Attraction effects in Nungon reveal independent mechanisms for agreement and switch-reference marking

Adam M. Morgan[1*], Ismael Dono[2,3,4], Lyn Ögate [2,5],
Stanly Girip [2,5], Hannah Sarvasy[6]

[1]Department of Neurology, NYU Grossman School of Medicine, New York 10016, USA.

[2]Towet Village Community, Uruwa Ward 1, Morobe Province, Papua New Guinea.

[3]Papua New Guinea University of Technology, Lae 411, Morobe Province, Papua New Guinea.

[4]Meptain Agricultural Organization, Uruwa Ward 1, Morobe Province, Papua New Guinea.

[5]Simebai Community Organization, Uruwa Ward 1, Morobe Province, Papua New Guinea.

[6]MARCS Institute for Brain, Behaviour and Development, Penrith 2751, Australia.

[*]Corresponding author. Email: adam.morgan@nyulangone.org

Like the majority of languages, Nungon, spoken by ∼1,000 people in Papua New Guinea, has subject-verb agreement. However, it additionally requires speakers to mark verbs for whether the subject of the *next*, as-yet-unspoken, sentence will be the same or different from the current sentence's subject. Here, we elicited narratives from Nungon speakers, manipulating properties of distractor nouns, to investigate this rare property known as *switch-reference marking*. We observed a strikingly high degree of agreement attraction errors (17%), but an even more striking number of "switch-reference attraction" errors: *Zero*. These findings reveal distinct processes for agreement and switch-reference marking, and resolve long-standing questions about agreement processing, favoring retrieval and feature-based accounts of agreement.

In the history of psycholinguistics, the greatest insights have sometimes come not from investigating how language functions, but from how it doesn't. Cases where language processing goes awry – tongue twisters (*1, 2*), illusions (*3, 4*), and violations (*5–7*) – have led to some of the most critical advances in understanding underlying mechanisms. Perhaps the most famous example is *agreement attraction* (*8*), as in *"The key to the cabinets* ***are*** *on the table,"* where the verb *are* mistakenly agrees with an "attractor" noun, *cabinets*, rather than the true subject, *key*.

Agreement attraction has revealed important facts about language (*9–11*) – and about human cognition more broadly (*12–14*). Yet despite the wealth of research on agreement attraction, two fundamental questions remain unresolved: (i) *when* agreement is computed, and (ii) exactly *what* is agreed with.

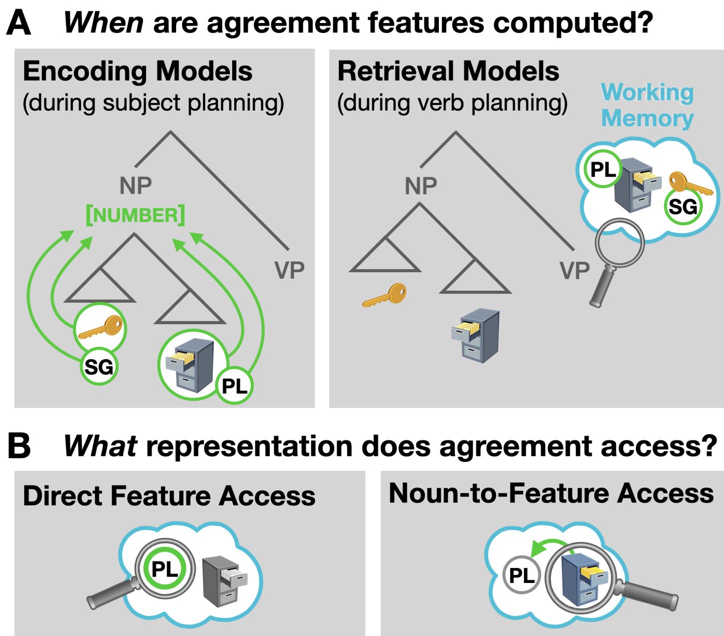

Figure 1: **Competing theories for the two open questions about agreement processing.** The figure depicts the canonical example subject, '*the key to the cabinets*.' (A) A schematic of two families of models addressing *when* agreement is computed in production. Encoding Models (left) posit that a number feature of the whole subject NP is computed during subject planning. Agreement Attraction errors occur when other elements being planned interfere with this process (green arrows). Retrieval Models (right) hold that agreement is computed during verb planning, when the production system executes a memory search for something to agree with. Exactly *what* is searched for is the second open question. Panel (B) depicts two possibilities (assuming a Retrieval Model for ease of depiction): the number feature itself (left) or a representation of the subject (right), which is then used to retrieve the bound number feature (green arrow).

Regarding *when*, models broadly fall into two categories (Fig. 1a). According to *Encoding Models* [e.g., (*8, 15–18*)], attraction features (things like NUMBER and PERSON) are computed early on during production – specifically, during subject planning. Attraction errors stem from interference in this process, leading to incorrect encoding of features. This error is then carried forward and surfaces on the verb. In contrast, *Retrieval Models* posit that during verb planning, the production system searches for something to agree with (*13, 19, 20*). These latter models attribute attraction errors to a faulty cue-based retrieval mechanism, which occasionally picks out the wrong element from working memory, resulting in an agreement error.

Exactly *what* that element is is the second unresolved question. Words like *scissors* and *pants*, which are notionally singular but grammatically plural, trigger plural agreement (scissors *are*, not *is*), indicating that agreement ultimately relies on grammatical features rather than meaning. However, it remains unclear whether the production system targets these features directly, or if they are extracted from the subject representation (Fig. 1b). These questions have been difficult to address in more commonly studied languages, where it is hard to disentangle words' features from their reference.

Figure 2: **Towet Village, located in a mountain valley in Morobe Province of Papua New Guinea, is home to ~300 Nungon speakers.** The village has a dedicated research building with several solar panels and batteries, which local teams have been using to run child language and language processing studies since 2015.

Nungon, a Papuan language spoken by ~1,000 people in remote villages in Papua New Guinea (Fig. 2), has rare grammatical properties that naturally circumvent these limitations. In a stark departure from more familiar grammatical systems, Nungon not only marks verbs for features of their own subject – ordinary agreement – but also for whether their subject is the same as or different from the subject of the *next* clause the speaker will say (*21*). Thus, before they are able to complete one clause, Nungon speakers must plan the next clause – a non-trivial task (see Supplementary Text and Ex. 3) that requires significant advance planning (*22*). This rare grammatical property, *switch-reference marking*, precisely dissociates referential and grammatical properties of the subject, offering an unusually valuable opportunity to disentangle what kinds of information are available to the production system and when and making Nungon an ideal test bed for the core mechanisms underlying agreement.

Switch-reference marking in Nungon occurs in structures known as *clause chains*. Unlike more commonly studied languages like English, where strings of clauses are grammatically independent (e.g., *The girl tripped; she cried out; her parents came running; etc.*), in clause chains only the final verb is marked for tense/mood, leaving the listener to wait until the end of what can be 20 or more clauses to know when the events took place (*21, 23*). Within these clause chains, all non-final verbs are marked for whether the subject of the next, as-yet-unspoken clause will be the same as or different from the subject of the current clause, as in Ex. 1:

(1)	*amna*	*boop*	*ongo-nga,*	*ewek*	*a-una,*	*yi-una,*	*ut-do-k.*
	man	forest	go-SAME	snake	see-DIFF.3SG	3SG.OBJ.bite-DIFF.3SG	cry-REMOTE.PAST-3SG

‘A man went into the forest, he saw a snake, it bit him, he cried.’

Here the SAME morpheme on the verb *go* indicates that the subject of the next clause – what we will refer to as the *upcoming subject* – will be the same as the subject of the current clause – the *current subject* (i.e., *man*). The DIFF morpheme on the verb *see* indicates that the following subject (*snake*) will differ from the current one. Note that, unlike SAME markers, DIFF markers agree with the current subject, changing depending on its person and number features. (See Supplementary Text and (*21*) for detail.)

