Author's version. Please cite the authoritative version in: M.N.Jones (ed.), Big Data in Cognitive Science. New York, NY: Routledge, 2017.

Who Aligns With Whom in Web Forum Dialogue: Studies in Big-Data Computational Psycholinguistics

David Reitter
The Pennsylvania State University

Abstract

Studies on linguistic alignment explain the emergence of mutual understanding in dialogue. They connect to psycholinguistic models of language processing. Recently, more computational cognitive models of language processing in the dialogue setting have been published, and they increasingly are based on observational, large datasets rather than hypothesis-driven experimentation. I review literature in data-intensive computational psycholinguistics and demonstrate the approach in a study that elucidates the mechanisms of alignment. I describe consistent results on data from large online forums, Reddit, and the Cancer Survivors Network, which suggest that linguistic alignment is a consequence of the architecture of the human memory system and not an explicit, consciously controlled process tailored to the recipient of a speaker's contributions to the dialogue.

In this chapter, I will survey studies on linguistic alignment in order to connect the inferences we make from large datasets to the psycholinguistics of dialogue. This work is motivated by a high-level question: How does our mind select and combine words in a way that reliably communicates our intentions to a dialogue partner? The search for a computational answer to this question has been revitalized in recent years by the advent of large datasets. Large data give us a window into an individual's mind and a cooperative process between minds. It allows us to look at how dialogue partners gradually converge in their choices of words and sentence structure, thereby creating a shared language.

This process relies on an implied contract among the people engaging in dialogue. It specifies long-term and temporary conventions that establish the meaning of words, idioms, syntactic structure, or the general topics of conversation. The rules governing convention-forming are subject to debate: how does the contract evolve, and how do people accommodate the linguistic needs of their dialogue partners? What are the cognitive mechanisms that access memory and produce sentences in order to comprehend and produce contextualized language? The studies I discuss in the following aim to shed light on these mechanisms.

The availability of data and new methods from information science has given researchers the tools they need to answer them in reference to "language in the wild": real-world and large-scale language use as opposed to hand-picked examples or carefully constructed experimental materials. However, information and network science as well as modern-era computational linguistics have all been somewhat agnostic to the psychological processes that produce the data that they study. Yet, many of their methods are useful in the the context of cognitive science. In the following, I will use linguistic and structural cues to identify syntactic repetition, but also to characterize an interlocutor's role in contributing novelty to the conversation.

Integrating psycholinguistics and cognitive modeling

Thus far, models of language production have used representations that were either too specialized or too generic. Grammar formalisms are representations that describe syntax at a high level, or that provide a computational account of the syntactic process (e.g., Pollard and Sag, 1994, Joshi and Schabes, 1997, Steedman, 2000). However, these representations leave open many computational questions. They may fall short of explaining all permissible sentences, or they can over-generate by permitting too many sentences. Connectionnist representations (e.g., Dell, Chang, and Griffin, 1999; Elman, 1990) are often focused on specific aspects of language processing, although the machine learning and A.I. literature has advanced far beyond (e.g., Mnih and Hinton, 2009; Mikolov, Sutskever, Chen, Corrado, and Dean, 2013). While models of language processing have been connected to the psychological literature, their assumptions about memory use are intentionally more modest (e.g., E. Gibson, 1998).To move forward, we need a tighter integration of models of language processing and cognitive architectures. This raises several questions, which address the core of cognitive science.

1. ***Cognitive plausibility*** Given the available information, and computational-cognitive resources, which accounts of linguistic representation can be learned and processed in real-time by the mind?

2. ***Representations*** Which learned mental representations guide language processing?

3. ***Specialization*** Which computational operations and memory components are specific to language, rather than being shared with general cognition?

An integrated account can take a stance with respect to each of these questions. In following Newell's call for models that "must have all the details" and describe "how the gears clank and how the pistons go" (J. R. Anderson, 2007), the model should be a computational account that actually carries out language acquisition, language comprehension and language production. Further, Newell's call for functional models means that we need to cover the broad range of linguistic constructions present in a corpus. To achieve this objective, we must use the large-scale language resources that are standard in computational linguistics. They reflect language use in the wild.

The converation about such approaches has been taking place in a relatively new field, *computational psycholinguistics*, which discovered a range of phenomena that may form the basis for how we think about the mechanisms of human language acquisition and processing.

Linguists have asked provocative questions using these methods. To name a few: how is information density distributed throughout text, and why? When is language production incremental? How is working memory used in language processing? Computational psycholinguistics has pushed the boundaries to cover the broad expressive range found in corpora.

The field discovers how humans learn, produce and comprehend natural language, and the models are informed by observations from contemporary language use. Standard psycholinguistic methods examine human language performance by collecting data on comprehension and production speed, eye movements while reading (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; Demberg & Keller, 2008; Henderson & Ferreira, 2004), or specific processing difficulties (e.g., self-embedding or general center-embedding: Chomsky and Miller, 1963; E. A. F. Gibson, 1991). These methods are productive but require data collection, while more can be learned from unannotated data. The machine learning field of semi-supervised learning has developed computational accounts that describe successful learning from small portions of annotated and large portions of unannotated data (e.g., Chapelle, Schölkopf, and Zien, 2006; Ororbia II, Giles, and Reitter, 2015).

Large datasets have been used to first verify experimental results in naturalistic language, and now they allow us to find more fine-grained support for theoretical models of dialogue. The subject that exemplifies the use of big data is *alignment*, the tendency for people in a conversation to conform to each other in syntax, word choice, or other linguistic decision levels. The studies relating to alignment are particularly interesting, as they relate low-level repetition effects such as *syntactic priming* to high-level dialogue strategies and even up to the social roles of those participating in dialogue. Big data has been contributing to our understanding of this process. I will bridge the range of work from traditional controlled experimentation to a new analysis of a very large internet dataset, illustrating not only results, but also challenges associated with such datasets.

Alignment

I focus on a set of well-known memory-related phenomena around alignment (Bock, 1986; Pickering & Ferreira, 2008). This is an effect of gradual convergence throughout dialogue. Alignment describes a range of related phenomena that cause speakers to repeat themselves or others, to gradually adapt to someone else, or to become more consistent in their own speech. This tendency affects not just the words speakers use; it also affects their sentence structure and even some aspects of semantics. Indeed, alignment[1] is claimed to be based on adaptation effects at several linguistic levels (Pickering & Garrod, 2004): lexical priming, syntactic (structural) priming, or priming at even higher, behavioral levels.

