No scope for planning – language pre-planning as mixture process

Jens Roeser

Mark Torrance Mark Andrews Thom Baguley

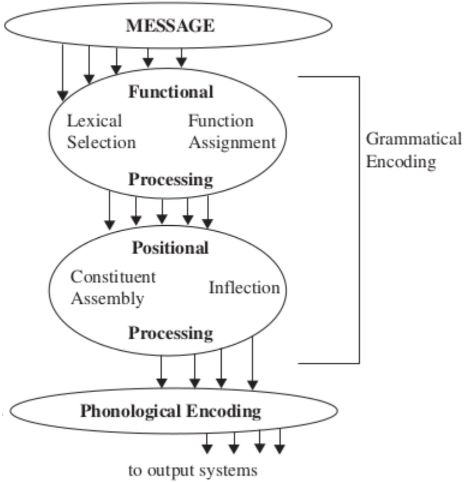
Nottingham Trent University, UK
jens.roeser@ntu.ac.uk

26th AMLaP, University of Potsdam

Sept 3, 2020

Turning ideas into language (Bock & Levelt, 1994)

- Message units are unordered.
- Output requires linearisation of words.
- Linearisation is subject to pragmatic, lexical and / or syntactic factors.

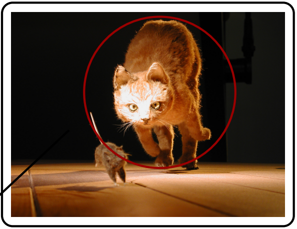


To what extent does syntax affect the linearisation of the output?

Syntax in language production (Bock & Ferreira, 2014)

1. Syntax is an emergent property of lexically-driven planning.
2. Syntactic relations guide lexical retrieval.
   a. **Deterministic:** syntax determines size of planning unit.
   b. **Non-deterministic:** multiple candidate structures (Kempen & Hoenkamp, 1987).
3. Either route (relational and non-relational) is available (at the message level; see Konopka & Meyer, 2014).

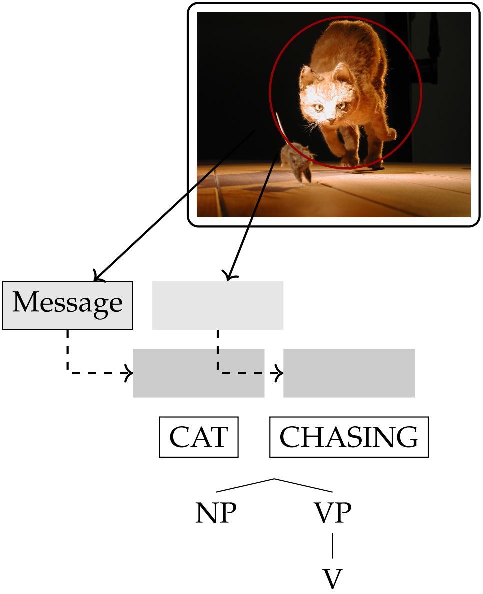
Consider the following evidence for possibility (2a).

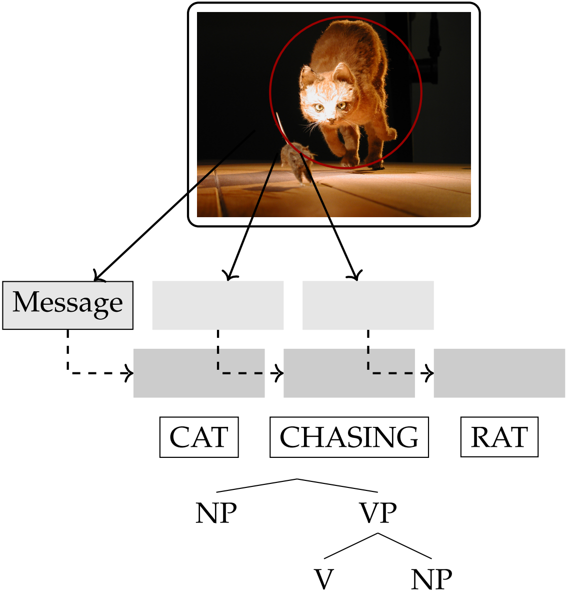


Message

CAT

NP

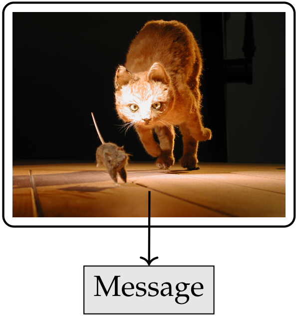

Message
CAT
CHASING
NP
VP
V

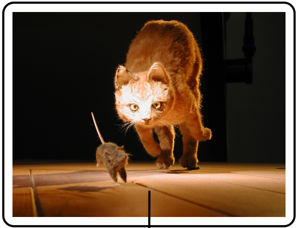

Message
CAT
CHASING
RAT
NP
VP
V
NP

Syntax in language production (Bock & Ferreira, 2014)

1. Syntax is an emergent property of lexically-driven planning.
2. Syntactic relations guide lexical retrieval.
   a. **Deterministic:** syntax determines size of planning unit.
   b. **Non-deterministic:** multiple candidate structures (Kempen & Hoenkamp, 1987).
3. Either route (relational and non-relational) is available (at the message level; see Konopka & Meyer, 2014).

Consider the following evidence for possibility (2a).