These properties provide a natural way of identifying *when* agreement is computed. Specifically, Nungon grammar imposes logical boundaries on when the upcoming subject can be planned. Clauses are planned in the order they are produced (*24, 25*), meaning that the upcoming subject should not be planned until after the current subject, defining the beginning of the window. Similarly, because switch-reference forces speakers to plan the upcoming subject before they can produce the verb, the onset of verb articulation marks the end of the planning window. Thus, the upcoming subject should in general be planned after current subject planning and before verb production – a window which is entirely after the subject planning window, when Encoding Models hold that agreement is computed, but which overlaps with the verb planning window, where Retrieval Models locate agreement processing. These two model families therefore make different predictions for whether the upcoming subject should be available to interfere with agreement processing. Under Retrieval Models, where agreement and upcoming subject planning overlap, we should expect upcoming subjects to be viable attractors for agreement, increasing the number of agreement errors when their features differ from those of the current subject. In contrast, such errors should be impossible under Encoding Models, which hold that agreement is computed before the upcoming subject is planned (*24,25*). A key puzzle piece, then, is whether upcoming subjects trigger agreement attraction in Nungon.

Nungon grammar also provides a natural way of addressing *what* information the agreement system accesses. Switch-reference marking depends on the referent of the subject, independent of its grammatical features (i.e., a sequence of "cat" then "mouse" subjects would trigger a DIFF switch-reference morpheme, even though both nouns would trigger the same third-person singular agreement). If agreement attraction errors stem from identifying the wrong noun as the subject (rather than the wrong features), then this should also have consequences for switch-reference. For instance, if an attractor noun has the same referent as the upcoming subject, then would-be instances of agreement attraction should instead result in a SAME morpheme (which does not come with agreement morphology in Nungon; Fig. 3D). However, if agreement attraction reflects misidentified features, then it should be possible for switch-reference marking to correctly identify a switch in subjects, even when agreement is with the wrong feature (Fig. 3C). Thus, the critical question is whether attraction errors in agreement and switch-reference are ever independent of one another, which would be evidence that agreement searches for features, not the subject.

By manipulating properties of attractor nouns in the current and upcoming clauses, we were able to address two central yet unresolved questions about the production of agreement morphology: *when* it is computed and *what* information about the subject it accesses. We first show that attraction effects can be induced by the upcoming subject, evidence which is consistent with Retrieval Models but inconsistent with Encoding Models. Next, we investigated whether agreement relies on the same representations as switch-reference marking by introducing *reference* attractors in addition to agreement attractors. Strikingly, despite high rates of agreement attraction in Nungon production, we found that switch-reference marking is completely impervious to such effects, showing no attraction errors whatsoever. The distinct error patterns for agreement and switch-reference reveal that the two systems do not share underlying mechanisms. Taken together, our results support a picture of agreement processing which involves a retrieval mechanism that directly targets grammatical features, bypassing other information about the subject like its referential properties. They also paint a picture of Nungon production as a remarkably complex system, which

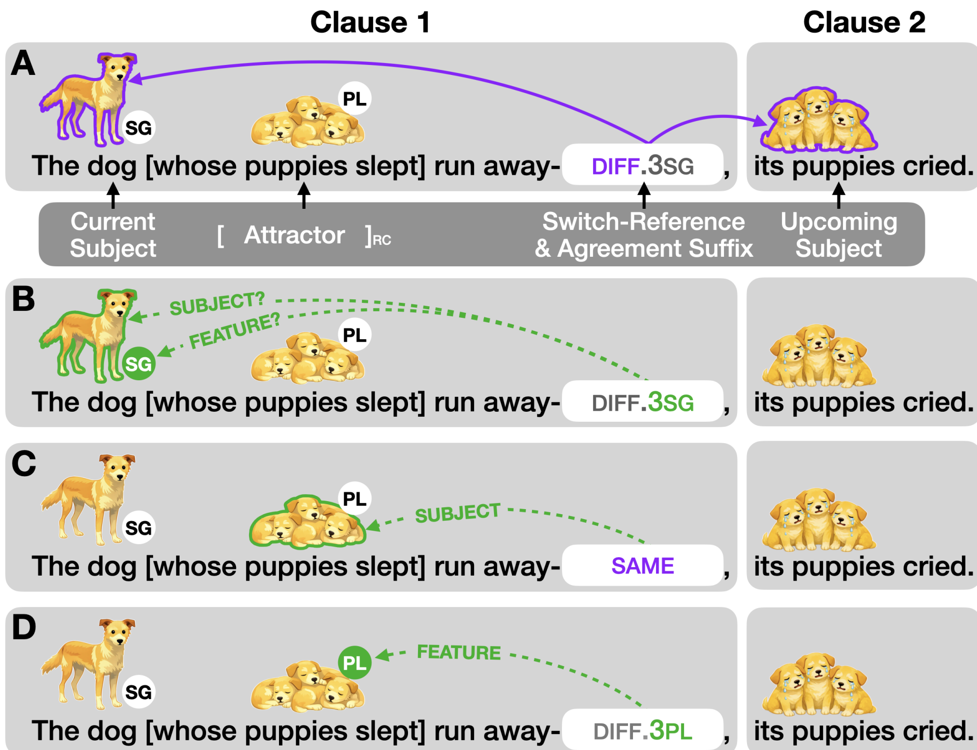

Figure 3: **Schematizing the possibility space for *what* representations agreement and switch-reference marking access.** The example is a pseudo-gloss of a clause chain taken from our stimuli. Clause 1 contains a relative clause that introduces the noun *puppies*, which serves as an attractor for both agreement (plural number) and switch-reference (same referent as the upcoming subject). (A) A schematization of switch-reference processing, which relies on a comparison of the current and upcoming subjects (outlined in purple). (B) Competing hypotheses for agreement processing: While agreement ultimately relies on an abstract number feature, it remains an open question whether it targets this information directly (bottom dashed line) or if it arrives at the features by first accessing the subject (top dashed line). (C) What agreement attraction would look like assuming a model in which the system accesses the subject first. In agreement attraction, the production system mistakenly identifies the attractor noun. Here, because the attractor and the upcoming subject are both 'puppies,' this results in a SAME morpheme (which does not come with agreement in Nungon). (D) What agreement attraction would look like assuming the system accesses features directly. Here, the system mistakenly identifies the PLURAL feature on the attractor noun *puppies*, resulting in incorrect agreement. Switch-reference is computed over subjects (not features), so a separate mechanism would be required to identify the subject here. Here, we depict the case where this mechanism correctly identifies subject, resulting in a DIFF morpheme along with the incorrect PL agreement morpheme.

Figure 4: **Experimental setup.** The experiment was run in the team's dedicated research building, administered in Nungon by Author S. G.

requires independent systems for tracking and comparing several distinct kinds of information about both current and upcoming subjects, all in order to produce the verb. This finding has implications not just for morphological processing, but more broadly for theories of computational economy, where efficiency is thought to be a primary driving force in shaping language processing. These implications are explored in more detail in the Discussion.

Results

We ran a production experiment in Towet, a Nungon-speaking village in Papua New Guinea, eliciting clause chains from 43 native Nungon speakers (see Fig. 4). Stimuli were clause chains consisting of two clauses (*Clause 1* and *Clause 2*) and were presented auditorily, with grammatically correct switch-reference and agreement marking. Participants were instructed to repeat the stimuli after hearing the entire stimulus, a method commonly used to elicit complex utterances [e.g., (*26*)] as it engages message-driven production (*25, 27*) rather than mere repetition of surface-level representations (*28, 29*). Two sub-experiments were embedded into the stimuli, one to address each outstanding theoretical question about agreement.

