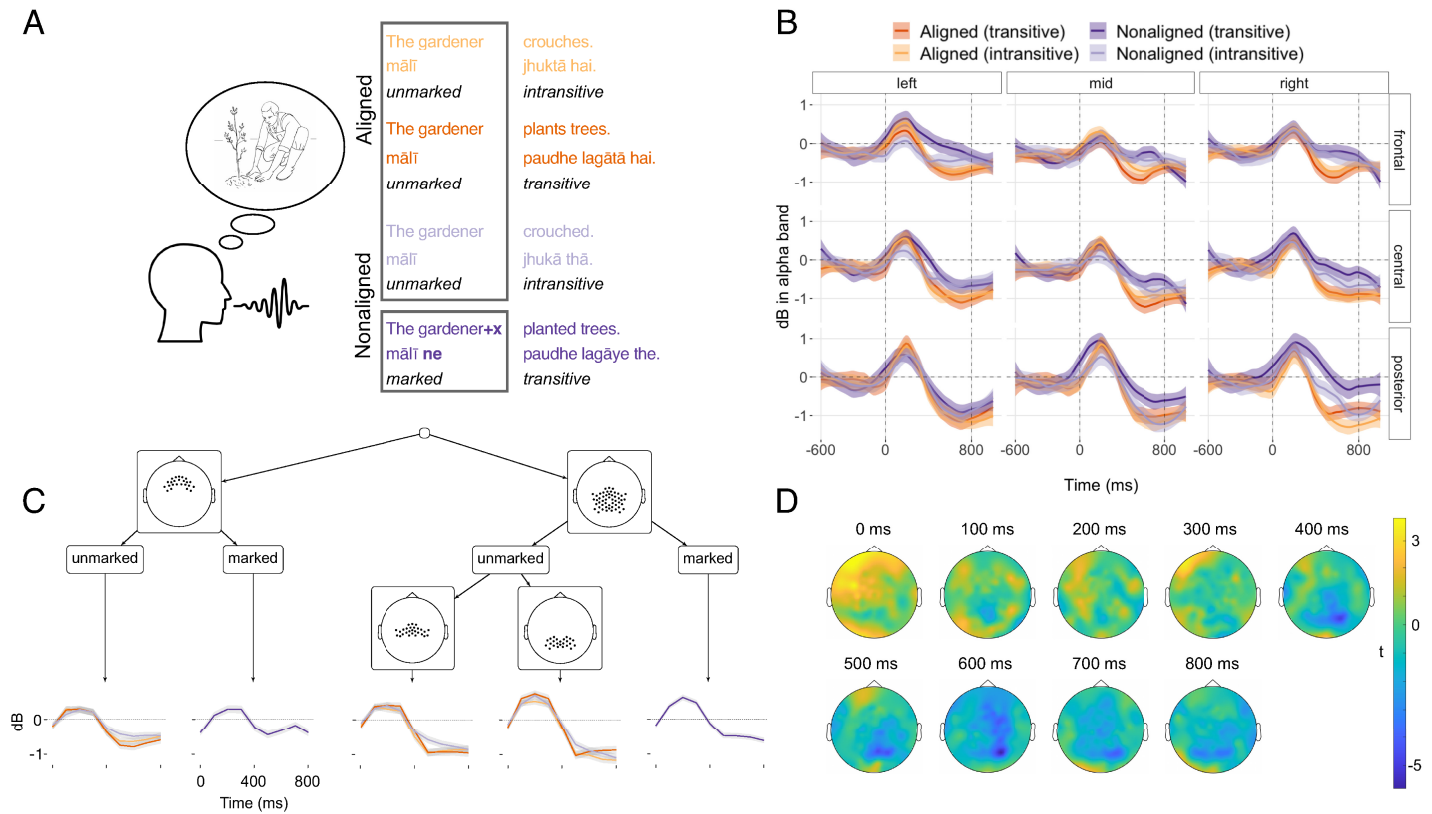


Fig 3. (A) Illustration of the contrast between unmarked and marked noun phrases (which crosses over aligned and nonaligned expressions) in Hindi: Aligned expressions are always unmarked (“The gardener . . .”), as are nonaligned noun phrases in intransitive sentences (“The gardener crouched”), while nonaligned noun phrases in transitive sentences are marked (“The gardener+x planted trees”). (B) Smoothed grand averages of event-related power changes (dB relative to a baseline period of –600 to –200 ms) in individually defined alpha frequency bands. (C) Regression model tree for power changes in individually defined alpha frequency bands between 0 and 800 ms. The model also identified lower-order splits which are subsidiary to the main difference between marked and unmarked expression (see S5 Fig, S6 Table). (D) Topographic maps of power differences (in *t*-values) between sentences with marked and unmarked expressions for a fixed alpha frequency band (8–12 Hz). (A–D) All other conventions are identical to the conventions in Fig 2. (Underlying data and scripts are available from <https://osf.io/uhtcn/> and in the Supporting information file S1 Data.) Overall, the planning of sentences with unmarked expression (intransitive and transitive aligned sentences and intransitive nonaligned sentences) reveals ERD (more negative dB) in the alpha band. ERD, event-related desynchronization.

<https://doi.org/10.1371/journal.pbio.3001038.g003>

processes set in only close to actual speech onset [57,60], which occurred at least 700 ms after our analysis window (see Materials and methods), and dominantly affect frequencies that are higher (>20 Hz [61]) than the alpha and theta bands we target. Finally, in single word production, theta band ERS and alpha band ERD specifically index the retrieval of conceptual and lexical information from memory [60,62] also when the phonological form of words is held constant.

Discussion

Our results suggest that the planning of different types of sentences is neurally implemented in distinct ways. This is reflected by differences in the temporal dynamics of visual attention allocation and neural oscillatory power changes in the theta and alpha frequency bands. We propose that the observed pattern of results stems from 2 separate processes.

First, speakers directed more visual attention to agents and exhibited greater theta ERS when planning aligned than when planning nonaligned sentences (Fig 2). This suggests that they used different planning strategies for the preparation of aligned and nonaligned sentences. Speakers need to commit to a sentence plan earlier in the nonaligned condition because

intransitive and transitive sentences already differ in the expression of the first noun phrase (with or without the additional agent marker “ne”). This requires speakers to prioritize early relational (and possibly structural) planning processes to encode who does what to whom [25,26]. This early (structural-)relational planning is reflected in looks that are more distributed over the whole picture and thus less focused on agents. In aligned sentences, by contrast, speakers do not need to decide on the full sentence plan as early [7]. The formal overlap between intransitive and transitive agent expressions allows speakers to delay their decision for 1 structure or the other. Speakers were thus able to primarily fixate on the agent as the action initiator [63] and as the referent of the first noun phrase in the sentence. In this way, speakers delayed the completion of relational processing [46] and held available all (aligned) structures that were compatible with a sentence starting with an unmarked agent noun phrase [10]. This scenario yields a natural interpretation of the increased theta ERS in the aligned condition as reflecting the simultaneous and noncommittal engagement of multiple compatible structures.

Intriguingly, sentence comprehension research has revealed that apart from being modulated by semantic and syntactic processing demands [32,33,37,38,64], theta oscillations play a role in maintaining working memory representations [34–37] and are linked to information retrieval from working memory [50,65]. Specifically, theta ERS is sensitive to the number of alternatives to be retrieved [50,66,67], in line with our proposed scenario of how aligned sentences are planned. The timing and frontal–parietal topography of theta ERS in our data also matches with frontal midline theta effects observed during the retrieval of syntactic structures in sentence comprehension [36].

