Sentence planning and production in Murrinhpatha, an Australian 'free word order' language

Rachel Nordlinger, Gabriela Garrido Rodriguez, Evan Kidd

SENTENCE PLANNING AND PRODUCTION IN MURRINHPATHA, AN AUSTRALIAN 'FREE WORD ORDER' LANGUAGE

RACHEL NORDLINGER   GABRIELA GARRIDO RODRIGUEZ   EVAN KIDD

University of Melbourne and ARC Centre of Excellence for the Dynamics of Language

University of Melbourne, Max Planck Institute for Psycholinguistics, and ARC Centre of Excellence for the Dynamics of Language

Max Planck Institute for Psycholinguistics, The Australian National University, and ARC Centre of Excellence for the Dynamics of Language

Psycholinguistic theories are based on a very small set of unrepresentative languages, so it is as yet unclear how typological variation shapes mechanisms supporting language use. In this article we report the first on-line experimental study of sentence production in an Australian free word order language: Murrinhpatha. Forty-six adult native speakers of Murrinhpatha described a series of unrelated transitive scenes that were manipulated for humanness (±human) in the agent and patient roles while their eye movements were recorded. Speakers produced a large range of word orders, consistent with the language having flexible word order, with variation significantly influenced by agent and patient humanness. An analysis of eye movements showed that Murrinhpatha speakers' first fixation on an event character did not alone determine word order; rather, early in speech planning participants rapidly encoded BOTH event characters and their relationship to each other. That is, they engaged in RELATIONAL ENCODING, laying down a very early conceptual foundation for the word order they eventually produced. These results support a WEAKLY HIERARCHICAL account of sentence production and show that speakers of a free word order language encode the relationships between event participants during earlier stages of sentence planning than is typically observed for languages with fixed word orders.*

Keywords: sentence planning, sentence production, Australian languages, free word order, conceptual accessibility, eye-tracking, typology

1. INTRODUCTION. The world's 7,000 or so languages exhibit a large degree of typological diversity (Evans & Levinson 2009, Dryer & Haspelmath 2013), reflecting both the architectural constraints of the language faculty and millennia of language evolution and change across the history of our species. While languages significantly differ, we are born with the same neural structure to support language acquisition and use, which must be flexible enough to navigate typological differences (Bornkessel-Schlesewsky & Schlesewsky 2016). One major theoretical question in psycholinguistic research therefore is: to what extent does the grammatical structure of an individual language impact the way it is processed (Norcliffe, Konopka, et al. 2015)? In the current article we investigate one prominent typological feature of a number of Australian Indigenous languages—so-called 'free' word order (Hale 1983, Austin & Bresnan 1996)—and ask whether speaking a free word order language has implications for sentence planning and production. In doing so we present the first sentence-production study of an Aus-

tralian Indigenous language and begin to address the persistent sampling bias in psycholinguistic research (Jaeger & Norcliffe 2009) and in cognitive science in general (Henrich et al. 2010, Schulz et al. 2018). Our study thus contributes to the growing body of psycholinguistic research that explores the relationship between language typology and language production (e.g. Norcliffe, Harris, & Jaeger 2015, Sauppe 2017a).

1.1. SENTENCE PRODUCTION. Sentence production involves a complex process of mapping an abstract event representation to a linear signal via language-specific grammatical machinery (Levelt 1989, Bock & Levelt 1994, Bock & Ferreira 2014). While there is broad agreement that the process is INCREMENTAL, theoretical debate concerns the degree of incrementality in the system and the information used to guide structural choice. Specifically, there is significant debate about whether speakers build sentences one concept and word at a time, such that production proceeds in a sequential fashion, or whether speakers are guided by a holistic construal of an event that then guides the mapping of concepts to syntactic roles. These two proposals have been commonly referred to as LINEAR INCREMENTAL and HIERARCHICAL INCREMENTAL planning, respectively (e.g. Griffin & Bock 2000, Gleitman et al. 2007, Konopka & Brown-Schmidt 2014).

Linear incremental planning contends that production can proceed in small conceptual and linguistic units, such that the selection of a starting point can have significant effects on structure building. For example, Tomlin (1997) showed English-speaking participants short cartoons of fish eating one another. Participants were asked to track one fish marked by an arrow, an overt cue that participants used when describing eating events. Notably, the cue served to increase the accessibility of the referent, such that it was more likely to serve as the sentence subject even in trials where the highlighted fish was the patient, thus leading to the production of a dispreferred passive (e.g. *the red fish was eaten by …*). Such perceptual cues need not be overt. In a picture-description task utilizing eye-tracking, Gleitman et al. (2007) subliminally primed English-speaking participants' attention to one referent in two-referent scenes using a briefly presented cue, and found that the cued referent was significantly more likely to be mentioned first across a range of structures (e.g. conjoined NPs, active-passive alternations). These data suggest that nonlinguistic perceptual processes can influence referent accessibility and, in languages such as English, can influence sentence formulation—speakers systematically place more accessible referents in more prominent sentence locations, and may begin an utterance by encoding a single accessible referent before they have a full conceptual representation of an event (see also Myachykov et al. 2018).

In eye-tracking paradigms, evidence in favor of linear and hierarchical planning comes from specific signatures in the eye-movement record. Correlations between first fixations and the first-mentioned NP have been claimed to be consistent with linear incremental planning (Gleitman et al. 2007, Konopka & Meyer 2014). In contrast, early RELATIONAL ENCODING, where fixations are distributed across referents within the early EVENT APPREHENSION window (approximately 0–600 ms) and during the beginnings of LINGUISTIC ENCODING (approximately 600 ms onward), is consistent with hierarchical planning, since it suggests that agents, patients, and their interrelationship are encoded in an early holistic representation of the event. Eye-tracking evidence for linear incremental planning comes almost exclusively from SVO languages that mark grammatical relations primarily through word order: that is, English (Gleitman et al. 2007) and Dutch (Konopka & Meyer 2014). Even within these languages, the degree to which speakers engage in linear or hierarchical planning appears to depend on contextual factors: Konopka and Meyer (2014) reported evidence for a flexible use of both strategies in Dutch-speaking individuals (for evidence of relational encoding in Dutch see Konopka 2019).

Eye-tracking studies of dependent-marking and verb-initial languages, which place higher demands on the need for encoding all core event roles before production begins, provide evidence in favor of hierarchical planning, although in these languages the evidence for relational encoding has typically been observed after event apprehension, during linguistic encoding (Sauppe et al. 2013, Hwang & Kaiser 2014, Norcliffe, Konopka, et al. 2015, Sauppe 2017a,b). Norcliffe, Konopka, et al. (2015) conducted a sentence-production experiment with Tseltal-speaking adults, comparing the data to those from the same experiment with Dutch-speaking adults. The participants were asked to produce descriptions of transitive events, where the agent and patient roles were manipulated for humanness (i.e. ±human), while their eye movements to the images were recorded. Human entities are more accessible than referents lower on the animacy hierarchy, and thus are typically given syntactic prominence, such as being chosen as a sentence subject and/or being placed early in the sentence (e.g. Silverstein 1976, Aissen 1999). A hierarchical planning account predicts that human agents should appear more frequently as subjects in active sentences and human patients should appear more frequently as subjects in passives, whereas a linear incremental account predicts that more prominent (i.e. human) entities should appear in early sentence positions (see Tanaka et al. 2011). In languages like Dutch, subjecthood and sentence position are confounded. However, Tseltal has VOS canonical word order, but allows a more infrequent SVO option, with both word orders allowing passivization. Thus, the influence of conceptual accessibility (i.e. referent humanness) on production speaks directly to the hierarchical vs. linear planning debate. Additionally, because Tseltal is canonically both verb-initial and head-marking, the authors predicted a pattern of relational encoding for verb-initial sentences, such that the participants' fixations should be more evenly distributed between the agent and patient referents during event apprehension and the early stages of linguistic encoding than in SVO sentences.

In both cases the results were consistent with hierarchical planning. First, the conceptual accessibility of referents influenced subject selection independent of sentence position. Specifically, Tseltal speakers had a preference for human subjects. Second, whereas planning processes for SVO sentences resembled planning processes in other SVO languages like English and Dutch—namely sustained looks to the agent followed by looks to the patient around speech onset (Griffin & Bock 2000, Gleitman et al. 2007, Kuchinsky & Bock 2010)—planning in VOS sentences showed more distributed looks to the agent and patient during linguistic encoding, suggesting that relational encoding necessarily underlies planning when the grammar requires the early production of a predicate.

These results are consistent with results reported by Sauppe et al. (2013), who conducted the same study as Norcliffe, Konopka, et al. (2015) in another verb-initial language, Tagalog (see also Sauppe 2017a). Tagalog has a symmetrical voice system (Foley 2008), such that sentences in the agent and patient/undergoer voice are equally transitive, with thematic roles marked via verbal and nominal morphology. Predicates are morphologically marked for voice through affixation, with voice type identifying a syntactically prominent argument marked on nouns (the 'privileged syntactic argument'). Sauppe et al. (2013) reported strong evidence for the influence of conceptual accessibility on voice marking, such that human entities were typically chosen as the privileged syntactic argument. The eye-movement data were again suggestive of an early influence of hierarchical structure on sentence planning: while there were early looks to the agent in all word orders produced, sentences in patient voice had comparatively greater looks to the patient, suggesting that message encoding is linguistically guided by the need to mark voice and to link the event construal to the privileged syntactic argument.

Overall, the small amount of research on lesser-studied languages suggests an early influence of language-specific features on sentence planning and production. Languages

that require argument marking, like Tagalog (Sauppe et al. 2013) and Tseltal (Norcliffe, Konopka, et al. 2015; see also Hwang & Kaiser 2014), show evidence for hierarchical planning. Additionally, those languages that allow multiple word orders show a clear relationship between word-order selection and sentence planning. For example, Norcliffe, Konopka, et al. (2015) observed that Tseltal speakers engaged in varying amounts of relational encoding post event apprehension depending both on whether they produced an SVO or VOS sentence and on whether these structures were passivized, suggesting that 'within and across languages, differences in the linear order of words in sentences affect the order of encoding operations throughout formulation' (Norcliffe, Konopka, et al. 2015:1202).

In our study we ask whether these same effects of word order and the encoding of operations throughout sentence formulation are also found in a language that has no fixed, basic word order. There is only one published study that has investigated sentence production in a free word order language: Christianson & Ferreira 2005 on Odawa, an Algonquian language spoken in North America. Odawa allows major constituents to be ordered freely within the clause and has a three-way verb split (direct, inverse, and passive) that distinguishes grammatical role from discourse topic. Christianson and Ferreira showed participants pictures of transitive actions and asked questions that focused agents or patients across all verb forms. Their results showed that the topicalization of referents led to the use of different syntactic frames in their descriptions, which the authors interpreted to be consistent with a (weakly) hierarchical account (see below for discussion of the distinction between weak and strong hierarchical accounts). However, as an off-line study, Christianson & Ferreira 2005 does not address the question of whether a relationship between the linear order of arguments and the encoding of operations in sentence formulation—a relationship shown to exist in both SVO and V-initial languages (Sauppe et al. 2013, Norcliffe, Konopka, et al. 2015)—likewise holds in languages with grammatically free word order, such as Odawa and Murrinhpatha.

In the next section we provide a brief overview of relevant linguistic facts about the Murrinhpatha language, before turning to a discussion of our study.

1.2. MURRINHPATHA OVERVIEW. Murrinhpatha is a non-Pama-Nyungan language of the Daly region of the Northern Territory of Australia. It is spoken by more than 2,500 people in and around the township of Wadeye (Port Keats), approximately 400 km southwest of Darwin, in northern Australia (see Figure 1). Murrinhpatha is one of a small number of Australian languages that continues to be learned as a first language by children, and it is the primary language of daily communication in Wadeye and surrounding communities (Marmion et al. 2014).

Like most of the non-Pama-Nyungan languages of Australia, Murrinhpatha is head-marking and polysynthetic, with complex verbal morphology and very minimal case marking. It is usually described as having relatively free word order at the clausal level (Walsh 1976a). A small corpus study found that SV and OV are the most prevalent orders, but that it is difficult to posit a default or basic word order (Mujkic 2013:53). Apart from these observations, however, there is no detailed analysis of word order in Murrinhpatha clauses. Argument NPs are regularly unexpressed, as in 2–4: Mujkic (2013) found that 29% of verbal clauses in the corpus had no argument NPs at all, consisting only of a verb, as in 4, and only 3% had both subject and object NPs overtly expressed (Mujkic 2013:54–55), as in 1.[1]

---

[1] Murrinhpatha examples in this article are given in a morphemic representation (abstracting over morphophonetic changes) and use the following glosses: ANIM: animate, CLF: noun classifier, DU: dual, F: feminine, FUT: future, HUM: human, IRR: irrealis mood, M: masculine, NFUT: nonfuture, O: object, OBL: oblique, PC: pau-
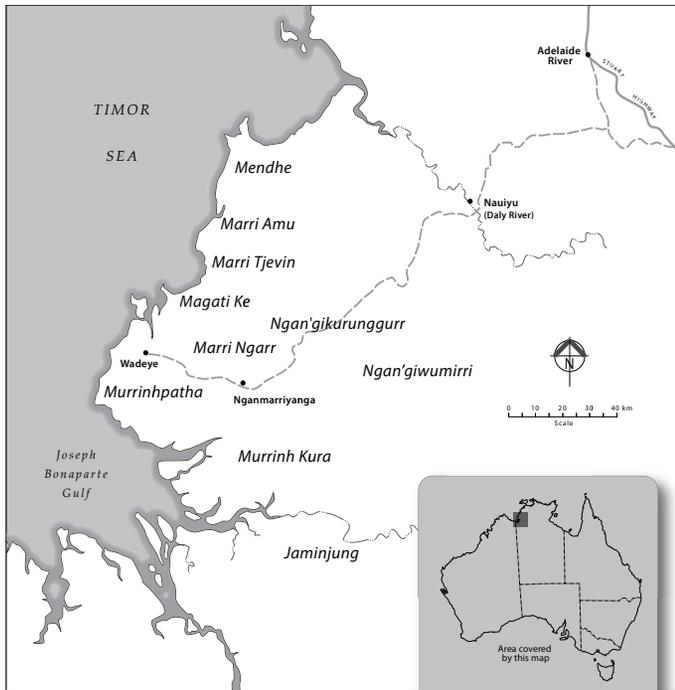
FIGURE 1. Murrinhpatha and surrounding languages. Map drawn by Brenda Thornley.