Sub-Experiment 1 pitted the Encoding and Retrieval Models against one another by asking *when* agreement errors occur. Critically, Retrieval Models predict that upcoming subjects should be able to attract agreement – an error that should be impossible under Encoding Models. To test this, we manipulated the number feature of the Clause 2 subject so that it was either the same as that of the Clause 1 subject (baseline condition) or different (attraction condition; see Fig. 5). As expected, we observed no errors in agreement marking in the baseline condition (Fig. 6). However, changing the number feature on the upcoming subject led to a statistically significant increase in agreement errors (p = .002, generalized linear mixed effects model; see Supplementary Text for full statistics). This result is consistent with the predictions of Retrieval Models, but inconsistent with Encoding Models.

In Sub-Experiment 2, we addressed *what* is retrieved by comparing the distribution of errors in agreement and switch-reference marking. We crossed the presence/absence of an agreement attractor and what we termed a *switch-reference attractor*, resulting in a 2×2 design (Fig. 7). A

Agreement Attractor?	Clause 1					Clause 2	
	Subject	Relative Clause	Verb	(Correct) SR Suffix	*Target Error*	Subject	Verb
No	Dog	who its-3SG child-SG slept-3SG	run.away-	SWITCH.3SG	—	Its-3SG owner-SG	cried-3SG-PAST
Yes	Dog	who its-3SG child-SG slept-3SG	run.away-	SWITCH.3SG	*SWITCH.3PL*	Its-3SG owners-**PL**	cried-3SG-PAST

Figure 5: **Stimuli in the first sub-experiment manipulated the number feature on the Clause 2 (i.e., *upcoming*) subject (green).** Retrieval Models predict that upcoming subjects should lead to agreement attraction errors (in orange) on the Clause 1 verb "run.away," while Encoding Error models predict that this should not be possible.

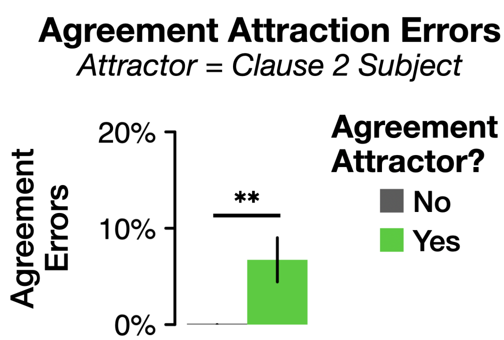

Figure 6: **Sub-Experiment 1 results: Agreement errors by condition.** Participants produced more agreement errors when the number features of the *upcoming* and *current* subjects differed (green; 8 agreement errors out of 119 trials) than when they were the same (gray; 0 agreement errors out of 174 trials). This statistically significant increase is consistent with Retrieval-Error Models, and inconsistent with Encoding-Error Models.

Reference Attractor?	Agreement Attractor?	Clause 1: Subject	Clause 1: Relative Clause	Clause 1: Verb	Clause 1: (Correct) SR Suffix	*Target Error*	Clause 2: Subject	Clause 2: Verb
No	**No**	Dog	who its-3SG child-SG slept-3SG	run.away-	SWITCH.3SG	—	Its-3SG owner-SG	cried-3SG-PAST
No	**Yes**	Dog	who its-3SG children-**PL** slept-3PL	run.away-	SWITCH.3SG	*SWITCH.3PL*	Its-3SG owner-SG	cried-3SG-PAST
Yes	**No**	Dog	who its-3SG **child**-SG slept-3SG	run.away-	SWITCH.3SG	*STAY*	Its-3SG **child**	cried-3SG-PAST
Yes	**Yes**	Dog	who its-3SG **children**-**PL** slept-3PL	run.away-	SWITCH.3SG	*STAY or SWITCH.3PL*	Its-3SG **children**-PL	cried-3PL-PAST

Figure 7: **Sub-Experiment 2 stimuli appeared in a** 2×2 **design.** We crossed the presence/absence of attractors for switch-reference (first column) and agreement (second column). Each stimulus was a clause chain consisting of two clauses. Clause 1 contained a relative clause (RC) that introduced a distractor noun (*child*) between the matrix subject and verb. By manipulating the distractor's number (agreement attraction; green) and whether it was the same as the Clause 2 subject (switch-reference attraction; purple), we aimed to elicit the incorrect morphemes in the *Target Error* column (orange).

relative clause (RC) modifying the subject of Clause 1 introduced the attractor noun (in the sample item: *child*, meaning 'puppy'). In AGREEMENT ATTRACTOR conditions, the RC noun's number was different from the number of the Clause 1 subject's head noun (*dog*). (Note that this contrasts with Sub-Experiment 1, where the agreement attractor appeared in Clause 2 rather than Clause 1.) In REFERENCE ATTRACTOR conditions, the RC noun was manipulated to be identical to the Clause 2 (i.e., upcoming) subject so as to induce attraction effects in switch-reference marking. In all conditions, an image of the relative clause noun (i.e., one or multiple people or animals, according to the particular item and condition; see Fig. S1) appeared on the screen throughout the trial to draw attention to the attractor noun.

If agreement relies on the same subject representation as switch-reference marking, then when retrieval processes erroneously pick out the attractor noun to agree with, as in cases of agreement attraction, we should see corresponding errors in switch-reference marking. Concretely, in the condition with both switch-reference and agreement attractors, participants should produce either the correct DIFF.3SG marker or an incorrect SAME marker (see Fig. 3), but never the DIFF.3PL marker, which would reflect agreement attraction without reference attraction. However, if agreement targets abstract features like *number*, then we should observe both SAME and DIFF.3PL errors in the final condition, as speakers should sometimes make agreement errors but get switch-reference right (i.e., DIFF.3PL), and other times the reverse (SAME).

In agreement, we observed an unusually high number of attraction errors (Fig. 8). Cross-

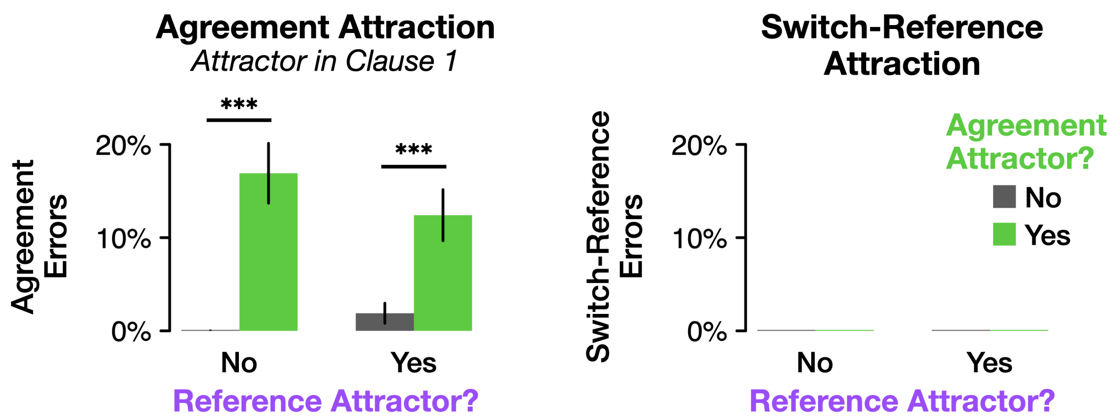

Figure 8: **Results of Sub-Experiment 2.** Agreement attraction errors are plotted in the left panel. Participants produced more agreement errors when number attractors were present (green bars) than when they were not (gray). The presence of a number attractor increased the likelihood of an error compared to the baseline condition ($p < .001$). Pairwise tests confirmed that this was true both when a reference attractor was present ($p < .001$) and absent ($p < .001$). The presence of a reference attractor (right two bars) did not significantly impact rates of agreement errors ($p = .686$). The right panel shows switch-reference errors, of which there were none in any condition.

linguistically, agreement attraction is typically observed in between 5% and 10% of trials (*12, 17, 19, 30–34*). Here the presence of an agreement attractor significantly increased the likelihood of an agreement error ($p < .001$, generalized linear mixed effects model; see Supplementary Materials for full statistics) from 0 in the baseline condition (174 included trials) to 17% in the agreement attractor condition (23 agreement errors in 136 trials).