The difference we detect in speakers’ commitment to sentence plans is consistent with behavioral findings on the importance of syntactic dependencies for the time course of planning: Like in the Hindi aligned condition, speakers of English and Japanese (where agent noun phrases are always aligned) plan agent noun phrases without an initial commitment to a verb and a sentence structure [68–70]. This contrasts with the earlier commitment that is required by sentences where the first noun phrase opens a strong dependency with the remaining sentence. In Hindi, we found this effect with marked, nonaligned noun phrases, and it is paralleled by the early commitment speakers need to make when planning patient noun phrases in English and Japanese which structurally depend on the verb [68–70]. More generally, the contrast we find between structures in Hindi expands previous findings that different dependencies afford different patterns in sentence planning [25,27,45,46,71–74].

Second, the planning of sentences with unmarked noun phrases induced larger ERD in the alpha band (Fig 3). We interpret this effect to reflect speakers’ need to keep distinct structures that share the same form (namely an unmarked noun phrase) at the beginning of sentences. This can be achieved by increased active neural information processing [29,75,76], e.g., in the form of increased engagement of cortical networks that are involved in processing syntactic information [40,42]. Converging evidence for our interpretation of the alpha ERD effect comes from reaction time findings that suggest higher processing loads when speakers need to separate alternative plans during sentence planning [9,10]. The need to keep distinct overlapping unmarked structures might also have contributed to the increased fixations to unmarked agents (Fig 2A) as longer fixation durations can indicate increased processing [49,77].

Our interpretation of the alpha ERD effect as reflecting the processing of syntactic alternatives is further supported by its central–posterior topography and its latency, with the largest differences setting in around 600 ms. The topography and latency are similar to alpha ERD effects associated with syntactic processes in sentence comprehension [34,35,40–42]. The current topographies are at the same time consistent with neuromodulatory evidence on the role of theta ERS and alpha ERD in controlling working memory–related processes [47]. Moreover,

in conversation, alpha ERD is associated with shifting from comprehending interlocutors' turns to preparing one's own production [78,79]. Our results show that the observed alpha ERD effects go beyond mere attention shifts to production and are also sensitive to differences in syntactic planning processes between sentence types.

Our finding that theta power increases are associated with alpha power decreases fits with previous reports on memory processes outside language [80]. However, our results also suggest a small timing difference: While alpha ERD was most pronounced toward the end of our analysis time window, around 700 ms, theta ERS was most pronounced around 300 ms. This could reflect that speakers decided on which structure to produce in aligned contexts probably between 300 and 700 ms after the start of sentence planning. Thus, while eye tracking allows the detection of distinct steps in the planning process, our concurrent multimodal approach demonstrates that the time course of sentence planning can be characterized in greater detail with additional evidence from highly time-resolved electrophysiological recordings. The combination of eye gaze and neural oscillatory power changes opens up the possibility of additional insights into the fine-grained structure of sentence planning in future studies.

The origin of oscillatory activity can be exogenous or endogenous for the comprehension and perception of language [81–83]. For language planning and production, however, there is no exogenous source that would obviously give rise to power changes in the theta and alpha bands. Participants were “externally stimulated” by pictures that were to be described—the same ones for aligned and nonaligned sentences. Theta ERS and alpha ERD in the planning of different sentence types therefore most likely reflect the internal processing of sentence structures, building on memory-related control processes [62].

The increased neural activity in the planning of aligned and unmarked syntax (in terms of theta ERS and alpha ERD) contrasts with the apparent simplicity and frequency of systems with aligned expressions among the world's languages. This challenges theories that propose sentence production processes as the key drivers of language form and the distribution of variants in language use [84,85]. These theories are based on the observation that speakers shape language use through choosing sentence structures that cause the least production difficulty, specifically, by placing easier to retrieve words first [86–88] and by reducing interference from similar words and structures [84,89,90]. Other structures would then be less likely to be produced and could only be used in restricted contexts or disappear from a language over time [84].

Our findings suggest either that (i) the relevant notion of production ease is independent of the neural activity we found and is instead grounded in more global constraints on sentence planning or that (ii) the distribution of aligned expressions in the world's languages is not driven by sentence planning processes.

Under the first possibility, what matters most for speakers is fluency in sentence planning. Delaying the commitment to 1 grammatical structure is arguably beneficial for speakers' fluency as it allows more flexibility when creating and implementing sentence plans incrementally [88,91]. Aligned expression facilitates this. The resulting benefit, e.g., for rapid response planning in conversation [78,92,93], would then be more important for language evolution than the increase in neural activity we found under aligned expression and the interference effects that this might generate.

The second possibility derives the prevalence for aligned expression from comprehension, specifically from a general preference of the sentence comprehension system to expect sentences to start with an unmarked noun phrase in the agent role [12–16]. Intriguingly, this expectation extends to languages with nonaligned syntax. For example, the nonaligned syntax of Hindi leads to unmarked noun phrases that regularly denote patients instead of agents in transitive sentences (e.g., the patient “pau^{dh}e” “trees” in Fig 1 is unmarked). When these noun

phrases occur initially in a sentence (as they often do), the comprehension parser nevertheless assigns them an agent role transiently and then shows an electrophysiologically detectable effect of reanalysis and prediction failure at the position of the verb [3]. This incurs an additional processing demand [94,95] that might explain the bias against nonaligned syntax in the world's languages.

Under this view, differences in planning processes between types of sentences have much less of an effect on linguistic distributions than processes during sentence comprehension. This is in line with theories of sound change that locate such effects chiefly in perception rather than production [96]. Prediction failure and reanalysis effects in comprehension would thus exert a vastly greater pressure on languages to abandon these sentence structures over time than the increased processing demands during the planning of aligned structures. This view is consistent with proposals that processing differences related to working memory engagement are key constraints of language evolution [20,21,97,98], to the extent that these differences stem from comprehension.

Alternatively, comprehension and production might form a trade-off: While the comprehension system favors unmarked, aligned syntax, production favors marked, nonaligned syntax, although to a much weaker extent. This would explain why nonaligned systems exist at all, despite their lower probability to emerge and persist during the evolution of languages [3].

More research is now needed to resolve these questions and to probe contrasts in alignment across diverse syntactic structures and languages. Our results demonstrate how new insights on the relationship between language production and the simplicity and prevalence of specific language structures critically rely on evidence from neural processes that underlie different syntactic variants in the earliest stages of sentence planning. This opens new avenues for research on the neural processes in language planning and speech production that go beyond the level of individual words [62,99,100] and beyond the small set of languages that have dominated the field so far [22,71,74,101–103].

Materials and methods

Participants

Fifty healthy, right-handed students at the Indian Institute of Technology Ropar with normal or corrected-to-normal vision (6 female, mean age = 20.47 years, SD = 3.35 years) participated in the experiment for payment and gave written informed consent. All participants were native speakers of Hindi, grew up in Uttar Pradesh, Madhya Pradesh, or Delhi, and reported to speak or grew up speaking Hindi to their parents on a daily basis. The study was approved by the Institutional Ethics Committee of the Indian Institute of Technology Ropar (approval number: 01–2016), and all procedures adhered to the Declaration of Helsinki. Sample size was determined based on previous studies on sentence planning [26,27,46,104].

Stimulus materials and experimental procedure

Participants described 55 line drawings depicting 2-participant (transitive) events, interspersed among 62 line drawings of 1-participant (intransitive) events. [S3 Fig](#) shows an example of a 2-participant picture. The order of stimulus presentation was randomized for each participant at runtime. To counterbalance the direction of agent/patient fixations, pictures were presented in 2 lists with vertically mirrored versions and lists contained roughly the same number of pictures with agents on the right or left.