(1) kardu   palngun ku   yagurr
    CLF:HUM woman CLF:ANIM goanna
        pan-punku-<u>yitthit</u>-nintha
        3SG.S.SLASH(23).NFUT-3DU.O-control-DU.M
    'The woman is holding captive the two goannas.' (SOV)   (MP31-3:23)
(2) ku   pigipigi **ka**-<u>warl</u>-u
    CLF:ANIM pig   3SG.S.POKE(19).FUT:IRR-spear-FUT
    'He's going to spear the pig.' (OV)   (MP32-1:25)
(3) ku   bapula **mangan**-<u>tha</u>=wurran
    CLF:ANIM buffalo 3SG.S.SNATCH(9).NFUT-chase=3SG.S.GO(6).NFUT
    'The buffalo is chasing her.' (SV)   (MP36-3:26)
(4) **nungam**-<u>rirda</u>
    3SG.S.FEET(7).NFUT-kick
    'He kicked him.' (V)   (MP09-4:05)

The main locus of grammatical complexity in Murrinhpatha is in the verb. Verbs in Murrinhpatha are (mostly) complex predicates, made up of two discontinuous stem elements that together define the predicational semantics and argument structure. In related

---

cal, RR: reflexive/reciprocal, S: subject, SG: singular. Verbs are complex predicates, the first part of which (the classifier stem, given in bold in these examples) plays a role in classifying the type of event and has traditionally been glossed with a number (Blythe et al. 2007). In this article we provide an indicative gloss, but readers should be aware that these are merely placeholders for a proper semantic analysis. We also include the number of the paradigm in parentheses, for example, 'SLASH(23)', '(do with) HANDS(8)', '(do with) FEET(7)'. The other part of the complex predicate, the lexical stem, is underlined in these examples. Each example is identified by speaker number and test picture, so that MP31-3:23 refers to speaker number 31, response to test picture number 23 in block 3.

verb-coverb constructions in many other northern Australian languages (Wilson 1999, Schultze-Berndt 2000, McGregor 2002, Bowern 2014) these two stems are separate words; in Murrinhpatha they are morphologically composed into a single word, as two distinct verbal stems—the classifier stem (analogous to the 'finite verb' in analytic systems) and the lexical stem (analogous to the 'coverb' of other systems). In the examples in 1–4 above, the classifier stems are given in bold, and the lexical stems are underlined. Classifier stems form thirty-eight distinct paradigms and are (synchronic) portmanteaux, marking also tense/mood and subject person and number.

The basic templatic structure of the Murrinhpatha verb is shown in Table 1 (see Nordlinger 2010, 2015, and Mansfield 2019 for more detailed discussion). The two stem elements are given here in italics and, as is clear from the template, can be separated by other morphological elements such as incorporated body parts, as in 5, and object and oblique argument marking, as shown in 1 and 6, respectively. The only truly obligatory slot in the verb is slot 1, since some verbs can consist of just a classifier stem, as shown in 7. Subject person and number are obligatorily marked as part of the classifier stem in slot 1, as is tense/mood. Objects are also marked in the verb, in slot 2, but third singular objects are unmarked, as can be seen in 2–4 above. Thus, the order of arguments encoded in the verb is always subject-object.

1	2	3	4	5	6	7	8	9
CS.SUBJ.TAM	SUBJ.NUM/OBJ/OBL	RR	IBP/APPL	*LEXS*	TAM	ADV	SUBJ.NUM/OBJ.NUM	ADV

TABLE 1. Murrinhpatha verbal template.[2]

(5) bangam-**pe**-kat=dim   kardu   olman
3SG.S.BASH(14).NFUT-head-cut.hair=3SG.S.SIT(1).NFUT CLF:HUM old.man
      'He's cutting the old man's hair.'   (MP08-3:14)
(6) kardu   mam-**nge**-met
CLF:HUM 3SG.S.HANDS(8)-3SG.F.OBL-plait.hair
      'The woman is plaiting her hair for her.'   (MP32-3:35)
(7) ngunungam
1SG.S.FEET(7).NFUT
      'I'm going.'

There is very little case marking on NPs in Murrinhpatha. Walsh (1976b) describes an ergative marker, but this is not obligatory, and transitive subject NPs generally appear with no case marking, as shown in 1 and 3 above (e.g. Blythe 2009, Nordlinger 2015, Mansfield 2019). The lack of case marking combined with generally free clausal word order means that many transitive structures involving third singular participants are ambiguous outside of context, as in 8, since there is no grammatical way to determine the argument roles of the NPs. The language has no passive constructions; patient-focus is achieved through word-order variation (as in the second interpretation of 8) or through the use of reflexive verbs (as shown in 9).

---

[2] Additional abbreviations used in this table are: CS: classifier stem, SUBJ: subject marking, TAM: tense/aspect/mood, NUM: number marking, OBJ: object marking, IBP: incorporated body part, APPL: applicative, LEXS: lexical stem, ADV: adverbial marker. Mansfield (2019) notes some variation in morph order on the right edge of this template, especially in the speech of younger speakers, but this variation does not affect the data being discussed in this article.

(8) ku   pigipigi ku   pujicat mangan-tha
    CLF:ANIM pig   CLF:ANIM cat   3SG.S.SNATCH(9).NFUT-chase
        'The pig chased the cat.' (SOV) *or* 'The cat chased the pig.' (OSV)
                                                                    (MP44-3:11)

(9) kardu   kigay wakal thungku-dhangunu dem-ma-bath
    CLF:HUM youth small fire-SOURCE   3SG.S.POKE:RR(21)-hand-burn
        'The young man was burnt on the hand by the fire.' (lit: 'The young man
        burnt-REFL hand from the fire.')   (MP38-1:22)

1.3. THE CURRENT STUDY. The current study investigated the planning and production of Murrinhpatha transitive sentences using eye-tracking. The study is novel in several respects. It is the first empirical study of word-order production in an Australian Indigenous language, languages that have been important in debates regarding nonconfigurationality in syntactic theory (e.g. Hale 1983, Jelinek 1984, Austin & Bresnan 1996). Additionally, it is the first study to investigate sentence formulation using eye-tracking in any free word order language, providing a typologically important data point bearing upon current theoretical approaches to sentence production. To this end, the study had three aims.

First, we aimed to investigate the distribution of word orders produced by speakers of Murrinhpatha in a picture-description task. Since this is the first empirical study of word-order production in Murrinhpatha, we had no specific hypothesis regarding the number and relative frequency of word-order types, so this component of the study was exploratory. Second, following Norcliffe, Konopka, et al. (2015) and Sauppe (2017a), we aimed to analyze how the conceptual accessibility of referents influenced the production of different word orders (see also Tanaka et al. 2011). Assuming that the animacy hierarchy influences sentence production in ways that are similar to those seen in other languages, we expected more agent-initial word orders when agents were human and more patient-initial word orders when nonhuman agents acted on human patients.

Finally, we aimed to investigate the sentence-planning process across different transitive word orders by analyzing participants' eye movements to agents and patients before they began to produce speech, and to contrast the Murrinhpatha findings with those of fixed word order languages. Here we contrast the predictions of the linear vs. hierarchical incremental approaches. Linear incrementality predicts that early looks to referents influences the word-order choice (Gleitman et al. 2007). In Murrinhpatha the prediction of this approach is that the first fixated referent is mentioned first, thus determining word order (e.g. a first fixation on the patient results in a patient-first word order). The hierarchical approach, by contrast, predicts early relational encoding, but exactly how that plays out in a language like Murrinhpatha is unclear. A strong version of hierarchical incrementality predicts that relational encoding always involves lemma selection (Ferreira 2000). Since in Murrinhpatha the order of argument roles within the verb complex is always subject-object, the prediction is a uniform pattern of fixations across all word-order types. In contrast, a weak version of hierarchical incrementality predicts early relational encoding that is conceptual in nature, and that further relational processing occurs prior to verb encoding after this initial phase (Hwang & Kaiser 2014, Sauppe 2017b). Thus, this approach predicts variable degrees of relational encoding depending on word-order type: that is, that the distribution of attention to agent and patient would systematically vary depending on the word order produced. Additionally, a weakly hierarchical account predicts that fixation patterns during linguistic encoding would differ depending on the location of the verb in the sentence. How early this appears in the eye-tracking record is unclear.

2. METHOD.

2.1. PARTICIPANTS AND LANGUAGE ECOLOGY. Forty-six adult first-language speakers of Murrinhpatha were recruited from the Wadeye community in northern Australia. Three participants were removed because they suffered potential vision problems and could not be calibrated (two) or because they were unable to complete the task (one). The final sample ($N = 43$, thirty-three females) ranged in age from seventeen to sixty-three years old ($M = 31.49$, $SD = 10.74$).

The overwhelming majority of the Indigenous members of the Wadeye community acquire Murrinhpatha as their first language and do not acquire much English until they begin primary school at around five or six years old, after which Murrinhpatha is still the preferred language in the community. English is used only in institutional contexts and with members of the Wadeye community who are not proficient in Murrinhpatha (e.g. shopkeepers, health professionals, police, visitors). Our sample varies in terms of contact with and competency in English; some participants could be classed as sequential Murrinhpatha-English bilinguals, whereas others have more limited English knowledge.[3] The commonality is that all participants are first-language speakers of Murrinhpatha, and all use the language in almost every aspect of their daily life.

2.2. MATERIALS. Forty-eight images depicting transitive events served as test pictures (see Figure 2). The test pictures fully crossed thematic role (agent, patient) with humanness (human, nonhuman), such that twelve pictures had a human agent and a human patient (e.g. 'girl pushes boy'), twelve had a human agent and a nonhuman patient (e.g. 'man catches fish'), twelve had a nonhuman agent and a human patient (e.g. 'crocodile chases children'), and twelve had a nonhuman agent and a nonhuman patient (e.g. 'kangaroo boxes cow'). All images in the materials depicted concepts known to Murrinhpatha speakers and could be described using either existing words in the language or Murrinhpatha borrowings from English (e.g. *pigipigi* 'pig/wild boar'). Twenty-nine of the images were used in Norcliffe, Konopka, et al. 2015; a further nineteen were constructed anew (all test pictures are accessible from the project's Open Science Framework (OSF) page: https://osf.io/2j3nu/). An additional ninety-six pictures served as fillers, which mostly depicted one-argument events (e.g. 'boy swimming', 'frog jumping').
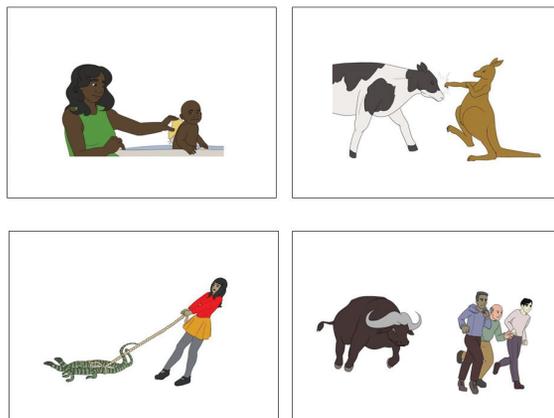
FIGURE 2. Example stimuli pictures depicting the different event combinations.

[3] Kriol, the English-lexified contact language spoken across many Indigenous communities in the north of Australia, is not commonly spoken in Wadeye.

2.3. PROCEDURE. Participants were tested in Wadeye in a single experimental session, in a quiet room with constant lighting conditions. The entire data set was collected over three field trips in May and October 2016, and in July 2018. Participants were told that they would see a series of pictures on a laptop screen and would be asked to describe what was happening in the picture. A portable eye-tracker attached immediately below the screen recorded their eye movements. The procedure was explained in Murrinhpatha to participants, who then provided verbal consent to continue.[4] The session began with a practice session that contained seven pictures. The practice session followed the same format as the test session blocks, and as such served to familiarize participants with the procedure. Each picture was preceded by a blank white screen that lasted for 1000 ms. At the top of the screen a small black dot appeared, which either was centered or appeared to the left or right of center. Participants were told to look at the dot, which functioned to direct participants' attention to the top of the screen so that they were not fixating on any of the test image characters as soon as the image came on screen. At the end of the blank screen a short (130 ms) high-pitched tone sounded and the image appeared on screen. The presentation of items was controlled by the experimenter using the space bar on an external keyboard connected to the eye-tracking computer.

It is important to note that participants were simply told to describe what they saw using Murrinhpatha. If in the practice session they did not use verbs in their description, a Murrinhpatha-speaking research assistant, who was naive to the aims of the experiment, spontaneously modeled a description containing one. Beyond requiring a full clause as a response, however, we did not require that participants describe every character in the scene using full NPs. This deviates from similar past research (Sauppe et al. 2013, Norcliffe, Konopka, et al. 2015) and was done for several reasons. First, we did not want to create expectations in our participants to produce or prefer particular word orders. Second, we did not want to artificially force the presence of argument NPs, given their frequent absence in natural speech. Since Murrinhpatha is a polysynthetic language with obligatory verbal argument marking, we were interested to see whether there was a preference for allowing the verbal morphology to function as the sole argument expression. It is notable that in this experimental task, while many responses in the data set (17.03%) omit one argument NP, only a small proportion (3.85%) omit both. This is very different from naturalistic speech in Murrinhpatha, where verb-only clauses are common and clauses with overt subject and object NPs are rare, as discussed in §1.2.

The test session was broken up into four blocks of thirty-six items (twelve test, twenty-four filler). Members of the Wadeye community do not regularly participate in behavioral tasks like ours; dividing the experiment into blocks provided the opportunity to allow short breaks, which also allowed recalibration of participants' eyes at the beginning of each block. The test images were pseudo-randomized within blocks, and the blocks were pseudo-randomized across participants, resulting in sixteen different test orders. Half of the orders contained mirror-reversed versions of the test images to control for the possibility that scanning paths influenced event conceptualization. In any one order, the agent was located to the left of the patient in half of the test items, and vice versa.