Function in dialogue: alignment as a driver of dialogue

The Interactive Alignment Model (IAM) (Pickering & Garrod, 2004) suggests that the convergence of situation models in dialogue is the result of an interactive process. It is based on mechanistic repetition at a number of linguistic levels.

This mechanistic repetition that forms the basic building blocks of alignment may be due to a known memory effect called *priming*. Pickering and Garrod (2004) suggest that a cascade of priming effects at different levels is a driver of mutual understanding and cooperation. Indeed, one can find priming or a priming-like effect in everything from

[1] Alignment in this sense is distinct from alignment between texts or sentences as used in machine translation. It is also different from alignment as used in constituent substitutability, as by Van Zaanen (2000)

phonetic reductions (Bard et al., 2000; Jaeger, 2006) to a joint interpretation of the dialogue situation (Garrod & Anderson, 1987). The notion of alignment in dialogue also applies to the distributional similarity and contrast of verbal and non-verbal features (Paxton, Abney, Kello, & Dale, 2014), and also the use of high-frequency words (Nenkova, Gravano, & Hirschberg, 2008). Empirical evidence also underlines the *function* of alignment in discourse. Speakers tailor the amount of their alignment to both the perceived needs of the interlocutor (Branigan, Pickering, Pearson, McLean, & Brown, 2011) and dialogue objective (Reitter, Keller, & Moore, 2006b). The adaptation at several linguistic levels is predictive of task success (Reitter & Moore, 2007; Fusaroli et al., 2012; Reitter & Moore, 2014).

Alignment can be seen as a default approach to ensuring mutual understanding. The converse would require speakers to model their intelocutors much more explicitly, tracking what they understand and what information they agree on (*grounding*, Brennan and Clark, 1996). However, there is only limited evidence that alignment has a possible function as a signal of agreement between speakers, as it seems to be mostly automatic. Danescu-Niculescu-Mizil and Lee (2011) examined a corpus of movie scripts, finding alignment of function word use in dialogue between movie characters. Even though the authors of these scripts did not receive any potential social benefits of alignment, they still created aligned dialogue.

Syntactic Priming

Syntactic priming is one adaptation effect that contributes to alignment. Speakers mainly choose different words and grammatical structures to express their ideas. However, when people can choose between several alternative grammatical structures, their choice tends to be influenced by what has been already been said in the conversation. Speakers tend to repeat previously encountered grammatical structures, a pattern of behavior that is referred to as syntactic (or structural) priming.[2]

Syntactic priming is interesting in the context of the present discussion because corpus-based studies of this effect have demonstrated that the big-data inquiry can reveal mental representation and processes in dialogue. It is an informative case study in connecting experimental evidence to observational data.

Priming occurs when comprehending or producing language material alters the likelihood of future linguistic choices. *Syntactic* or *structural* priming applies this definition to syntactic choices, while *semantic priming*) refers to the priming of words. A descriptive measure of the magnitude of priming effects has been a challenge to define. Szmrecsanyi (2005), Gries (2005), Jaeger and Snider (2007), Reitter (2008) look at either repetition counts or use logistic regression to examine priming's effect size or decay. These are not commensurable measures that would be suitable to compare magnitudes across studies. Work is under way to define a statistically useful, robust metric with a sensitive measuring apparatus (S. Jones, Cotterill, Dewdney, Muir, & Joinson, 2014; Fusaroli et al., 2012; Xu & Reitter, 2015).

As a result of *syntactic priming*, speakers have a tendency to prefer one syntactic construction over an available alternative shortly after having used this structure or having

[2]For a review, see Pickering and Ferreira (2008); the terms syntactic priming and structural priming are used more or less interchangeably in the literature.

heard an interlocutor use it (Bock, 1986).

Alignment can be used as a method to infer information about language processing from data, if we agree on the following paradigm: *structures that do not exist during language production do not prime*. These structures refer to the actual content of our memories: in this way, alignment can actually serve as a window into the mind. This paradigm can be applied to more or less detailed process models. However, the focus of the study I will present is to go beyond simple short-term repetition effects. I use lexical and syntactic alignment to analyze memory effects that are the precursor to permanent language change in the individual, which in turn are precursor to language change within a cultural community.

Syntactic priming effects are particularly interesting because they reflect implicit decisions. Speakers do not consciously choose syntactic structure, as they might with some words. The effect has been studied extensively in psychological literature. In a now-classical study, Bock (1986) found that adults who listened to and repeated a sentence in a passive form (*The boy was kissed by the girl*) were more likely to describe an image about something completely different using the passive form (*The cat was chased by the dog* as opposed to the active form (*The dog chased the cat*). Many other syntactic choices have been shown to exhibit priming as well. Some others include prepositional objects (The painter showed his drawing to the customer) versus direct objects (The painter showed the customer his drawing) (Bock, 1986; Pickering & Branigan, 1998), complex noun phrases (Cleland & Pickering, 2003), the order of main verbs and auxiliary verbs (Hartsuiker, Bernolet, Schoonbaert, Speybroeck, & Vanderelst, 2008), and a range of other grammatical structures in various languages. Not surprisingly, syntactic priming applies to both dialogues and monologues. Nonetheless, some studies suggest that priming effects are stronger in dialogue than in monologue (Kootstra, van Hell, & Dijkstra, 2010; Cleland & Pickering, 2003).

The tendency to repeat a particular structure when it has been recently encountered has also been observed in corpus studies of spoken language (Reitter, 2008; Szmrecsanyi, 2005; Travis, 2007) and internet forum conversations (Wang, Reitter, & Yen, 2014). Alignment applies to other levels of linguistic analysis, including referring expressions such as pronouns (Brennan & Clark, 1996) and style (Niederhoffer & Pennebaker, 2002).

Corpus studies have since provided evidence for syntactic priming outside of carefully controlled laboratory settings. Speakers adapt in situated, realistic dialogue. For instance, the Map Task corpus (A. H. Anderson et al., 1991; McKelvie, 1998) shows syntactic priming-like repetition effects (Reitter, Moore, & Keller, 2006). The cited corpus study also model priming as an effect that applies to syntactic rules in general, rather than specific alternations such as those in the above examples.