Participants were assigned alternately to 1 of 2 groups (25 participants per group) and were instructed to describe the pictures with sentences either in imperfective aspect (as if the event was ongoing, aligned condition) or in perfective aspect (as if the event was completed,

nonaligned condition). We chose this between-subject design to avoid priming effects when switching conditions and because we expected planning and encoding processes to be influenced by the repeated presentation of the same picture stimuli [105]. (Participants groups did not differ in working memory capacity and executive function measures, cf. [S1 Text](#), [S1 Fig](#), [S2 Table](#)). Participants were instructed to describe the pictures spontaneously, to start speaking as early as possible, and to mention all depicted event participants. The spontaneous elicitation of picture descriptions [25,26] aimed at ensuring that participants planned their sentences in the most natural way possible. While the overall semantic content of responses in such a paradigm is guided by what was depicted, participants' freedom in how to respond also often leads to responses that do not conform to the target structures. The experiment proceeded in a self-paced manner so that participants initiated the start of the each trial by button press. Participants were given short breaks after approximately every 25 trials. The experiment lasted approximately 35 minutes. Eye tracking data were recorded using a Tobii TX-300 eye tracker (Tobii AB, Stockholm, Sweden) at a sampling rate of 250 Hz. EEG data were recorded with a 129 channel (128 + VREF) HydroCel Geodesic Sensor Net (Electrical Geodesics, Eugene, Oregon, United States of America), at a sampling rate of 500 Hz and amplified by a NET AMPS 400 amplifier.

Eye tracking analysis

Eye tracking data were processed in R [106]. Fixations were extracted from the eye tracker's raw samples [107], consecutive fixations to agent and patients in the pictures were subsumed into gazes [108] and, for the analyses, aggregated into 100-ms bins for each trial to reduce temporal autocorrelation [109]. Only transitive sentences in which participants named both agent and patient characters and used the targeted sentence structures were included for the analysis because in intransitives, there is only 1 character to fixate. Trials with response latencies larger than 6 seconds or longer than 2.5 SD from a participant's mean, as well as trials with first fixations to agent or patient characters later than 500 ms after stimulus onset or with track loss, were excluded. On balance, 1,552 trials were included (57.4% of all trials with transitive picture stimuli). The full exclusion criteria are described in [S1 Text](#).

Based on the literature [25,26,45,46] and visual inspection of the eye tracking graph, we determined the time window of interest to span from 200 to 800 ms after stimulus onset. As expected, very few language-related eye movements were observed before 200 ms [49,110]. Eye fixations to agent characters were analyzed on the single-trial level with logistic mixed-effects growth curve regression [52,111,112]. Based on visual inspection of the number of inflection points of the fixation curves in the analysis time window [52], linear, quadratic, cubic time terms, and their interactions with aligned versus nonaligned conditions were included as predictors. Additional variables and their interactions with the time terms were also included to control for their potential influence on the fixation time course. These were speech onset, naming agreement/codability, noun phrase length, trial number, as well as a number of stimulus properties (full details are given in [S1 Text](#)). The models also included random (group-level) effects for participants (intercept and slopes for time terms) and pictures (intercept and slopes for time terms, alignment type, and their interaction). The full model structure is described in [S1 Text](#). Inferences on the statistical significance of predictors are based on likelihood ratio tests. Fixation proportions in [Fig 2A](#) were smoothed with simple moving averages over 12 epochs (48 ms).

Electroencephalography (EEG) analysis

EEG data were preprocessed in EEGLAB (version 14.1.2, [113]), FieldTrip (version 20190107, [114]), and R (version 3.6.0, [106]). Continuous EEG data were re-referenced to the average of

the left and right mastoids offline, filtered (0.3 to 40 Hz band-pass filter), and down-sampled to a sampling rate of 250 Hz. Pauses between blocks were removed, and artifactual channels were automatically identified (by deviation of more than 5 SD from the mean of all channels in kurtosis or probability). The data were then decomposed into independent components (excluding rejected channels) and epoched from $-1,000$ to $1,500$ ms relative to stimulus picture onset. The SASICA and ADJUST algorithms [115,116] were used to identify artifactual components (number of rejected independent components: mean = 32.10, SD = 8.00). Previously rejected channels were spherically interpolated after artifactual components were subtracted from the EEG data.

The single-trial data were then transformed into a time–frequency representation (in 0.5-Hz steps) using wavelet decomposition (multi-taper method convolution) with Hanning-tapered time windows using a length of 3 cycles and advancing in 100-ms steps. Power was averaged separately for each frequency across channels into ROIs to reduce spatial autocorrelation. S7 Table shows the grouping of electrodes into ROIs. Power spectra were then transformed into dB relative to median power in a baseline period of -600 to -200 ms before stimulus picture onset in order to mitigate the $1/f$ problem and to make power in all frequencies directly comparable [117]. Finally, power was averaged within frequency bands. Frequency bands were defined individually, based on each participant's individual peak alpha frequency (IAF) [80]. We extracted IAFs [118] from resting state EEG recordings in which participants sat still with their eyes open and closed for 2 minutes each, recorded once directly before and once directly after the picture description experiment. The theta band was defined as ranging from IAF-6 to IAF-4 (on average 4.12 to 6.12 Hz, $w = 2.63$ to 8.07 Hz). The alpha band was defined as ranging from IAF-4 to IAF+2 (on average 6.12 to 12.12 Hz, $w = 4.63$ to 14.07 Hz). The IAF of 1 participant could not be computed due to excessive artifacts and was imputed with the median IAF of all other participants.

Throughout all analyses, we time-locked to the onset of the picture stimulus because time-locking the neural signals in a way to single out specific steps in the planning of sentences is not feasible. The average timing of these steps is potentially too variable, unlike the average timing of planning steps for the production of single words in isolation, which is well known [119]. This is in contrast to studies on sentence comprehension where the timing of external events (such as the presentation of individual words) can be precisely determined and thus allows the study of evoked potentials.

We included transitive and intransitive sentences in which participants named all picture characters and used the targeted sentence structures. We excluded epochs that were found to be contaminated by heavy artifacts (by visual inspection after independent components correction). To avoid muscle artifacts resulting from movements of the articulators during the 0- to 800-ms analysis time window, only epochs of trials in which participants began speaking later than 1,500 ms after picture onset and that did not contain any “pre-speech noises” (such as smacking lips or saying “uh”) were included. Thus, speech onset occurred at least 700 ms after the end of our analysis time window in all of the included trials. Spoken word encoding requires around 600 ms [119], speakers retrieve depicted characters' names immediately before uttering them [57] and preparations for speaking affect oscillatory power roughly up to 400 to 500 ms before speech onset [60]. Our analysis time window can therefore be expected to cover primarily relational-structural planning processes of utterances in which participants started speaking at 1,500 ms or later. On balance, 3,447 trials were included in the EEG analysis (58.9% of all trials).

The time course of power changes in the EEG theta and alpha frequency bands was analyzed with linear mixed-effects regression model trees as implemented in the R packages `glmertree` [53] and `lme4` [112]. Regression model trees recursively partitioned the data

into subsets to find the best-fitting model in each cell of these partitions, based on which subgroups showed statistically similar effects of ROI or sentence type (Figs 1 and 3A) or their interaction. The fixed effects structure was a fourth-order growth curve model on the time course of EEG power in a time window from 0 to 800 ms, analogously to the time window in the eye tracking analysis. The EEG time window started at 0 ms because electrophysiological responses to stimulus onset can occur almost immediately (unlike eye fixations). The choice of the order of polynomials was based on visual inspection of the shapes of the power curves. Subgroups in the tree were identified by their behavior with respect to the estimates for the regressors (intercept and polynomial time terms). On each iteration, the regression tree algorithm searched for differences in the robustness of model estimates (polynomial time courses, across ROIs and sentence types, conditioned by random effects; technically known as a parameter instabilities [53,120]). The data were then partitioned along the variable for which a split led to the greatest increase in fit on that iteration [120], i.e., for which the largest significant differences between subgroups were identified.