However, the most striking feature of our data was that despite the unusually high rate of agreement attraction errors, there was a complete absence of switch-reference errors (Fig. 8, right). The distinct error patterns for switch-reference and agreement demonstrate that these two morphological systems involve distinct underlying processes for accessing information about the subject. This is consistent with agreement involving a search that directly targets the features of the subject, while switch-reference marking depends on a representation of the subject itself, which is presumably less susceptible to interference than abstract features.

Discussion

Nungon, a language spoken by ~1,000 people in a remote mountain valley in Papua New Guinea, requires speakers to mark verbs for whether the subject of the upcoming, as-yet-unspoken clause will be the same or different from the current subject. Here, we leveraged this complex system to manipulate whether and when different kinds of distractor information are available to the production system. Three striking patterns emerged.

First, because switch-reference marking entails planning the next clause before producing the verb, we were able to induce agreement attraction by manipulating the number feature on the *upcoming* subject. With few exceptions [e.g., (*35*)], previous reports of agreement attraction effects come from attractors that precede the verb. Indeed, previous evidence has suggested that it might in fact not be possible for elements outside the clause to attract agreement (*36*). This evidence, however, comes from languages in which the clause is a grammatically independent unit. Here, we demonstrate that in Nungon clause chains, where verb forms depend on nouns outside the local clause, agreement attraction effects can be induced by a noun that comes not only after the verb, but after the whole clause.

This finding highlights the importance of interpreting processing patterns with respect to language-specific grammatical properties, adding to growing evidence that cautions against generalizing based on data from similar languages. Second, Nungon speakers were highly susceptible to agreement attraction effects, producing far more agreement errors (~17% of trials) than observed in other languages. Yet, despite this susceptibility in agreement, the third surprising result was that Nungon speakers were completely immune to attraction effects in switch-reference marking.

These results speak to two long-debated questions about the core mechanisms of agreement. First is *when* agreement is computed: either during subject encoding, as in Encoding Models, or later on during verb planning, as in Retrieval Models. Our observation of attraction effects from upcoming subjects places the locus of agreement errors in the verb-planning stage, supporting Retrieval Models over Encoding Models. This interpretation is consistent with broader trends in the agreement attraction literature, where, despite early evidence supporting Encoding Models [e.g., *Marking and Morphing*; (*8, 15–18*)], experimental work increasingly favors retrieval-based accounts (*13, 19, 20*).

Regarding the second question – *what* information is accessed for computing agreement – the literature delineates two possibilities: the subject itself, or abstract features (e.g., PERSON, NUMBER). A shared retrieval mechanism for switch-reference and agreement would indicate that agreement, like switch-reference, relies on a representation of the subject. However, our data show that errors in agreement were independent of errors in switch-reference, meaning that these two morphological systems cannot rely on the same underlying process and agreement likely relies on features rather than the subject itself. This conclusion finds support in the literature, with previous work demonstrating that retrieval processes are insensitive to surface form (*37*) and that agreement features can be incrementally updated in memory independent of the noun's discourse-level representation (*38*). (See also (*13, 39*) for mechanistic accounts of agreement in a cue-based retrieval framework.)

The complexity of Nungon production. Our data also pose an interesting problem for linguistic theories that rely heavily on principles of computational economy. For one, growing evidence supports the idea that efficiency is a powerful driving force in the evolution of grammatical systems (*40, 41*). For instance, languages can encode syntactic roles (who did what to whom) in a few ways: with fixed word order (e.g., English), case marking (e.g., Japanese), or verbal morphology (e.g., Q'anjob'al). But it is rare for languages to simultaneously employ multiple of these approaches, and when languages change to gain one of these systems, they tend to lose the others (*42–45*).

Regarding diachronic language change, Nungon provides an important data point. Nungon has flexible word order, argument dropping, and no case marking, meaning that neither nouns nor their order provide clear cues to their syntactic roles. Thus, a functional role for agreement would be highly sensible: it would serve to convey which noun is the subject. But a functional account of switch-reference marking is much harder to formulate. It is simultaneously unnecessary – evidenced by its rarity among the world's languages – and computationally demanding, requiring advance planning that appears to far exceed what is required by other languages (*22*). Furthermore, the results of Sub-Experiment 1 (Fig. 6) show that it introduces additional sources of interference for agreement, making agreement even more difficult and error-prone than in languages without switch-reference marking.

Efficiency is not only a guiding force diachronically, but also synchronically in language processing (*46*). For instance, speakers strive to maintain a fixed rate of information per unit time (*47*), facilitating production and mitigating comprehension difficulty (*48*). However, our findings indicate that here too efficiency cannot be the whole story. A purely efficiency-based processing model would predict that Nungon speakers should cut corners where possible – for instance, by using a shared search mechanism for both agreement and switch-reference, or by staggering these processes in time to minimize complex parallel computations. However, Sub-Experiment 2 demonstrated that switch-reference and agreement rely on distinct subject access mechanisms, and Sub-Experiment 1 showed that both computations occur in the same narrow time window between subject planning and verb production. From a processing perspective, the 17% agreement attraction rate clearly confirms that Nungon's agreement system is computationally burdensome, and the fact that the production system does not take advantage of such opportunities for streamlining presents a challenge to the primacy of efficiency in modern processing models (*46*). Indeed, these facts highlight that efficiency is not the only guiding force in language processing and change. Our findings support more multifaceted theories that leave room for other influences such as learnability biases (*49*), social cohesion (*50–52*), and even potentially aesthetics (*53*), although it remains unclear which (combination) of these might account for switch-reference marking in Nungon.

Limitations. Future work should aim to validate two assumptions we relied on in interpreting our data. First, we argued that the upcoming subject should not be active in working memory when the current subject is being planned. Under ordinary production circumstances, where clauses are planned in the order they are uttered, this is likely a safe assumption (*24, 25*). However, in our task, there remains the possibility that the upcoming subject remains active in working memory throughout production as an artifact of the verbatim repetition task (*54–56*). A different elicitation paradigm like scene description, while less controlled, would be elucidating in this regard. The second assumption we relied on is that if both agreement and switch-reference rely on accessing the subject noun, this would be carried out in the same operation. However, there also exist other logically possible scenarios. For instance, one possibility is that agreement and switch-reference might both involve a search for the subject noun, just not the same search. Such redundancy seems unlikely given the system's propensity for efficiency, not to mention the fact that one would additionally need to account for why one of these searches is so much more error prone (agreement) than the other (switch-reference).

Conclusion. Nungon's switch-reference system forces speakers to plan the next clause's subject before they finish the current one, a fact which we leveraged to directly probe when and how agreement is computed. Our data showed that, while highly susceptible to agreement attraction, Nungon speakers were completely impervious to parallel errors in the arguably more complex switch-reference marking system. These findings indicate that speakers were simultaneously tracking multiple kinds of information about the subject, revealing a remarkable degree of complexity in Nungon production. Our findings further address two long-standing questions about agreement: *when* it is computed in the course of sentence production, and what information about the subject it accesses. Regarding *when*, our data support a family of models that rely on a memory search mechanism during verb planning. Regarding *what*, our data indicate that this search mechanism does not seek out the subject itself, but instead an associated abstract grammatical feature (in the present case, NUMBER).

Our findings underscore the importance of typologically rare languages in addressing questions that are effectively untestable in English and related languages, and invite a re-evaluation of cross-linguistic processing claims based on that narrow set. Finally, by highlighting the complexity of verbal morphology processing in Nungon, our findings sharpen a broader puzzle about human cognition: why so many languages develop elaborate systems of verbal morphology when these systems are unnecessary and highly error-prone. The present data help define the constraints that any answer to this puzzle must satisfy.

References and Notes

1. S. A. Frisch, R. Wright, The phonetics of phonological speech errors: An acoustic analysis of slips of the tongue. *Journal of Phonetics* **30** (2), 139–162 (2002).