2.4. APPARATUS. Eye movements were recorded using an SMI REDn Professional portable remote eye-tracker (spatial resolution: RMS 0.05°, tracking range 50 × 30 cm

---

[4] This research is covered by University of Melbourne Human ethics approval—Project Ethics ID: 1237988.

at 65 cm distance, distance to participant ~57–60 cm), with a sampling rate of 60 Hz. The eye-tracker was connected to a Lenovo ThinkPad laptop computer with a 13" monitor, which was placed on a monitor stand so that the center of the monitor was approximately at eye level. The experiment was presented using SMI Experimenter Centre 3.6. Participants' picture descriptions were recorded using a Zoom H4n recorder.

2.5. SENTENCE TRANSCRIPTION AND CODING. Participants' descriptions of target trials were transcribed using the ELAN linguistic annotation program (ELAN 5.7, retrieved from https://archive.mpi.nl/tla/elan) by the first author in conjunction with a Murrinhpatha speaker to ensure accuracy of transcription and interpretation.[5] The Murrinhpatha speaker who worked to transcribe and translate the responses was naive to the specific research goals of the project. Each transcribed item was then coded for word order (e.g. agent-verb-patient, or AVP) by a single coder, with approximately 25% of the coded corpus cross-checked by at least one other coder. The decision to code order based on thematic roles (AVP) rather than grammatical function (e.g. SVO) was made in order to account for responses such as 10 and 11, where the expressions of agent and patient do not correlate with subject and object, respectively. In 10 the intended patient of the test item is expressed as an oblique argument, and in 11 the intended patient is expressed as the subject of a verb meaning something like 'to shy away from', with the intended agent expressed as the object. In examples such as these, it was clear that the participant had interpreted the scene as intended, as captured by the thematic role-based coding scheme.

(10) TEST ITEM: young man chases kangaroos   (coded AVP)
kardu   kigay nungam-winharart=dim   ku
CLF:HUM youth 3SG.S.FOOT(7).NFUT-run=3SG.S.SIT(1).NFUT CLF:ANIM
    kumpit-nu
    kangaroo-DAT
  'The young man is running after the kangaroos.'   (MP48-4:26)

(11) TEST ITEM: bird of prey swoops two boys   (coded VA)
mem-nintha-ngkaywey=dim   ku
3SG.S.HANDS:RR(10).NFUT-DU.M-decline=3SG.S.SIT(1).NFUT CLF:ANIM
    murrirrbe
    bird
  'The two boys are shying away from the bird.'   (MP05-1:31)

A small number of responses (approximately 2%) were included with a greater difference between thematic role order and grammatical function order, as shown in 12.

(12) TEST ITEM: one person throws a bucket of water on a group of three people
    [kardu   perrkenkuneme pibimka-neme]   kardu
    CLF:HUM three.M   3DU.S.STAND(3).NFUT-PC.M CLF:HUM
        numi-kathu   dam-wunku-we-wu-neme
        one-this.way 3SG.S.POKE(19).NFUT-3PC.O-head-throw.water-PC.M
      'Three people are standing; one person throws (water) on them.'
                                                                    (MP01-1:01)

As is clear from the English translation in 12, the Murrinhpatha response contains two clauses: the first is a presentational clause that identifies three people, and the second is the transitive clause within which these people have the thematic role of patient (and

grammatical role of object). This relationship is established by the morphology of the transitive verb in the second clause in which the object marker *-wunku-* clearly expresses that the object of the verb is of nonsingular number. Thus, it is clear that the respondent has interpreted the overall event as intended, with the three people as patient of the transitive verb. However, when the relevant NP is first mentioned in the first clause, it is expressed as a subject of the verb *pibimkaneme* 'stand', not as an object. Examples such as this were coded patient-agent-verb (PAV, with the relevant V being the transitive verb).

Responses in which it was clear that the participant interpreted the event structure differently from intended were excluded. For example, if the intended event was 'crocodile (A) chases children (P)' but the participant described the scene as 'children run ahead of crocodile' or 'children and crocodile are running together', then the response was excluded from analysis due to the fact that the intended mapping between thematic roles and event characters was not realized.

Other coding issues arose from the fact that Murrinhpatha allows discontinuous NPs. A very common construction in the data set involves expression of the argument's noun classifier marker before the verb and then the full NP (including repeated noun classifier) after the verb. An example is shown in 13.[6]

(13) **ku**   mangan-tha=wurran   **ku**
     CLF:ANIM 3SG.S.SNATCH(9).NFUT-chase=3SG.S.GO(6).NFUT CLF:ANIM
        thiken
        chicken
     'She is chasing the chicken.' (coded PV)   (MP02-3:05)

In these, and in all cases involving discontinuous NPs, we coded the first mention as the relevant argument (including body parts). So 13, for example, was coded as PV since the *ku* before the verb refers to the patient.

Since Murrinhpatha also freely allows the omission of NPs, the noun classifier before the verb need not be doubled by a full NP as it is in 13. For test items that involve humans acting on humans, or animals acting on animals, this sometimes resulted in ambiguous constructions, as in 14.

(14) TEST ITEM: a woman pushes a man
        kardu   wurdam-rurt=pirrim
        CLF:HUM 3SG.S.SHOVE(29).NFUT-push.over=3SG.S.STAND(3).NFUT
        'She is pushing him.'   (MP21-1:16)

Given the absence of any case marking, the argument role of the noun classifier *kardu* in 14 is ambiguous: it could refer to either the agent or the patient. Examples such as these were coded as ambiguous and excluded from the analyses discussed in this article.

Finally, we also excluded all intransitive clauses from the analyses. Determining transitivity in Murrinhpatha can be difficult, however, since there is no case marking or word order that clearly identifies objects (Nordlinger 2011). The most reliable indicator of clausal transitivity is the presence of object marking in the verb, but since third singular objects are unmarked this indicator was not relevant for many of the test items. Thus, where there was no object marking in the verb to verify transitivity, responses were coded as transitive, and therefore included in the analyses, if they (i) included verbs that were known to occur elsewhere with object marking and thus could be as-

[6] Note that the *ku* classifier can refer only to an animal here, not to a human. The classifier used for humans in the data set is *kardu*.

sumed to be transitive, as in 15, or (ii) included two characters that were functioning as agent and patient consistent with the intended thematic roles of the test item, shown in 16 and also 10 and 11 above.

(15) TEST ITEM: horse drags man on the ground by a rope   (coded V)
kanthin-wurr=wurran
3SG.S.TAKE(22).NFUT-drag=3SG.S.GO(6).NFUT
    'It's dragging him.' (lit: '3SG-dragging')   (MP02-3:17)

(16) TEST ITEM: a dog is pulling a bunch of logs tied with a rope   (coded APV)
ku   were thay log help
CLF:ANIM dog   stick log help
    mam-na=dim
    3SG.S.DO(34).NFUT-3SG.M.OBL=3SG.S.SIT(1).NFUT
    'The dog is helping (carry) the logs.'   (MP05-2:05)

3. Data preprocessing.

3.1. PRODUCTION DATA. The experiment could yield a maximum of 2,064 observations (48 stimuli * 43 participants = 2,064). A small proportion of the data (2.28%) was missing due to equipment failure or experimenter error, problems with calibration, or the participant electing not to respond. This resulted in 2,017 descriptions for analysis. Of these, 431 trials were ruled out because they were not a transitive clause, participants described a different transitive event from the one depicted (e.g. 'the bird is carrying a stick' instead of 'the bird is carrying a grasshopper on a stick'), or their production contained self-corrections and hesitations. A further thirty trials were excluded because of speech onsets longer than 6500 ms, and another seventy-eight were removed from analysis due to being ambiguous, as in 14. Therefore, the final data set consisted of 1,478 sentences, consisting of ten different word orders. This data set was used for the conceptual accessibility analysis (§4.1).

3.2. EYE-MOVEMENT DATA. As is typical of sentence-production studies that use eye-tracking (e.g. Griffin & Bock 2000, Norcliffe, Konopka, et al. 2015, Sauppe 2017b), our eye-movement data constitute a subset of the total number of possible valid trials. Of our 1,478 bivalent clauses, 219 trials (14.8%) were excluded due to contradictory information provided by the eye-tracker regarding the behavior of the two eyes (i.e. when one eye is executing a saccade and the other is blinking). A further thirty-four trials (2.3%) were removed because of tracker loss. This left us with 1,225 trials suitable for analysis, distributed across the ten different word orders: APV ($n = 188$), AVP ($n = 551$), AV ($n = 76$), PAV ($n = 72$), PVA ($n = 126$), PV ($n = 100$), VAP ($n = 6$), VA ($n = 20$), VP ($n = 31$), and V ($n = 55$). This data set was used for the time-course analyses (§4.3). A subset of this data was used to investigate perceptual accessibility, explained below (§4.2).

4. RESULTS. We divide our analyses into three categories that map onto our research questions. We first test how the conceptual accessibility of agent and patient referents influences word-order selection (§4.1). We then test whether participants' first fixations influence word-order selection in our analysis of perceptual accessibility (§4.2). Finally, we analyze participants' eye movements across our most frequent word orders to determine the nature of sentence planning in Murrinhpatha (§4.3). Our data and analysis scripts are available at the project's OSF repository: https://osf.io/2j3nu/. All statistical analyses were conducted in R version 3.6.1 (R Core Team 2014).

4.1. WORD-ORDER VARIABILITY AND CONCEPTUAL ACCESSIBILITY. Murrinhpatha speakers produced ten different word orders, and all participants used multiple word or-

ders across the experiment (average = 5.45). Agent-initial word orders were the most common ($n = 1,014$), followed by patient-initial ($n = 344$) and verb-initial ($n = 120$). The full breakdown of word orders is shown in Table 2.

WORD ORDER		FREQUENCY	%
Agent-initial	AVP	695	47.02
	APV	223	15.09
	AV	96	6.50
Patient-initial	PVA	150	10.15
	PAV	79	5.35
	PV	115	7.78
Verb-initial	VAP	6	0.41
	VPA	0	0.00
	VA	22	1.49
	VP	32	2.17
	V	60	4.06
TOTAL		*1,478*	*100.00*

TABLE 2. Frequency distribution of word orders produced.

This range of word orders is consistent with Christianson and Ferreira's (2005) findings for Odawa and supports the analysis of Murrinhpatha as having flexible word order. Note that we consider the absence of VPA to be an accidental gap here, rather than reflecting an ungrammatical word-order option.

The large range of word orders forces us to make some arbitrary decisions when analyzing how conceptual accessibility influences word-order production. We thus present different analyses of the data by collapsing across word order in two ways. The first set of analyses compared agent-initial (A-INITIAL: AVP, APV, AV = 1,014 observations), patient-initial (P-INITIAL: PVA, PAV, PV = 344 observations), and verb-initial (V-INITIAL: VAP, VA, VP, V = 120 observations) word orders (see Table 3). We refer to these as the INITIAL-NP analyses, which capture a speaker's choice to mention a given element in sentence-initial position, but collapse across word orders with different numbers of expressed arguments. The second set of analyses compared agent-before-patient (A-BEFORE-P: AVP, APV, VAP = 924 observations) vs. patient-before-agent (P-BEFORE-A: PVA, PAV = 229 observations), and agent-only (A-ONLY: AV, VA = 118 observations) vs. patient-only (P-ONLY: PV, VP = 147 observations) word orders (see Table 3). We refer to these as the NP-MATCHED analyses.

WORD ORDER	FREQUENCY (%)	
Data for initial-NP analyses		
Agent-initial (AVP, APV, AV)	1,014	(68.61)
Patient-initial (PVA, PAV, PV)	344	(23.27)
Verb-initial (VAP, VA, VP, V)	120	(8.12)
TOTAL	*1,478*	*(100.00)*
Data for NP-matched analyses		
Agent-before-patient (AVP, APV, VAP)	924	(65.16)
Patient-before-agent (PVA, PAV)	229	(16.15)
Agent-only (AV, VA)	118	(8.32)
Patient-only (PV, VP)	147	(10.37)
TOTAL	*1,418*	*(100.00)*

TABLE 3. Data used for the conceptual accessibility analyses.

The data were analyzed using GENERALIZED LINEAR MIXED MODELS (Baayen et al. 2008, Barr 2008, Jaeger 2008), estimated using the glmer function in the 'lme4' package in R (version 1.1-21; Bates et al. 2015). We modeled the associations between agent

and patient humanness and word order. Agent and patient humanness and their interaction were included as fixed effects (i.e. sum-coded as human = 0.5, nonhuman = −0.5, computed using the MASS package in R: Venables & Ripley 2002; see also Schad et al. 2020). The dependent measure (i.e. word order) was defined as a dichotomous variable, differentiating between a specific word order produced (e.g. A-initial sentences, coded as 1) compared to other sentence types (e.g. P- and V-initial sentences, coded as 0). We entered each predictor separately in a series of stepwise mediation analyses (evaluated via forward model comparison). The maximal random-effects structure that was justified by design and that allowed the models to converge was used (Barr 2013, Barr et al. 2013). Confidence intervals (95%) are provided for the regression coefficients in order to assess the significance of the effects.

INITIAL-NP ANALYSES. We first present the initial-NP analyses, which compare the A-initial, P-initial, and V-initial word orders. Table 4 shows the frequency distribution of bivalent utterances produced by Murrinhpatha speakers as a function of word-order type and event characters. Table 5 shows the output for each regression model; we outline the findings below.

	EVENT DESCRIPTION				
WORD ORDER	HUM-HUM	HUM-NONHUM	NONHUM-HUM	NONHUM-NONHUM	TOTAL
	(%)	(%)	(%)	(%)	(all 100%)
A-initial	170 (16.77)	305 (30.08)	221 (21.79)	318 (31.36)	*1,014*
P-initial	73 (21.22)	97 (28.20)	118 (34.30)	56 (16.28)	*344*
V-initial	68 (56.67)	20 (16.67)	21 (17.50)	11 (9.17)	*120*
TOTAL	*311 (21.04)*	*422 (28.55)*	*360 (24.36)*	*385 (26.05)*	*1,478*

TABLE 4. Frequency distribution of word orders with respect to event characters
(e.g. hum-nonhum = human agent acting on nonhuman patient).