Analyses of spoken language corpora generally showed that the probability of repeating a structure decreases as the amount of time between priming object and the primed object increases (Gries, 2005; Reitter et al., 2006b; Szmrecsanyi, 2006). This would suggest priming effects decay over time. However, these earlier corpus studies did not control for the characteristics of the language between prime and target. To control for this possible confound, I modeled decay of repetition probability as the variable that quantified priming (Reitter, 2008).

There are two clearly separate syntactic priming effects: (a) fast, short-term and short-lived priming, and (b) slow, long-term adaptation that is likely a result of implicit learning (see V. S. Ferreira and Bock, 2006; Pickering and Ferreira, 2008). Long-term adaptation is a learning effect that can persist, at least, over several days (Bock, Dell, Chang, & Onishi, 2007; Kaschak, Kutta, & Schatschneider, 2011). Recent work has proposed models that explain the mechanisms of these effects within the context of language acquisition (Bock & Griffin, 2000; Kaschak, Kutta, & Jones, 2011; Chang, Dell, & Bock, 2006) and general memory retrieval (Reitter, Keller, & Moore, 2011). These studies suggests that priming is the precursor to persistent language change.

Characteristics of syntactic priming

Priming means that the use of a syntactic construction increases the probability of its future occurrence overall, but initially, the probability decays rapidly (Branigan, Pickering, & Cleland, 1999). This decay has been shown to follow a logarithmic function of time or linguistic activity (Reitter et al., 2006b). Since this decay is a side effect of priming, we can use it to quantify the strength of the priming: a stronger priming effect will decay more quickly. This way, we can distinguish it from other potential sources of increased repetition, such as text genre or the clustering of topics. Figures 1 and 4 illustrate the decay effect using different methods and datasets. The repetition rate of linguistic material can be modeled as a function of either time (Fig. 1) or linguistic activity (Fig. 4). Indeed, decay can be observed at much larger time-scales than the one found in spoken dialogue, which has been observed in Internet forum conversations (Fig. 4, Wang et al., 2014). We operationalize decay as follows: *Decay* occurs when the probability of a linguistic choice declines systematically over time, or declines while processing linguistic material. After priming material, the probability of that choice occurring again first increases, and then decays rapidly.

Syntactic priming is also cumulative (Hartsuiker & Westenberg, 2000; Jaeger & Snider, 2008; Kaschak, Loney, & Borregine, 2006). While cumulativity has not been taken into account by previous corpus analyses, it is included in the proposed effort. Models that account for cumulativity and decay in a cognitively plausible manner will make more precise predictions about structure use, processing principles, and parameters defining alignment strength.

Priming is evident in corpus data

Previous work showed that syntactic priming and several related interactions can be observed in corpora of spoken dialogue (Gries, 2005; Reitter et al., 2006b; Reitter, 2008; Szmrecsanyi, 2006). In a series of experiments, we developed logistic linear mixed-effects regression models (logistic GLMMs) that predict repetition probability as a function of the prime-target distance, which represents the decay of short-term effects, and prior exposure, which represents the effect of longer-term learning (Reitter et al., 2006b; Reitter & Moore, 2014). We found priming resulted from both comprehending and producing sentences. Interestingly, we find not only syntactic priming, but also convergence of syntactic complexity (Xu & Reitter, 2016) and general communicative (pragmatic) intent (Wang, Yen, & Reitter, 2015).

Note that GLMMs have become the statistical model of choice to describe corpus data. GLMMs are models that predict variables as functions of discrete factors, continu-

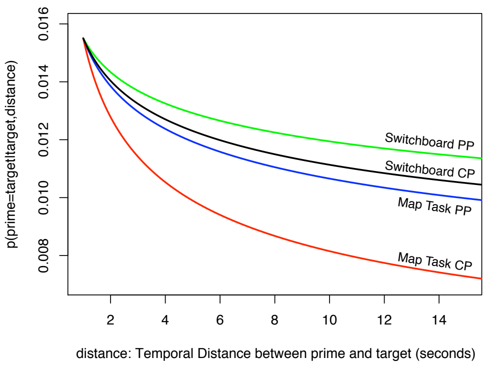

Figure 1. Logistic regression models of syntactic repetition fitted to data from two corpora: Map Task (task-oriented dialogue) and Switchboard (spontaneous conversation by phone). CP: priming between speakers, PP: self-priming.

ous predictors, their interactions, and additional random variables. The purpose of these additional random variables is to control for repeated items or several data-points from a single speaker. They remain, however, linear with the goal of fitting response variables that can be described as the sum of the covariates and their interactions, in contrast with non-statistical process models formulated in connectionist paradigms or cognitive frameworks such as ACT-R.

The work shows that memory effects can be studied in large, observational datasets and that big data methodology can go beyond the replication of known effects. For instance, we have modeled syntactic and lexical priming to predict the successful outcome of conversations (Reitter & Moore, 2014). This work also shows that linear mixed-effects regression can be used to contrast syntactic priming in different conditions. The regression models approximate the decay described by cognitive psychology (J. R. Anderson, 1993; J. R. Anderson & Lebiere, 1998).

How mechanistic is the effect?

The observation that syntactic priming can differ in strength brings into question how mechanistic the adaptation effect actually is. Do people have some form of control over how much they adapt? Is it strategic as opposed to automatic?

Both copying and contrasting syntactic structure are common rhetorical devices. The sentence *Guess what, I went to the Waffle Shop for breakfast this morning!* could be answered with *Guess what, I went to the gym for a workout* to demonstrate a contrast. By comparison, *Well, I did a workout at the gym today* does not convey quite the same pragmatic implicature. How much control do speakers exercise over sentence structure? Is this what controls interactive alignment in dialogue?

A range of modulation effects could address this set of questions. For one, under an account of alignment-for-a-purpose we would expect increased adaptation when needed, such as in task-oriented dialogue. There could also be social modulation of adaptation, such as by social status. There could be modulation according to a speaker's role in the conversation, and there could be implications for adaptation by and among people with social communication disorders and/or Autism Spectrum Disorders.