2. G. M. Oppenheim, G. S. Dell, Inner speech slips exhibit lexical bias, but not the phonemic similarity effect. *Cognition* **106** (1), 528–537 (2008), doi:10.1016/j.cognition.2007.02.006.

3. F. Ferreira, K. G. Bailey, V. Ferraro, Good-enough representations in language comprehension. *Current Directions in Psychological Science* **11** (1), 11–15 (2002), doi:10.1111/1467-8721.0015.

4. C. Phillips, M. W. Wagers, E. F. Lau, Grammatical illusions and selective fallibility in real-time language comprehension. *Experiments at the Interfaces* **37**, 147–180 (2011), doi:10.1163/9781780523750_006.

5. F. Ferreira, J. M. Henderson, Recovery from misanalyses of garden-path sentences. *Journal of Memory and Language* **30** (6), 725–745 (1991), doi:10.1016/0749-596X(91)90034-H.

6. K. Christianson, A. Hollingworth, J. F. Halliwell, F. Ferreira, Thematic roles assigned along the garden path linger. *Cognitive Psychology* **42** (4), 368–407 (2001), doi:10.1006/cogp.2001.0752.

7. M. Kutas, S. A. Hillyard, Event-related brain potentials to grammatical errors and semantic anomalies. *Memory & Cognition* **11** (5), 539–550 (1983).

8. K. Bock, C. A. Miller, Broken agreement. *Cognitive Psychology* **23** (1), 45–93 (1991).

9. H. Lorimor, K. Bock, E. Zalkind, A. Sheyman, R. Beard, Agreement and attraction in Russian. *Language and Cognitive Processes* **23** (6), 769–799 (2008), doi:10.1080/01690960701774182.

10. J. C. Acuña-Fariña, Agreement, attraction and architectural opportunism1. *Journal of Linguistics* **48** (2), 257–295 (2012), doi:10.1017/S0022226712000084.

11. G. Vigliocco, One or more labels on the bottles? Notional concord in Dutch and French. *Language and Cognitive Processes* **11** (4), 407–442 (1996), doi:10.1080/016909696387169.

12. J. Franck, G. Lassi, U. H. Frauenfelder, L. Rizzi, Agreement and movement: A syntactic analysis of attraction. *Cognition* **101** (1), 173–216 (2006), doi:10.1016/j.cognition.2005.10.003.

13. M. W. Wagers, E. F. Lau, C. Phillips, Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language* **61** (2), 206–237 (2009).

14. A. C. Bleotu, B. Dillon, Romanian (subject-like) DPs attract more than bare nouns: Evidence from speeded continuations. *Journal of Memory and Language* **134**, 104445 (2024), doi:10.1016/j.jml.2023.104445.

15. N. J. Pearlmutter, S. M. Garnsey, K. Bock, Agreement processes in sentence comprehension. *Journal of Memory and Language* **41** (3), 427–456 (1999).

16. K. M. Eberhard, J. C. Cutting, K. Bock, Making syntax of sense: number agreement in sentence production. *Psychological Review* **112** (3), 531 (2005).

17. J. Franck, G. Vigliocco, J. Nicol, Subject-verb agreement errors in French and English: The role of syntactic hierarchy. *Language and Cognitive Processes* **17** (4), 371–404 (2002), doi: 10.1080/01690960143000254.

18. A. Staub, On the interpretation of the number attraction effect: Response time evidence. *Journal of Memory and Language* **60** (2), 308–327 (2009).

19. W. Badecker, F. Kuminiak, Morphology, agreement and working memory retrieval in sentence production: Evidence from gender and case in Slovak. *Journal of Memory and Language* **56** (1), 65–85 (2007).

20. L. A. Jäger, F. Engelmann, S. Vasishth, Similarity-based interference in sentence comprehension: Literature review and Bayesian meta-analysis. *Journal of Memory and Language* **94**, 316–339 (2017).

21. H. Sarvasy, *A grammar of Nungon: A Papuan language of northeast New Guinea*, vol. 4 (Brill) (2017).

22. H. S. Sarvasy, A. M. Morgan, J. Yu, V. S. Ferreira, S. Momma, Cross-clause planning in nungon (papua new guinea): Eye-tracking evidence. *Memory & Cognition* **51** (3), 666–680 (2023), doi:10.3758/s13421-021-01253-3.

23. P. Söderström, H. Sarvasy, Processing clause chains in Nungon, in *Crosslinguistic Perspectives on Processing and Learning Workshop* (2023).

24. V. M. Holmes, Sentence planning in a story continuation task. *Language and Speech* **27** (2), 115–134 (1984), doi:10.1177/002383098402700202.

25. K. Bock, W. J. Levelt, Language production: Grammatical encoding, in *Handbook of psycholinguistics* (Academic Press), pp. 945–984 (1994).

26. L. Giglio, P. Hagoort, M. Ostarek, Neural encoding of semantic structures during sentence production. *Cerebral Cortex* **34** (12) (2024).

27. L. R. Slevc, Grammatical encoding, in *Language production* (Routledge), pp. 4–31 (2023).

28. M. C. Potter, L. Lombardi, Regeneration in the short-term recall of sentences. *Journal of Memory and Language* **29** (6), 633–654 (1990), doi:10.1016/0749-596X(90)90042-X.

29. L. Lombardi, M. C. Potter, The regeneration of syntax in short term memory. *Journal of Memory and Language* **31** (6), 713–733 (1992), doi:10.1016/0749-596X(92)90036-W.

30. K. Bock, J. Nicol, J. C. Cutting, The ties that bind: Creating number agreement in speech. *Journal of Memory and Language* **40** (3), 330–346 (1999), doi:10.1006/jmla.1998.2616.

31. J. Franck, G. Soare, U. H. Frauenfelder, L. Rizzi, Object interference in subject–verb agreement: The role of intermediate traces of movement. *Journal of Memory and Language* **62** (2), 166–182 (2010), doi:10.1016/j.jml.2009.11.001.

32. M. Gillespie, N. J. Pearlmutter, Against structural constraints in subject–verb agreement production. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **39** (2), 515 (2013), doi:10.1037/a0029005.

33. A. Veenstra, D. J. Acheson, A. S. Meyer, Keeping it simple: Studying grammatical encoding with lexically reduced item sets. *Frontiers in Psychology* **5**, 783 (2014), doi:10.3389/fpsyg.2014.00783.

34. M. Kandel, C. Phillips, Number attraction in verb and anaphor production. *Journal of Memory and Language* **127**, 104370 (2022), doi:10.1016/j.jml.2022.104370.

35. G. Vigliocco, J. Nicol, Separating hierarchical relations and word order in language production: Is proximity concord syntactic or linear? *Cognition* **68** (1), B13–B29 (1998).

36. D. Mertzen, A. Laurinavichyute, B. W. Dillon, R. Engbert, S. Vasishth, Crosslinguistic evidence against interference from extra-sentential distractors. *Journal of Memory and Language* **137**, 104514 (2024).

37. Z. Schlueter, A. Williams, E. Lau, Exploring the abstractness of number retrieval cues in the computation of subject-verb agreement in comprehension. *Journal of Memory and Language* **99**, 74–89 (2018).

38. M. Keshev, A. Meltzer-Asscher, The representation of agreement features in memory is updated during sentence processing: Evidence from verb-reflexive interactions. *Journal of Memory and Language* **135**, 104495 (2024), doi:10.1016/j.jml.2023.104495.

39. R. L. Lewis, S. Vasishth, An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science* **29** (3), 375–419 (2005).

40. S. T. Piantadosi, H. Tily, E. Gibson, Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences* **108** (9), 3526–3529 (2011), doi:10.1073/pnas.1012551108.

41. M. Fedzechkina, T. F. Jaeger, E. L. Newport, Language learners restructure their input to facilitate efficient communication. *Proceedings of the National Academy of Sciences* **109** (44), 17897–17902 (2012), doi:10.1073/pnas.1215776109.