A-INITIAL VS. OTHER. There was no significant effect of agent humanness and word-order choice ($\beta = -0.35$, $SE = 0.32$, 95% CI [−0.98, 0.29]), but there was a significant main effect of patient humanness ($\beta = -1.50$, $SE = 0.27$, 95% CI [−2.04, −0.97]). Thus, compared to other word orders, speakers were more likely to produce A-initial sentences when the patient was nonhuman, regardless of agent humanness.

P-INITIAL VS. OTHER. There was a significant main effect of patient humanness ($\beta = 0.82$, $SE = 0.33$, 95% CI [0.16, 1.48]), with a preference for P-initial sentences when the patient was human, but no significant effect of agent humanness ($\beta = 0.03$, $SE = 0.21$, 95% CI [−0.38, 0.44]) on word-order choice. In addition, there was a significant interaction between event character combinations ($\beta = -1.30$, $SE = 0.42$, 95% CI [−2.13, −0.47]), which was driven by the fact that the effect of patient humanness changed depending on the humanness of the agent. Thus, speakers were more likely to produce P-initial sentences compared to other types of sentences to describe events featuring human agents and nonhuman patients and events depicting nonhuman agents acting on human patients.

V-INITIAL VS. OTHER. Both main effects of character humanness were significant: V-initial sentences were more likely to be produced with human agents ($\beta = 2.85$, $SE = 0.99$, 95% CI [0.89, 4.80]) and human patients ($\beta = 1.68$, $SE = 0.34$, 95% CI [1.02, 2.34]). There was also a significant interaction ($\beta = 1.54$, $SE = 0.66$, 95% CI [0.24, 2.84]), which was driven by the preference to produce V-initial sentences when the test pictures contained both human agents and patients.

VARIABLE	A-INITIAL VS. OTHER			P-INITIAL VS. OTHER		
	REGRESSION COEFF [95% CI]	SE	z	REGRESSION COEFF [95% CI]	SE	z
FIXED EFFECTS						
(intercept)	1.07 [0.62, 1.53]	0.23	4.62 *	−1.59 [−1.94, −1.24]	0.18	−8.93 *
Agent	−0.35 [−0.98, 0.29]	0.32	−1.07	0.03 [−0.38, 0.44]	0.21	0.13
Patient	−1.50 [−2.04, −0.97]	0.27	−5.54 *	0.82 [0.16, 1.48]	0.33	2.45 *
Agent * Patient				−1.30 [−2.13, −0.47]	0.42	−3.08 *
RANDOM EFFECTS	VARIANCE	SD		VARIANCE	SD	
Participant						
(intercept)	1.81	1.35		0.78	0.88	
Agent	2.64	1.62				
Patient	1.18	1.09		2.62	1.62	
Item						
(intercept)	0.44	0.66		0.51	0.71	
Agent						
Patient						

VARIABLE	V-INITIAL VS. OTHER		
	REGRESSION COEFF [95% CI]	SE	z
FIXED EFFECTS			
(intercept)	−4.78 [−6.06, −3.49]	0.66	−7.28 *
Agent	2.85 [0.89, 4.80]	0.99	2.86 *
Patient	1.68 [1.02, 2.34]	0.34	5.00 *
Agent * Patient	1.54 [0.24, 2.84]	0.66	2.32 *
RANDOM EFFECTS	VARIANCE	SD	
Participant			
(intercept)	6.16	2.48	
Agent	2.21	1.49	
Patient			
Item			
(intercept)	0.44	0.66	
Agent			
Patient			

TABLE 5. Multilevel logistic regression results of conceptual accessibility analyses given a first-mentioned character. Note: * $p < 0.05$. Agent = agent humanness, Patient = patient humanness.

NP-MATCHED ANALYSES. We next analyzed the NP-matched data, which tested the different word orders with two overt NPs (encoding agent and patient) produced by Murrinhpatha speakers and sentences where only one argument together with the verb was encoded separately (agent or patient). The descriptive statistics are shown in Table 6. Table 7 shows the output for the regression models comparing (i) A-before-P vs. P-before-A and (ii) A-only vs. P-only word orders.

	EVENT DESCRIPTION				
WORD ORDER	HUM–HUM (%)	HUM–NONHUM (%)	NONHUM–HUM (%)	NONHUM–NONHUM (%)	TOTAL (all 100%)
A-before-P	147 (15.91)	300 (32.47)	172 (18.61)	305 (33.01)	924
P-before-A	34 (14.85)	29 (12.66)	116 (50.66)	50 (21.83)	229
A-only	36 (30.51)	7 (5.93)	61 (51.69)	14 (11.86)	118
P-only	58 (39.46)	77 (52.38)	2 (1.36)	10 (6.80)	147
TOTAL	275 (19.39)	413 (29.13)	351 (24.75)	379 (26.73)	1,418

TABLE 6. Frequency distribution of word orders with respect to event characters. Word orders are collapsed across sentences with two NPs and one NP.

VARIABLE	A-BEFORE-P VS. P-BEFORE-A			A-ONLY VS. P-ONLY		
	REGRESSION COEFF [95% CI]	SE	z	REGRESSION COEFF [95% CI]	SE	z
FIXED EFFECTS						
(intercept)	1.96 [1.52, 2.40]	0.22	8.76 *	0.17 [−0.55, 0.89]	0.37	0.47
Agent	1.06 [0.39, 1.74]	0.34	3.09 *	−4.63 [−6.48, −2.78]	0.94	−4.91 *
Patient	−1.64 [−2.29, −0.99]	0.33	−4.92 *	3.33 [1.69, 4.97]	0.84	3.98 *
RANDOM EFFECTS	VARIANCE	SD		VARIANCE	SD	
Participant						
(intercept)	0.93	0.96		0.34	0.58	
Agent	1.26	1.12				
Patient	0.99	1.00				
Item						
(intercept)	0.62	0.79		3.14	1.77	

TABLE 7. Multilevel logistic regression results of conceptual accessibility analyses in utterances with two NPs and in those where only one character was mentioned. Note: * $p < 0.05$. Agent = agent humanness, Patient = patient humanness.

A-BEFORE-P VS. P-BEFORE-A. There were significant main effects of agent and patient humanness. Participants were more likely to produce A-before-P than P-before-A with human compared to nonhuman agents ($\beta = 1.06$, $SE = 0.34$, 95% CI [0.39, 1.74]) and with nonhuman compared to human patients ($\beta = -1.64$, $SE = 0.33$, 95% CI [−2.29, −0.99]).

A-ONLY VS. P-ONLY. There were significant main effects of agent and patient humanness in the opposite direction of the previous analysis. Participants were more likely to produce A-only than P-only utterances when the events featured nonhuman agents compared to human ones ($\beta = -4.63$, $SE = 0.94$, 95% CI [−6.48, −2.78]) and when events depicted human patients compared to nonhuman patients ($\beta = 3.33$, $SE = 0.84$, 95% CI [1.69, 4.97]). Overall, speakers were more likely to produce sentences with a single expressed argument when the omitted argument was human.

SUMMARY OF CONCEPTUAL ACCESSIBILITY ANALYSES. Different configurations of agent and patient humanness influenced word-order selection. Our initial expectation was that A-initial sentences would be produced with human (and thus accessible) agents, but in fact the results suggested an effect of patient humanness, such that A-initial sentences were more likely when the patient was nonhuman, irrespective of the humanness of the agent. However, our NP-matched analyses showed that these initial results were confounded by an interaction between humanness and NP omission, whereby human participant roles were more likely to be omitted than nonhuman roles. This tendency is best demonstrated if we compare the analyses of the data in Table 7, where two-argument and one-argument sentences were analyzed separately. In the analysis of the two-argument sentences, participants produced more A-before-P than P-before-A sentences when stimulus pictures contained human agents and nonhuman patients, thus conforming to the animacy hierarchy (Silverstein 1976) and its interactions with argument structure and argument encoding, where agents tend to be human (or animate) and syntactic subjects (Van Valin & La Polla 1997, de Swart et al. 2008). Conversely, P-before-A sentences were more likely to be produced with human patients and nonhuman agents. Thus the overarching generalization is that humans are privileged and appear in prominent positions in the sentence, consistent with the fact that they have high conceptual accessibility (Branigan et al. 2008).

However, because human entities are prominent and highly predictable, they are also more likely to be omitted (cf. Everett 2009, Haig & Schnell 2016). This is evident in the

analysis of the A-only vs. P-only sentences (Table 7), where we see a complementary pattern of results. Specifically, in A-only responses we find a greater tendency to express nonhuman agents and omit human patients, whereas in P-only sentences there was a greater tendency to express nonhuman patients and omit human agents. Further evidence for the role of humanness in NP omission comes from V-initial sentences (Table 2), a full 50% of which were verb-only ($n = 60$), and which most commonly occurred when stimulus pictures contained both human agents and patients ($n = 36$, 60%). Thus, in responses with two expressed arguments, initial agents are more likely to be human, but in responses with only an agent argument expressed, initial agents are more likely to be nonhuman.

The Murrinhpatha results regarding NP omission are consistent with past work by Christianson and Cho (2009) concerning the role of topicality and humanness in argument realization. In a comprehension experiment with speakers of Odawa, they found that participants preferred sentences that omitted topical and human NPs, and conversely expected nontopical arguments to be lexically expressed. Odawa grammaticalizes topicality via an obviation hierarchy that interacts with a direct-inverse system in the verb, and thus Christianson and Cho (2009) analyzed their results in terms of expectations of canonical alignment across multiple hierarchies (e.g. Aissen 1997). These hierarchies do not play a role in Murrinhpatha grammar in the same way, and moreover our results do not show any effect of hierarchical interaction such that NP omission results from unexpected alignment (e.g. human patients in the presence of nonhuman agents). In contrast, in our results human NPs are more likely to be omitted across the board, whether as agents or patients and irrespective of the humanness of the other participant.

4.2. Perceptual accessibility analyses. We next analyzed whether participants' early fixations to characters influenced word-order choice. Past researchers have varied in their analytic approaches to this question. For instance, Gleitman et al. (2007) subliminally cued English-speaking participants to look at individual characters and showed that cueing the patient character resulted in a greater preference to produce a passive (i.e. P-initial). In contrast, Norcliffe, Konopka, et al. (2015), who did not use cues, analyzed whether the participants' first fixation to an entity resulted in a higher likelihood that the entity would be first mentioned, thus determining word order. One problem with this latter approach is that it was unclear how a fixation was defined and over what duration. Since we did not provide cues to event participants, we developed a principled way to define first fixation in order to determine whether it had an effect on word-order selection.

Although there is evidence that event roles are rapidly encoded in visual scenes (Dobel et al. 2007, Hafri et al. 2018), we set a conservative window of 400 ms from picture onset within which to define a first fixation, consistent with arguments that gist extraction occurs in this time window (Griffin & Bock 2000, Gerwien & Flecken 2016). Within the time window, we followed Gerwein & Flecken 2016 in defining the first fixation as 'the event that followed the first saccade after stimulus onset, as registered by the eye-tracker' (p. 2635). Fixations, blinks, and saccades were detected by the algorithm implemented in BeGaze 3.7 (SensoMotoric Instruments 2017). Thus, we operationalized first fixations as those fixations that started after the first saccade after picture onset, within a 400 ms time window, and that lasted at least 100 ms, to ensure that the fixated object was adequately encoded during event apprehension. This resulted in the exclusion of 776 trials (see §3.2), since we excluded trials where (i) the first fixation after a saccade occurred later than 400 ms after picture onset or (ii) the first fixation after a saccade fell within the 400 ms time window but did not last 100 ms. We created one data set containing all of the valid trials (449 observations) that detailed

where the participant first looked (agent, patient, or white space), the starting time of the fixation, its duration, our independent variables (agent and patient humanness), and our dependent variable (word order).

We again performed two sets of analyses of the data by collapsing across word order in two ways (see Table 8): (i) A-initial (290 sentences) compared to P-initial (118 sentences) word orders (V-initial utterances were removed from the analyses, given the small number of only forty-one observations), and (ii) A-before-P (270 sentences) compared to P-before-A (80 sentences) word orders. P-only (55), A-only (33), and verb-only (23) utterances were not analyzed due to the small number of observations. The data were analyzed using generalized linear mixed models, estimated using the glmer function in the 'lme4' package in R (version 1.1-21; Bates et al. 2015). All models predicted word-order type (i.e. categorical dependent variable where we modeled the success of producing: A-initial vs. P-initial utterances, and A-before-P vs. P-before-A sentences) as a function of first fixated character (i.e. categorical independent variable sum-coded as agent = 0.5, patient and white space = −0.5), agent and patient humanness (i.e. categorical independent variable sum-coded as human = 0.5, nonhuman = −0.5), and the interactions between first fixated character and character humanness. Since the results of the two analyses (i.e. A-initial vs. P-initial and A-before-P vs. P-before-A) do not qualitatively differ, we report only the results from the comparison between A-initial and P-initial word orders here (see our OSF repository for the full set of NP-matched analyses: https://osf.io/2j3nu/).

WORD ORDER	FREQUENCY (%)	
Data for initial-NP analyses		
Agent-initial (AVP, APV, AV)	290	(71.08)
Patient-initial (PVA, PAV, PV)	118	(28.92)
TOTAL	*408*	*(100.00)*
Data for NP-matched analyses		
Agent-before-patient (AVP, APV, VAP)	270	(77.14)
Patient-before-agent (PVA, PAV)	80	(22.86)
TOTAL	*350*	*(100.00)*

TABLE 8. Data used for the perceptual accessibility analyses.

First fixations landed most often on the agent after picture onset (45%), followed by the patient (36%) and then white space (19%). Table 9 shows the frequency distribution of A-initial and P-initial sentences produced given a first fixation to the agent, patient, or white space.

	FIRST FIXATIONS			
WORD ORDER	AGENT (%)	PATIENT (%)	WHITE SPACE (%)	TOTAL (all 100%)
A-initial	135 (46.55)	101 (34.83)	54 (18.62)	*290*
P-initial	49 (41.53)	44 (37.29)	25 (21.19)	*118*
TOTAL	*184 (45.10)*	*145 (35.54)*	*79 (19.36)*	*408*

TABLE 9. Word orders produced as a function of the first fixated character.