Indeed, we have observed an increased amount of syntactic priming in task-oriented dialogue compared to spontaneous conversation (Reitter & Moore, 2007). A follow-up cognitive model of syntactic priming (Reitter et al., 2011) may have a mechanistic explanation for the differences we observed between task-oriented dialogue and spontaneous conversation. According to the model, working memory serves as cues to the retrieval of syntactic material. That means that attention to concepts discussed in the conversation plays a role in making syntactic material available. By learning associations of concepts with syntactic decisions, remaining within a specific topic would yield stronger priming effects, because the semantics and associations to syntactic construction is still available, while switching topics would reduce priming effects. This does not disallow any control over adaptation, but it argues for *mechanistic* adaptation as the default in conversation. However, more empirical work is clearly needed to back up the account.

Examining social modulation of alignment

Interactive alignment suggests an alternative theory to the deep cognitive processes suggested by explicit grounding(Brennan & Clark, 1996). The new theory assumed a cascade of simple, mechanistic priming effects at all linguistic levels that led to a shared language (Pickering & Garrod, 2004). At least at the syntactic level, the priming effects can be explained in terms of general cue-based memory retrieval and decay (Reitter et al., 2011). The model is formulated within an independently validated cognitive architecture, ACT-R (J. R. Anderson, 2007), and it occupies a middle ground between psychological theories that argue syntactic priming is purely the result of either implicit learning or residual activation. To summarize, a complex, explicit thought process was largely replaced by a fast, intuitive heuristic default. This parallels a general trend in behavioral science. This heuristic may not always produce the normatively correct answer, but it seems to generally work well enough.

What this theoretical commitment does not mean is that speakers lack control over their choices of words and sentence structure. It does not mean that alignment cannot be modulated under any circumstance. But there are explanations of this modulation that arise out of mechanistic adaptation effect. For instance, I explain stronger priming in task-oriented dialogue as the result of increased persistence in working memory, leading to more associative activation of syntactic constructions. Danescu-Niculescu-Mizil and Lee (2011) examined a corpus of movie scripts, finding alignment of function word use in dialogue between movie characters. Even though the authors of these scripts did not receive any potential "social benefits" of alignment, they still created linguistically coordinated dialogue. According to the authors of the paper, this means that alignment has been engrained in communicative patterns and is removed from its functional role. They even compare linguistic alignment to what was called the *Chameleon Effect*, namely the "nonconscious mimicry of the postures, mannerisms, facial expressions, and other behaviors of one's interaction partners" (Chartrand & Bargh, 1999). Modulation of alignment levels may be linked to affect: del Prado Martin and Bois (2015) find that positive attitudes are linked to more alignment. It is possible that attitudes are linked to engagement and attention, which affects alignment.

This brings up the question of just how does alignment really depend on attention. Do we align primarily with attended-to speakers? This would support a more mechanistic theory of alignment rooted in basic cognitive processes. On the other hand, it may be the case that alignment is an audience-design effect, which causes us to align strategically with those speakers we address as opposed to those speakers to whose language we were most recently exposed.

Branigan, Pickering, McLean, and Cleland (2007) studied the latter hypothesis using lab-based, staged interactions among experimental participants. The results were mixed. While there was some alignment to speakers that weren't addressed, the effect was smaller. This leaves an attention-based explanation as well as social modulation on the table. Observational studies with large datasets allow us to look beyond primary effects of repetition. They cannot replace studies that establish causality, but with enough high-quality data, we can examine more fine-grained interactions that, in this case, would reveal the effects of social modulation on decay. Such observations would have consequences for the architecture of both language production and language acquisition.

The proposed memory-based explanation for alignment suggests that alignment relies on the same mechanisms as language learning. This has the empirically verifiable consequence that alignment is a precursor to permanent language change in individuals and among members of a cultural group. We can consider this possibility in the context of two scientific realms: psycholinguistics and information science. From a psycholinguistic perspective, we seek confirmation of the mechanisms of language production. A sample of pertinent questions: is it attention or intention that modulates alignment? Is memory the driver of alignment, and if so, which kind of memory: declarative or procedural? Answering these psycholinguistic questions faces challenges, as big data in the form of online forums comes with both variability and confounds. Variability occurs due to a of lack of control over which messages an author actually read before writing a reply; however, we can counter such error with more data. Confounds occur, for instance, because the selection of words is inextricably linked to topics, which shift systematically throughout dialogue. However, we can avoid this confound by measuring alignmthe lateent on other linguistic decisions, such as sentence structure.

From the perspective of information science, I am interested in how choices of words and topics propagate through a community. How does discourse change as a consequence of an individual's contributions? How does a speaker's role (as initiator, moderator, or information provider) determine his or her influence in the larger-scale process of language change?

Data and Methods

We make use of data from two web forums that allow us to study the case of asynchronous, written dialogue. The first forum is the *Cancer Survivors Network* (CSN) web forum data. This forum represents about ten years of online conversation (5GB) in a community of cancer patients and survivors, which focuses on providing peer support on informational and emotional levels. The second is the *Reddit* web forum, which is one of the largest discussion platforms on the Internet, consisting of 2TB of uncompressed text by early 2015. Reddit comprises years of discussions, each started by a single question or submission. The CSN conversations, or *threads*, are treated as a flat sequence of replies. The user interface of the web forum made it easier to reply to a post with a message for the whole thread than it did to reply to a specific message on that thread. By contrast, the Reddit conversations are hierarchical, giving us an additional criterion to identify structure in language change. We first published our analysis of the CSN dataset in Wang et al. (2014), and I owe gratitude to Yafei Wang for preprocessing the dataset, and her and John Yen for our collaboration that led to the intial methodology published in Wang at al..

Both datasets are, in principle, publicly available through scraping the respective websites. However, the CSN dataset was obtained through an agreement with the American Cancer Society. The Reddit dataset was curated by a Reddit user and augmented through use of the Reddit interfaces. We use a distributed NoSQL database, which provides high performance in exchange for computational limitations. (NoSQL is a type of database that avoids relational storage and queries, which would result in worse-than constant-time behavior. It is well-suited to scalable operations over very large datasets.)

In the studies presented in this chapter, I will sample from these datasets to infer statistical models rather than use the complete data for two reasons. The first is computational convenience (I admit), the second is predictive validity: novel hypotheses should be tested on a fresh dataset to prevent data-fishing and the associated risk of non-replication.