42. A. Baugh, T. Cable, *A history of the English language* (Routledge) (1993).

43. K. Sinnemäki, Complexity trade-offs in core argument marking, in *Language complexity: Typology, contact, change* (John Benjamins Publishing Company), pp. 67–88 (2008), doi:10.1075/slcs.94.06sin.

44. M. Jaeggli, K. Safir, *The null subject parameter*, vol. 15 (Springer Science & Business Media) (2012).

45. M. Fedzechkina, E. L. Newport, T. F. Jaeger, Balancing effort and information transmission during language acquisition: Evidence from word order and case marking. *Cognitive Science* **41** (2), 416–446 (2017), doi:10.1111/cogs.12346.

46. E. Gibson, *et al.*, How efficiency shapes human language. *Trends in Cognitive Sciences* **23** (5), 389–407 (2019), doi:10.1016/j.tics.2019.02.003.

47. A. F. Frank, T. F. Jaeger, Speaking rationally: Uniform information density as an optimal strategy for language production, in *Proceedings of the Annual Meeting of the Cognitive Science Society*, vol. 30 (2008).

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

49. M. H. Christiansen, N. Chater, Language as shaped by the brain. *Behavioral and Brain Sciences* **31** (5), 489–509 (2008), doi:10.1017/S0140525X08004998.

50. A. Wray, G. W. Grace, The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua* **117** (3), 543–578 (2007), doi:10.1016/j.lingua.2005.05.005.

51. G. Lupyan, R. Dale, Language structure is partly determined by social structure. *PloS One* **5** (1), e8559 (2010), doi:10.1371/journal.pone.0008559.

52. D. Nettle, Social scale and structural complexity in human languages. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367** (1597), 1829–1836 (2012), doi:10.1098/rstb.2011.0216.

53. M. Dingemanse, D. E. Blasi, G. Lupyan, M. H. Christiansen, P. Monaghan, Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences* **19** (10), 603–615 (2015), doi:10.1016/j.tics.2015.07.013.

54. R. Rummer, J. Engelkamp, Phonological information contributes to short-term recall of auditorily presented sentences. *Journal of Memory and Language* **45** (3), 451–467 (2001), doi:10.1006/jmla.2000.2788.

55. R. Rummer, J. Engelkamp, Phonological information in immediate and delayed sentence recall. *The Quarterly Journal of Experimental Psychology Section A* **56** (1), 83–95 (2003), doi:10.1080/02724980244000002.

56. K. Polišenská, S. Chiat, A. Comer, K. McKenzie, Semantic effects in sentence recall: The contribution of immediate vs delayed recall in language assessment. *Journal of Communication Disorders* **52**, 65–77 (2014), doi:10.1016/j.jcomdis.2014.08.002.

57. K. Neumann, New Guinea: A cradle of agriculture. *Science* **301** (5630), 180–181 (2003), doi:10.1126/science.1086677.

58. P. Mühlhäusler, External History of Tok Pisin, in *Handbook of Tok Pisin (New Guinea Pidgin)*, S. A. Wurm, P. Mühlhäusler, Eds., no. C-70 in Pacific Linguistics (Dept. of Linguistics, Research School of Pacific Studies, Australian National University, Canberra), pp. 35–64 (1985).

59. H. Sarvasy, E. Ögate, Early writing in Nungon in Papua New Guinea. *Teaching Writing to Children in Indigenous Languages* pp. 185–201 (2019).

60. H. S. Sarvasy, Ethical budgets in (psycho-)linguistic fieldwork. *Linguistics* **63** (2), 487–510 (2025), doi:10.1515/ling-2023-0137.

61. Audacity Team, Audacity (2024), `https://www.audacityteam.org/`, free, open-source audio editing software.

62. P. Boersma, D. Weenink, Praat: Doing Phonetics by Computer (2022), `https://www.praat.org/`.

63. Midjourney, Inc., Midjourney (2023), `https://www.midjourney.com/`, AI-based image generation system.

64. J. Peirce, *et al.*, PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods* **51** (1), 195–203 (2019).

65. A. J. Milne, E. A. Smit, H. S. Sarvasy, R. T. Dean, Evidence for a universal association of auditory roughness with musical stability. *PLoS One* **18** (9), e0291642 (2023), doi:10.1371/journal.pone.0291642.

66. K. E. Mulak, H. S. Sarvasy, A. Tuninetti, P. Escudero, Word learning in the field: Adapting a laboratory-based task for testing in remote Papua New Guinea. *PLoS One* **16** (9), e0257393 (2021), doi:10.1371/journal.pone.0257393.

67. E. A. Smit, A. J. Milne, H. S. Sarvasy, R. T. Dean, Emotional responses in Papua New Guinea show negligible evidence for a universal effect of major versus minor music. *PLoS One* **17** (6), e0269597 (2022), doi:10.1371/journal.pone.0269597.

68. B. MacWhinney, *The CHILDES project: The database*, vol. 2 (Psychology Press) (2000).

69. D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* **67** (1), 1–48 (2015), doi:10.18637/jss.v067.i01.

70. R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria (2021), `https://www.R-project.org/`.

71. D. J. Barr, R. Levy, C. Scheepers, H. J. Tily, Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* **68** (3), 255–278 (2013).

Acknowledgments

We would like to acknowledge the Towet village community for expertly managing and executing this experiment, and for their gracious hosting. We are grateful to those who participated in the experiments for helping further the scientific study of the Nungon language. We also thank: the cooks who provided food for the experimenters and participants, the security detail who provided security and ensured that there was no noise or disturbance around the experiments, the teachers, police, politicians, community leaders, and community members who attended opening and closing ceremonies, and all community members who contributed their farm produce and chickens to the communal food. In particular, thanks to Mr. Dono Ögate and Ms. Yaunggao Gietine for their staunch support of research into the Nungon language. Lemuel Dom of the PNG University of Technology gave us crucial assistance in navigating Morobe political circles to support our research visa application, and Ms. Georgia Kaipu and Mr. Nou Korema of the PNG National Research Institute also assisted greatly. SIL Aviation accommodated our flight requests and flew us safely into and out of the Uruwa area; SIL-PNG also hosted us graciously in Ukarumpa. Finally, thanks to Nat Livingston Johnson for photos of Towet and the experimental setup.

Funding: A. M. was funded by the National Institutes of Health Grant F32DC019533 (to A. M.). H. S., L. Ö., I. D., H. S. were funded by the Australian Research Council Grant DE180101609 (to H. S).

Author contributions: A. M. contributed to conceptualization, formal analysis, methodology, software, visualization, writing – original draft, and writing – review & editing. I. D. contributed to resources. L. Ö. contributed to investigation. S. G. contributed to investigation. H. S. contributed to conceptualization, data curation, funding acquisition, investigation, methodology, project administration, resources, supervision, writing – original draft, and writing – review & editing.

Competing interests: There are no competing interests to declare.

Data and materials availability: All materials needed to reproduce the experiments and analyses are available on OSF at

`https://osf.io/ncved/overview?view_only=1b657f70d49140b991c3c427068e3054`

(DOI: TO BE ADDED UPON DE-ANONYMIZATION). The repository includes all auditory stimulus files (.wav), a stimulus spreadsheet with English pseudo-glosses and approximate Nungon translations (which may differ slightly from the recorded stimuli due to translation difficulty), all visual stimulus images, the PsychoPy experiment itself, the coding spreadsheet (the version used in the analysis script as well as a "clean" version for visual inspection), the R script used to generate stimulus lists, and the R script used to analyze the data and generate all plots and summary statistics reported in the manuscript. Raw participant audio recordings are not publicly shared; all shared data are de-identified.