Speakers produced more A-initial sentences when they first fixated on the agent than on the patient or white space. The pattern was less clear for P-initial sentences. Table 10 shows the output for the final regression model predicting the production of A-initial vs. P-initial sentences given first fixation to an event character and patient humanness. The model that best described the data did not include agent humanness as predictor because this variable did not improve model fit.

VARIABLE	A-INITIAL VS. P-INITIAL		
	REGRESSION COEFF [95% CI]	SE	z
FIXED EFFECTS			
(intercept)	0.86 [0.53, 1.20]	0.17	5.08 *
First fixation	0.01 [−0.47, 0.49]	0.25	0.03
Patient	−0.81 [−1.30, −0.32]	0.25	−3.23 *
First fixation * Patient	−1.28 [−2.23, −0.32]	0.49	−2.62 *
RANDOM EFFECTS	VARIANCE	SD	
Participant			
(intercept)	0.37	0.61	
Item			
(intercept)	0.04	0.21	

TABLE 10. Multilevel logistic regression results for perceptual accessibility analyses.
Note: * $p < 0.05$. Patient = patient humanness.

Table 10 shows that the main effect of first fixation was not significant ($\beta = 0.01$, $SE = 0.25$, 95% CI [−0.47, 0.49]). There was a significant effect of patient humanness ($\beta = −0.81$, $SE = 0.25$, 95% CI [−1.30, −0.32]), with a preference for A-initial word orders with nonhuman patients, replicating the analysis presented in §4.1. In addition, there was a significant interaction of first fixation by character humanness, whereby first fixations to the agent resulted in greater A-initial sentences when the patient was nonhuman ($\beta = −1.28$, $SE = 0.49$, 95% CI [−2.23, −0.32]; see the OSF files for visualization and frequency distributions). Thus, we have evidence for an influence of perceptual accessibility on Murrinhpatha speakers' word-order production, but only in concert with conceptual properties of the patient character. This result contrasts with the effect of perceptual accessibility reported by Gleitman et al. (2007); rather than demonstrating that the Murrinhpatha speakers follow purely bottom-up cues in sentence production, it suggests that such cues interact with higher-level cues, in this case semantics, to jointly determine word order. This suggests that Murrinhpatha speakers engage in significant relational encoding early in sentence conceptualization, providing initial support for hierarchical planning. We test further for the existence of hierarchical planning in our time-course analyses in §4.3.

4.3. TIME-COURSE ANALYSIS OF SENTENCE PLANNING. Our final set of analyses investigated the time course of fixations to the agent and patient characters during sentence planning. We investigated sentence planning across the most frequent two-argument word orders, from picture onset to 1000 ms post (average) sentence onset. Our goal was to determine the degree of relational encoding that exists in prominent Murrinhpatha word orders, and where in the eye-tracking record it occurs. This addresses the degree to which sentence planning in Murrinhpatha is best described as a linear or hierarchical incremental process, with greater and earlier degrees of relational encoding providing evidence in support of the latter. Notably, if Murrinhpatha is best described as a weakly hierarchical process, we expected to observe word-order-specific variation in participants' attention to agents and patients either during event apprehension or early in linguistic encoding, as observed in Tseltal (Norcliffe, Konopka, et al. 2015) and Tagalog (Sauppe et al. 2013). In contrast, if it is a strongly hierarchical process, we expected to observe eye movements more uniformly following the subject-object order contained in the polysynthetic verb, and thus be less influenced by word-order variation. Our perceptual accessibility analysis already provides evidence against the linear incremental account (§4.2); this account would predict little to no relational encoding during the earliest

stages of sentence planning, such that in A-initial word orders participants attend primarily (or exclusively) to the agent, and in P-initial word orders primarily to the patient.

It was not practical to analyze the fixation patterns for every word order produced, since to do so would be analytically challenging, and more importantly most word orders did not result in enough tokens to analyze robustly. We also did not analyze utterances with unexpressed arguments because the location of implicit arguments is unclear (e.g. an AV could be AV(P), (P)VA, or A(P)V). We thus present an analysis of the four most frequent word orders that contain two overt NPs: AVP, APV, PVA, and PAV. We calculated the proportion of fixations to the agent, patient, and white space starting at picture onset (time 0 ms) until 1000 ms after average speech onset. The data were analyzed across four time windows. The first time window is from 100–600 ms following picture onset, corresponding to event apprehension (Griffin & Bock 2000).[7] Similar past studies have defined the second time window as 600–1800 ms, corresponding to linguistic encoding leading to speech onset. However, average speech onset in the Murrinhpatha data is significantly later, at approximately 2300 ms or later depending on the word order, and analyzing the time course over such a large time window would potentially result in substantial information loss.[8] We therefore divided the linguistic encoding period of 600–2300 ms into two windows: (i) 600–1600 ms, and (ii) 1600–2300 ms (i.e. average speech onset of our most common word order, AVP). We chose to divide the window at the 1600 ms time point because this constitutes the point at which fixations to agent and patient in our most common word order (AVP) cross, suggesting the boundary of a processing event. Our final time window analyzed eye movements for 1000 ms after average speech onset (i.e. 2300–3300 ms).[9]

Figure 3 shows the time-course graph of the proportion of fixations to the agent and patient referents from picture onset until 3800 ms for AVP, APV, PVA, and PAV sentences. We point out two notable features of the eye-movement records. First, there is clear relational encoding between agent and patient characters in all word orders during the event apprehension window (100–600 ms), which is earlier than has been observed in similar studies of languages with fixed word orders (e.g. Sauppe et al. 2013, Norcliffe, Konopka, et al. 2015). Second, the pattern of attention to the agent and patient during this same time window varies across each word order, which results in variability in the distribution of attention to the event characters in subsequent time windows.

We used GENERALIZED ADDITIVE MIXED MODELS (GAMMs; Hastie & Tibshirani 1990, Zuur et al. 2009, Wood 2017, Porretta et al. 2018) to analyze the time course of fixations to event characters during sentence planning. The major advantage of

---

[7] This differs from past research (e.g. Norcliffe, Konopka, et al. 2015) in which this window is analyzed from t = 0. However, it is unlikely that fixations in the first 100 ms are in response to the picture stimuli (Duchowski 2007), so excluding them removes random noise.

[8] While the reason for long speech latencies is difficult to determine, it may well be due to the heavy linguistic burden placed on speakers by the polysynthetic nature of the verb complex and/or the flexibility of word-order choice.

[9] Note that mean speech onsets differed across word orders: AVP (M = 2329.68, SD = 886.8), APV (M = 2446.78, SD = 942.05), PVA (M = 2423.04, SD = 858.63), and PAV (M = 2718.36, SD = 839.67). Variable speech-onset latencies are an inevitable consequence of the method. We chose to follow past studies in determining the speech-onset time window as beginning at the average of the dominant word order. While it is true that the variable onset latencies mean that some individual observations may not have been produced within this time frame, calibrating the time windows for each word order would require us to make even more difficult decisions about earlier time windows at different planning stages, which would make comparisons across the word orders difficult. Thus, we must acknowledge that there may be some information loss associated with this way of treating the data, but this does not have a significant impact on our main findings.
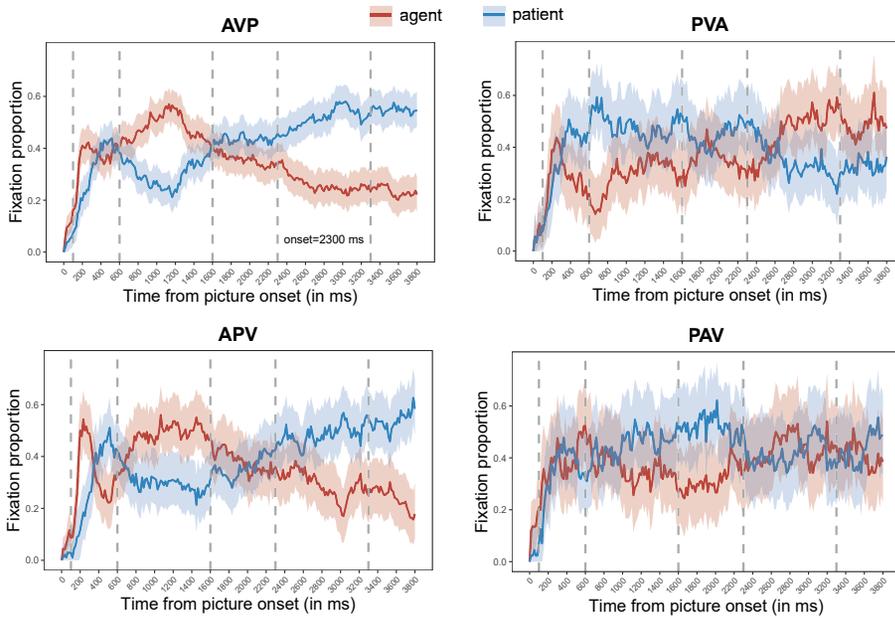
FIGURE 3. Proportion of agent- and patient-directed fixations in A-initial and P-initial utterances in Murrinhpatha. Ribbons indicate confidence intervals (95%), calculated for each sampling step. The dashed lines indicate the four analysis time windows. Speech onset time is marked on AVP (reference).

GAMMs is that they estimate nonlinear (smoothed) relationships between predictors (categorical or continuous) and dependent variables, thus allowing us to model the nonlinear development of fixation patterns across time. For our purposes, the technique allows us to determine whether the eye-movement patterns for each word order in Fig. 3 differ statistically, and how, thus providing tests of the different accounts of sentence planning. We report analyses for each time window in Table 11, but in our discussion here spend most of our time on the event apprehension window (time window 1), since it is in this window where we immediately find evidence in favor of one account to the exclusion of the others.

The analyses were conducted using the 'bam' function in the 'mgcv' package (version 1.8-28; Wood 2017). The models included fixations to the agent (dependent variable coded as 1 = yes, 0 = no), smooth interactions for word order (four levels: AVP, APV, PVA, PAV) and time (continuous variable), and random smooth interactions for time by participants and items (random effects). Thus, since the dependent variable is dichotomous, we used a binomial distribution and specified a logit link function. We used fast RESTRICTED MAXIMUM LIKELIHOOD (REML) for the smoothing parameter estimation; approximate *p*-values are provided for each smooth term to determine the significance of the curve (Zuur et al. 2009, Zuur et al. 2014). Model fit was determined using the AKAIKE INFORMATION CRITERION (AIC) value (Akaike 1998).

Table 11 contains the results from the final models at each time window. The results consist of two sections: (i) the parametric coefficients, which show the linear terms of the model—here, the *p*-values of the different word orders indicate whether and how overall looks to the agent for a given word order differ from the baseline word order (AVP, which we set as the baseline because it was the most frequent)—and (ii) the smooth terms defined in the model, which indicate the functional form, either linear or

nonlinear, of the relationship between looks to the agent across time for each word order. The first column shows the EFFECTIVE DEGREES OF FREEDOM (*edf*, range of $0, \infty$), which represent the complexity of the smooth. If the *edf* are equal to 1, there is a linear relationship between fixations to the agent and time for that word order (i.e. it is modeled as a straight line). The higher the *edf*, the more 'wiggly' (nonlinear) the smoother (Zuur et al. 2009), meaning that more complex looking behaviors are better captured by a nonlinear function. Thus, looks to the agent across time could change linearly (i.e. increase or decrease at a stable rate, or stay the same, thus capturing one behavior) or nonlinearly (i.e. multiple looking behaviors are captured). The *p*-values indicate the approximate significance of the smooth terms: specifically, whether the fitted curve modeling looks to the agent across time is significantly different from zero. There are no accepted standards for interpreting the approximate *p*-values of smooth functions (e.g. Zuur et al. 2009, Zuur et al. 2014), with a rule of thumb being that they should be sufficiently low to ensure a reliable effect. While acknowledging the arbitrary nature of *p*-value cut-offs, we mark all values where $p < 0.01$ as suggesting a meaningful interaction between word order and time, and invite readers to make their own inferences.

TIME WINDOW 1 (EVENT APPREHENSION): 100–600 MS. Two theoretically important results emerge in time window 1. First, the plots in Fig. 3 show a degree of relational encoding in all word orders. This result is particularly notable for the A-initial orders, standing in sharp contrast to data using the same paradigm from English (Griffin & Bock 2000), Dutch (Norcliffe, Konopka, et al. 2015, although see Konopka 2019), German (Sauppe 2017b), and Tseltal (Norcliffe, Konopka, et al. 2015), all of which reported very little relational encoding for A-initial orders, with participants devoting most attention to the agent in this early time window. This suggests that Murrinhpatha speakers engage in rapid early conceptual encoding of both characters in the event. Importantly, this encoding appears to be sequentially fixed, whereby speakers have an agent-first preference regardless of the word order they eventually produce. This is consistent with other crosslinguistic work (Sauppe et al. 2013, Bickel et al. 2015), supporting arguments for the prominence of agents in sentence processing (Bornkessel-Schlesewsky & Schlesewsky 2009).

Figure 4 plots the model predictions for looks to both the agent and patient for each word order across time window 1.[10] These plots reveal the same result: that is, an initial preference for the agent gives way to a subsequent rise in looks to the patient across all word orders. Such relational encoding is consistent with hierarchical planning; however, the second important feature of the results in this time window distinguishes between the weak and strong versions of the hypothesis. Notably, the data show that eye-movement patterns differed across all word orders in either overall looks to one character over another, or in how looks to characters changed across time. To interpret these we direct the reader to the model output in Table 11 and the plots in Figs. 3 and 4. For A-initial word orders, we see that, while AVP and APV do not differ in their overall looks to the agent in this window (i.e. the parametric coefficient is not significant), looks to the agent in the APV change nonlinearly with time, as indicated by the significant smooth term and the more prominent maximum and minimum in Figs. 3 and 4. P-initial word orders were also different. Compared to the most frequent AVP word order, fixations to the agent were significantly higher in PAV sentences, but signifi-

---

[10] To obtain the patient data we ran separate models with looks to the patient as the dependent measure. The output of these models can be found in our OSF repository: https://osf.io/2j3nu/.