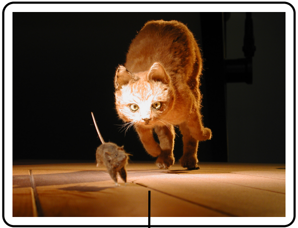

Message



Message

NP VP

CAT V NP

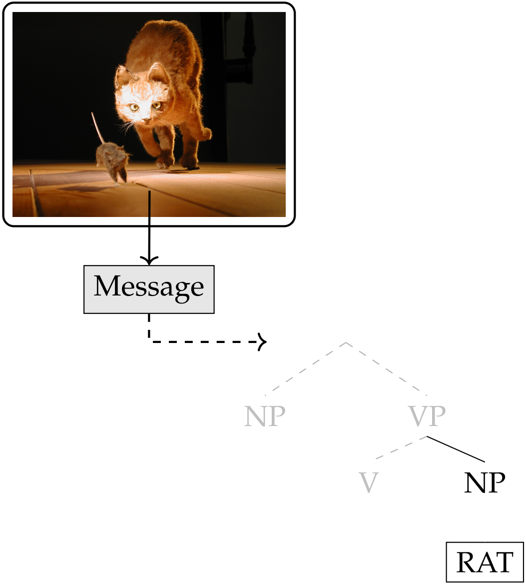


Message

NP VP

V NP

CHASE


Message
NP
VP
V
NP
RAT

Syntax in language production (Bock & Ferreira, 2014)

1. Syntax is an emergent property of lexically-driven planning.
2. Syntactic relations guide lexical retrieval.
   a. **Deterministic:** syntax determines size of planning unit.
   b. **Non-deterministic:** multiple candidate structures (Kempen & Hoenkamp, 1987).
3. Either route (relational and non-relational) is available (at the message level; see Konopka & Meyer, 2014).

Consider the following evidence for possibility (2a).

Syntax in language production (Bock & Ferreira, 2014)

1. Syntax is an emergent property of lexically-driven planning.
2. Syntactic relations guide lexical retrieval.
   a. **Deterministic:** syntax determines size of planning unit.
   b. **Non-deterministic:** multiple candidate structures (Kempen & Hoenkamp, 1987).
3. Either route (relational and non-relational) is available (at the message level; see Konopka & Meyer, 2014).

Consider the following evidence for possibility (2a).

Syntax in language production (Bock & Ferreira, 2014)

1. Syntax is an emergent property of lexically-driven planning.
2. Syntactic relations guide lexical retrieval.
   a. **Deterministic:** syntax determines size of planning unit.
   b. **Non-deterministic:** multiple candidate structures (Kempen & Hoenkamp, 1987).
3. Either route (relational and non-relational) is available (at the message level; see Konopka & Meyer, 2014).

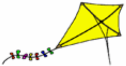
Consider the following evidence for possibility (2a).

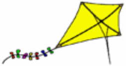
Syntax in language production (Bock & Ferreira, 2014)

1. Syntax is an emergent property of lexically-driven planning.
2. Syntactic relations guide lexical retrieval.
   a. **Deterministic:** syntax determines size of planning unit.
   b. **Non-deterministic:** multiple candidate structures (Kempen & Hoenkamp, 1987).
3. Either route (relational and non-relational) is available (at the message level; see Konopka & Meyer, 2014).

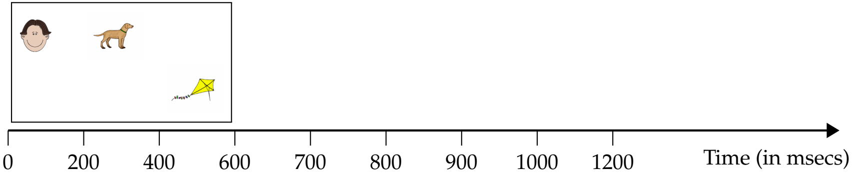
Consider the following evidence for possibility (2a).

