

# Grammar as logic, processing as deduction, actions as theorems, states as propositions

Shinnosuke Isono<sup>1</sup> (s-isono@nijal.ac.jp) Kohei Kajikawa<sup>2</sup>

<sup>1</sup>National Institute for Japanese Language and Linguistics (NINJAL) <sup>2</sup>Georgetown University

## In short

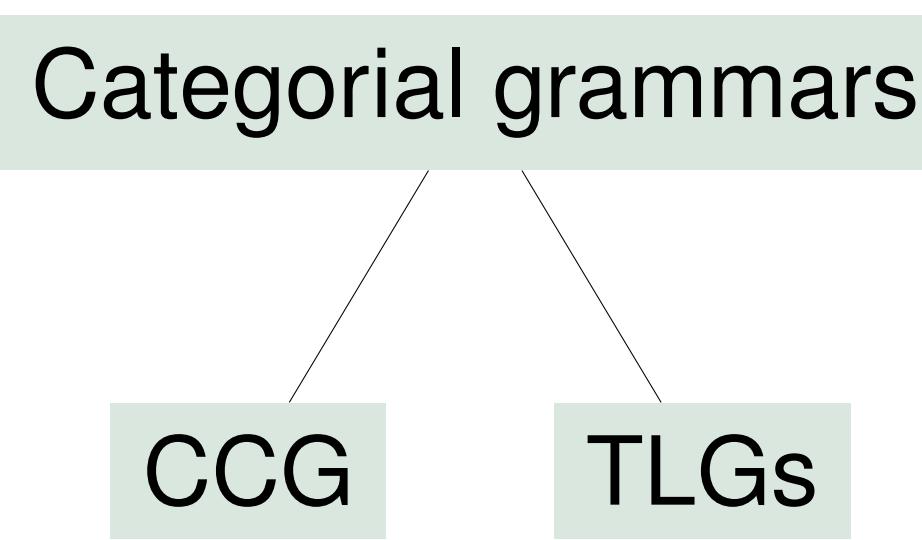
It's time to adopt the **categorial** approach as a **unifying** framework for analyzing competence and performance!

## Background

- Formal grammar is still important in the era of LLM [1]
  - But Chomsky's approach has serious problems
    - Lack of formality [2]
    - Incompatibility with real-time processing [3]
  - Recent success of CCG [4] in psycholinguistic modeling [5,6,7]
- ⇒ Time for a more general theory of the categorial approach to sentence processing!

## Going type-logical

- While CCG is implementation-friendly, **type-logical grammars (TLGs)** [8,9] are more elegant and illuminating
- TLGs view **grammar as logic**
- Extensively studied as a competence theory
- Linguistic expressions are assigned **types**
  - Basic types: S, NP, etc.
  - Derived types: A/B, B\A (cf. fractions), etc.
- Basic rules of inference (Lambek calculus [7])



$$\frac{B \quad B \setminus A}{A} \setminus E \quad \frac{A/B \quad B}{A} / E \quad \frac{\begin{matrix} [B]_i \dots \\ \vdots \\ A \end{matrix}}{B \setminus A} \setminus I_i \quad \frac{\begin{matrix} \dots [B]_i \\ \vdots \\ A \end{matrix}}{A/B} / I_i$$

- A basic example:

$$\frac{\begin{matrix} \text{saw} & \text{Kim} \\ \text{Pat} & \frac{(\text{NP} \setminus \text{S})/\text{NP}}{\text{NP}} \text{NP} \end{matrix}}{\text{S}} \setminus E$$

★ Can account for key characteristics of human sentence processing

## Incrementality, predictivity, grammar-guidedness

- Incremental parse just by mechanisms equipped in the grammar (strict competence hypothesis; [4])
- Parsing as deduction**; hypothetical reasoning as prediction

$$\frac{\begin{matrix} \text{Pat} \Rightarrow \text{Pat} \frac{\text{saw}}{(\text{NP} \setminus \text{S})/\text{NP}} \Rightarrow \frac{\begin{matrix} \text{saw} \\ \text{Pat} \frac{(\text{NP} \setminus \text{S})/\text{NP}}{\text{NP}} \end{matrix}}{\frac{\begin{matrix} \text{NP} \quad \text{NP} \setminus \text{S} \\ \text{S} \end{matrix}}{\text{S}/\text{NP}}} \setminus I_1 \end{matrix}}{(\text{S if NP follows})} / E$$