Supplementary materials

Supplementary Materials for Attraction effects in Nungon reveal independent mechanisms for agreement and switch-reference marking

Adam Morgan*, Lyn Ögate, Ismael Dono, Stanly Girip, Hannah Sarvasy
*Corresponding author. Email: adam.morgan@nyulangone.org

This PDF file includes:

Other Supplementary Materials for this manuscript:

All data and materials are publicly available on OSF and can be accessed by DOI at `[anonymized_for_peer_review]` or anonymously at:

`https://osf.io/ncved/overview?view_only=1b657f70d49140b991c3c427068e3054`

Materials and Methods

Participants

A total of 43 native Nungon speakers were recruited via word of mouth from Towet Village in the Uruwa River Valley, Morobe Province, Papua New Guinea. Participants were at least 18 years old (exact age is not possible to obtain for most Nungon speakers) and were screened for normal vision and hearing. Locals who were not native speakers of the Towet dialect of Nungon, such as several women who married Nungon-speaking men and learned Nungon as adults, were excluded from participation. Participants were paid PGK 50 for participation in the two experiments.

Field site and sociolinguistic context

The experiment was run in Towet in a dedicated research building that the community constructed in 2018. The building contains rooms for running experiments and is equipped with solar panels and batteries which powered the equipment used in the present study. Towet is a village of about 30 households located at ~1500m elevation in cloud forest. Towet villagers have land holdings up to over 4,000m. Papua New Guinea is one of the world's cradles of agriculture (*57*), and Nungon speakers grow, forage, and hunt all they need to survive. All houses, including the building in which the present experiment was run, are constructed by hand by local people, using locally sourced and hand-milled lumber.

There is no electricity supply to this remote area, and no access by road. Materials such as roofing iron must be either flown in by small plane or helicopter to a small local airstrip, or shipped to the Bismarck Sea coast, a day's hike away for locals, and then carried up by hand. The area gained cell phone coverage with the construction of a cell tower in the Uruwa River Valley in 2015, but as the tower is powered only by solar-powered batteries, it is unreliable. Local people largely sustain themselves as smallholder farmers; there are few opportunities for salaried work in the region, apart from a few jobs with a conservation NGO, teaching positions in the local schools, or health workers in the aid post further down the Uruwa River valley (about a three-hour hike from Towet).

Within the greater Nungon-speaking region, each village community has a distinct dialect. These largely share grammar, but differ in phonetics, phonology and lexicon (*21*). All our participants were part of the Towet village community. Some, however, had one parent who was from a different Nungon village community. Most adults under 50 are literate both in the Nungon language, which uses a Latin-based orthography, and in the regional lingua franca, Tok Pisin (which originated in a form of Pacific Pidgin English (*58*)). The area has had a government elementary school (grades K-3) since 1998, with literacy training in Nungon and then Tok Pisin instruction, and a primary school (grades 4-8), with instruction in Tok Pisin and English, since 2007 (*59*). Participants had varying degrees of familiarity with other Nungon dialects, with Tok Pisin, and with other regional languages.

A key feature of this research was that Nungon speakers were closely involved in all aspects of implementation, and all aspects of the experiments were conducted using the Nungon language as meta-language, following the general practice in our long-standing work with the Nungon-speaking community [described in (*60*)]. This means that all communications among the field team and between the team and community participants was strictly in Nungon, making information about the research accessible and comprehensible by all community members. Records of participants, file backup names, and other data sheets were all in Nungon.

Materials

Stimuli. Stimuli were clause chains consisting of two clauses, exemplified by the pseudo-gloss in Figs. 5 and 7, and images of characters/animals/objects in the stimuli. Glosses for a sample item set appear in Ex. 2b (see Supplementary Text for gloss key). All 35 item sets are provided in the supplementary materials. Sub-Experiment 1 involved a comparison of the conditions in Ex. 2a and 2e. Sub-Experiment 2 involved conditions 2a–2d.

(2) Sample stimulus item set:

a. ***No attractors (Baseline):***

oesit niip-no=dek yonggut-do-k=ma=ho mö-una, maam-no
girl [cousin-3SG.POSS=LOC laugh-PAST-3SG=REL=FOC]$_{RC}$ fall-DIFF.3SG aunt-3SG.POSS

yonggut-do-k.
laugh-PAST-3SG

'The girl who laughed at her cousin fell, her aunt laughed.'

b. ***Reference attractor:***

oesit niip-no=dek yonggut-do-k=ma=ho mö-una, niip-no
girl [cousin-3SG.POSS=LOC laugh-PAST-3SG=REL=FOC]$_{RC}$ fall-DIFF.3SG cousin-3SG.POSS

yonggut-do-k.
laugh-PAST-3SG

'The girl who laughed at her **cousin** fell, her **cousin** laughed.'

c. ***Agreement attractor:***

oesit niip-n-in=dek yonggut-do-k=ma=ho möuna,
girl [cousin-3SG.POSS-DU=LOC laugh-PAST-3SG=REL=FOC]$_{RC}$ fall-DIFF.3SG

maam-no yonggut-do-k.
aunt-3SG.POSS laugh-PAST-3SG

'The girl who laughed at her (two) **cousins** fell, her aunt laughed.'

d. ***Reference and agreement attractors:***

oesit niip-n-in=dek yonggut-do-k=ma=ho mö-una,
girl [cousin-3SG.POSS-DU=LOC laugh-PAST-3SG=REL=FOC]$_{RC}$ fall-DIFF.3SG

niip-n-in yonggut-do-morok.
cousin-3SG.POSS-DU laugh-PAST-3SG

'The girl who laughed at her (two) **cousins** fell, her (two) **cousins** laughed.'

e. ***Agreement attractor in Clause 2:***

oesit niip-no dek yonggut-do-k=ma=ho
girl [cousin-3SG.POSS=LOC laugh-PAST-3SG=REL=FOC]$_{RC}$ fall-DIFF.3SG

mö-una, maam-n-i yonggut-du-ng.
aunt-3SG.POSS-PL laugh-PAST-3PL

'The girl who laughed at her cousin fell, her (three or more) **aunts** laughed.'

Clause 1 consisted of a subject noun phrase (NP) and a VP ending in verb with a DIFF morpheme. The DIFF morpheme (correctly) agreed with the head noun of the subject NP (*oesit*, ‘girl’). To introduce an attractor NP, we included a relative clause (RC) modifying the Clause 1 subject (in Ex. 2a: *niip-no=dek yonggut-do-k=ma*, ‘who mocked her cousin’).

Clause 2 consisted of a subject NP and a VP. Because Clause 2 was the end of the clause chain, the verb was inflected for tense (always the Remote Past tense) and agreed with the Clause 2 subject in person and number.

We aimed to make the stimuli similar to clause chains that Nungon speakers might hear in daily life. They therefore involved events related to daily activities in Towet like harvesting, cooking, building, eating, etc. They also involved familiar objects (e.g., yams, arrows, retaining walls, etc.), animals (echidnas, tree kangaroos, dogs, etc.), and types of people (children, cannibals, family members, etc.).

Audio recordings of stimuli were created by splicing together recordings of the stimuli as read aloud, roughly half by Author H. S. (a non-native speaker) and half by Author I. D. (a native speaker). Splicing was performed in Audacity (*61*) and Praat (*62*) such that within an item set, all words were acoustically identical except for the critical (manipulated) morphemes and immediately adjacent phonemes that were not separated by a pause in the original recordings (so as to remove co-articulation effects). For each item, the clearest sounding segments across all recordings were used. Within an item set, all resulting stimulus audio files were spoken by only one of the two speakers.

We navigated a number of language-specific constraints in generating these stimuli. First, only a closed set of prototypically human nouns, mostly kin terms, are marked for number in Nungon, and this number marking, which distinguishes between three number values (singular, dual, and plural), only occurs when these nouns are possessed by a singular possessor, as seen in the distinction between *niip-no* ‘his/her cousin,’ *niip-n-in* ‘his/her two cousins,’ and *niip-n-i* ‘his/her three or more cousins.’ Nungon grammar also provides multiple other ways of conveying a noun’s number, including agreement on verbs, possessive markers, and anaphoric pronouns, and four different markers of associative plural categories, (see Supplementary Text for elaboration). We leveraged these tools to ensure that each noun’s number was made explicit at least twice in each stimulus. For example, in Ex. 2, the *-no* suffix on *niip* ‘cousin’ indicates that the possessor, *oe* ‘girl’ is singular, as does the *-una* suffix on the clause’s main verb *mö* ‘fall’. (The number of *niip* ‘cousin’ is also made clear twice, first by the *-no* (the possessive suffix) and second by the relative clause finite verb.)

