When the food arrives before the menu
Modeling Event-driven Surprisal in Language Comprehension

Online language comprehension is strongly guided by our knowledge about the world, including our knowledge about structured events [e.g., 1, 2]. One way to think about this event knowledge is in terms of the typical ordering of basic everyday events (cf. the idea of *scripts* [3] or *schemata* [4]). For instance, our knowledge about *going to a restaurant* typically consists of a sequence of events, including 'getting the menu', 'ordering food', 'eating' and 'paying'. During linguistic processing, this kind of event knowledge is readily used by interpreters, for instance to predict upcoming words, as illustrated in the following example: *"Rita entered a restaurant and asked for … [the menu/ the bill]"*. In this example, the preference for 'the menu' as opposed to 'the bill' critically depends on our knowledge about the temporal ordering of events in a restaurant setting, and cannot be deduced from the linguistic context alone. However, due to the difficulty in representing and incorporating world knowledge in a computational model of linguistic processing, this information is generally disregarded [5]. A case in point is a recent strand of models that estimate word processing difficulty in online language processing through information theoretic metrics such as *surprisal* [6,7] and *entropy reduction* [8]. The core idea of these models is that the difficulty in processing a word w_t can be estimated as a function of its probability given its prior linguistic ($w_1 \ldots w_{t-1}$) and extra-linguistic context (ELC): e.g., $surprisal(w_t) = -\log P(w_t | w_1 \ldots w_{t-1}, \text{ELC})$. The probability of w_t given the linguistic context can be straightforwardly derived from language models, but its probability given the extra-linguistic context is generally left out of implemented models.

In order to improve upon this situation, we present a neurocomputational—recurrent artificial neural network—model of language processing that integrates linguistic knowledge and world/event knowledge, and that produces word surprisal estimates that take into account both. Our model constructs a cognitively motivated situation model of the state-of-the-affairs as described by a sentence [9]. Critically, these situation model representations inherently encode world/event knowledge. We will show that the surprisal estimates that our model produces reflect both linguistic surprisal as well as surprisal that is driven by knowledge about structured events. We will outline how we can employ the model to explore the interaction between these types of knowledge in online language processing.

Situation model representations encoding event structure

Modeling the interaction between linguistic and event knowledge within a recurrent neural network model, requires a representational scheme for meaning that inherently encodes knowledge about event structure. To this end, we employ the Distributed Situation Space (DSS) model: a psychologically motivated, distributed representational scheme for meaning. DSS representations have been successfully employed to model aspects of both story comprehension [10] and world-knowledge driven semantic processing [9]. In the DSS model, an observation of a state of affairs is represented as a vector encoding which events are the case in that state of affairs. A large sample of such observations creates a high-dimensional "situation-state space", which encodes all knowledge about the world by means of event co-occurrences. To render this feasible, the situation-state space encodes a *microworld*: a confined world consisting of a limited set of events and conditions on the co-occurrence of these events.

The DSS model defines a microworld in terms of a finite set of atomic events (e.g., *enter(rita,restaurant)*). Observations in a microworld are described in terms of these atomic events, which can be combined to form complex events (e.g.,*enter(rita,restaurant)* ∧ *leave(rita,restaurant)*). Critically, world knowledge dictates that not all combinations of events are possible (hard constraints) or equally likely (probabilistic contraints). A situation-