Time window 1: Event apprehension (100–600 ms)

Parametric coefficients:	EST [95% CI]	SE	z-VALUE	p-VALUE
(intercept)	−0.96 [−1.39, −0.52]	0.22	−4.32	< 0.001 ***
PAV	0.30 [0.14, 0.47]	0.08	3.61	< 0.001 ***
APV	−0.10 [−0.22, 0.01]	0.06	−1.76	0.078
PVA	−0.26 [−0.38, −0.13]	0.06	−4.08	< 0.001 ***
Smooth terms:	edf	ref.df	CHI.SQ	p-VALUE
Time * AVP	3.97	4.69	6.57	0.169
Time * PAV	1.00	1.00	6.85	0.008 **
Time * APV	4.71	5.66	30.93	< 0.001 ***
Time * PVA	2.85	3.46	5.88	0.240
Random effect: Subjects	185.33	386.00	1828.96	< 0.001 ***
Random effect: Items	482.11	845.00	4391.89	< 0.001 ***

Time window 2: Linguistic encoding I (600–1600 ms)

Parametric coefficients:	EST [95% CI]	SE	z-VALUE	p-VALUE
(intercept)	−0.22 [−0.63, 0.21]	0.21	−1.03	0.305
PAV	−0.56 [−0.66, −0.45]	0.05	−10.66	< 0.001 ***
APV	−0.09 [−0.16, −0.01]	0.04	−2.31	0.021 *
PVA	−1.15 [−1.23, −1.07]	0.04	−27.78	< 0.001 ***
Smooth terms:	edf	ref.df	CHI.SQ	p-VALUE
Time * AVP	3.81	4.58	36.75	< 0.001 ***
Time * PAV	1.76	2.15	2.94	0.236
Time * APV	2.71	3.30	6.12	0.103
Time * PVA	1.00	1.00	0.61	0.434
Random effect: Subjects	241.54	386.00	3289.53	< 0.001 ***
Random effect: Items	559.06	845.00	7100.44	< 0.001 ***

Time window 3: Linguistic encoding II (1600–2300 ms)

Parametric coefficients:	EST [95% CI]	SE	z-VALUE	p-VALUE
(intercept)	−0.57 [−0.93, −0.21]	0.18	−3.14	0.001 **
PAV	−0.40 [−0.52, −0.28]	0.06	−6.59	< 0.001 ***
APV	0.11 [0.02, 0.19]	0.04	2.56	< 0.010 *
PVA	−0.27 [−0.36, −0.18]	0.05	−5.91	< 0.001 ***
Smooth terms:	edf	ref.df	CHI.SQ	p-VALUE
Time * AVP	1.00	1.00	8.43	0.003 **
Time * PAV	1.00	1.00	0.54	0.460
Time * APV	1.00	1.00	8.62	0.003 **
Time * PVA	2.17	2.68	3.24	0.281
Random effect: Subjects	156.13	377.00	1721.18	< 0.001 ***
Random effect: Items	393.79	845.00	3843.71	< 0.001 ***

Time window 4: Speech onset (2300 ms)–1 s after

Parametric coefficients:	EST [95% CI]	SE	z-VALUE	p-VALUE
(intercept)	−1.13 [−1.48, −0.78]	0.18	−6.41	< 0.001 ***
PAV	0.17 [0.06, 0.27]	0.05	3.17	0.001 **
APV	−0.04 [−0.11, 0.04]	0.04	−0.96	0.339
PVA	−0.91 [−0.98, −0.83]	0.04	23.25	< 0.001 ***
Smooth terms:	edf	ref.df	CHI.SQ	p-VALUE
Time * AVP	1.11	1.15	4.56	0.051
Time * PAV	2.57	3.16	5.43	0.142
Time * APV	4.63	5.59	23.62	< 0.001 ***
Time * PVA	3.01	3.70	23.28	< 0.001 ***
Random effect: Subjects	211.41	384.00	2376.04	< 0.001 ***
Random effect: Items	471.33	845.00	5324.46	< 0.001 ***

TABLE 11. Results of the generalized additive mixed models reporting parametric coefficients and estimated degrees of freedom (*edf*), reference degrees of freedom (*ref.df*), chi-square, and *p*-values for smooth terms, comparing fixations to the agent over time in AVP (reference), PAV, APV, and PVA word orders.
Note: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

cantly lower in PVA sentences, as indicated by the significant parametric coefficients. The smooth curve of the interaction between time and word order was significant for PAV ($p = 0.008$), which can be interpreted to mean that fixations to the agent changed in a linear fashion across time ($edf = 1$). This can be seen in Fig. 4, where the model predictions show a straight line that increases across time. We interpret the unique eye-movement patterns for each individual word order as evidence for a weakly hierarchical explanation of sentence planning in Murrinhpatha.
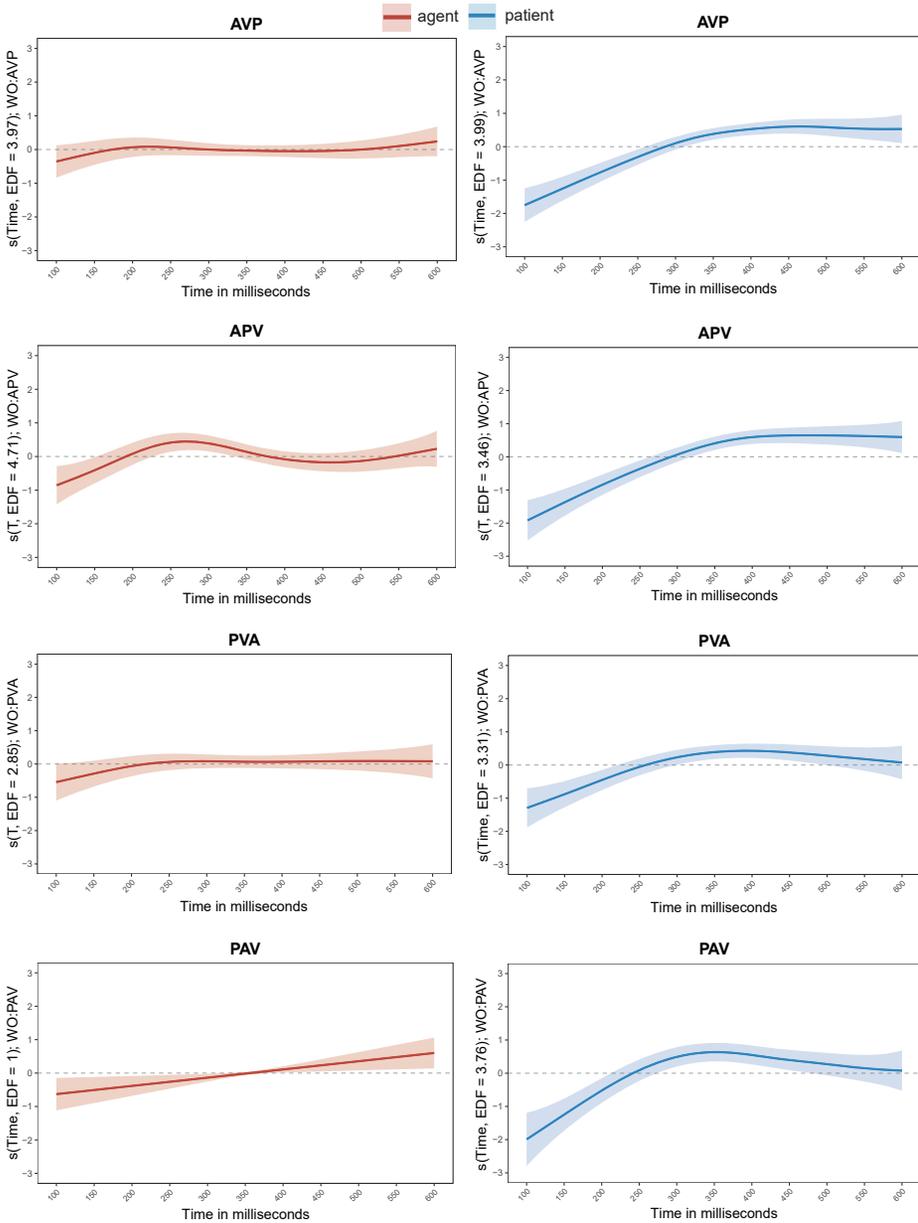

FIGURE 4. Model smoothers plots of the generalized additive mixed models of the interaction between time and word order. The plots show the changes in the odds of fixating the agent (red lines) and of fixating the patient (blue lines) over time (i.e. during event apprehension) when a specific type of sentence is produced. Solid lines represent the smooth function of time by word order (on logit scale) predicted by the model. Ribbons indicate CI 95%.

TIME WINDOW 2 (LINGUISTIC ENCODING I): 600–1600 MS. Linguistic encoding typically follows an incremental order-of-mention pattern, which is what we see across the next two time windows. Thus, in time window 2 we see that there were significantly fewer looks to the agent in the P-initial orders compared to the AVP order (i.e. the parametric coefficients are significant and negative). This is also the case for the APV order. The smooth interaction between time and AVP order resulted in a significant ($p < 0.001$) nonlinear effect (edf = 3.81), suggesting that fixations to the agent varied significantly across time. This effect captures the sharp rise in looks to the agent in this word order, which peaks at around 1200 ms and rapidly descends thereafter.

Based on findings from Dutch-speaking adults' production of AVP sentences, Konopka (2019) argued that looks in this time window reflect not only the linguistic encoding of the first NP but also encoding of the verb. With German speakers, Sauppe (2017b) reported differences in looking time to AVP and APV sentences that were argued to reflect the location of the verb. Thus, what we are likely observing here is differences in fixations to the agent across word orders and time driven by both the role of the first NP (agent vs. patient) and the location of the verb (medial vs. final). For A-initial word orders the difference likely reflects the early encoding of the verb in AVP order in comparison to APV: since both agent and patient must be marked in the Murrinhpatha verb, the point at which fixations to the agent and patient cross over is a possible signature of verb encoding in the eye-tracking record. Note that this occurs early for AVP (at around 1600 ms, the endpoint of time window 2) in comparison to APV, where it occurs some 500 ms later. This same effect occurs later in P-initial word orders, which overall appear to be more effortful to encode.

TIME WINDOW 3 (LINGUISTIC ENCODING II): 1600–2300 MS. In time window 3 we again see significant differences across all word orders. In comparison to AVP sentences, there were significantly fewer looks to the agent in P-initial orders. In contrast, there were significantly greater looks to the agent in APV sentences compared to AVP. The smooth terms show that there was a significant linear interaction between time and A-initial sentences (edf = 1, p = 0.003), capturing the fact that looks to the agent in these word orders are decreasing at a constant rate across time. During this time window speakers are likely encoding the patient in the AVP word order, since there are increased looks to the patient from 1600 ms onward. In contrast, in APV sentences this process is delayed by approximately 400 ms before speech onset. Here, participants are likely still encoding the verb, since the agent-patient crossover point, which we have interpreted as a signature of verb encoding, occurs at around 2100 ms in this word order.

TIME WINDOW 4: 2300–3300 MS. In this time window we find a different pattern of results from time window 3, with significantly more looks to the agent in P-initial orders than in AVP. The smooth interaction between time and APV order was significant ($p < 0.001$), with fixations to the agent decreasing at a nonlinear rate over time (edf = 4.63). This is different from the pattern of results observed in PVA order, where there was also a significant nonlinear interaction (edf = 3.01, $p < 0.001$), but where looks to the agent increased and where the rate of change slowed across time. The tendency in the P-initial word orders to look more at the agent in this time window likely reflects a tendency for speakers to focus on late-appearing arguments following speech onset (Griffin & Bock 2000, Norcliffe, Konopka, et al. 2015). The nonlinear relationship between looks to the agent and time in the APV order may be a residual effect deriving from the later encoding of the verb in the verb-final order.

5. DISCUSSION. Our results are novel in a number of respects and have significant implications for understanding free word order and the way in which it influences sentence planning and production. Consistent with Christianson & Ferreira 2005, we found that

Murrinhpatha speakers produced a wide range of word-order combinations, with all speakers producing multiple word orders, thus demonstrating the extreme word-order flexibility allowed in the language. This basic result is striking given data from a similar experimental study on German, a language that in principle also allows relatively free word order, where speakers overwhelmingly use one dominant word order (>75%) and one or two other infrequent word orders in specific circumstances, such as a passive in cases where a human patient is affected by an inanimate entity (Sauppe 2017b).

Nevertheless, the word orders were not randomly distributed: approximately two thirds were A-initial, a quarter were P-initial, and the remaining were V-initial. The predominance of A-initial orders may reflect agent prominence common to natural languages (Jackendoff 1972, Dik 1978, Riesberg et al. 2019). We observed signals of this prominence in our eye-movement data, where we saw rapid early looks to the agent regardless of the word order produced, consistent with the argument that agents are universally more prominent in language processing (Bornkessel-Schlesewsky & Schlesewsky 2009, Alday et al. 2014, Bickel et al. 2015). Encoding agents early in sentence production is crucial, since lower arguments are defined with respect to their relationship to the agent; thus this effect suggests that agent prominence may represent a universal property of event conceptualization and could explain why, in free word order languages like Murrinhpatha, speakers produce proportionally more A-initial sentences.

However, a third of speakers' transitive clauses were not A-initial. Some of the differences in distribution can be attributed to semantic properties of the NP arguments: consistent with similar past studies and with our predictions (Christianson & Ferreira 2005, Tanaka et al. 2011, Norcliffe, Konopka, et al. 2015), we found a clear influence of conceptual accessibility, such as the tendency to front human patients when they were affected by nonhuman agents, and a preference for A-before-P orders when agents were human and patients were nonhuman. This interacted with patterns of NP omission, which showed a preference for human NPs to be omitted over nonhuman NPs. Thus, as is common crosslinguistically, prominent predictable arguments, which are typically human, are more likely to be omitted (Du Bois 1987, Christianson & Cho 2009, Everett 2009, Gennari et al. 2012, Haig & Schnell 2016). Our results, which involve isolated sentences independent of discourse context, support recent work which argues that crosslinguistic patterns of NP omission are more likely explained by semantic features such as humanness than by discourse considerations (Everett 2009, Haig & Schnell 2016; cf. Du Bois 1987). However, our results suggest that the effect of humanness on NP omission is not just a feature of agents (Everett 2009) or subjects (Haig & Schnell 2016), but is found with patients as well.