Visual stimuli were created using Midjourney (*63*), an AI-based image generator (Fig. S1). Images were in the style of grayscale sketches, and were meant to be easily recognizable. The image generator was prompted to ensure all objects, animals, and people in the stimuli would be familiar-looking to Nungon speakers so as to mitigate any effects related to how unnatural computer-based tasks are for this community. For each stimulus, an image of the relative clause subject (the attractor noun) appeared on the screen for the duration of the trial so as to draw attention to this element. Nungon grammar distinguishes between singular, dual (2), and plural (3+). Where the number of the attractor was dual, two of the elements were shown on the screen; where plural, 3 or more were shown.

Equipment. The experiment was run in PsychoPy (*64*) on a Dell laptop running Windows 10. The computer’s internal microphone was used to record participants’ productions. Equipment was

Figure S1: **Sample visual stimuli.** Single items were shown for attractors with singular number (left), two for dual number (middle), and three or more for plural (right).

powered by several 12V batteries and power inverters, which are charged by three 100W solar panels installed on the research building.

Procedure

The experiments were run and managed by community members James Jio, author Stanly Girip, and author Lyn Ögate, all of whom have prior experience running psychology experiments with human participants in the region (*22, 65–67*), and have also managed a long-term child language acquisition study since 2015 [available on CHILDES (*68*) at `https://talkbank.org/childes/access/Other/Nungon/Sarvasy.html`].

Informed consent was obtained before participation. The experimenter then gave instructions orally, after which participants began the experiment. Participants listened to audio stimulus presentation using headphones and then repeated the narrative out loud. The experiment took approximately 30 minutes to complete.

The experimenter remained present with the participant throughout the experiment. Once the participant finished speaking, it was the experimenter who pressed the space bar to advance to the next trial. This was done because most participants lacked experience with computer keyboards, and we did not want to create additional distractions for them in making them control the experiment themselves using the keyboard.

Ethics Approval

Research was approved by the Western Sydney University Human Research Ethics Committee, approval H13536.

Data collection and coding

Data (audio recordings) were collected for each trial by the experiment software using the laptop's built-in microphone. Each recording was then manually coded by Author H. S. as one of the target responses (either verbatim repetition or repetition with a target error) or "non-target." For non-target responses, semantic and structural deviations from the original stimulus were excluded. Lexical substitutions were permitted where they did not result in a change to event interpretation (e.g., near-synonym lexical substitutions). A small number of unintelligible trials (due to, e.g., mumbling

Table S1: **Results of a mixed-effects logistic regression analyzing agreement attraction errors.** The presence of an agreement attractor led to a significant increase in agreement attraction errors, both when it appeared before the verb (in a relative clause modifying the subject; $p < .001$) or as the subject of the next clause ($p = .003$).

	Estimate	Standard Error	z-value	p-value	
Intercept	-4.9431	0.6555	-7.541	< .001	***
Clause 1 Number Attractor	3.0278	0.6176	4.903	< .001	***
Reference Attractor	-0.1392	0.3440	-0.405	0.6858	
Clause 2 Number Attractor	2.1337	0.7167	2.977	0.0029	**

or background noise) were also excluded. Based on these criteria, we removed 871 of the original 1604 trials (54.3%); the remaining 733 trials were included in the analysis (the number of trials per cell ranged from 114 to 174).

During the period in which experimentation took place, there was heavy rain at times. When the sound of heavy rain on the corrugated iron roof of the research building was overpowering, experimentation was paused.

For a small number of participants, the experiment crashed early on. When this happened, the experiment was re-started from scratch using the same list (same stimulus order), but only the first instance of each stimulus item was included in analysis.

Analysis

Agreement attraction errors were analyzed in a mixed effects logistic regression (*69*) in R [Version 4.5.0; (*70*)]. We used treatment coding to compare each condition to the baseline condition (no attractors for agreement or switch-reference). The full model, which had three fixed effects (NUMBER ATTRACTOR, REFERENCE ATTRACTOR, and CLAUSE 2 NUMBER ATTRACTOR), random intercepts for participants and items, and random slopes within each of those for the three fixed effect terms, did not converge. To achieve convergence, we followed (*71*) and removed terms from the random effects structure one at a time (in order from least to most variance accounted for) until the model converged. The final model, reported in Table S1, converged with intercepts for participants and items, but no nested slopes.

Pairwise comparisons confirmed that agreement attractors significantly increased the rate of agreement errors both when reference attractors were present and absent. Without reference attractors, error rates were significantly higher in the number-attractor condition (errors on 23 out of 136 trials) than in the no-attractor condition (errors on 0 out of 174 trials; Fisher's exact test: $p < .001$, odds ratio = ∞, lower bound of 95% CI = 3.8). With reference attractors, error rates were still significantly higher in the number-attractor condition (errors on 18 out of 145 trials) than in the no-attractor condition (errors on 3 of 159 trials; Fisher's exact test: $p < .001$, odds ratio = 7.328, 95% CI = [2.072, 39.704]).

Switch-reference attraction was not statistically analyzed because there were no observations of a switch-reference attraction error anywhere in our results. Across all participants, only one error in switch-reference marking was observed, but it was in the control condition (i.e., not an *attraction* error). This trial was independently excluded due to other deviations from the original stimulus.

All analysis code and de-identified data are available on OSF at `https://osf.io/ncved/`

`overview?view_only=1b657f70d49140b991c3c427068e3054` (DOI: [to be added here after peer review when repository is de-anonymized]).

Supplementary Text

Gloss key

Below are all morpheme glosses that occur in examples in this manuscript:

- SG: singular
- DU: dual (i.e., two)
- PL: plural (i.e., three or more)
- DIFF: a switch-reference morpheme indicating that the upcoming subject is different from the current subject (also commonly appears in the literature glossed as SWITCH or DS for 'different subject')
- SAME: a switch-reference morpheme indicating that the upcoming subject is the same as the current subject (also commonly appears in the literature glossed as STAY or SS for 'same subject')
- OBJ: object (for when agreement morphemes refer to features of the object)
- PAST: past tense
- POSS: possessor
- REL: relativizer (at the right edge of a relative clause)
- FOC: focus
- RC: relative clause

Switch-reference tracks subjects

A nuanced but highly consequential point is that switch-reference marking tracks grammatical subjects, not, for example, notional subjects or focused/topicalized nouns. To understand the impressive implications of this, consider the pair of clause chains in Ex. 3, which Nungon speakers report have the same meaning:

(3) a. *ketket kombut to-nga. mab=o-go-k.*
boy anger do-SAME. shout-PAST-3SG.

'The boy was angry. He shouted.'

b. *ketket iik i-m-una. mab=o-go-k.*
boy anger 3SG.OBJ-give-DIFF.3SG. shout-PAST-3SG.

'The boy was angry (lit.: Anger afflicted him). He shouted.'

In Ex. 3a, “boy” is the subject of both clauses, so the non-final verb “to” (do) is marked with a “stay” morpheme. In Ex. 3b, however, “the boy was angry” is constructed differently (literally translating to something like “anger afflicted the boy”). Here, even though the notional subject is still “boy,” the grammatical subject is now “anger.” And because the subject of the next clause is still “boy,” Nungon grammar requires a switch marker on the non-final verb.

It is therefore not merely that Nungon speakers must plan a single word. Because the identity of the subject depends on the particular syntax with which they choose to encode the upcoming clause, they must plan both the meaning and syntax of the next clause to produce the current one.