Random effects (intercepts and slopes for time terms by participant and by picture) and fixed effects of additional control variables were fitted globally. Analogously to the eye tracking model, the model included as fixed effects control variables of speech onset, naming agreement/codability, noun phrase length, trial number, as well as a number of stimulus properties (see S1 Text). Inferences on the statistical significance of splits were assessed through parameter instability tests during the tree fitting process [53,120].

Time courses of power changes in Figs 2B and 3B are smoothed with local polynomial regression (loess) with a span of 0.4. To visualize the extent of the observed differences in event-related power changes between aligned and nonaligned sentences for the theta band and between sentences with marked and unmarked noun phrases for the alpha band, we also plotted topographic maps with *t*-values (Figs 2D and 3D) across all scalp electrodes. For this plot, mean power during -600 ms to -200 ms relative to picture onset served as baseline.

Supporting information

S1 Fig. Individual differences measures for the 2 experimental groups, instructed to describe pictures in imperfective aspect (transitive sentences are aligned with intransitive sentences) vs. perfective aspect (transitive sentences are not aligned with intransitive sentences). (A) Mean partial-credit load scores for automated complex span tasks. (B) Mean congruency effect for Flanker task. (Underlying data and scripts are available from <https://osf.io/uhtcn/> and in the Supporting information file S1 Data.)

(TIFF)

S2 Fig. Distribution of speech onset latencies for intransitive (1 participant) and transitive (2 participants) sentences in the aligned (ongoing events in imperfective aspect) and non-aligned (completed events in perfective aspect) conditions. Speech onset latencies did not significantly differ between conditions (all *ps* > 0.13, see S3 Table). (Underlying data and scripts are available from <https://osf.io/uhtcn/> and in the Supporting information file S1 Data.)

(TIFF)

S3 Fig. Example 2-participant (transitive) stimulus picture. (A) Line drawing showing an agent (“gardener”) performing an action (“planting”) on a patient (“tree”); however, picture versions with gray backgrounds were presented during the experiment. (B) Scrambled gray background version of the stimulus picture with all pixels randomly redistributed over the screen display and black fixation square. (C) Gray background version of example stimulus.

(TIFF)

S4 Fig. Model tree for power changes (in dB) in individually defined theta frequency bands between 0 and 800 ms, showing the subsidiary splits. Scalp maps show electrode positions of ROIs included in the respective grouping (see [S5 Table](#) for details). All splits were statistically significant with $p < 0.003$. Tree tips show model fits for the respective grouping; ribbons represent 95% CIs. (Underlying data and scripts are available from <https://osf.io/uhtcn/> and in the Supporting information file [S1 Data](#).) ROI, region of interest.

(TIFF)

S5 Fig. Model tree for power changes (in dB) in individually defined alpha frequency bands between 0 and 800 ms, showing the subsidiary split between intransitive and transitive sentences with nominative marking in posterior ROIs. Scalp maps show electrode positions of ROIs included in the respective grouping (see [S5 Table](#) for details). All splits were statistically significant with $p < .02$. Tree tips show model fits for the respective grouping; ribbons represent 95% CIs. (Underlying data and scripts are available from <https://osf.io/uhtcn/> and in the Supporting information file [S1 Data](#).) ROI, region of interest.

(TIFF)

S1 Table. Example sentences from Hindi illustrating the split-ergative case marking system. More literal translations are “The gardener is crouching/had crouched” and “The gardener is planting a tree/had planted a tree.” Sentence types were chosen so that the overall syntactic structure matched as much as possible across conditions and all sentences consisted of 1 or 2 noun phrases, a verb, and an auxiliary. AUX, auxiliary; ERG, ergative case; IPFV, imperfective aspect; NOM, nominative case; PFV, perfective aspect; \emptyset , null (phonologically empty) morpheme.

(PDF)

S2 Table. Mean partial-credit load scores for automated complex span tasks and mean congruency effect for Flanker task (standard deviations in parentheses). (Underlying data and scripts used to generate graphs are available from <https://osf.io/uhtcn/>.)

(PDF)

S3 Table. Generalized gamma linear mixed-effects regression results modeling speech onset latencies (inverse link function). Statistical significance of predictors was assessed using likelihood ratio tests; significance was only assessed for the predictors of interest (sentence transitivity and alignment condition). (Underlying data, scripts, and models are available from <https://osf.io/uhtcn/>.)

(PDF)

S4 Table. Generalized (binomial) linear mixed-effects regression results modeling fixations on agent characters (successes) vs. fixations elsewhere in the pictures (failures) between 200 and 800 ms relative to picture stimulus onset. Statistical significance of predictors was assessed using likelihood ratio tests; significance was only assessed for predictors of interest. (Underlying data, scripts, and models are available from <https://osf.io/uhtcn/>.)

(PDF)

S5 Table. Linear mixed-effects regression tree results modeling EEG power dynamics (in dB) in individually defined theta frequency bands. Splits and predicted power time courses are shown in [S4 Fig](#). (Underlying data, scripts, and models are available from <https://osf.io/uhtcn/>.) EEG, electroencephalography.

(PDF)

S6 Table. Linear mixed-effects regression tree results modeling EEG power dynamics (in dB) in individually defined alpha frequency bands. Splits and predicted power time courses are shown in [S5 Fig](#). (Underlying data, scripts, and models are available from <https://osf.io/uhtcn/>). EEG, electroencephalography.
(PDF)

S7 Table. Grouping of electrodes of the HydroCel Geodesic Sensor Net into ROIs used for analysis of EEG event-related synchronization/de-synchronization, based on locations represented in FieldTrip; cf. the channel map provided by the manufacturer for another representation of channel locations (<https://drive.google.com/file/d/0B388xdH0Vxl2T3U4aS1oWDdPdTg/view>). EEG, electroencephalography; ROI, region of interest.
(PDF)

S1 Data. Data underlying the plots in Figs 2 and 3 and S1–S5 Figs.
(XLSX)

S1 Text. Supporting methods.
(PDF)

Acknowledgments

We thank Cyril Bosshard, Samina Fayyaz, and André Müller for help with processing the audio recordings; Anna Jancso and Lei He for programming help; and Prabir Sarkar for providing access to the eye tracker.

Author Contributions

Conceptualization: Sebastian Sauppe, Ina Bornkessel-Schlesewsky, Martin Meyer, Balthasar Bickel.

Data curation: Sebastian Sauppe.

Formal analysis: Sebastian Sauppe, Nathalie Giroud, Damián E. Blasi, Balthasar Bickel.

Funding acquisition: Sebastian Sauppe, Ina Bornkessel-Schlesewsky, Martin Meyer, Balthasar Bickel.

Investigation: Sebastian Sauppe, Kamal K. Choudhary, Shikha Bhattamishra, Mahima Gulati.

Methodology: Sebastian Sauppe.

Project administration: Sebastian Sauppe, Martin Meyer, Balthasar Bickel.

Resources: Kamal K. Choudhary, Martin Meyer, Balthasar Bickel.

Software: Sebastian Sauppe, Nathalie Giroud.

Supervision: Martin Meyer, Balthasar Bickel.

Validation: Sebastian Sauppe.

Visualization: Sebastian Sauppe, Balthasar Bickel.

Writing – original draft: Sebastian Sauppe, Martin Meyer, Balthasar Bickel.

Writing – review & editing: Sebastian Sauppe, Kamal K. Choudhary, Nathalie Giroud, Damián E. Blasi, Elisabeth Norcliffe, Aitor Egurtzegi, Ina Bornkessel-Schlesewsky, Martin Meyer, Balthasar Bickel.