In both online forums, threads are structured into an initial submission, followed by a tree hierarchy of replies. Each *post*, a synonym for a written message on a forum, replies to exactly one other post. The submission can be thought of as the first post, though it can be empty and only contain a picture or hyperlink. In the case of CSN, the subsequent posts after the submission are flat, rather than hierarchical, due to the presentation of the messages on the website. Therefore, we can only utilize the dependency information in Reddit. Posts are written by authors using pseudonyms, so the authors can be identified throughout the discourse.

The two corpora were processed using natural-language processing tools: they were parsed with the Stanford CoreNLP PCFG parser (Manning et al., 2014); syntax trees for each sentence were converted into sets of rules such as NP→DET N. Then, measures for syntactic (SILLA) and lexical repetition (LILLA) (Fusaroli et al., 2012) were calculated over pairs of posts:

$$S/LILLA(target, origin) = \frac{\sum_{word_i \in target_i} \delta(word_i)}{len(origin)len(target)}$$
$$\delta(word_i) = \{1\ if\ word_i \in origin; 0\ otherwise\}$$

The Reddit dataset reflects a sample of three million origin-target pairs which stem from 2,200 different discussion threads that occurred in 2014 and 2015. No outliers were removed, and messages were not selected to come from specific, potentially more interactive subreddits. The CSN data is a sample of three thousand conversation threads containing 23,045 posts. The SILLA/LILLA measures tend to have favorable distributional properties (Xu & Reitter, 2015) that allow for parametric inference.

Research questions

To determine social influence on alignment, we ask whether messages of different status in a thread can be more or less linguistically influential. We determine the amount of repetition of words and of syntactic constructions for pairs of messages: consisting of an origin and a target message. The earlier "origin" message can take on one of several roles. The origin message can be the first reply to the topic (**F**), messages by the initial author (**I**), self-replies (**S**), and any other messages (**A**).

Under the hypothesis of social modulation of the memory effect, we would expect differences in strength of adaptation regardless of the distance. However, we would also expect differences in decay. Specifically, we would expect to see more adaptation, and thus more decay, in important origin messages or those origin posts authored by someone deemed important. For the purposes of this study, we assume that the initiator of the conversation is important, as is his or her first message, as well as the first reply to the initiator. Under the alternative hypothesis of a purely mechanistic effect, we would see no difference in decay in this scenario or possibly even the opposite relationship.

In order to examine decay of alignment, we first need a system for measuring the distance between the origin and target. One metric for measuring decay in forums is the *reply distance*. For an example, if post P_A replies to post P_B, which in turn is a reply to post P_C, then we would say the reply distance from P_A to P_C is 2. As an alternative measure of distance, we can use the actual time that the post was written. This can be useful because information in a relatively uncontrolled conversation becomes stale and loses influence as time passes by. Notably, there is no single, correct measure of distance when it comes to the analysis of linguistic expression in a study of adaptation. As is typical for a study of Internet-mediated dialogue, we do not have information about whether and when exactly an author of a target message has read the origin message. Likewise, we have no information about how closely she or he has paid attention to that message or any intervening material. Distance in replies and in time are simply proxies for how much material has intervened, and how much time has passed between consuming the messages. The lack of control and the inaccuracy of the measurement proxy is counteracted by the sheer amount of data.

Results and Discussion

In the Reddit dataset, we observe some differences in the decay of syntactic or lexical alignment (Figs 2) when comparing different types of sources for the origin of each origin-target pair. Messages show the most lasting alignment with the initial post, while experiencing stronger decay when aligning with just any given parent[3] post. Importantly, lexical and syntactic alignment with the thread initiator is consistently lower and shows no characteristic decay in the case of lexical alignment. This is incompatible with the hypothesis of alignment as audience design: messages do not align more with the messages of the person who initiated the thread, who should be the most socially important person.

In a secondary analysis (Figure 3), we show decay over time rather than as a function of intervening messages. Early decay is observed for all classes of origin-target pairs including for those where the origin was written by the thread initiator. For longer time

[3]A parent is defined as being a message to which the target message responds, or as being the parent of such any parent message.

spans between origins and targets beyond about 30,000 seconds, repetitions may actually increase with increasing distance to the origin message, which is not alignment. Only the syntax analysis shows a mild repetition increase, which is then followed by a strong decay. Overall, these data do not seem to support an audience-design hypothesis. However, I would caution that the time between writing origin and target posts is a measure that confounds an individual authors memory with a form of externalized, networked memory. This can be appropriate from the perspective of an ecological, high-level model of multi-party dialogue, but is more problematic when interpreting these data from the psychologist's perspective.

The CSN dataset in an analysis first discussed elsewhere tells a similar story (Wang et al., 2014). Lexical alignment in CSN decreases with post distance. Alignment with posts written by the thread initiator is lower initially and also decays less (Fig. 4). For syntactic priming, we even observe an increase in alignment over time, speaking against the audience design hypothesis.

The evidence we find, overall, confirms the results by Branigan et al. (2007) for the cases of both naturalistic multi-party dialogue and lexical alignment. Any memory mechanisms underlying alignment seem to have little sensitivity to the role of the source (origin message), as decay is not greater for such roles. Further, absolute repetitions are initially lower for origin messages by the thread initiator than for other origin messages. The observed differences in lexical similarity can be interpreted as the result of the pragmatic consequences of addressing one another's messages throughout the conversation, rather than as a sign of a lower-level mechanistic process.

The Cancer Survivors Network and Reddit are very different communities: In CSN, members aim to provide emotional and social support. Reddit is a forum where a writer does not necessarily address another individual. Replies respond to specific questions and comments, but we cannot assume that the author of a reply has a specific addressee in mind. Using natural-language processing and perhaps the analysis of pronouns, we may be able to infer better information about addressees in the future. This caveat implies that one needs to analyze more than one corpus to draw conclusions about audience design, for we cannot always determine who a writer's *audience* actually is.

With the parallel presentation of lexical and syntactic adaptation data, I would like to draw attention to some problems with the use of multi-level alignment in corpus data for psycholinguistic modeling. While syntactic priming lends itself to corpus study, lexical adaptation may not be due to a priming effect in the process of lexical choice. Lexical repetition has much to do with the topic structure of text: lexical choices are of course, in part, a consequence of shifting topic clusters. Topic shifting similarly interacts with information distribution in discourse Qian and Jaeger, 2011. Short of modeling topic flow in an attempt to subtract topic effects from the observed lexical alignment, we have to rely on additional syntactic analyses to draw conclusions about language production.