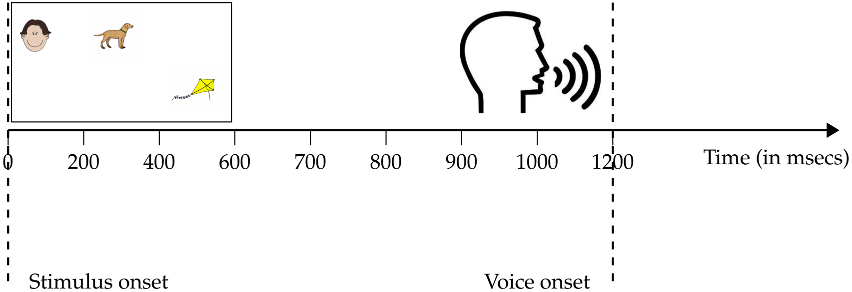
+

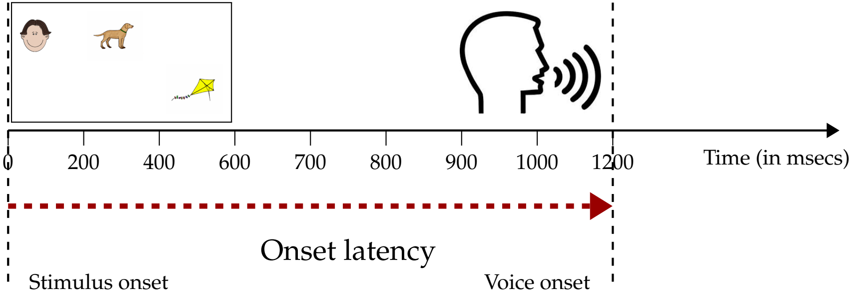


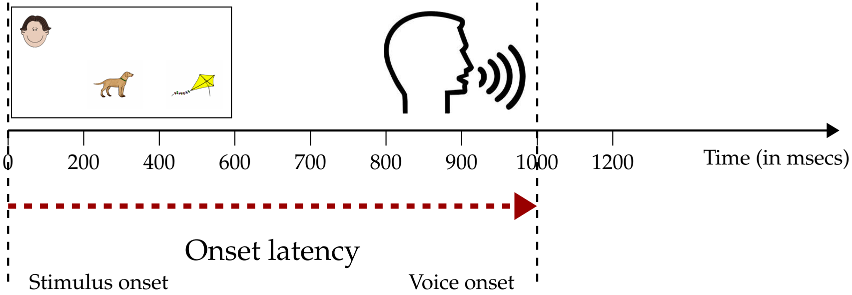


The boy and the dog moved above the kite.


0
200
400
600
700
800
900
1000
1200
Time (in msecs)


0
200
400
600
700
800
900
1000
1200
Time (in msecs)
Stimulus onset
Voice onset


0
200
400
600
700
800
900
1000
1200
Time (in msecs)
Onset latency
Stimulus onset
Voice onset


0
200
400
600
700
800
900
1000
1200
Time (in msecs)
Onset latency
Stimulus onset
Voice onset

Preplanning scope involves syntax

1. **The boy and the dog** moved above the kite.
2. **The boy** moved above the dog and the kite.

- Frequently reproduced effect (e.g. Martin et al., 2014; Smith & Wheeldon, 1999; Wagner et al., 2010).
- "Phrase as default planning scope" (Martin et al., 2010)
- NP syntax is planned before production onset.
- Lexical scope is smaller (Griffin, 2001) and flexible (Wheeldon et al., 2013).

Implication of the standard statistical treatment

- Statistical models used (LMM, ANOVA) map onto a deterministic syntax-driven model.
- Systematic difference between simple and conjoined NPs.
- Under these statistical models, the following alternative hypothesis couldn't be tested.

Implication of the standard statistical treatment

- Statistical models used (LMM, ANOVA) map onto a deterministic syntax-driven model.
- Systematic difference between simple and conjoined NPs.
- Under these statistical models, the following alternative hypothesis couldn't be tested.

Implication of the standard statistical treatment

- Statistical models used (LMM, ANOVA) map onto a deterministic syntax-driven model.
- Systematic difference between simple and conjoined NPs.
- Under these statistical models, the following alternative hypothesis couldn't be tested.

Alternative hypothesis

- **Preplanning beyond the first noun is more likely but not obligated by the phrase syntax** because, for example, …

1. Fluency pressure requires preplanning of B in *The A and the B moved …* if there is not enough time to plan B in parallel to articulation (Allum & Wheeldon, 2007; Griffin, 2003).
2. Activation of phrase syntax or use of the syntactic route is non-deterministic.

Research focus

- Direct comparison of two hypothesis.

i. Phrase scope obligated by the production system, leading to a systematic slowdown for conjoined NPs.
ii. Preplanning beyond the first noun is more likely in conjoined NPs but not obligated by the production system.

Pooled re-analysis

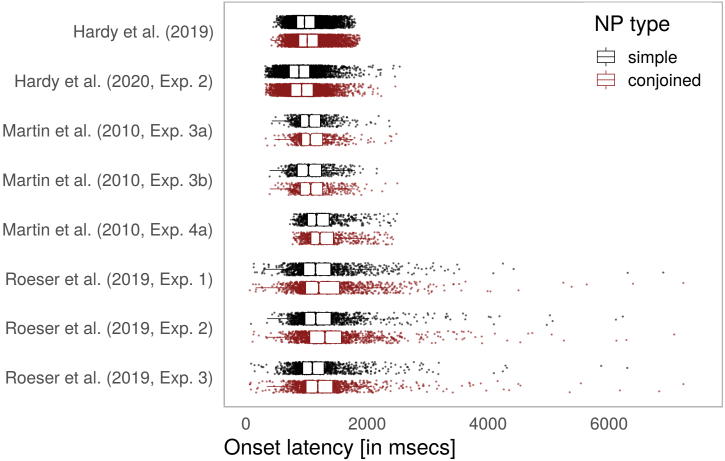
- Stimulus-to-onset latencies

a. **Conjoined NPs:** *The boy and the dog moved above the kite.*

b. **Simple NPs:** *The boy moved above the dog and the kite.*

- Hardy et al. (2019): 90 ppts; 36 items
- Hardy et al. (2020): 105 ppts; 80 items
- Martin et al. (2010): 3×12 ppts; 2×48 and 1×64 items
- Roeser et al. (2019): 3×32 ppts; 96 items

Data overview



Model overview

1. Null LMM
2. **LMM (NP effect)**
3. LMM (unequal variance)
4. Null mixture model
5. **Mixture model**

- Stan code based on Sorensen et al. (2016) and Vasishth, Chopin et al. (2017); also Vasishth, Jäger et al. (2017).

- *LogNormal* distribution with mean μ and error variance σ_e^2
- Random intercepts
  - participants: $u_i \sim Normal(0, \sigma_u^2)$
  - items: $w_j \sim Normal(0, \sigma_w^2)$
- Weakly informative priors (McElreath, 2016)

Model overview

1. Null LMM
2. **LMM (NP effect)**
3. LMM (unequal variance)
4. Null mixture model
5. **Mixture model**

- Stan code based on Sorensen et al. (2016) and Vasishth, Chopin et al. (2017); also Vasishth, Jäger et al. (2017).