- Incremental processing of long-distance dependencies, which can be a headache for mainstream-based approaches [10], is also possible just within grammar
- While CCG poses a fixed limit on incrementality (by the choice of rules) [11], TLGs offer a scale of complexity

## Resource sensitivity

- TLG-based processing is resource-efficient in two ways, overcoming limitations in cognitive resources [12]
  - Actions as theorems**: Multiple inference steps can be chunked as a theorem, e.g.,

$$\frac{\text{NP } (\text{NP} \setminus \text{S})/\text{NP}}{\text{S}/\text{NP}} \text{ Theorem}$$

\* CCG can be seen as a set of theorems

- States as propositions (types)**: Types provide an austere summary of the past input; once a type is inferred for the string, no need to maintain what was in between, e.g.:

$$\frac{\text{Pat saw whatever}}{\text{S}/\text{NP}}$$

- [6,7] show (CCG-based) chunking failure predicts slowdown in English and Japanese reading time datasets

$$\frac{\begin{matrix} \text{the article that the student} \\ \text{NP}/(\text{S} \mid \text{NP}) \end{matrix}}{\text{NP}} \text{ ***No rule to apply***}$$

## Resilience and vulnerability to structural ambiguity

- A puzzling contrast in the magnitude of garden-path effect:
  - “NP/S” ambiguity: *Pat knows the baby cried.* **not too hard**
  - “NP/Z” ambiguity: *When Pat dressed the baby cried.* **very hard**
- TLGs’ conjunction operators,  $\sqcap$  ‘and’ and  $\sqcup$  ‘or’, offer a new approach to ambiguity, accounting for the puzzle
- Lexical type assignments based on semantic considerations:
  - know* →  $(\text{NP} \setminus \text{S})/(\text{NP} \sqcup \text{S})$
  - dressed* →  $((\text{NP} \setminus \text{S})/\text{NP}) \sqcap (\text{NP} \setminus \text{S})$
- “NP/S” ambiguity requires **reanalysis** of the NP only

$$\frac{\begin{matrix} \text{Pat knows the baby} \\ \vdots \\ \text{S}/(\text{NP} \sqcup \text{S}) \end{matrix}}{\text{S}} \frac{\text{NP} \sqcup \text{S}}{\text{NP}} \sqcup I_1 / E$$

- “NP/Z” ambiguity requires **reanalysis** of the verb and the NP

$$\frac{\begin{matrix} \text{When Pat} \frac{\text{dressed}}{((\text{NP} \setminus \text{S})/\text{NP}) \sqcap (\text{NP} \setminus \text{S})} \end{matrix}}{\text{(S/S)}/\text{NP}} \frac{\text{(NP} \setminus \text{S})/\text{NP}}{\text{NP}} \sqcap E_1 \quad \frac{\text{the baby}}{\text{NP}} / E$$

## References

- [1] Futrell & Mahowald 2025 *arXiv* [2] Borsley & Müller 2021 In Müller et al. (eds) *HPSG handbook* [3] Sag & Wasow 2011 In Borsley & Börjars (eds) *Non-transformational syntax* [4] Steedman 2000 *The syntactic process* [5] Stanojević et al. 2023 *Cog Sci* [6] Isono 2024 *Cogn* [7] Isono et al. 2025 *PsyArXiv* [8] Lambek 1958 *Am Math Mon* [9] Moortgat 1997 In van Benthem & ter Meulen (eds) *Handbook of Logic and Lang* [10] Hunter 2019 In Berwick & Stabler (eds) *Minimalist Parsing* [11] Demberg 2012 *TAG+11* [12] Christiansen & Chater 2016 *Behav & Brain Sci*