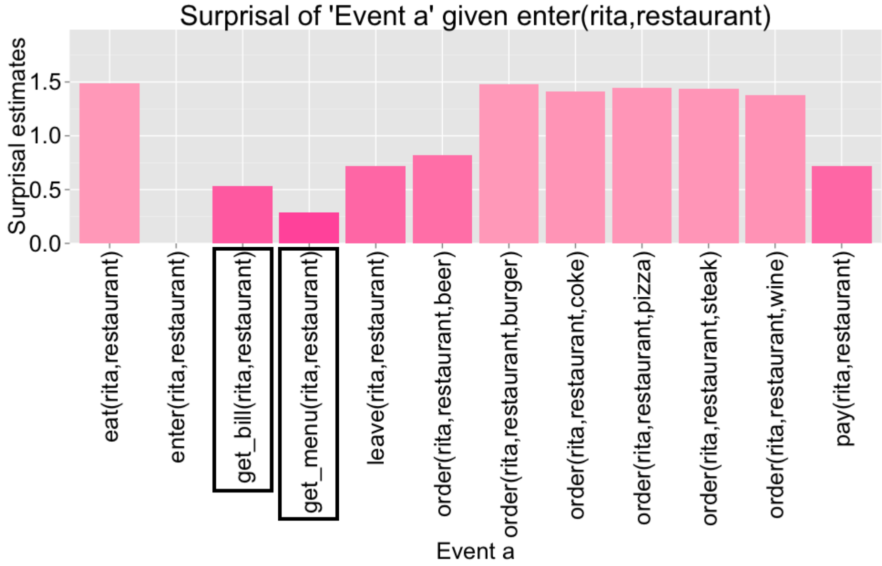

Figure 1: Surprisal estimates from the DSS model for a set of events given 'enter(rita,restaurant)'.

state space, then, is a large sample of independent, (mostly) complex observations. As such, there is no temporal dependence between observations. In order to incorporate temporal ordering between events within DSS representations, then, we assume that an observation must contain the temporally preceding events for any given event; i.e., an observation that contains the event *get_menu(rita,restaurant)* must also contain the event *enter(rita,restaurant)*. Critically, the "situation-state space" inherently carries information about the likelihood of events in the world relative to other events [see 10]. This allows for estimating the surprisal of an event given any other (possibly complex) event. Figure 1 illustrates the surprisal estimates for a set of events in our microworld, given the event '*enter(rita,restaurant)*'. This figure shows that the temporally closer event '*get_menu(rita,restaurant)*' is more expected than '*get_bill(rita,restaurant)*' in this context.

Modeling the online interaction between linguistic and event knowledge

In order to obtain an online measure of surprisal, which is driven by both the linguistic context and event knowledge encoded in the DSS, we employ a Simple Recurrent Neural network (SRN) of the Elman type [11]. The model is trained to take sequences of word representations as input (*linguistic knowledge)*, and to produce a situation situation model of the state-of-the-affairs—a DSS representation—described by a sentence on a word-by-word basis (*event knowledge*). Hence, the model naturally combines linguistic and situational knowledge, which allows for estimating the surprisal of a word w_t directly from the situation model sm_t that is constructed after processing w_t, and the previous situation model $sm_{1\ldots t-1}$: $surprisal(w_t) = -\log P(sm_t|sm_{1\ldots t-1})$. Using these estimates, we can model the effect of different phenomena in which event knowledge has been shown to influence expectations above and beyond purely linguistic surprisal. These phenomena include discourse level event representations [2], the influence of event knowledge on verb complement expectations [12], the processing of coercion phenomena [13], and the effects of temporal violations in event sequences [14].

References. [1] Zacks & Tversky (2001), *Psychol. Bull.*; [2] Metusalem et al. (2012), *JML*; [3] Schank & Abelson (1977); [4] Rumelhart (1980), *In: Spiro et al. (eds)*; [5] Frank et al. (2008), *Discourse Process.*; [6] Hale (2001), *NAACL*; [7] Levy (2008), *Cognition*; [8] Hale (2006), *Cognitive Sci.*; [9] Frank et al. (2009), *Cognition*; [10] Frank et al. (2003), *Cognitive Sci.*; [11] Elman (1990), *Cognitive Sci.*; [12] Bicknell et al. (2010), *JML*; [13] Kuperberg et al. (2010), *JOCN*; [14] Raisig et al. (2010), *Int. J. Psychophysiol.*