- *LogNormal* distribution with mean μ and error variance σ_e^2
- Random intercepts
  - participants: $u_i \sim Normal(0, \sigma_u^2)$
  - items: $w_j \sim Normal(0, \sigma_w^2)$
- Weakly informative priors (McElreath, 2016)

Null model (null hypothesis)

$$y_{ij} \sim LogNormal(\mu_{ij}, \sigma_e^2)$$
$$\mu_{ij} = \alpha + u_i + w_j$$

Meta null model (null hypothesis)

$$
\begin{aligned}
y_{ijk} \sim LogNormal(\mu_{ijk}, \sigma^2_{e_k}) \\
\mu_{ijk} = \alpha_k + u_i + w_j \\
\alpha_k = \alpha_\mu + \alpha_\tau \cdot \alpha_{\eta_k}
\end{aligned}
$$

- For $k = 1, \ldots, K$ where K is the number of studies.
- α_k is the latency coefficient for the kth study.
- α_μ is the pooled latency coefficient.
- Non-centred parametrisation for α_k (Gelman et al., 2014).

Meta LMM (standard analysis)

$$
\begin{aligned}
y_{ijk} &\sim LogNormal(\mu_{ijk}, \sigma^2_{e_k}) \\
\mu_{ijk} &= \alpha_k + \beta_k \cdot x_{[0,1]} + u_i + w_j \\
\alpha_k &= \alpha_\mu + \alpha_\tau \cdot \alpha_{\eta_k} \\
\beta_k &= \beta_\mu + \beta_\tau \cdot \beta_{\eta_k}
\end{aligned}
$$

- $x = 0$ for simple NPs; $x = 1$ for conjoined NPs.
- β_k is the latency change for conjoined NPs for the kth study.
- β_μ is the pooled latency change for conjoined NPs.

Mixture model (alternative hypothesis)

$$
\begin{aligned}
y_{ij} \sim \theta_{NP} \cdot LogNormal(\mu_{ij} + \delta, \sigma^2_{e'}) + \\
(1 - \theta_{NP}) \cdot LogNormal(\mu_{ij}, \sigma^2_{e}) \\
\mu_{ij} = \alpha + u_i + w_j \\
\text{constraint: } \delta > 0 \\
\sigma^2_{e'} > \sigma^2_{e}
\end{aligned}
$$

- Probability of long latencies θ by NP type.
- μ and σ^2 constant across NP type.

Mixture model (alternative hypothesis)

$$
\begin{aligned}
y_{ij} \sim \theta_{NP} \cdot LogNormal(\mu_{ij} + \delta, \sigma^2_{e'}) + \\
(1 - \theta_{NP}) \cdot LogNormal(\mu_{ij}, \sigma^2_{e}) \\
\mu_{ij} = \alpha + u_i + w_j \\
\text{constraint: } \delta > 0 \\
\sigma^2_{e'} > \sigma^2_{e}
\end{aligned}
$$

- Probability of long latencies θ by NP type.
- μ and σ^2 constant across NP type.

Meta mixture model (alternative hypothesis)

$$
\begin{aligned}
y_{ijk} \sim \theta_{NP_k} \cdot LogNormal(\mu_{ijk} + \delta_k, \sigma^2_{e'_k}) + \\
(1 - \theta_{NP_k}) \cdot LogNormal(\mu_{ijk}, \sigma^2_{e_k}) \\
\mu_{ijk} = \alpha_k + u_i + w_j \\
\theta_{NP_k} = Logit^{-1}(\phi_{NP_k}) \\
\phi_{NP_k} \sim Normal(\phi_{\mu_{NP}}, \phi^2_{\tau}) \\
\delta_k \sim Normal(\delta_{\mu}, \delta^2_{\tau}) \\
\text{constraint: } \delta_k > 0
\end{aligned}
$$

- α_k, $\sigma^2_{e'_k}$, $\sigma^2_{e_k}$ defined as before.
- Pooled coefficient for parameters: θ_{NP}, δ

Meta LMM (unequal variance)

$$y_{ijk} \sim \begin{cases} LogNormal(\mu_{ijk}, \sigma^2_{e_k}), & \text{if NP}_{ijk} = \text{simple} \\ LogNormal(\mu_{ijk} + \beta_k, \sigma^2_{e'_k}) & \text{else if NP}_{ijk} = \text{conjoined} \end{cases}$$

$$\mu_{ijk} = \alpha_k + u_i + w_j$$

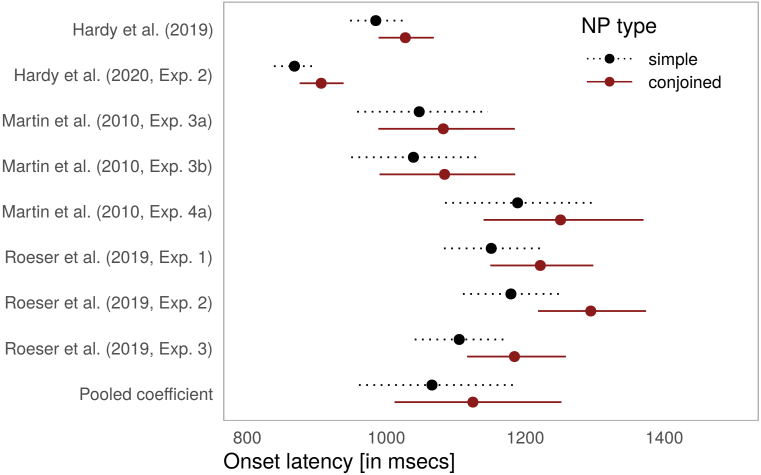
$$\alpha_k = \alpha_\mu + \alpha_\tau \cdot \alpha_{\eta_k}$$