References

1. Dixon RMW. Ergativity. *Language*. 1979; 55(1):59–138.
2. Bickel B. Grammatical relations typology. In: Song JJ, editor. *The Oxford Handbook of Linguistic Typology*. Oxford: Oxford University Press; 2010. p. 399–444.
3. Bickel B, Witzlack-Makarevich A, Choudhary KK, Schlesewsky M, Bornkessel-Schlesewsky I. The Neurophysiology of Language Processing Shapes the Evolution of Grammar: Evidence from Case Marking. *PLoS ONE*. 2015; 10(8):e0132819. <https://doi.org/10.1371/journal.pone.0132819> PMID: 26267884
4. Müller S. Grammatical theory: From transformational grammar to constraint-based approaches. 2nd ed. No. 1 in *Textbooks in Language Sciences*. Berlin: Language Science Press; 2018.
5. Levelt WJM. Spoken word production: A theory of lexical access. *Proc Natl Acad Sci U S A*. 2001; 98(23):13464–13471. <https://doi.org/10.1073/pnas.231459498> PMID: 11698690
6. Lee DK, Fedorenko E, Simon MV, Curry WT, Nahed BV, Cahill DP, et al. Neural encoding and production of functional morphemes in the posterior temporal lobe. *Nat Commun*. 2018; 9(1877). <https://doi.org/10.1038/s41467-018-04235-3> PMID: 29760465
7. Gleitman LR, January D, Nappa R, Trueswell JC. On the *give* and *take* between event apprehension and utterance formulation. *J Mem Lang*. 2007; 57(4):544–596. <https://doi.org/10.1016/j.jml.2007.01.007> PMID: 18978929
8. Kempen G, Hoenkamp E. An Incremental Procedural Grammar for Sentence Formulation. *Cogn Sci*. 1987; 11(2):201–258. [https://doi.org/10.1016/S0364-0213\(87\)80006-X](https://doi.org/10.1016/S0364-0213(87)80006-X)
9. Myachykov A, Scheepers C, Garrod S, Thompson D, Fedorova O. Syntactic flexibility and competition in sentence production: The case of English and Russian. *Q J Exp Psychol*. 2013; 66(8):1601–1619. <https://doi.org/10.1080/17470218.2012.754910> PMID: 23286507
10. Stallings LM, MacDonald MC, O'Seaghdha PG. Phrasal ordering constraints in sentence production: Phrase length and verb disposition in Heavy-NP Shift. *J Mem Lang*. 1998; 39(3):392–417. <https://doi.org/10.1006/jmla.1998.2586>
11. Comrie B. Alignment of Case Marking of Full Noun Phrases. In: Dryer MS, Haspelmath M, editors. *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology; 2013. Available from: <http://wals.info/>.
12. Haupt FS, Schlesewsky M, Roehm D, Friederici AD, Bornkessel-Schlesewsky I. The status of subject–object reanalyses in the language comprehension architecture. *J Mem Lang*. 2008; 59(1):54–96. <https://doi.org/10.1016/j.jml.2008.02.003>
13. Wang L, Schlesewsky M, Bickel B, Bornkessel-Schlesewsky I. Exploring the nature of the 'subject'-preference: Evidence from the online comprehension of simple sentences in Mandarin Chinese. *Lang Cogn Process*. 2009; 24(7):1180–1226. <https://doi.org/10.1080/01690960802159937>
14. Demiral ŞB, Schlesewsky M, Bornkessel-Schlesewsky I. On the universality of language comprehension strategies: Evidence from Turkish. *Cognition*. 2008; 106(1):484–500. <https://doi.org/10.1016/j.cognition.2007.01.008> PMID: 17336956
15. Bornkessel-Schlesewsky I, Schlesewsky M. Reconciling time, space and function: A new dorsal-ventral stream model of sentence comprehension. *Brain Lang*. 2013; 125(1):60–76. <https://doi.org/10.1016/j.bandl.2013.01.010> PMID: 23454075
16. Krebs J, Malaia E, Wilbur RB, Roehm D. Subject preference emerges as cross-modal strategy for linguistic processing. *Brain Res*. 2018; 1691:105–117. <https://doi.org/10.1016/j.brainres.2018.03.029> PMID: 29627484
17. Pickering MJ, Garrod S. An integrated theory of language production and comprehension. *Behav Brain Sci*. 2013; 36(4):329–347. <https://doi.org/10.1017/S0140525X12001495> PMID: 23789620
18. Silbert LJ, Honey CJ, Simony E, Poeppel D, Hasson U. Coupled neural systems underlie the production and comprehension of naturalistic narrative speech. *Proc Natl Acad Sci U S A*. 2014; 111(43):E4687–E4696. <https://doi.org/10.1073/pnas.1323812111> PMID: 25267658
19. Segaert K, Menenti L, Weber K, Petersson KM, Hagoort P. Shared Syntax in Language Production and Language Comprehension—An fMRI Study. *Cereb Cortex*. 2012; 22(7):1662–1670. <https://doi.org/10.1093/cercor/bhr249> PMID: 21934094
20. Hawkins JA. *Cross-Linguistic Variation and Efficiency*. Oxford: Oxford University Press; 2014.
21. Futrell R, Levy RP, Gibson E. Dependency locality as an explanatory principle for word order. *Language*. 2020; 96(2):371–412. <https://doi.org/10.1353/lan.2020.0024>
22. Sinnemäki K. Cognitive processing, language typology, and variation. *Wiley Interdiscip Rev Cogn Sci*. 2014; 5(4):477–487. <https://doi.org/10.1002/wcs.1294> PMID: 26308657

