david **reitter**

Research Statement

Overview

As a computational cognitive scientist, my goal is to understand how the mind works. I study how humans produce language and how they engage in dialogue. I study how they process and store information, how they make decisions, and how these processes interact. Computational and observational methods on large-scale data are key to my agenda. I would like to use our understanding of the cognitive processes involved in language processing and dialogue to make natural language processing systems more robust, dialogue systems more effective, and human-computer interfaces more natural. My theoretical goal is to have an integrated model of the mind that is computational, predictive, and evaluated on naturalistic data. The model will tell us how we conceptualize and learn. It will link language processing to general cognition, and it will be formulated within a computational cognitive architecture (Newell, 1991) that can generate new predictions and that can contribute to major challenges in artificial intelligence.

I have studied dialogue extensively: how people converge towards a common language through *alignment* (Reitter and Moore, 2014), and how speakers opportunistically distribute information (Xu and Reitter, 2018) in dialogue. I have examined and modeled *syntactic priming* as a specific alignment effect (Reitter et al., 2011). We have modeled language using connectionist architectures that learn based on prediction (Ororbia II et al., in press). Generally, I see language processing as a task that utilizes general cognitive capabilities such as declarative, procedural and working memory, with some components that are separate, optimized towards tasks useful in language processing, but not fundamentally different from general cognitive function. Small local changes and convergence in dialogue lead to long-term language change (cf., Christiansen and Chater, 2008). In pursuit of an architectural descriptions of the human language faculty, we now study how distributed representations for memory can account for syntax and associated psycholinguistic effects (on the basis of Kelly et al., 2013; Kelly et al., 2017). Other, recent work extends my data-driven but theoretically guided approach to decision-making.

My work in computational psycholinguistics has led to follow-up experiments and new insights by other researchers. Collaborations with social scientists have allowed them to apply natural-language processing and cognitive modeling to new tasks.

My research endeavor is difficult, important, and opportune. It is *difficult* because linguistic and behavioral-economic theories do not yet cover the breadth of human behavior: real-world behavior is complex and ambiguous. It is *important* because cognition and its dynamics in networks are still not fully understood (i.e., we lack predictive computational models). My agenda is *opportune* because computational and data-driven methods such as the ones outlined above *now* provide the means to test hypotheses in naturalistic data. Cognitive science, and especially deep, integrated, computational models will remain fundable and intellectually rewarding.

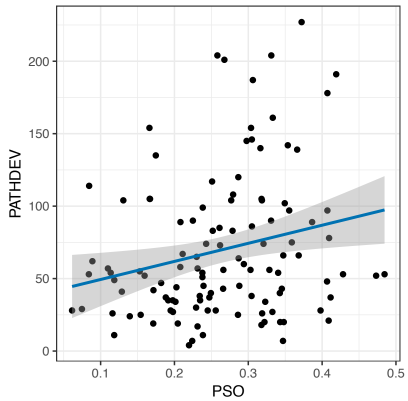

Figure 1: Overlap between information density in frequency space correlates with task success in the HCRC Map Task corpus (shown, 95% C.I.) and in the Danish DJD corpus. Combining alignment features (Reitter and Moore, 2007) and information density spectrum overlap, our can predict 28% of the variance in task success in Map Task (Xu and Reitter, 2018).

How speakers align in dialogue

Communication and cooperation are fascinating human abilities. To solve problems, people organize in teams, develop communication systems and exchange information effortlessly. How do innate, quite general capabilities and learned processes combine to allow us to do so? Studies of language-specific grammar and of linguistic universals provide theoretical constraints, while psycholinguistics contributes findings based on human behavior and neurophysiology. These fields provide a mature scientific basis for the formalization of cognitive processes in models.

I have employed powerful data analysis methods to understand how humans process, store and exchange information. A notable line of inquiry was on *alignment*, the property of dialogue that leads participants to converge in their languages. This is based on an experimentally established adaptation effect, *syntactic priming* (Pickering and Ferreira, 2008). Its cognitive underpinnings have been unclear, and its generality and relevance for language use beyond specific experimental designs are not well understood. First, I generalized the syntactic priming effect to all syntactic constructions and quantified it through statistical analysis of spontaneous speech in dialogue corpora. Turning to alignment, my empirical work was the first to test a hypothesis that links linguistic adaptation to mutual understanding in dialogue (Pickering and Garrod, 2004); I tested it using large corpora of spoken and written language (Reitter and Moore, 2007; Reitter and Moore, 2014). This result is not just of theoretical, but also of practical importance; it demonstrates how a computer program can predict a team's task success from linguistic adaptation (Reitter and Moore, 2007).