$$\beta_k = \beta_\mu + \beta_\tau \cdot \beta_{\eta_k}$$

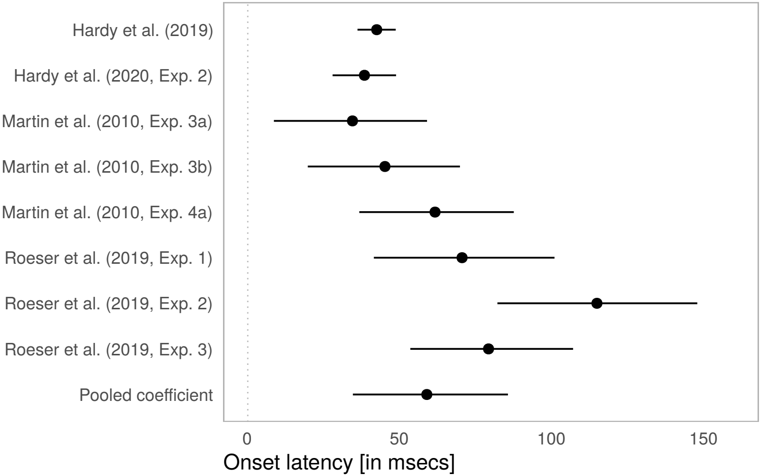
$$\text{constraint: } \sigma^2_{e_k} > 0$$

$$\sigma^2_{e'_k} > \sigma^2_{e_k}$$

NP effect (LMM)



NP effect (LMM)



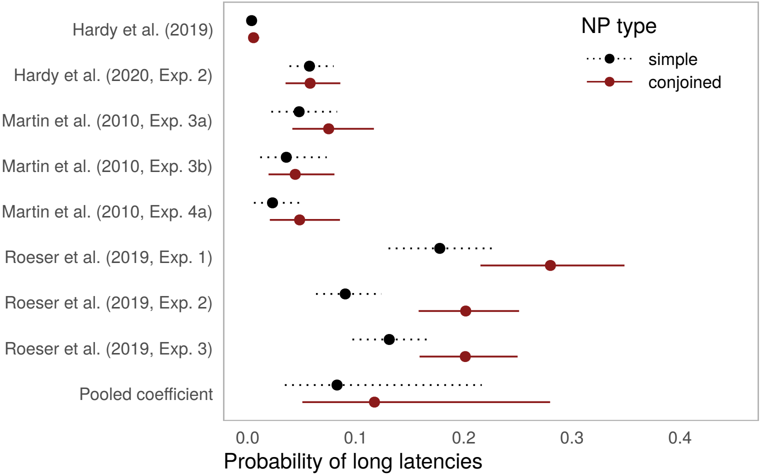
Model comparisons

Predictive performance estimated as the *expected log pointwise predictive density* ($\widehat{elpd}$) (Vehtari et al., 2015, 2017). Models are ordered by predictive performance (model with highest predictive performance in top row). Standard error in parentheses.

Models	$\Delta\widehat{elpd}$	$\widehat{elpd}$	Description
MoG-1	–	-201,486 (176)	Mixing proportions by NP
MoG-0	-15 (8)	-201,500 (176)	Null model
LMM-1	-1,006 (97)	-202,492 (214)	NP effect
LMM-0	-1,192 (92)	-202,678 (212)	Null model
LMM-2	-3,537 (214)	-205,022 (307)	Unequal variance

Note. LMM = Linear mixed effects model; MoG = Mixture of Gaussians

Probability of long latencies (MoG)



Summary

- The frequently observed slowdown for conjoined NPs is better explained by a larger, yet relatively small, probability of long latencies.
- Different pattern for Hardy et al. (2019, 2020) compared to Martin et al. (2010) and Roeser et al. (2019); possibly because of data trimming threshold.
- Most studies in our pool included other manipulations.

Follow-up experiments

Experiment 1:

- Reproduce analysis after …

  i. Reducing the manipulation to simple and conjoined NPs.

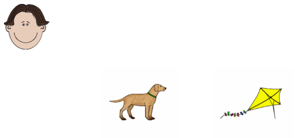
  ii. Controlling image names

Experiment 2:

- Assess the impact of the visual manipulation independently of utterance sytnax.
- Elicit name lists instead of sentences (as in Martin et al., 2010)

Methods

Condition 1: A and B moved above C



Condition 2: A moved above B and C

- **Experiment 1:**

  1a. **Peter and the dog** moved above the kite

  1b. **Peter** moved above the dog and the kite

  - 78 ppts (after cleaning)

- **Experiment 2:**

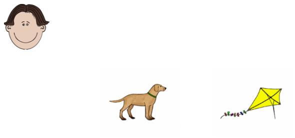
  2. Peter, the dog, the kite

  - 45 ppts (after cleaning)

- First noun: *Peter* or *Tania*
- Movement: up or down
- 48 items; 96 fillers; 6 practice trials
- Image names: high frequency and naming agreement.