23. Gibson E, Futrell R, Piantadosi ST, Dautriche I, Mahowald K, Bergen L, et al. How Efficiency Shapes Human Language. *Trends Cogn Sci*. 2019; 23(5):389–407. <https://doi.org/10.1016/j.tics.2019.02.003> PMID: 31006626
24. Christiansen MH, Chater N. Language as shaped by the brain. *Behav Brain Sci*. 2008; 31(5):489–558. <https://doi.org/10.1017/S0140525X08004998> PMID: 18826669
25. Norcliffe E, Konopka AE. Vision and Language in Cross-Linguistic Research on Sentence Production. In: Mishra RK, Srinivasan N, Huettig F, editors. *Attention and Vision in Language Processing*. New Delhi: Springer; 2015. p. 77–96.
26. Griffin ZM, Bock K. What the eyes say about speaking. *Psychol Sci*. 2000; 11(4):274–279. <https://doi.org/10.1111/1467-9280.00255> PMID: 11273384
27. Sauppe S. Word order and voice influence the timing of verb planning in German sentence production. *Front Psychol*. 2017; 8(1648). <https://doi.org/10.3389/fpsyg.2017.01648> PMID: 29018379
28. Hwang H, Kaiser E. The role of the verb in grammatical function assignment in English and Korean. *J Exp Psychol Learn Mem Cogn*. 2014; 40(5):1363–1376. <https://doi.org/10.1037/a0036797> PMID: 24884649
29. Pfurtscheller G, Lopes da Silva FH. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin Neurophysiol*. 1999; 110(11):1842–1857. [https://doi.org/10.1016/s1388-2457\(99\)00141-8](https://doi.org/10.1016/s1388-2457(99)00141-8) PMID: 10576479
30. Buzsáki G, Draguhn A. Neuronal Oscillations in Cortical Networks. *Science*. 2004; 304(5679):1926–1929. <https://doi.org/10.1126/science.1099745> PMID: 15218136
31. Roehm D, Schlesewsky M, Bornkessel I, Frisch S, Haider H. Fractionating language comprehension via frequency characteristics of the human EEG. *Cogn Neurosci Neuropsychol*. 2004; 15(3):409–412. <https://doi.org/10.1097/00001756-200403010-00005> PMID: 15094493
32. Bastiaansen M, van Berkum JJA, Hagoort P. Syntactic Processing Modulates the θ Rhythm of the Human EEG. *Neuroimage*. 2002; 17(3):1479–1492. <https://doi.org/10.1006/nimg.2002.1275> PMID: 12414287
33. Bastiaansen MCM, van Berkum JJA, Hagoort P. Event-related theta power increases in the human EEG during online sentence processing. *Neurosci Lett*. 2002; 323(1):13–16. [https://doi.org/10.1016/s0304-3940\(01\)02535-6](https://doi.org/10.1016/s0304-3940(01)02535-6) PMID: 11911979
34. Bastiaansen M, Magyari L, Hagoort P. Syntactic Unification Operations Are Reflected in Oscillatory Dynamics during On-line Sentence Comprehension. *J Cogn Neurosci*. 2010; 22(7):1333–1347. <https://doi.org/10.1162/jocn.2009.21283> PMID: 19580386
35. Kielar A, Panamsky L, Links KA, Meltzer JA. Localization of electrophysiological responses to semantic and syntactic anomalies in language comprehension with MEG. *Neuroimage*. 2015; 105:507–524. <https://doi.org/10.1016/j.neuroimage.2014.11.016> PMID: 25463470
36. Meyer L, Grigutsch M, Schmuck N, Gaston P, Friederici AD. Frontal–posterior theta oscillations reflect memory retrieval during sentence comprehension. *Cortex*. 2015; 71:205–218. <https://doi.org/10.1016/j.cortex.2015.06.027> PMID: 26233521
37. Weiss S, Mueller HM, Schack B, King JW, Kutas M, Rappelsberger P. Increased neuronal communication accompanying sentence comprehension. *Int J Psychophysiol*. 2005; 57(2):129–141. <https://doi.org/10.1016/j.ijpsycho.2005.03.013> PMID: 15935501
38. Röhm D, Klimesch W, Haider H, Doppelmayr M. The role of theta and alpha oscillations for language comprehension in the human electroencephalogram. *Neurosci Lett*. 2001; 310(2–3):137–140. [https://doi.org/10.1016/s0304-3940\(01\)02106-1](https://doi.org/10.1016/s0304-3940(01)02106-1) PMID: 11585586
39. Cross ZR, Kohler MJ, Schlesewsky M, Gaskell MG, Bornkessel-Schlesewsky I. Sleep-Dependent Memory Consolidation and Incremental Sentence Comprehension: Computational Dependencies during Language Learning as Revealed by Neuronal Oscillations. *Front Hum Neurosci*. 2018; 12(). <https://doi.org/10.3389/fnhum.2018.00018> PMID: 29445333
40. Davidson DJ, Indefrey P. An inverse relation between event-related and time–frequency violation responses in sentence processing. *Brain Res*. 2007; 1158:81–92. <https://doi.org/10.1016/j.brainres.2007.04.082> PMID: 17560965
41. Lam NHL, Schoffelen JM, Uddén J, Hultén A, Hagoort P. Neural activity during sentence processing as reflected in theta, alpha, beta, and gamma oscillations. *Neuroimage*. 2016; 142:43–54. <https://doi.org/10.1016/j.neuroimage.2016.03.007> PMID: 26970187
42. Vassileiou B, Meyer L, Beese C, Friederici AD. Alignment of alpha-band desynchronization with syntactic structure predicts successful sentence comprehension. *Neuroimage*. 2018; 175:286–296. <https://doi.org/10.1016/j.neuroimage.2018.04.008> PMID: 29627592
43. David O, Kilner JM, Friston KJ. Mechanisms of evoked and induced responses in MEG/EEG. *Neuroimage*. 2006; 31(4):1580–1591. <https://doi.org/10.1016/j.neuroimage.2006.02.034> PMID: 16632378

44. Schneider JM, Maguire MJ. Identifying the relationship between oscillatory dynamics and event-related responses. *Int J Psychophysiol.* 2018; 133:182–192. <https://doi.org/10.1016/j.ijpsycho.2018.07.002> PMID: 29981766
45. Sauppe S, Norcliffe EJ, Konopka AE, Van Valin RD, Levinson SC. Dependencies First: Eye Tracking Evidence from Sentence Production in Tagalog. In: Knauff M, Pauen M, Sebanz N, Wachsmuth I, editors. *Proceedings of the 35th Annual Conference of the Cognitive Science Society.* Austin, TX: Cognitive Science Society; 2013. p. 1265–1270.
46. Norcliffe E, Konopka AE, Brown P, Levinson SC. Word order affects the time course of sentence formulation in Tzeltal. *Lang Cogn Neurosci.* 2015; 30(9):1187–1208. <https://doi.org/10.1080/23273798.2015.1006238>
47. Riddle J, Scimeca JM, Cellier D, Dhanani S, D'Esposito M. Causal Evidence for a Role of Theta and Alpha Oscillations in the Control of Working Memory. *Curr Biol.* 2020; 30(9):1748–1754.E4. <https://doi.org/10.1016/j.cub.2020.02.065> PMID: 32275881
48. Griffin ZM. Why Look? Reasons for Eye Movements Related to Language Production. In: Henderson JM, Ferreira F, editors. *The interface of language, vision, and action: Eye movements and the visual world.* New York / Hove: Psychology Press; 2004. p. 213–247.
49. Rayner K. Eye movements and attention in reading, scene perception, and visual search. *Q J Exp Psychol (Hove).* 2009; 62(8):1457–1506. <https://doi.org/10.1080/17470210902816461> PMID: 19449261
50. Krause CM, Sillanmäki L, Koivisto M, Saarela C, Häggqvist A, Laine M, et al. The effects of memory load on event-related EEG desynchronization and synchronization. *Clin Neurophysiol.* 2000; 111(11):2071–2078. [https://doi.org/10.1016/s1388-2457\(00\)00429-6](https://doi.org/10.1016/s1388-2457(00)00429-6) PMID: 11068244
51. Klimesch W, Doppelmayr M, Pachinger T, Russegger H. Event-related desynchronization in the alpha band and the processing of semantic information. *Cogn Brain Res.* 1997; 6(2):83–94. [https://doi.org/10.1016/s0926-6410\(97\)00018-9](https://doi.org/10.1016/s0926-6410(97)00018-9) PMID: 9450602
52. Mirman D, Dixon JA, Magnuson JS. Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *J Mem Lang.* 2008; 59(4):475–494. <https://doi.org/10.1016/j.jml.2007.11.006> PMID: 19060958
53. Fokkema M, Smits N, Zeileis A, Hothorn T, Kelderman H. Detecting treatment-subgroup interactions in clustered data with generalized linear mixed-effects model trees. *Behav Res Methods.* 2018; 50(5):2016–2034. <https://doi.org/10.3758/s13428-017-0971-x> PMID: 29071652
54. Pickering MJ, Ferreira VS. Structural Priming: A Critical Review. *Psychol Bull.* 2008; 134(3):427–459. <https://doi.org/10.1037/0033-2909.134.3.427> PMID: 18444704
55. Segaert K, Wheeldon LR, Hagoort P. Unifying structural priming effects on syntactic choices and timing of sentence generation. *J Mem Lang.* 2016; 91:59–80. <https://doi.org/10.1016/j.jml.2016.03.011>
56. Levelt WJM. Producing spoken language: a blueprint of the speaker. In: Brown CM, Hagoort P, editors. *The Neurocognition of Language.* Oxford: Oxford University Press; 1999. p. 83–122.
57. Griffin ZM. Gaze durations during speech reflect word selection and phonological encoding. *Cognition.* 2001; 82(1):B1–B14. [https://doi.org/10.1016/s0010-0277\(01\)00138-x](https://doi.org/10.1016/s0010-0277(01)00138-x) PMID: 11672707
58. Jaeger TF, Furth K, Hilliard C. Incremental phonological encoding during unscripted sentence production. *Front Psychol.* 2012; 3(481). <https://doi.org/10.3389/fpsyg.2012.00481> PMID: 23162515
59. Wheeldon LR, Konopka AE. Spoken Word Production: Representation, retrieval, and integration. In: Rueschemeyer SA, Gaskell MG, editors. *The Oxford handbook of psycholinguistics.* 2nd ed. Oxford: Oxford University Press; 2018. p. 335–371.
60. Piai V, Roelofs A, Maris E. Oscillatory brain responses in spoken word production reflect lexical frequency and sentential constraint. *Neuropsychologia.* 2014; 53:146–456. <https://doi.org/10.1016/j.neuropsychologia.2013.11.014> PMID: 24291513
61. Whitham EM, Pope KJ, Fitzgibbon SP, Lewis TW, Clark CR, Loveless S, et al. Scalp electrical recording during paralysis: Quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG. *Clin Neurophysiol.* 2007; 118(8):1877–1888. <https://doi.org/10.1016/j.clinph.2007.04.027> PMID: 17574912
62. Piai V, Zheng X. Speaking waves: neuronal oscillations in language production. In: Federmeier KD, editor. *Psychology of Learning and Motivation.* vol. 71. Academic Press; 2019. p. 265–302.
63. Webb A, Knott A, MacAskill MR. Eye movements during transitive action observation have sequential structure. *Acta Psychol (Amst).* 2010; 133(1):51–56. <https://doi.org/10.1016/j.actpsy.2009.09.001> PMID: 19854431
64. Meyer L. The neural oscillations of speech processing and language comprehension: state of the art and emerging mechanisms. *Eur J Neurosci.* 2018; 48(7):10.1111/ejn.13748. <https://doi.org/10.1111/ejn.13748> PMID: 29055058