KEY PUBLICATIONS: Reitter and Moore (2007) and Reitter and Moore (2014)

FUNDED BY: NSF Robust Intelligence

Information distribution in dialogue

Dialogue is about information exchange – but speakers do not contribute information equally. Information has been of great interest to many researchers, and it is a question that spans the boundaries of psycholinguistics, neuroscience, computer science and linguistics. We know that the predictability of words or syntactic choices as a proxy for information has empirical effects on performance in sentence production and comprehension, and this effect has profound implications for our understanding of how language processing works and interacts with general cognitive function (e.g., Hale, 2003). Recently, we discovered how a principle of information distribution, commonly found in monologue, applies to spoken dialogue (Xu and Reitter, 2016): speakers display an unexpected convergence pattern (Fig. 2) that suggests that dialogue partners are to

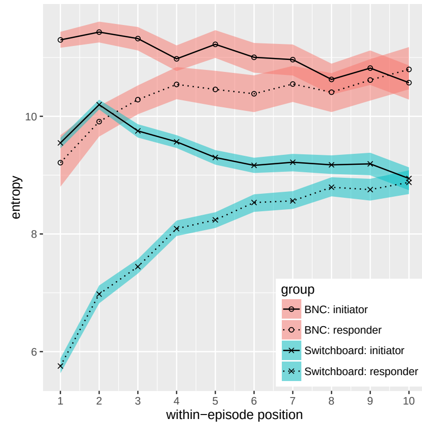

Figure 2: Sentence information vs. sentence position within a topic segment in two-party dialogue (in British National Corpus and Switchboard/Penn Treebank). Shadowed areas: 95% bootstrapped confidence intervals.

be treated as a collaborative system rather than as mere individuals, who aim to balance information density. Information can then be used to improve outcome prediciton in task-oriented dialogue (Xu and Reitter, 2018).

Exploring these questions requires empirical and computational, interdisciplinary methods. My research program engages in a cycle of statistical analysis of large-scale datasets, incremental and computational model-building, and experimental validation. The level of the individual is reflected in cognitive models of language processing and decision making, informed by cognitive science and linguistics and based on empirically established effects.

Such work has produced an emerging research area, *computational psycholinguistics*, which explains human linguistic behavior (qualitatively) and performance (quantitatively) through computational and statistical models. On the basis of my leadership as area chair and chair of conferences central to the area (e.g., ACL, ICCM, CMCL, etc.), and having edited issues of *Topics in Cognitive Science*, I see the field of computational psycholinguistics moving towards large-scale data and advanced statistical inference for scientific gain.

KEY PUBLICATIONS: Xu and Reitter (2018)

FUNDED BY: NSF Robust Intelligence (CRII)

Integrated models of language production

A theoretical understanding of the alignment and syntactic priming phenomena is important, which is why I pursue cognitive models of language production. These integrate with psychological frameworks that account for the general function of the mind, such as how fast and accurate memory is. A core idea of this is "forgetting"; the decay of information held in memory allows us to contextualize our thinking, adapting to an ever-changing environment. In a cognitive model of language production, I explained syntactic priming as a learning effect that follows from basic memory retrieval principles, including decay. My model relies on a broad-coverage grammar to reflect linguistic processes. Borrowing from ACT-R's (Anderson, 2007) concept of memory, I develop a unified theory of syntactic adaptation in language production (Reitter et al., 2011). It reconciles several known, but counterintuitive modulators of syntactic priming. I demonstrate how two assumed-to-be-disparate effects are fully explained by cue-based memory retrieval. The cognitive architecture and my model form a joint account of how general cognition and innate and acquired linguistic devices combine. Predictions arising from this model have since been affirmatively tested in human-subjects experiments by other labs (Bernolet et al., 2016; Kaschak et al., 2011; Tooley and Traxler, 2018) and have differenciated the model from other ones (e.g., Chang et al., 2006; Jaeger and Snider, 2013).

We have evaluated audience design and social factors (power status) as covariates of alignment in Internet forum discourse, for which we have data spanning more a decade. We are now working on an unsupervised approach that will allow us to use vast amounts of data to validate such a model (NSF proposal under review). This would allow evolutionary model building, a new, radical research design for computational psycholinguistics.

KEY PUBLICATIONS: Reitter et al. (2011)

FUNDED BY: NSF Linguistics, Air Force Research Lab, in review: NSF CAREER

Deep insights? Connectionist models of memory.