Therefore, while Murrinhpatha speakers have all word orders available to them, their production is semantically and functionally constrained in much the same way and by the same variables as has been observed in other languages. At the same time, the amount of word-order variation in Murrinhpatha is notable because the language lacks obligatory case marking, which has traditionally been associated with higher word-order variability (e.g. Siewierska 1998, Blake 2001, Futrell et al. 2015). Levshina (2019) analyzed word-order variation across sixty typologically diverse languages by computing Shannon entropy (Shannon 1948), reporting that those languages with the highest variability tend to have formal means of distinguishing between arguments via case marking (e.g. Lithuanian, Latvian, Ancient Greek). Murrinhpatha and many other polysynthetic languages such as Bininj Gun-wok (Evans 2003), Mohawk (Baker 1996), and Odawa (Christianson & Ferreira 2005) provide clear counterexamples to this trend. Interestingly, while Mur-

rinhpatha once had obligatory ergative case (Walsh 1976b), ergative marking is no longer common (Nordlinger 2015). Thus, word-order variability has persisted in the language despite the absence of obligatory core grammatical case marking.

Turning to the eye-tracking results, our data suggest that sentence planning in Murrinhpatha is best categorized as a weakly hierarchical incremental process (Griffin & Bock 2000, Hwang & Kaiser 2014, Konopka & Meyer 2014, Sauppe 2017a,b). Two features of the eye-tracking record support this conclusion. First, we observed clear signatures of hierarchical planning during event conceptualization, which differed according to the word orders speakers produced. Second, like in Sauppe's (2017b) German data (see also Hwang & Kaiser 2014), we observed differences in the timing of verb planning depending on whether the verb was sentence-medial or sentence-final, which was clearest for A-initial word orders. Together, these two results suggest that Murrinhpatha speakers rapidly generate a conceptual representation of an event through extensive early relational encoding, which subsequently guides their linguistic encoding and production downstream.

In this respect, the Murrinhpatha results are strikingly different from those of other languages. It is particularly notable that we see evidence for early relational encoding as well as an early signature of the different word orders during event apprehension (time window 1), which we interpret to derive from typological properties of the language. Head marking necessitates the need to plan event roles early, as Norcliffe, Konopka, et al. (2015) observed for Tseltal V-initial utterances, although they observed relational encoding somewhat later in the eye-tracking record during the linguistic encoding stage. Word-order differences between the languages may explain this difference in timing: Tseltal does not have free word order, and thus the restricted range of word-order options may remove the necessity to rapidly encode event roles. This is most evident in Tseltal speakers' formulation of SVO sentences, which patterned almost exactly like that of Dutch speakers (Norcliffe, Konopka, et al. 2015) with a high proportion of fixations to the agent in event apprehension, whereas Murrinhpatha speakers pattern differently, with clear relational encoding at this early stage. Thus, it appears that the option to freely order constituents places additional pressure on Murrinhpatha speakers to develop a conceptual representation of the event relatively early in sentence planning, with relational encoding during event apprehension (time window 1) laying down a template for the linearization of the message.

This rapid encoding of event roles was also evident in our perceptual accessibility analysis. In this analysis we found that participants' first fixation to agents influenced word-order choice only in concert with semantic properties of the patient. In Murrinhpatha, it appears, sentence production cannot proceed without detailed relational encoding very early in the planning process (i.e. in the first 600 ms).

Taking into account the accumulated work on sentence planning across several languages (Gleitman et al. 2007, Kuchinsky et al. 2011, Hwang & Kaiser 2014, Konopka & Meyer 2014, Norcliffe, Konopka, et al. 2015, Sauppe 2017a,b, Konopka 2019), we argue that a speaker's language plays a moderating role across all stages of sentence production, including event apprehension (Gerwien & Flecken 2016, Sauppe & Flecken 2021). Thus, in languages like English and Dutch, which have little to no dependent marking and rigid word-order requirements, planning can proceed in linear fashion, as observed by Gleitman et al. (2007) and Konopka and Meyer (2014), but can also involve hierarchical planning (Konopka 2019). However, typologically different languages place different requirements on the speaker; once we move away from com-

monly studied languages we see an additional need for hierarchical planning that is driven by word-order requirements (or lack thereof) and obligatory argument marking. Thus, speakers will vary crosslinguistically as to how sentence planning proceeds, with the timing of planning adapting to the requirements of the language at hand.

We end by considering an important question in work on nonconfigurational languages: does Murrinhpatha have a basic word order? Whereas some grammatical theories argue that word order in nonconfigurational languages is generated from a basic underlying word order (Baker 1996, Legate 2001), other theories such as LEXICAL-FUNCTIONAL GRAMMAR do not make this assumption (Bresnan 2001), instead allowing free word order to be base-generated and grammatical functions to be assigned via morphological processes outside of constituency (Nordlinger & Bresnan 2011). On the basis of our word-order data, one could suggest that Murrinhpatha has a basic AVP word order, since this is the most common order in our data set (although found in only 47.02% of utterances). However, we argue that our eye-movement data provide evidence against this conclusion, because sentence planning in AVP sentences in Murrinhpatha looks different from that of speakers of other languages that have AVP (SVO) word order (i.e. English, Dutch; Griffin & Bock 2000, Norcliffe, Konopka, et al. 2015). Whereas speakers of these other languages overwhelmingly fixate on the agent in early event planning, consistent with their AVP basic word order, Murrinhpatha speakers show clear evidence of relational encoding in this first 600 ms. If AVP were likewise a basic word order in Murrinhpatha, this processing difference would be unaccounted for. Thus, it appears that Murrinhpatha speakers generate word order on-line rather than deferring to a default basic order.

6. CONCLUSION. This study makes a number of contributions to our understanding of the interaction between grammatical structure and language processing in sentence production. First, it is the first sentence-production study of an Australian Indigenous language, and the first on-line production study of a free word order language, and therefore contributes to broadening the scope of psycholinguistic studies beyond the current (circa) 0.5% of the world's languages (Jaeger & Norcliffe 2009). Moreover, the findings have interesting implications for models of crosslinguistic processing. Murrinhpatha speakers are consistent with speakers of other typologically diverse languages in showing an agent-first preference early in sentence planning, and in the interaction of conceptual accessibility with word-order choice. However, in other respects the Murrinhpatha speakers are unusual in showing early relational encoding across all word-order types before preferencing fixations to the first-mentioned arguments. In effect, our data suggest that Murrinhpatha speakers conceptualize the whole event AND begin selecting a word order within the first 600 ms, approximately 1700 ms before they begin to speak. These findings suggest that the sentence-planning and sentence-production mechanism is 'softly assembled' and adaptive to typological variation across languages. This results in noticeably different sentence-planning strategies for a free word order language as opposed to a language with more fixed word order, thus supporting the growing body of research revealing significant crosslinguistic differences in sentence production that are linked to grammatical properties of languages (e.g. Sauppe et al. 2013, Hwang & Kaiser 2014, Norcliffe, Konopka, et al. 2015).

<div align="center">APPENDIX</div>

Here we provide a list of our experimental events ordered alphabetically by agent and patient combinations. The objects in parentheses refer to instruments carried by agents. The full set of pictures can be found at https://osf.io/2j3nu/. * indicates materials used with permission from Norcliffe, Konopka, et al. 2015.

Human agent acting on human patient
1.   army general kicking boy*
2.   barber cutting man's beard (with scissors)*
3.   boxer punching man*
4.   doctor vaccinating boy (with needle)*
5.   girl wetting boys (with bucket of water)
6.   girl pushing boy*
7.   girl tripping construction worker*
8.   man throwing child up in the air*
9.   nun braiding girl's hair*
10. policeman stopping men
11. soldier shooting man (with gun)*
12. woman washing baby (with sponge)

Human agent acting on nonhuman patient
13. boy chasing kangaroos
14. farmer whipping donkey (with whip)*
15. girl chasing dogs
16. man throwing boomerang
17. man catching fish (with fishing rod)*
18. man hunting pig (with spear)*
19. priest pulling donkey (with lasso)*
20. man roasting pig (on a spit)
21. child poking lizard (with stick)*
22. woman dragging goannas (with rope)
23. woman chasing chicken*
24. woman petting sheep*

Nonhuman agent acting on human patient
25. buffalo chasing men
26. bull charging girl*
27. cart hitting street vendor*
28. cat scratching girl*
29. crocodile chasing kids
30. crocodile biting man*
31. dog licking boy*
32. fire burning man*
33. horse dragging man*
34. magpie swooping boys
35. monkey painting boy (with paintbrush)*
36. rock falling on men

Nonhuman agent acting on nonhuman patient
37. bird catching insect (with stick)*
38. dog chasing lizards
39. dog catching butterfly (with net)*
40. dog chasing car
41. dog chasing squirrel*
42. dog pulling logs (with rope)
43. eagle grabbing rabbit*
44. kangaroo boxing cow
45. lightning hitting tree
46. monkey hooking snake (with stick)
47. pig sniffing cat*
48. tree falling on cars

REFERENCES

AISSEN, JUDITH. 1997. On the syntax of obviation. *Language* 73(4).705–50. DOI: 10.1353 /lan.1997.0042.

AISSEN, JUDITH. 1999. Markedness and subject choice in optimality theory. *Natural Language & Linguistic Theory* 17(4).673–711. DOI: 10.1023/A:1006335629372.

AKAIKE, HIROTUGU. 1998. Information theory and an extension of the maximum likelihood principle. *Selected papers of Hirotugu Akaike*, ed. by Emmanuel Parzen, Kunio Tanabe, and Genshiro Kitagawa, 199–213. New York: Springer. DOI: 10.1007/978-1-4612-1694-0_15.

ALDAY, PHILLIP M.; MATTHIAS SCHLESEWSKY; and INA BORNKESSEL-SCHLESEWSKY. 2014. Towards a computational model of actor-based language comprehension. *Neuroinformatics* 12(1).143–79. DOI: 10.1007/s12021-013-9198-x.

AUSTIN, PETER, and JOAN BRESNAN. 1996. Non-configurationality in Australian Aboriginal languages. *Natural Language & Linguistic Theory* 14(2).215–68. DOI: 10.1007/BF00133684.

BAAYEN, R. HARALD; DOUGLAS J. DAVIDSON; and DOUGLAS M. BATES. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59(4).390–412. DOI: 10.1016/j.jml.2007.12.005.

BAKER, MARK C. 1996. *The polysynthesis parameter*. Oxford: Oxford University Press.

BARR, DALE J. 2008. Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language* 59(4).457–74. DOI: 10.1016/j.jml.2007.09.002.

BARR, DALE J. 2013. Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology* 4:328. DOI: 10.3389/fpsyg.2013.00328.

BARR, DALE J.; ROGER LEVY; CHRISTOPH SCHEEPERS; and HARRY J. TILY. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3).255–78. DOI: 10.1016/j.jml.2012.11.001.

BATES, DOUGLAS; MARTIN MÄCHLER; BEN BOLKER; and STEVE WALKER. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1).1–48. DOI: 10.18637/jss.v067.i01.

BICKEL, BALTHASAR; ALENA WITZLACK-MAKAREVICH; KAMAL K. CHOUDHARY; MATTHIAS SCHLESEWSKY; and INA BORNKESSEL-SCHLESEWSKY. 2015. The neurophysiology of language processing shapes the evolution of grammar: Evidence from case marking. *PLOS ONE* 10(8):e0132819. DOI: 10.1371/journal.pone.0132819.

BLAKE, BARRY J. 2001. *Case*. Cambridge: Cambridge University Press.

BLYTHE, JOE. 2009. *Doing referring in Murrinh-Patha conversation*. Sydney: University of Sydney dissertation.

BLYTHE, JOE; RACHEL NORDLINGER; and NICK REID. 2007. Murriny Patha finite verb paradigms. Wadeye, NT, MS.

BOCK, KATHRYN, and VICTOR FERREIRA. 2014. Syntactically speaking. *The Oxford handbook of language production*, ed. by Matthew Goldrick, Victor S. Ferreira, and Michele Miozzo, 21–46. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199735471.013.008.

BOCK, KATHRYN, and WILLEM J. M. LEVELT. 1994. Language production: Grammatical encoding. *Handbook of psycholinguistics*, ed. by Morton A. Gernsbacher, 945–84. London: Academic Press.

BORNKESSEL-SCHLESEWSKY, INA, and MATTHIAS SCHLESEWSKY. 2009. The role of prominence information in the real-time comprehension of transitive constructions: A cross-linguistic approach. *Language and Linguistics Compass* 3(1).19–58. DOI: 10.1111/j.1749-818X.2008.00099.x.

BORNKESSEL-SCHLESEWSKY, INA, and MATTHIAS SCHLESEWSKY. 2016. The importance of linguistic typology for the neurobiology of language. *Linguistic Typology* 20(3).615–21. DOI: 10.1515/lingty-2016-0032.

BOWERN, CLAIRE. 2014. Complex predicates in Australian languages. *The languages and linguistics of Australia: A comprehensive guide*, ed. by Harold Koch and Rachel Nordlinger, 263–94. Berlin: De Gruyter Mouton. DOI: 10.1515/9783110279771.263.

BRANIGAN, HOLLY P.; MARTIN J. PICKERING; and MIKIHIRO TANAKA. 2008. Contributions of animacy to grammatical function assignment and word order during production. *Lingua* 118(2).172–89. DOI: 10.1016/j.lingua.2007.02.003.

BRESNAN, JOAN. 2001. *Lexical-functional syntax*. Oxford: Blackwell.

CHRISTIANSON, KIEL, and HEE YOUN CHO. 2009. Interpreting null pronouns (*pro*) in isolated sentences. *Lingua* 119(7).989–1008. DOI: 10.1016/j.lingua.2008.11.005.

Christianson, Kiel, and Fernanda Ferreira. 2005. Conceptual accessibility and sentence production in a free word order language (Odawa). *Cognition* 98.105–35. DOI: 10.1016/j.cognition.2004.10.006.

de Swart, Peter; Monique Lamers; and Sander Lestrade. 2008. Animacy, argument structure, and argument encoding. *Lingua* 118(2).131–40. DOI: 10.1016/j.lingua.2007.02.009.

Dik, Simon C. 1978. *Functional grammar*. Amsterdam: North-Holland.

Dobel, Christian; Heidi Gumnior; Jens Bölte; and Pienie Zwitserlood. 2007. Describing scenes hardly seen. *Acta Psychologica* 125(2).129–43. DOI: 10.1016/j.actpsy.2006.07.004.