65. Sauseng P, Griesmayr B, Freunberger R, Klimesch W. Control mechanisms in working memory: A possible function of EEG theta oscillations. *Neurosci Biobehav Rev.* 2010; 34(7):1015–1022. <https://doi.org/10.1016/j.neubiorev.2009.12.006> PMID: 20006645
66. Karrasch M, Laine M, Rapinoja P, Krause CM. Effects of normal aging on event-related desynchronization/synchronization during a memory task in humans. *Neurosci Lett.* 2004; 366(1):18–23. <https://doi.org/10.1016/j.neulet.2004.05.010> PMID: 15265582
67. Jensen O, Tesche CD. Frontal theta activity in humans increases with memory load in a working memory task. *Eur J Neurosci.* 2002; 15(8):1395–1399. <https://doi.org/10.1046/j.1460-9568.2002.01975.x> PMID: 11994134
68. Momma S, Ferreira VS. Beyond linear order: The role of argument structure in speaking. *Cogn Psychol.* 2019; 114:101228. <https://doi.org/10.1016/j.cogpsych.2019.101228> PMID: 31421521
69. Momma S, Slevc LR, Phillips C. The timing of verb selection in Japanese sentence production. *J Exp Psychol Learn Mem Cogn.* 2016; 42(5):813–824. <https://doi.org/10.1037/xlm0000195> PMID: 26569434
70. Momma S, Slevc LR, Phillips C. Unaccusativity in sentence production. *Linguist Inq.* 2018; 49(1):181–194. https://doi.org/10.1162/LING_a_00271
71. Norcliffe E, Harris AC, Jaeger TF. Cross-linguistic psycholinguistics and its critical role in theory development: Early beginnings and recent advances. *Lang Cogn Neurosci.* 2015; 30(9):1009–1032. <https://doi.org/10.1080/23273798.2015.1080373>
72. Sauppe S. Symmetrical and asymmetrical voice systems and processing load: Pupillometric evidence from sentence production in Tagalog and German. *Language.* 2017; 93(2):288–313. <https://doi.org/10.1353/lan.2017.0015>
73. Brown-Schmidt S, Konopka AE. Little houses and casas pequeñas: Message formulation and syntactic form in unscripted speech with speakers of English and Spanish. *Cognition.* 2008; 109(2):274–280. <https://doi.org/10.1016/j.cognition.2008.07.011> PMID: 18842259
74. Jaeger TF, Norcliffe EJ. The Cross-linguistic Study of Sentence Production. *Lang Linguist Compass.* 2009; 3(4):866–887. <https://doi.org/10.1111/j.1749-818x.2009.00147.x>
75. Klimesch W, Sauseng P, Hanslmayr S. EEG alpha oscillations: The inhibition-timing hypothesis. *Brain Res Rev.* 2007; 53(1):63–88. <https://doi.org/10.1016/j.brainresrev.2006.06.003> PMID: 16887192
76. Jensen O, Mazaheri A. Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Front Hum Neurosci.* 2010; 4(186). <https://doi.org/10.3389/fnhum.2010.00186> PMID: 21119777
77. Just MA, Carpenter PA. A Theory of Reading: From Eye Fixations to Comprehension. *Psychol Rev.* 1980; 87(4):329–354. <https://doi.org/10.1037/0033-295X.87.4.329> PMID: 7413885
78. Bögels S, Magyari L, Levinson SC. Neural signatures of response planning occur midway through an incoming question in conversation. *Sci Rep.* 2015; 5(12881). <https://doi.org/10.1038/srep12881> PMID: 26242909
79. Bögels S, Casillas M, Levinson SC. Planning versus comprehension in turn-taking: Fast responders show reduced anticipatory processing of the question. *Neuropsychologia.* 2018; 109:295–310. <https://doi.org/10.1016/j.neuropsychologia.2017.12.028> PMID: 29269305
80. Klimesch W. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res Rev.* 1999; 29(2–3):169–195. [https://doi.org/10.1016/s0165-0173\(98\)00056-3](https://doi.org/10.1016/s0165-0173(98)00056-3) PMID: 10209231
81. Meyer L, Sun Y, Martin AE. Synchronous, but not entrained: exogenous and endogenous cortical rhythms of speech and language processing. *Lang Cogn Neurosci.* 2020; 35(9):1089–1099. <https://doi.org/10.1080/23273798.2019.1693050>
82. Giraud AL, Kleinschmidt A, Poeppel D, Lund TE, Frackowiak RSJ, Laufs H. Endogenous Cortical Rhythms Determine Cerebral Specialization for Speech Perception and Production. *Neuron.* 2007; 56(6):1127–1134. <https://doi.org/10.1016/j.neuron.2007.09.038> PMID: 18093532
83. Ding N, Melloni L, Zhang H, Tian X, Poeppel D. Cortical tracking of hierarchical linguistic structures in connected speech. *Nat Neurosci.* 2016; 19(1):158–164. <https://doi.org/10.1038/nn.4186> PMID: 26642090
84. MacDonald MC. How language production shapes language form and comprehension. *Front Psychol.* 2013; 4(226). <https://doi.org/10.3389/fpsyg.2013.00226> PMID: 23637689
85. Jaeger TF, Tily HJ. On language ‘utility’: processing complexity and communicative efficiency. *Wiley Interdiscip Rev Cogn Sci.* 2011; 2(3):323–335. <https://doi.org/10.1002/wcs.126> PMID: 26302080
86. Branigan HP, Pickering MJ, Tanaka MN. Contributions of animacy to grammatical function assignment and word order during production. *Lingua.* 2008; 118(2):172–189. <https://doi.org/10.1016/j.lingua.2007.02.003>