Experimental control for lexical repetition and topic clustering is possible to extent. And indeed, when comparing syntactic priming effects vanish or are overshadowed by negative priming effects where speakers avoid syntactic parallels (Healey, Purver, & Howes, 2014). These results, on non-task-oriented, spontaneous conversation, first underline the need to verify lab-derived studies on naturalistic data in order to get a picture of how ubiquitous a reported effect really is. Despite the observations I presented here as an argument against audience design, it is clear that alignment is still sensitive to modulation

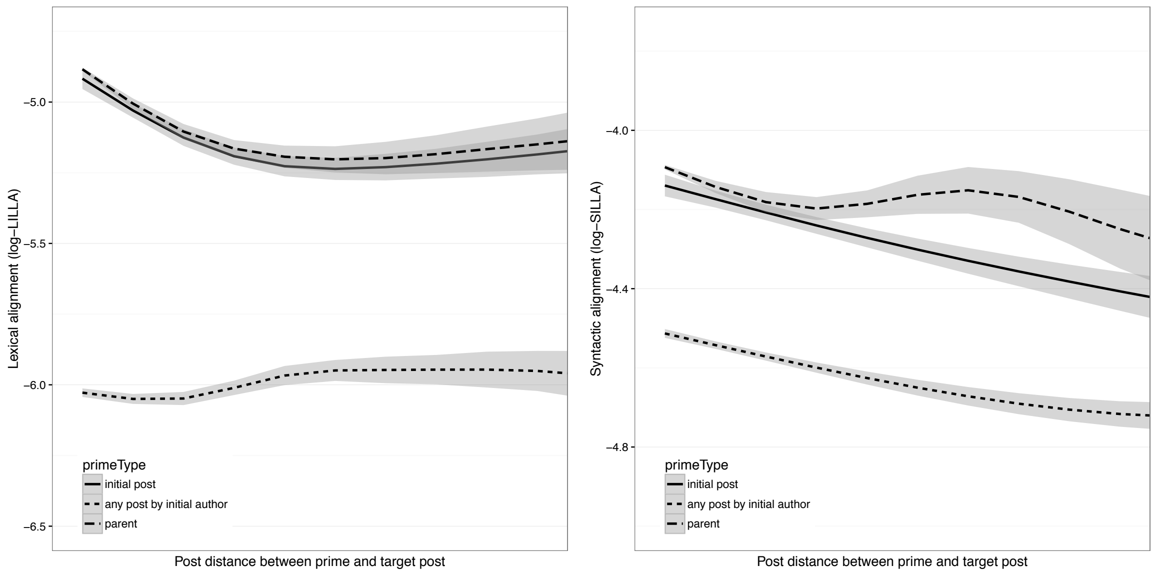

Figure 2. Lexical and syntactic alignment in Reddit as a function of origin-target distance measured in intervening replies (+1). Repetition between the first (F) post of a thread and any other, a parent (P) origin and a target post, and those posts by the thread initiator (I) and any target post. Shaded areas indicate approx. 95% confidence intervals assuming gaussian errors around LOESS regression (as in other graphs).

as a memory-based effect. High syntactic rule frequency reduces priming (Scheepers, 2003; Snider & Jaeger, 2009), and semantic repetition or relatedness (as in topic chains) is predicted to boost syntactic priming (Reitter et al., 2011).

Questions and challenges for data-intensive computational psycholinguistics

As I have shown, alignment is a phenomenon that can be examined using naturalistic language. I use datasets that reach from small, annotated corpora, such as *MapTask*, to huge web forums, such as *Reddit*. Oftentimes, the size of the dataset comes with a tradeoff regarding annotations. Small datasets offer deeper, more precise and hand-corrected annotations (e.g., Zeldes, 2016), while big data in the range of hundreds of megabytes in size are far too large for such annotations. In that case, interesting variables such as inter-speaker relationships have to be inferred from network structure or linguistic phenomena.

Existing datasets present attractive opportunity for psycholinguistics, and cognitive science as a whole. Small-scale data with reliable secondary information, such as the Dundee eye-tracking corpus (Kennedy & Pynte, 2005) or the Schilling corpus (Schilling, Rayner, & Chumbley, 1998) may give insights into syntax processing (Demberg & Keller, 2008; Bicknell & Levy, 2010). Such data may be used to evaluate and refine even complex models of reading (Reichle, Rayner, & Pollatsek, 2003). Larger-scale, naturalistic and relatively unannotated datasets similarly play a role, as in work on memorability (Danescu-Niculescu-Mizil, Cheng, Kleinberg, & Lee, 2012) or on alignment (Danescu-Niculescu-Mizil & Lee, 2011) using movie dialogue transcripts.

I will suggest three components of a vision for psycholinguistics that uses datasets to