Methods

Condition 1: A and B moved above C

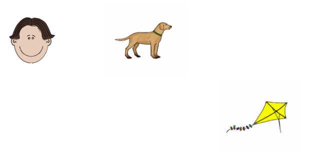


Condition 2: A moved above B and C

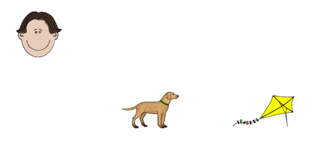
- **Experiment 1:**
  - 1a. **Peter and the dog** moved above the kite
  - 1b. **Peter** moved above the dog and the kite
  - 78 ppts (after cleaning)
- **Experiment 2:**
  - 2. Peter, the dog, the kite
  - 45 ppts (after cleaning)

- First noun: *Peter* or *Tania*
- Movement: up or down
- 48 items; 96 fillers; 6 practice trials
- Image names: high frequency and naming agreement.

Methods



Condition 1: A and B moved above C



Condition 2: A moved above B and C

- **Experiment 1:**
  - 1a. **Peter and the dog** moved above the kite
  - 1b. **Peter** moved above the dog and the kite
  - 78 ppts (after cleaning)
- **Experiment 2:**
  - 2. Peter, the dog, the kite
  - 45 ppts (after cleaning)

- First noun: *Peter* or *Tania*
- Movement: up or down
- 48 items; 96 fillers; 6 practice trials
- Image names: high frequency and naming agreement.

Methods

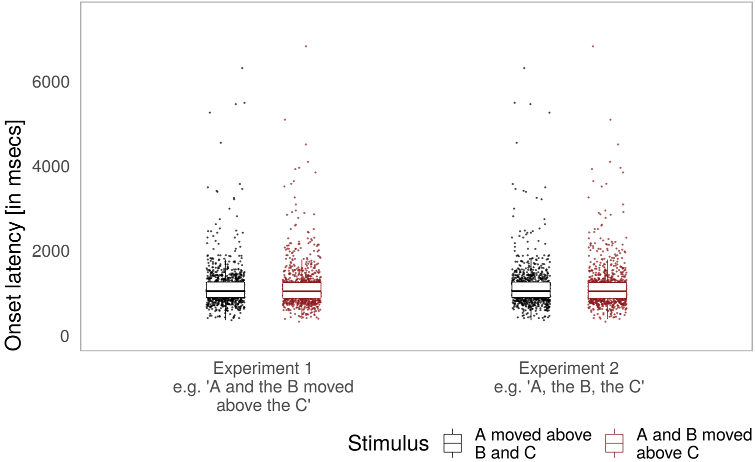
Condition 1: A and B moved above C

Condition 2: A moved above B and C

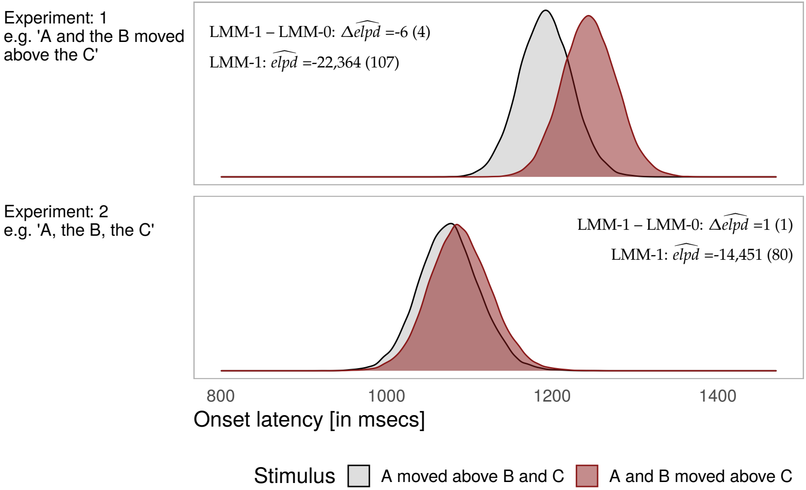
- **Experiment 1:**
  1a. **Peter and the dog** moved above the kite
  1b. **Peter** moved above the dog and the kite
  - 78 ppts (after cleaning)
- **Experiment 2:**
  2. Peter, the dog, the kite
  - 45 ppts (after cleaning)

- First noun: *Peter* or *Tania*
- Movement: up or down
- 48 items; 96 fillers; 6 practice trials
- Image names: high frequency and naming agreement.

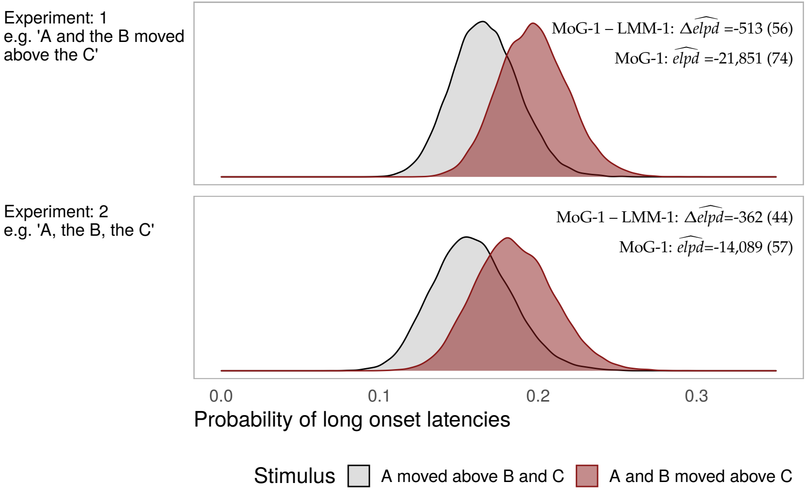
Onset latencies



NP-type effect (LMM)



Probability of long latencies (MoG)



Summary

- Evidence against the phrase-as-planning-unit hypothesis: preplanning NP syntax is not obligated by the language production system.
- Instead, the slowdown for conjoined NPs (as in Martin et al., 2010; Smith & Wheeldon, 1999) is better explained by a larger tendency to exhibit long onset latencies.