If memory serves such an important role in language processing, how are linguistic structures learned, routinized and represented? Connectionist modeling has reemerged in machine learning as the training of deep representations has become algorithmically feasible (Hinton et al., 2006). Deep learning may be a point at which models of memory and A.I. connect. Deep connectionist representations let the models discover latent structure in data (Bengio, 2009). Deep Learning has been successful in applications, but so far, it has not led to additional insights into human cognition. However, there is much interest in the community in that it has the potential to explain important cognitive questions by maximizing learning from sparse data (e.g., poverty-of-the-stimulus argument). As a past area chair at the meeting of the Association for Computational Linguistics (ACL) and regular contributor to the ACL family of conferences, I have seen a desire among reviewers to move beyond the applications of existing machine learning techniques to problems that can be formulated in terms of shallow features and defined outcomes. We seek *theory* as a way to formalize our understanding of cognitive and linguistic phenomena, and to use learning as a model of human behavior and human cognition rather than to merely achieve defined (often shallow) goals. In other words, while we consider networks with two hidden layers already as "deep", I anticipate a "deep learning" revolution where learning leads to "deep insight".

We study models in semi-supervised, online situations (where data are fed gradually); the work implements life-long learning rather than a traditional fulll-batch train-test approach. We devised scalable algorithms for restricted, multi-layer Boltzmann Machines with above-state-of-

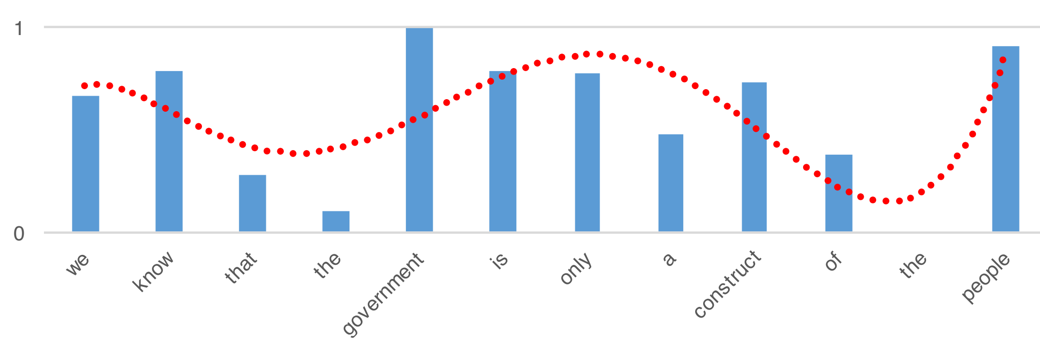

Figure 3: Distances measured in representational space between consecutive states of model trained on Penn Treebank for each word of a sentence, showing how much the model adjusts its representation at each step. This indirectly illustrates the level of surprisal of the model on the basis of its syntagmatic and paradigmatic expectation for each next word.

the-art outcomes(Ororbia II et al., 2015; Ororbia II et al., 2015). Recently, we developed the *Delta-RNN*, a recursive network that mimics the cognitive concept of *surprisal*. It works exceedingly well as a language model, being smaller by order of magnitude and thus more generalizable, and improving on all competing models before regularization(Ororbia II et al., in press). We are now seeing language models in terms of their explanatory potential in psychology in addition to their being an engineering goal.

We explore connectionnist, distributed models of higher-order relationships in memory. These can explain n-th order priming effects, and we are working on syntagmatic in addition ot the known paradigmatic representations. With this new NSF-funded project, we will elucidate how distributed models can represent syntactic knowledge.

KEY PUBLICATIONS: Ororbia II et al. (2015), Ororbia II et al. (in press), and Kelly et al. (2017)
FUNDED BY: NSF Perception, Action, Cognition and Robust Intelligence

Timing decision-making under uncertainty

Beyond successful communication, human information processing has further facets. My group explores risky decision-making under uncertainty in the context of security. This work explores non-rational decision-making in humans through large-scale empirical data (obtained via cost-effective experiments run world-wide in Amazon Mechanical Turk) and through computational-cognitive models. The scenarios we study concern continuous decision-making, which may be more ecologically valid than discrete decision-making studied by traditional game theory. We are able to link behavior in the experimental game to personality traits obtained via standardized instruments. Through behavioral experiments, we have started to successfully map the space of individual differences, pointing out how some aspects of personality influence human risk-taking and problem-solving. Motivated by pervasive problems in cyber security, this work is the result of an ongoing collaboration with Jens Grossklags (TU Munich). Recently, we developed a manipulation that affects individual's impatience (and/or time perception) to affect their material choices in a timing task (see Fig. 4).