Dryer, Matthew, and Martin Haspelmath. 2013. *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Online: http://wals.info.

Du Bois, John W. 1987. The discourse basis of ergativity. *Language* 63(4).805–55. DOI: 10.2307/415719.

Duchowski, Andrew. 2007. *Eye tracking methodology*. 2nd edn. London: Springer. DOI: 10.1007/978-1-84628-609-4.

Evans, Nicholas. 2003. *Bininj Gun-wok: A pan-dialectal grammar of Mayali, Kunwinjku and Kune*. (Pacific linguistics 541.) Canberra: The Australian National University. DOI: 10.15144/PL-541.

Evans, Nicholas, and Stephen C. Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32(5).429–48. DOI: 10.1017/S0140525X0999094X.

Everett, Caleb. 2009. A reconsideration of the motivations for preferred argument structure. *Studies in Language* 33(1).1–24. DOI: 10.1075/sl.33.1.02eve.

Ferreira, Fernanda. 2000. Syntax in language production: An approach using tree-adjoining grammars. *Aspects of language production*, ed. by Linda Wheeldon, 291–330. New York: Psychology Press.

Foley, William. 2008. The place of Philippine languages in a typology of voice systems. *Voice and grammatical relations in Austronesian languages*, ed. by Peter Austin and Simon Musgrave, 22–44. Stanford, CA: CSLI Publications.

Futrell, Richard; Kyle Mahowald; and Edward Gibson. 2015. Quantifying word order freedom in dependency corpora. *Proceedings of the Third International Conference on Dependency Linguistics (Depling 2015)*, 91–100. Online: https://aclanthology.org/W15-2112.

Gennari, Silvia P.; Jelena Mirković; and Maryellen C. MacDonald. 2012. Animacy and competition in relative clause production: A cross-linguistic investigation. *Cognitive Psychology* 65(2).141–76. DOI: 10.1016/j.cogpsych.2012.03.002.

Gerwien, Johannes, and Monique Flecken. 2016. First things first? Top-down influences on event apprehension. *Proceedings of the 38th annual meeting of the Cognitive Science Society (CogSci 2016)*, 2633–38. Online: https://cogsci.mindmodeling.org/2016/papers/0454/index.html.

Gleitman, Lila R.; David January; Rebecca Nappa; and John C. Trueswell. 2007. On the *give* and *take* between event apprehension and utterance formulation. *Journal of Memory and Language* 57(4).544–69. DOI: 10.1016/j.jml.2007.01.007.

Griffin, Zenzi M., and Kathryn Bock. 2000. What the eyes say about speaking. *Psychological Science* 11(4).274–79. DOI: 10.1111/1467-9280.00255.

Hafri, Alon; John C. Trueswell; and Brent Strickland. 2018. Encoding of event roles from visual scenes is rapid, spontaneous, and interacts with higher-level visual processing. *Cognition* 175.36–52. DOI: 10.1016/j.cognition.2018.02.011.

Haig, Geoffrey, and Stefan Schnell. 2016. The discourse basis of ergativity revisited. *Language* 92(3).591–618. DOI: 10.1353/lan.2016.0049.

Hale, Ken. 1983. Warlpiri and the grammar of non-configurational languages. *Natural Language & Linguistic Theory* 1(1).5–47. DOI: 10.1007/BF00210374.

Hastie, Trevor, and Robert Tibshirani. 1990. *Generalized additive models*. London: Chapman & Hall/CRC.

Henrich, Joseph; Steven J. Heine; and Ara Norenzayan. 2010. The weirdest people in the world? *Behavioral and Brain Sciences* 33(2–3).61–83. DOI: 10.1017/S0140525X0999152X.

Hwang, Heeju, and Elsi Kaiser. 2014. The role of the verb in grammatical function assignment in English and Korean. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 40(5).1363–76. DOI: 10.1037/a0036797.

Jackendoff, Ray S. 1972. *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

Jaeger, T. Florian. 2008. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* 59(4).434–46. DOI: 10.1016/j.jml.2007.11.007.

Jaeger, T. Florian, and Elisabeth J. Norcliffe. 2009. The cross-linguistic study of sentence production. *Language and Linguistics Compass* 3(4).866–87. DOI: 10.1111/j.1749-818X.2009.00147.x.

Jelinek, Eloise. 1984. Empty categories, case, and configurationality. *Natural Language & Linguistic Theory* 2(1).39–76. DOI: 10.1007/BF00233713.

Konopka, Agnieszka E. 2019. Encoding actions and verbs: Tracking the time-course of relational encoding during message and sentence formulation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 45(8).1486–1510. DOI: 10.1037/xlm0000650.

Konopka, Agnieszka E., and Sarah Brown-Schmidt. 2014. Message encoding. *The Oxford handbook of language production*, ed. by Matthew Goldrick, Victor S. Ferreira, and Michele Miozzo, 3–20. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199735471.013.010.

Konopka, Agnieszka E., and Antje S. Meyer. 2014. Priming sentence planning. *Cognitive Psychology* 73.1–40. DOI: 10.1016/j.cogpsych.2014.04.001.

Kuchinsky, Stefanie E., and Kathryn Bock. 2010. From seeing to saying: Perceiving, planning, producing. Paper presented at the 23rd meeting of the CUNY Human Sentence Processing Conference, New York.

Kuchinsky, Stefanie E.; Kathryn Bock; and David E. Irwin. 2011. Reversing the hands of time: Changing the mapping from seeing to saying. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 37(3).748–56. DOI: 10.1037/a0022637.

Legate, Julie Anne. 2001. The configurational structure of a nonconfigurational language. *Linguistic Variation Yearbook* 1(1).61–104. DOI: 10.1075/livy.1.05leg.

Levelt, Willem J. M. 1989. *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levshina, Natalia. 2019. Token-based typology and word order entropy: A study based on universal dependencies. *Linguistic Typology* 23(3).533–72. DOI: 10.1515/lingty-2019-0025.

Mansfield, John. 2019. *Murrinhpatha morphology and phonology*. Berlin: De Gruyter Mouton. DOI: 10.1515/9781501503306.

Marmion, Doug; Kazuko Obata; and Jakelin Troy. 2014. *Community, identity, wellbeing: The report of the Second National Indigenous Languages Survey*. Canberra: Australian Institute of Aboriginal and Torres Strait Islander Studies. Online: https://aiatsis.gov.au/publication/35167.

McGregor, William B. 2002. *Verb classification in Australian languages*. Berlin: De Gruyter.

Mujkic, Elena. 2013. Murrinh-Patha syntax: Clausal structure and the noun phrase. Melbourne: University of Melbourne honors thesis.

Myachykov, Andriy; Simon Garrod; and Christoph Scheepers. 2018. Attention and memory play different roles in syntactic choice during sentence production. *Discourse Processes* 55(2).218–29. DOI: 10.1080/0163853X.2017.1330044.

Norcliffe, Elisabeth; Alice C. Harris; and T. Florian Jaeger. 2015. Cross-linguistic psycholinguistics and its critical role in theory development: Early beginnings and recent advances. *Language, Cognition and Neuroscience* 30(9).1009–32. DOI: 10.1080/23273798.2015.1080373.

Norcliffe, Elisabeth; Agnieszka E. Konopka; Penelope Brown; and Stephen C. Levinson. 2015. Word order affects the time course of sentence formulation in Tzeltal. *Language, Cognition and Neuroscience* 30(9).1187–1208. DOI: 10.1080/23273798.2015.1006238.

Nordlinger, Rachel. 2010. Verbal morphology in Murrinh-Patha: Evidence for templates. *Morphology* 20(2).321–41. DOI: 10.1007/s11525-010-9184-z.

NORDLINGER, RACHEL. 2011. Transitivity in Murrinh-Patha. *Studies in Language* 35(3). 702–34. DOI: 10.1075/sl.35.3.08nor.

NORDLINGER, RACHEL. 2015. Inflection in Murrinhpatha. *The Oxford handbook of inflection*, ed. by Matthew Baerman, 491–519. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199591428.013.21.

NORDLINGER, RACHEL, and JOAN BRESNAN. 2011. Lexical-functional grammar: Interactions between morphology and syntax. *Non-transformational syntax: Formal and explicit models of grammar*, ed. by Robert D. Borsley and Kersti Börjars, 112–40. Oxford: Wiley-Blackwell.

PORRETTA, VINCENT; AKI-JUHANI KYRÖLÄINEN; JACOLIEN VAN RIJ; and JUHANI JÄRVIKIVI. 2018. Visual world paradigm data: From preprocessing to nonlinear time-course analysis. *Intelligent design technologies 2017: Proceedings of the 9th KES International Conference on Intelligent Decision Technologies (KES-IDT 2017)*, ed. by Ireneusz Czarnowski, Robert J. Howlett, and Lakhmi C. Jain, 268–77. New York: Springer. DOI: 10.1007/978-3-319-59424-8_25.

R CORE TEAM. 2014. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Online: http://www.R-project.org/.

RIESBERG, SONJA; KURT MALCHER; and NIKOLAUS P. HIMMELMANN. 2019. How universal is agent-first? Evidence from symmetrical voice languages. *Language* 95(3).523–61. DOI: 10.1353/lan.2019.0055.

SAUPPE, SEBASTIAN. 2017a. Symmetrical and asymmetrical voice systems and processing load: Pupillometric evidence from sentence production in Tagalog and German. *Language* 93(2).288–313. DOI: 10.1353/lan.2017.0015.

SAUPPE, SEBASTIAN. 2017b. Word order and voice influence the timing of verb planning in German sentence production. *Frontiers in Psychology* 8:1648. DOI: 10.3389/fpsyg.2017.01648.

SAUPPE, SEBASTIAN, and MONIQUE FLECKEN. 2021. Speaking for seeing: Sentence structure guides visual event apprehension. *Cognition* 206:104516. DOI: 10.1016/j.cognition.2020.104516.

SAUPPE, SEBASTIAN; ELISABETH NORCLIFFE; AGNIESZKA E. KONOPKA; ROBERT D. VAN VALIN, JR.; and STEPHEN C. LEVINSON. 2013. Dependencies first: Eye tracking evidence from sentence production in Tagalog. *Proceedings of the 35th annual meeting of the Cognitive Science Society (CogSci 2013)*, 1265–70. Online: https://cogsci.mindmodeling.org/2013/papers/0242/index.html.

SCHAD, DANIEL J.; SHRAVAN VASISHTH; SVEN HOHENSTEIN; and REINHOLD KLIEGL. 2020. How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language* 110:104038. DOI: 10.1016/j.jml.2019.104038.

SCHULTZE-BERNDT, EVA. 2000. *Simple and complex verbs in Jaminjung: A study of event categorisation in an Australian language*. Nijmegen: Radboud University Nijmegen dissertation. Online: https://hdl.handle.net/2066/147040.

SCHULZ, JONATHAN; DUMAN BAHRAMI-RAD; JONATHAN BEAUCHAMP; and JOSEPH HENRICH. 2018. The origins of WEIRD psychology. Available at SSRN. DOI: 10.2139/ssrn.3201031.

SENSOMOTORIC INSTRUMENTS. 2017. BeGaze Manual (version 3.7). Online: http://www.humre.vu.lt/files/doc/Instrukcijos/SMI/BeGaze2.pdf.

SHANNON, CLAUDE E. 1948. A mathematical theory of communication. *The Bell System Technical Journal* 27(3).379–423. DOI: 10.1002/j.1538-7305.1948.tb01338.x.

SIEWIERSKA, ANNA. 1998. Variation in major constituent order: A global and a European perspective. *Constituent order in the languages of Europe*, ed. by Anna Siewierska, 475–552. Berlin: De Gruyter. DOI: 10.1515/9783110812206.475.

SILVERSTEIN, MICHAEL. 1976. Hierarchy of features and ergativity. *Grammatical categories in Australian languages*, ed. by R. M. W. Dixon, 112–71. Canberra: Australian Institute for Aboriginal Studies.

TANAKA, MIKIHIRO N.; HOLLY P. BRANIGAN; JANET F. MCLEAN; and MARTIN J. PICKERING. 2011. Conceptual influences on word order and voice in sentence production: Evidence from Japanese. *Journal of Memory and Language* 65(3).318–30. DOI: 10.1016/j.jml.2011.04.009.

TOMLIN, RUSSELL S. 1997. Mapping conceptual representations into linguistic representations: The role of attention in grammar. *Language and conceptualization*, ed. by Jan

Nuyts and Eric Pederson, 162–89. Cambridge: Cambridge University Press. DOI: 10 .1017/CBO9781139086677.007.

VAN VALIN, ROBERT D., JR., and RANDY J. LAPOLLA. 1997. *Syntax: Structure, meaning, and function*. Cambridge: Cambridge University Press.

VENABLES, W. N., and B. D. RIPLEY. 2002. *Modern applied statistics with S*. (Statistics and computing.) New York: Springer. DOI: 10.1007/978-0-387-21706-2.

WALSH, MICHAEL. 1976a. *The Murinypata language of North-West Australia*. Canberra: The Australian National University dissertation. DOI: 10.25911/5d78d8be956e6.

WALSH, MICHAEL. 1976b. Ergative, locative and instrumental case inflections: Murinjpata. *Grammatical categories in Australian languages*, ed. by R. M. W. Dixon, 405–8. Canberra: Australian Institute for Aboriginal Studies.

WILSON, STEPHEN. 1999. *Coverbs and complex predicates in Wagiman*. Stanford, CA: CSLI Publications.

WOOD, SIMON N. 2017. *Generalized additive models: An introduction with R*. 2nd edn. Boca Raton: Chapman & Hall/CRC.

ZUUR, ALAIN; ELENA N. IENO; NEIL WALKER; ANATOLY A. SAVELIEV; and GRAHAM M. SMITH. 2009. *Mixed effects models and extensions in ecology with R*. New York: Springer. DOI: 10.1007/978-0-387-87458-6.

ZUUR, ALAIN; ANATOLY A. SAVEL; and ELENA N. IENO. 2014. *A beginner's guide to additive mixed models with R*. Newburgh: Highland Statistics.

[racheln@unimelb.edu.au]
[gabriela.garrido@mpi.nl]
[evan.kidd@mpi.nl]