Alternative explanations

- Preplanning scope is in-line with non-deterministic theories of language production.
- Together with Exp. 2, results suggest that the slowdown observed in a subset of trials origins during visual rather than grammatical encoding.
- Both a relational and non-relational route are available during high level encoding (Konopka & Meyer, 2014; Kuchinsky et al., 2011).

Thank you for listening!

jens.roeser@ntu.ac.uk

nottinghamtrent.academia.edu/JensRoeser

R and *Stan* code: github.com/jensroes/NP-effect

…and **Randi Martin**, **Jason Crowther**, and **Sophie Hardy** for sharing their data.

…and **Dora Kramar** and **Andra Tanasescu** for help with the data collection.

References I

Allum, P. H. & Wheeldon, L. R. (2007). Planning scope in spoken sentence production: The role of grammatical units. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(4), 791–810.

Bock, J. K. & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). Academic Press.

Bock, J. K. & Ferreira, V. S. (2014). Syntactically speaking. In M. Goldrick, V. S. Ferreira & M. Miozzo (Eds.), *The Oxford Handbook of Language Production* (pp. 21–46). Oxford University Press.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A. & Rubin, D. B. (2014). *Bayesian data analysis* (3rd ed.). Chapman; Hall/CRC.

Griffin, Z. M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition*, *82*(1), B1–B14.

Griffin, Z. M. (2003). A reversed word length effect in coordinating the preparation and articulation of words in speaking. *Psychonomic Bulletin & Review*, *10*(3), 603–609.

Hardy, S. M., Segaert, K. & Wheeldon, L. (2019). Age-related disruption in the use of lexical information during sentence production, despite preserved syntactic planning. *PsyArXiv*.

Hardy, S. M., Segaert, K. & Wheeldon, L. (2020). Healthy aging and sentence production: Disrupted lexical access in the context of intact syntactic planning. *Frontiers in Psychology*, *11*, 257.

References II

Kempen, G. & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive science*, *11*(2), 201–258.

Konopka, A. E. & Meyer, A. S. (2014). Priming sentence planning. *Cognitive Psychology*, *73*, 1–40.

Kuchinsky, S. E., Bock, K. & Irwin, D. E. (2011). Reversing the hands of time: Changing the mapping from seeing to saying. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(3), 748–756.

Martin, R. C., Crowther, J. E., Knight, M., Tamborello II, F. P. & Yang, C.-L. (2010). Planning in sentence production: Evidence for the phrase as a default planning scope. *Cognition*, *116*(2), 177–192.

Martin, R. C., Yan, H. & Schnur, T. T. (2014). Working memory and planning during sentence production. *Acta Psychologica*, *152*, 120–132.

McElreath, R. (2016). *Statistical rethinking: A bayesian course with examples in R and Stan*. CRC Press.

Roeser, J., Torrance, M. & Baguley, T. (2019). Advance planning in written and spoken sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(11), 1983–2009.

Smith, M. & Wheeldon, L. R. (1999). High level processing scope in spoken sentence production. *Cognition*, *73*, 205–246.

Sorensen, T., Hohenstein, S. & Vasishth, S. (2016). Bayesian linear mixed models using stan: A tutorial for psychologists, linguists, and cognitive scientists. *Quantitative Methods for Psychology*, *12*(3), 175–200.

References III

Vasishth, S., Chopin, N., Ryder, R. & Nicenboim, B. (2017). Modelling dependency completion in sentence comprehension as a Bayesian hierarchical mixture process: A case study involving Chinese relative clauses. *ArXiv e-prints*.

Vasishth, S., Jäger, L. A. & Nicenboim, B. (2017). Feature overwriting as a finite mixture process: Evidence from comprehension data. *arXiv preprint arXiv:1703.04081*.

Vehtari, A., Gelman, A. & Gabry, J. (2015). Pareto smoothed importance sampling. *arXiv preprint arXiv:1507.02646*.

Vehtari, A., Gelman, A. & Gabry, J. (2017). Practical bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413–1432.

Wagner, V., Jescheniak, J. D. & Schriefers, H. (2010). On the flexibility of grammatical advance planning during sentence production: Effects of cognitive load on multiple lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(2), 423–440.

Wheeldon, L. R., Ohlson, N., Ashby, A. & Gator, S. (2013). Lexical availability and grammatical encoding scope during spoken sentence production. *The Quarterly Journal of Experimental Psychology*, *66*(8), 1653–1673.

Meta LMM

$$
\begin{aligned}
y_{ijk} &\sim LogNormal(\mu_{ijk}, \sigma^2_{e_k}) \\
\mu_{ijk} &= \alpha_k + \beta_k \cdot x_{[0,1]} + u_i + w_j \\
\alpha_k &= \alpha_\mu + \alpha_\tau \cdot \alpha_{\eta_k} \\
\beta_k &= \beta_\mu + \beta_\tau \cdot \beta_{\eta_k} \\
\text{constraint: } \sigma^2_{e_k} &> 0
\end{aligned}
$$

- $x = 0$ for simple NPs; $x = 1$ for conjoined NPs.
- β_k is the latency change for conjoined NPs for the *k*th study.
- β_μ is the pooled latency change for conjoined NPs.

