Processing referring expressions: Accessibility is not predictability

Weijie Xu, Ming Xiang (The University of Chicago)

Introduction. The concept of *accessibility* is often assumed to be underlying factor in reference resolution. According to the Givenness Hierarchy (GH) Theory [1], a referent's accessibility in the mental state of a comprehender is encoded in the form of the reference (RF) as part of its lexical semantic representation. In the example in Table 1, therefore, pronouns encode the highest accessibility level, and definite descriptions the lowest. However, the current literature has not reached a consensus on what accessibility exactly means and how to best quantify it. The factors that modulate accessibility, however, show a great extent of overlap with another independently motivated concept of *predictability* [2-6], raising the possibility that the two could be unified. Unlike accessibility, there is a formalized metric of predictability: the likelihood that a given referent is to be mentioned next given the current discourse context. It is theoretically desirable if predictability could serve as the approximation of accessibility. In a self-paced reading study, the current study examines whether the two theoretical constructs are empirically equivalent.

Hypothesis. If accessibility in GH theory is exchangeable with predictability, each RF should encode a certain level of predictability, in the same order as the GH. For example, in Table 1, pronouns encode the highest predictability and definite descriptions the lowest. A plausibility-violation effect is therefore expected when the comprehender encounters a referent whose actual discourse predictability mismatches the predictability implied by the reference form.

Experiment. We evaluated whether each RF in Table 1 encodes a certain level of predictability in the same order as theorized by GH with a self-paced reading experiment. Given the hierarchy, from "the N" to the pronoun, the above mentioned violation effect should be gradually dampened for highly predictable referents and be enhanced for referents that are less predictable, resulting in an interaction effect between predictability and RF.

Method. Native English speakers recruited on Amazon MTurk (n=112) read a context passage and then self-paced read a one-sentence continuation, as in (1). We manipulated the form of the target referent in the continuation sentence (as shown in the curly bracket in (1)). Since the experimental materials were adapted from the corpus constructed by [7], the predictability of the target referents measured with a referent cloze game in the original study was available to us.

Results. LMEMs over log RTs were performed for the critical referent region and the spill-over region. The critical fixed effects predictors are the Reference Form (RF) and the Predictability of the referent. The regression model also controls for a number of other effects (see (2)). When comparing each RF in Table 1 with the previous RF on the GH, on neither the critical region nor the spill-over region, did we find step-by-step RF x Predictability interaction from the pronoun to "the N", indicating that the RFs are not forming a hypothesized "Predictability Hierarchy". However, in the spill-over region, there is a RF x Predictability interaction when comparing "the N" (Figure 1, Right) to the pronoun ($\beta = 0.188$, $p = 0.018$) and to "that N" ($\beta = 0.169$, $p = 0.034$). This provides some evidence that at least "the N" encodes a different degree of predictability of the referent, distinguishable from other reference forms

Conclusion. While there is no robust support to approximate the Givenness Hieracrchy with a "Predictability Hierarchy", there is some preliminary evidence for a partial correlation between the form of a referent and the predictability of a referent.

(1) **Sample Experiment Stimuli** (only the continuation sentence was read in the SPR paradigm), critical region in the curly bracket.

Context Passage: Today, in Rich's Kitchen we'll learn about the fine attributes of baking a cake. Since I am not a phenomenal baker we will be assisted by the use of Little Debbie in using one of their fine cake mixes.

Continuation: In order to/ properly make/ {it/this cake/that cake/the cake}/ we/ will/ need/ some vegetable oil/ and/ a couple of eggs.

in focus	>	activated	>	familiar	>	uniquely identifiable
{it}		{this N}		{that N}		{the N}

Table 1: The GH investigated in the current study. The hierarchy is in descending order: the simplex pronoun encodes the highest accessibility level; the proximal "this N" encodes the second highest accessibility level, followed by the distal "that N" and the definite "the N".

(2) LMEMs over logRT with the maximal random effects that allow the model to converge.

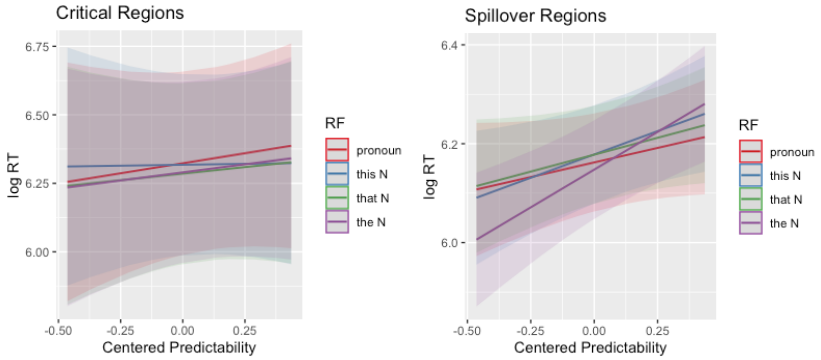
Fixed effects: Predictability * Reference.form + Word.length + Chunk.position + RT.previous + Phi.featured.ref + Recency + Frequency + Intervening.ref + Previous.ref + Gram.role + Previous.gram.role + If.in.SPR

Random effects: *Critical region:* (Predictability|participant) + (1|item)

Spill-over region: (1|participant) + (1|item)

Note: "RT.previous" is the logRT of the previous chunk; "Phi.featured.ref" is the number of referents with the same phi features as the target referent; "Recency" is the distance between the last antecedent and the target referent; "Frequency" is the number of mentions of the target referent in the discourse; "Intervening.ref" is the number of referents between the last antecedent and the target referent; "Previous.ref" is the number of referent appeared so far in the discourse; "Gram.role" is the grammatical role of the target referent; "Previous.gram.role" is the grammatical role of the most recent antecedent; "If.in.SPR" indicates whether the most recent antecedent is in the SPR sentence.

Figure 1. Model predicted interaction between Predictability and Reference Form



References. [1] Gundel et al. (1993) *Language* [2] Ariel (2001) *Text representation, linguistic and psycholinguistic aspect* [3] Kaiser & Trueswell (2008) *Language and Cognitive Processes* [4] Tily & Piantadosi (2009) *Proceedings of the workshop on the production of referring expressions* [5] Arnold (2001) *Discourse Processes* [6] Arnold & Zerkle (2019) *Language, Cognition and Neuroscience* [7] Modi et al. (2017) *Transactions of the Association for Computational Linguistics*