KEY PUBLICATIONS: Ghafurian and Reitter (2016)
FUNDED BY: Penn State Seed Funding

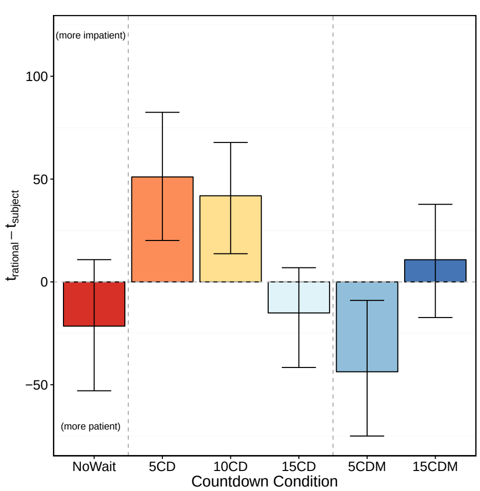

Figure 4: People become less impatient when they are shown faster-changing countdowns. In our manipulation, we slow down (5CD) or speed up (15CD) countdowns during a pause of equal duration to affect a person's time perception and their impatience. The timing task that follows shows systematic deviations of timing choices from the rational solution. (From Ghafurian and Reitter, 2016.)

References

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* Oxford University Press.

Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and Trends in Mach Learn*, *2*(1), 1–127.

Bernolet, S., Collina, S., & Hartsuiker, R. J. (2016). The persistence of syntactic priming revisited. *Journal of Memory and Language*, (91), 99–116. doi:http://dx.doi.org/10.1016/j.jml.2016.01.002

Chang, F., Dell, G. S., & Bock, J. K. (2006). Becoming syntactic. *Psych Review*, *113*(2), 234–272.

Christiansen, M. H. & Chater, N. (2008). Language as shaped by the brain. *Behavioral and brain sciences*, *31*(5), 489–509.

Ghafurian, M. & Reitter, D. (2016). Impatience induced by waiting: An effect moderated by the speed of countdowns. In *DIS '16: Proceedings of the Designing Interactive Systems*. Brisbane, Australia: ACM.

Hale, J. (2003). The information conveyed by words in sentences. *Journal of Psycholinguistic Research*, *32*(2), 101–123.

Hinton, G. E., Osindero, S., & Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. *Neural computation*, *18*(7), 1527–1554.

Jaeger, T. F. & Snider, N. (2013). Alignment as a consequence of expectation adaptation. *Cognition*, *127*, 57–83.

Kaschak, M. P., Kutta, T. J., & Jones, J. L. (2011). Structural priming as implicit learning: cumulative priming effects and individual differences. *Psychonomic Bulletin & Review*, *18*(6), 1133–1139.

Kelly, M. A., Blostein, D., & Mewhort, D. (2013). Encoding structure in holographic reduced representations. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *67*(2), 79.

Kelly, M. A., Reitter, D., & West, R. L. (2017). Degrees of Separation in Semantic and Syntactic Relationships. In *Proc 15th. International Conference on Cognitive Modeling*. Warwick, UK.

Newell, A. (1991, April). Desires and diversions. Lecture delivered at Carnegie Mellon University, School of Computer Science.

Ororbia II, A. G., Giles, C. L., & Reitter, D. (2015). Learning a deep hybrid model for semi-supervised text classification. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Lisbon, Portugal.

Ororbia II, A. G., Mikolov, T., & Reitter, D. (in press). Learning Simpler Language Models with the Differential State Framework. *Neural Computation*. arXiv preprint arXiv:1703.08864.

Ororbia II, A. G., Reitter, D., Wu, J., & Giles, C. L. (2015). Online learning of deep hybrid architectures for semi-supervised categorization. In *ECML PKDD*. European conference on machine learning and principles and practice of knowledge discovery in databases. Porto, Portugal: Springer.

Pickering, M. J. & Ferreira, V. S. (2008). Structural priming: a critical review. *Psychological Bulletin*, 427–459.

Pickering, M. J. & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*, 169–225.

Reitter, D., Keller, F., & Moore, J. D. (2011). A computational cognitive model of syntactic priming. *Cognitive Science*, *35*(4), 587–637. doi:10.1111/j.1551-6709.2010.01165.x

Reitter, D. & Moore, J. D. (2007). Predicting success in dialogue. In *Proc. 45th annual meeting of the Association of Computational Linguistics* (pp. 808–815). Prague, Czech Republic.

Reitter, D. & Moore, J. D. (2014). Alignment and task success in spoken dialogue. *Journal of Memory and Language*, *76*, 29–46. doi:10.1016/j.jml.2014.05.008

Tooley, K. M. & Traxler, M. J. (2018). Implicit learning of structure occurs in parallel with lexically-mediated syntactic priming effects in sentence comprehension. *Journal of Memory and Language*, *98*(Supplement C), 59–76. doi:https://doi.org/10.1016/j.jml.2017.09.004

Xu, Y. & Reitter, D. (2016). Entropy converges between dialogue participants: explanations from an information-theoretic perspective. In *Proc. 64th Annual Mtg. of the Association for Computational Linguistics*. Berlin, Germany.

Xu, Y. & Reitter, D. (2018). Information density converges in dialogue: Towards an information-theoretic model. *Cognition*, (170), 147–163.