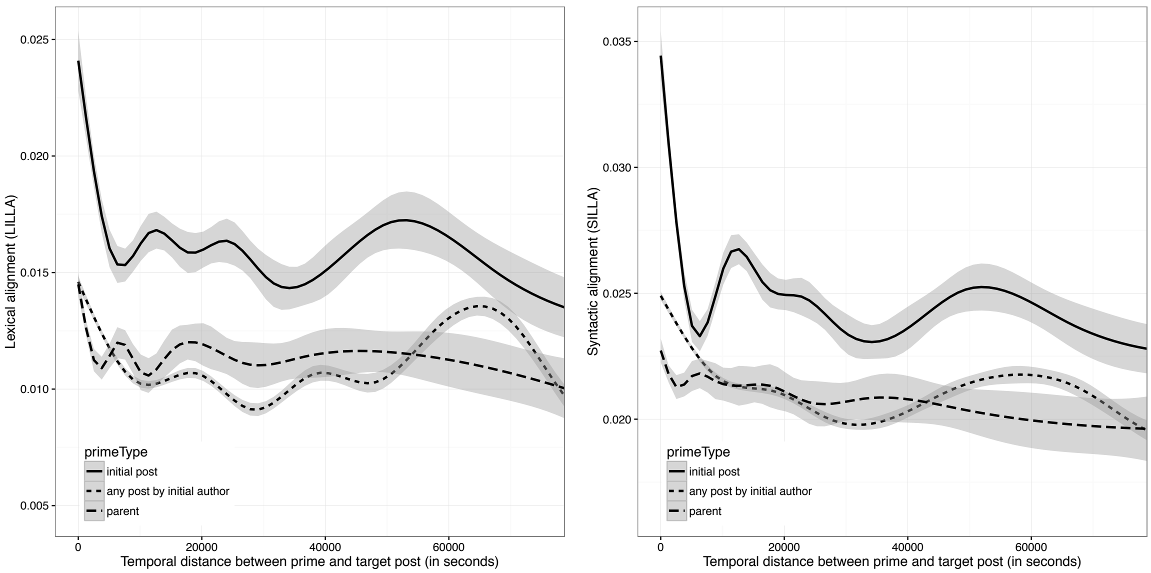

Figure 3. Lexical and syntactic alignment in Reddit as a function of origin-target distance measured in time of posting. Repetition between the first (F) post of a thread and any other, a parent (P) origin and a target post, and those posts by the thread initiator (I) and any target post.

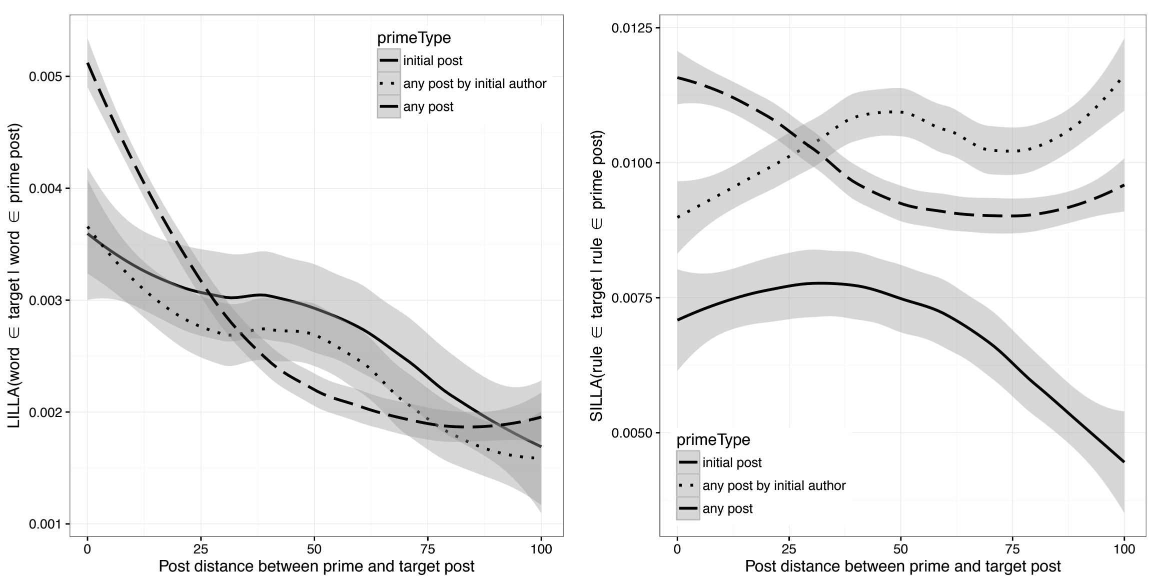

Figure 4. Lexical and syntactic alignment in the Cancer Survivors Network. Data from Wang, Reitter, and Yen (2014).

examine language processing and communication. Just like this chapter as a whole, they reach from internal representations and micro-processes to high-level models of text and dialgoue.

Adaptation as a window into the mind. Syntactic priming has been of great interest to psycholinguists. Its temporal decay and its associations with other linguistic representations give clues about the memory representations involved in language production and comprehension. Going beyond, syntactic priming and indeed alignment at multiple levels can be useful in identifying concrete processes and mental representations. To give an example: suppose we hypothesize a structure X that is part of a language's grammar, and that is representative of a cognitive process involved in speaking that language. The basic principle we follow is that another speaker will adjust his or her uses of X after hearing another speaker use X. However, for this adjustment of the speaker's language model to happen, structure X must be actually cognitively present. Otherwise, there would be no memory item to reinforce. That means that by identifying speakers adapt upon hearing X we find evidence in favor of X as a cognitive artifact. As long as we can cheaply determine, on a large dataset, where that structure applies, we can measure sensitivity to its use and thereby detect adaptation. As an example, we have done that in a small study with competing classes of representations (Reitter, Hockenmaier, & Keller, 2006a). The structures we looked at described either fully incremental or non-incremental syntactic processing. (The question here is whether new words and phrases are immediately adjoined to the semantics or syntactic type of the existing sentence, or if they are buffered in some form of working memory and combined out of order.) By looking at adaptation in a relatively small corpus, we found some hints that incrementality is actually flexible —although more work is necessary to robustly model incrementality on more data. Much of this work can be done cheaply on unannotated data once we have the computational means to induce grammar from data (c.f., Bod, 1992) based on weak adaptation effects.

Evaluation of integrated models of language acquisition. Large-scale datasets rarely come annotated with interpreted linguistic knowledge. However, using predefined deep linguistic knowledge should not be more than a temporary goal anyway. After all, we model how individuals can learn to process *surface form* into semantics. Syntactic structure, for example, is transient. It reflects a cognitive process rather than permanent mental representations. Which representations are learned from the data is the consequence of computational constraints for language processing, prior knowledge, and general-purpose learning algorithms. Inference from raw language data (in multiple languages) without prior constraints may well be possible. However, it is more likely that the integration of language and general cognitive architectures will provide useful priors. I see two complementary approaches. The first approach here is to integrate what cognitive psychology teaches us, in architecture as well as quantitatively, about memory and processing. ACT-R (J. R. Anderson, 2007) defines an independently validated set of principles that codify that knowledge computationally. In Reitter et al. (2011), I proposed a model of syntactic priming in ACT-R, and we are now extending its coverage to corpora. The second approach to integrating general cognitive architectures is to assume basic constraints of learnability. Does the data that a learner is exposed to hold enough information to acquire the hypothesized structures and processes when combined with a general-purpose learning mechanism? In this context, we should consider connectionist approaches, which have seen a remarkable resurgence, mostly prompted by Hinton, Osindero, and Teh's 2006 discovery of methods that facilitate the learning of multi-layered ("deep") networks, which lead to plausible real-world performance in machine-learning tasks. The underlying artificial neural networks are only loosely inspired by biological neural networks. However, at a higher level, some types of networks may serve as models of learning. Frameworks with the potential to enable not just language processing (Manning, 2015) but general artificial intelligence may be integrated into psycholinguistic models. This approach involves online, semi-supervised learning mechanisms (e.g., Ororbia II, Reitter, Wu, and Giles, 2015). This systems will acquire structural representations that allow us to make context-dependent processing decisions. They do so rapidly with just a few annotated and many unannotated examples. If complex syntactic and lexical representations *can* be learned from unannotated data, we may have a computational answer to the *Poverty of the Stimulus* argument (Chomsky, 1965; Pullum & Scholz, 2002), which suggests that children need to have substantive language acquisition device (e.g., Pinker, 1991).