Meta LMM (unequal variance)

$$y_{ijk} \sim \begin{cases} LogNormal(\mu_{ijk}, \sigma^2_{e_k}), & \text{if NP}_{ijk} = \text{simple} \\ LogNormal(\mu_{ijk} + \beta_k, \sigma^2_{e'_k}) & \text{else if NP}_{ijk} = \text{conjoined} \end{cases}$$

$$\mu_{ijk} = \alpha_k + u_i + w_j$$

$$\alpha_k = \alpha_\mu + \alpha_\tau \cdot \alpha_{\eta_k}$$

$$\beta_k = \beta_\mu + \beta_\tau \cdot \beta_{\eta_k}$$

$$\text{constraint: } \sigma^2_{e_k} > 0$$

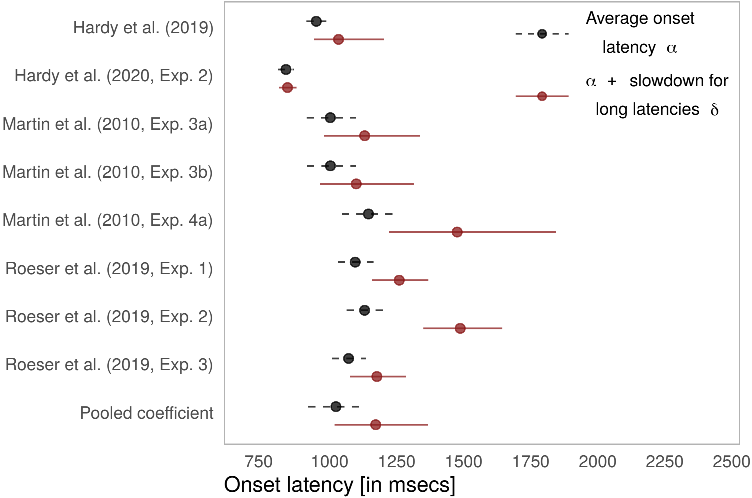
$$\sigma^2_{e'_k} > \sigma^2_{e_k}$$

Meta mixture model (alternative hypothesis)

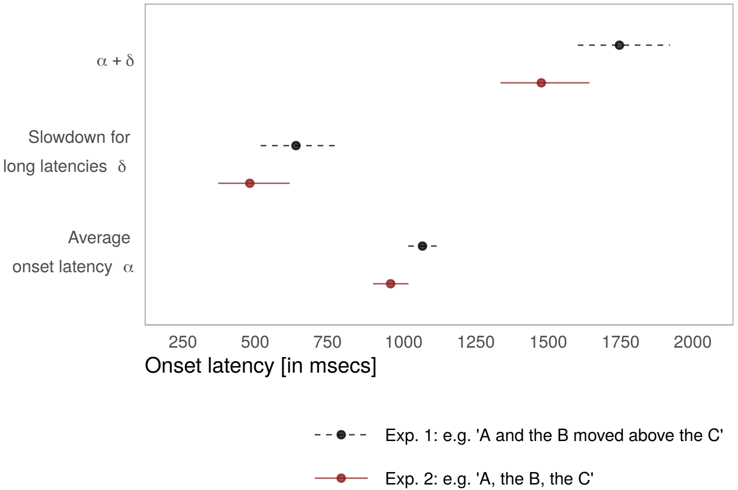
$$
\begin{aligned}
y_{ijk} \sim \theta_{NP_k} \cdot LogNormal(\mu_{ijk} + \delta_k, \sigma^2_{e'_k}) + \\
(1 - \theta_{NP_k}) \cdot LogNormal(\mu_{ijk}, \sigma^2_{e_k}) \\
\mu_{ijk} = \alpha_k + u_i + w_j \\
\alpha_k = \alpha_\mu + \alpha_\tau \cdot \alpha_{\eta_k} \\
\theta_{NP_k} = Logit^{-1}(\phi_{NP_k}) \\
\theta_{NP} = Logit^{-1}(\phi_{\mu_{NP}}) \\
\phi_{NP_k} \sim Normal(\phi_{\mu_{NP}}, \phi^2_\tau) \\
\delta_k \sim Normal(\delta_\mu, \delta^2_\tau) \text{constraint: } \delta_k, \sigma^2_{e_k}, \phi_\tau, \delta_\tau > 0 \\
\sigma^2_{e'_k} > \sigma^2_{e_k}
\end{aligned}
$$

- Pooled coefficient for parameters: θ_{NP}, δ

Onset-latency coefficients (meta mixture components)



Onset-latency coefficients (mixture components)



Model comparisons

Predictive performance estimated as the *expected log pointwise predictive density* ($\widehat{elpd}$). Models are ordered by predictive performance (model with highest predictive performance in top row). Standard error in parentheses.

	Model	$\Delta\widehat{elpd}$	$\widehat{elpd}$
Experiment 1	MoG-0	–	-21,851 (74)
	MoG-1	0 (2)	-21,851 (74)
	LMM-1	-513 (56)	-22,364 (107)
	LMM-2	-515 (57)	-22,366 (107)
	LMM-0	-519 (56)	-22,369 (107)
Experiment 2	MoG-0	–	-14,088 (57)
	MoG-1	-1 (1)	-14,089 (57)
	LMM-0	-362 (44)	-14,450 (80)
	LMM-1	-363 (44)	-14,451 (80)
	LMM-2	-364 (44)	-14,452 (81)

Note. LMM = Linear mixed effects model; MoG = Mixture of Gaussians