87. Christianson KT, Ferreira F. Conceptual accessibility and sentence production in a free word order language (Odawa). *Cognition*. 2005; 98(2):105–135. <https://doi.org/10.1016/j.cognition.2004.10.006> PMID: 16307955
88. Ferreira VS. Is It Better to Give Than to Donate? Syntactic Flexibility in Language Production. *J Mem Lang*. 1996; 35(5):724–755. <https://doi.org/10.1006/jmla.1996.0038>
89. Gennari SP, Mirković J, MacDonald MC. Animacy and competition in relative clause production: A cross-linguistic investigation. *Cogn Psychol*. 2012; 65(2):141–176. <https://doi.org/10.1016/j.cogpsych.2012.03.002> PMID: 22537914
90. Humphreys GF, Mirković J, Gennari SP. Similarity-based competition in relative clause production and comprehension. *J Mem Lang*. 2016; 89:200–221. <https://doi.org/10.1016/j.jml.2015.12.007>
91. Ferreira VS. Ambiguity, Accessibility, and a Division of Labor for Communicative Success. In: Ross BH, editor. *Advances in Research and Theory*. vol. 49 of *Psychology of Learning and Motivation*. London: Academic Press; 2008. p. 209–246.
92. Levinson SC, Torreira F. Timing in turn-taking and its implications for processing models of language. *Front Psychol*. 2015; 6(731). <https://doi.org/10.3389/fpsyg.2015.00731> PMID: 26124727
93. Barthel M, Sauppe S, Levinson SC, Meyer AS. The timing of utterance planning in task-oriented dialogue: Evidence from a novel list-completion paradigm. *Front Psychol*. 2016; 7(1858). <https://doi.org/10.3389/fpsyg.2016.01858> PMID: 27990127
94. Frisch S, Schlesewsky M, Saddy D, Alpermann A. The P600 as an indicator of syntactic ambiguity. *Cognition*. 2002; 85(3):B83–B92. [https://doi.org/10.1016/s0010-0277\(02\)00126-9](https://doi.org/10.1016/s0010-0277(02)00126-9) PMID: 12169415
95. Frazier L, Rayner K. Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cogn Psychol*. 1982; 14(2):178–210. [https://doi.org/10.1016/0010-0285\(82\)90008-1](https://doi.org/10.1016/0010-0285(82)90008-1)
96. Blevins J. *Evolutionary Phonology: The Emergence of Sound Patterns*. Cambridge: Cambridge University Press; 2004.
97. Hawkins JA. *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press; 2004.
98. Futrell R, Mahowald K, Gibson E. Large-scale evidence of dependency length minimization in 37 languages. *Proc Natl Acad Sci U S A*. 2015; 112(33):10336–10341. <https://doi.org/10.1073/pnas.1502134112> PMID: 26240370
99. Piai V, Roelofs A, Rommers J, Maris E. Beta oscillations reflect memory and motor aspects of spoken word production. *Hum Brain Mapp*. 2015; 36(7):2767–2780. <https://doi.org/10.1002/hbm.22806> PMID: 25872756
100. Eulitz C, Hauk O, Cohen R. Electroencephalographic activity over temporal brain areas during phonological encoding in picture naming. *Clin Neurophysiol*. 2000; 111(11):2088–2097. [https://doi.org/10.1016/s1388-2457\(00\)00441-7](https://doi.org/10.1016/s1388-2457(00)00441-7) PMID: 11068246
101. Bornkessel-Schlesewsky I, Schlesewsky M. The importance of linguistic typology for the neurobiology of language. *Linguist Typol*. 2016; 20(3):615–621. <https://doi.org/10.1515/lingty-2016-0032>
102. Henrich J, Heine SJ, Norenzayan A. The weirdest people in the world? *Behav Brain Sci*. 2010; 33:61–135. <https://doi.org/10.1017/S0140525X0999152X> PMID: 20550733
103. Seifart F, Strunk J, Danielsen S, Hartmann I, Pakendorf B, Wichmann S, et al. Nouns slow down speech across structurally and culturally diverse languages. *Proc Natl Acad Sci U S A*. 2018; 115(22):5720–5725. <https://doi.org/10.1073/pnas.1800708115> PMID: 29760059
104. Konopka AE, Meyer AS. Priming sentence planning. *Cogn Psychol*. 2014; 73:1–40. <https://doi.org/10.1016/j.cogpsych.2014.04.001> PMID: 24838190
105. Konopka AE, Kuchinsky SE. How message similarity shapes the timecourse of sentence formulation. *J Mem Lang*. 2015; 84:1–23. <https://doi.org/10.1016/j.jml.2015.04.003>
106. R Core Team. *R: A language and environment for statistical computing*; 2018. Available from: <http://www.R-project.org>.
107. Engbert R, Kliegl R. Microsaccades uncover the orientation of covert attention. *Vision Res*. 2003; 43(9):1035–1045. [https://doi.org/10.1016/s0042-6989\(03\)00084-1](https://doi.org/10.1016/s0042-6989(03)00084-1) PMID: 12676246
108. Griffin ZM, Davison JC. A technical introduction to using speakers' eye movements to study language. *Ment Lex*. 2011; 6(1):53–82. <https://doi.org/10.1075/ml.6.1.03gr>
109. Cho SJ, Brown-Schmidt S, Lee Wy. Autoregressive Generalized Linear Mixed Effect Models with Crossed Random Effects: An Application to Intensive Binary Time Series Eye-Tracking Data. *Psychometrika*. 2018; 83(3):751–771. <https://doi.org/10.1007/s11336-018-9604-2> PMID: 29417454
110. Allopenna PD, Tanenhaus MK. Tracking the Time Course of Spoken Word Recognition Using Eye Movements: Evidence for Continuous Mapping Models. *J Mem Lang*. 1998; 38(4):419–439. <https://doi.org/10.1006/jmla.1997.2558>

111. Donnelly S, Verkuilen J. Empirical logit analysis is not logistic regression. *J Mem Lang*. 2017; 94:28–42. <https://doi.org/10.1016/j.jml.2016.10.005>
112. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. *J Stat Softw*. 2015; 67(1):1–48. <https://doi.org/10.18637/jss.v067.i01>
113. Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods*. 2004; 134(1):9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009> PMID: 15102499
114. Oostenveld R, Fries P, Maris E, Schoffelen JM. FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Comput Intell Neurosci*. 2011; 2011(156869). <https://doi.org/10.1155/2011/156869> PMID: 21253357
115. Chaumon M, Bishop DVM, Busch NA. A practical guide to the selection of independent components of the electroencephalogram for artifact correction. *J Neurosci Methods* 2015; 250:47–63. <https://doi.org/10.1016/j.jneumeth.2015.02.025> PMID: 25791012
116. Mognon A, Jovicich J, Bruzzone L, Buiatti M. ADJUST: An Automatic EEG artifact Detector based on the Joint Use of Spatial and Temporal features. *Psychophysiology*. 2011; 48(2):229–240. <https://doi.org/10.1111/j.1469-8986.2010.01061.x> PMID: 20636297
117. Cohen MX. *Analyzing Neural Time Series Data: Theory and Practice*. Cambridge, MA / London: MIT Press; 2014.
118. Corcoran AW, Alday PM, Schlesewsky M, Bornkessel-Schlesewsky I. Towards a reliable, automated method of individual alpha frequency (IAF) quantification. *Psychophysiology*. 2018; 55(7):e13064. <https://doi.org/10.1111/psyp.13064> PMID: 29357113
119. Indefrey P. The spatial and temporal signatures of word production components: a critical update. *Front Psychol*. 2011; 2:255. <https://doi.org/10.3389/fpsyg.2011.00255> PMID: 22016740
120. Zeileis A, Hothorn T, Hornik K. Model-Based Recursive Partitioning. *J Comput Graph Stat*. 2008; 17(2):492–514. <https://doi.org/10.1198/106186008X319331>