Information-theoretic models of text and dialogue. The mention of learnability brings us to the question of computational and psychological plausibility of processing: which algorithms can recognize words, parse sentences and interpret meaning given the information available in text or dialogue (c.f., Lewis, Vasishth, and Van Dyke, 2006)? Computational psycholinguists have been particularly interested in the question of predictability, as it gives insights into probabilistic learning of, e.g., word meanings or syntactic choices. For instance, expectations that are guided by past experience can facilitate or burden online processing (e.g., Hale, 2003; Levy, 2008; Smith and Levy, 2013). These accounts may be seen as agnostic to the concrete algorithms that produce or comprehend language. Yet they do represent a higher-level description of informativity. Information content, or conversely, entropy, varies systematically throughout small and large units of text (Genzel & Charniak, 2003). It has been hypothesized that speakers striving to distribute information evenly among text are doing so in order to optimize the use of cognitive resources (Jaeger, 2010). To date, dialogue as a text genre has been under-studied with respect to entropy distribution, even though a model of entropy in dialogue may answer a key question: Which dialogue partner contributes information, and why?

Conclusion

In this article, I hoped to demonstrate a big-data approach to cognitive science that observes linguistic performance in the wild. The minimally controlled environment comes with obvious benefits and with some challenges. The benefits lie in the broad coverage of syntactic constructions, conversational styles, and communities. With the analysis of dialogue corpora such as *Switchboard*, *Maptask* and *Reddit*, we were able to not only show that alignment effects in real-world data were smaller than observed in the lab, but that they also varied in theoretically relevant ways: for example, with task success (Reitter & Moore, 2014), but not necessarily with the intended audience.

The challenges of the big-data approach, however, also illustrate where a carefully constructed experiment can produce more informative conclusions. The correlation between lexical and syntactic levels is an example for this problem. Work with large datasets in general comes with an inherent challenge: they are *observational*. While we can observe correlations, causal inference is much more difficult and requires more information, such as temporal relationships (what happens later cannot have caused what happened earlier). However, direct causal relationships without latent variables cannot be inferred. Further, with a large dataset, we can usually find some correlations that are deemed significant, numerically. As Adam Kilgariff put it: "Language is never, ever random" (Kilgarriff, 2005). We should not rely on a single dataset or at least not on a single sample of one to draw good conclusions. Most importantly, the opportunity to observe seemingly reliable correlations between variables emphasizes our obligation to always begin with a theoretical framework and clear hypotheses. For with hindsight, many models can explain observational data.

Acknowledgements

This research was supported by the National Science Foundation (IIS 1459300 and BCS 1457992). The analyses on the CSN dataset was made possible by a collaboration agreement between Penn State and the American Cancer Society. I would like to thank Yafei Wang, Yang Xu, John Yen for their discussions and work on the CSN dataset, Kenneth Portier and Greta E. Greer for creating and making available the CSN dataset, Lillian Lee for her comments on the Reddit dataset, as well as Jason M. Baumgartner (Pushshift.io) for curating the Reddit dataset, and Jeremy Cole for advice in preparing this manuscript.

References

Anderson, A. H., Bader, M., Bard, E., Boyle, E., Doherty, G. M., Garrod, S., ... Weinert, R. (1991). The HCRC Map Task corpus. *Language and Speech*, *34*(4), 351–366.

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.

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

Anderson, J. R. & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Erlbaum.

Bard, E. G., Anderson, A. H., Sotillo, C., Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000). Controlling the intelligibility of referring expressions in dialogue. *Journal of Memory and Language*, *42*, 1–22.

Bicknell, K. & Levy, R. (2010). A rational model of eye movement control in reading. In *Proceedings of the 48th annual meeting of the Association for Computational Linguistics (ACL)* (pp. 1168–1178). Uppsala, Sweden: Association for Computational Linguistics. Retrieved from http://dl.acm.org/citation.cfm?id=1858681.1858800

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, *18*, 355–387.

Bock, J. K., Dell, G. S., Chang, F., & Onishi, K. H. (2007). Persistent structural priming from language comprehension to language production. *Cognition*, *104*(3), 437–458.

Bock, J. K. & Griffin, Z. M. (2000). The persistence of structural priming: transient activation or implicit learning? *Journal of Experimental Psychology: General*, *129*(2), 177–192.

Bod, R. (1992). A computational model of language performance: data oriented parsing. In *Proceedings of the 14th conference on computational linguistics-volume 3* (pp. 855–859). Association for Computational Linguistics.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (1999). Syntactic priming in language production: evidence for rapid decay. *Psychonomic Bulletin and Review*, *6*(4), 635–640.

Branigan, H. P., Pickering, M. J., McLean, J. F., & Cleland, A. A. (2007). Syntactic alignment and participant role in dialogue. *Cognition*, *104*(2), 163–197. doi:http://dx.doi.org/10.1016/j.cognition.2006.05.006

Branigan, H. P., Pickering, M. J., Pearson, J., McLean, J. F., & Brown, A. (2011). The role of beliefs in lexical alignment: evidence from dialogs with humans and computers. *Cognition*, *121*(1), 41–57. doi:10.1016/j.cognition.2011.05.011

Brennan, S. E. & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(6), 1482.

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

Chapelle, O., Schölkopf, B., & Zien, A. (Eds.). (2006). *Semi-supervised learning*. Adaptive computation and machine learning. Cambridge, MA: MIT Press.

Chartrand, T. L. & Bargh, J. A. (1999). The chameleon effect: the perception–behavior link and social interaction. *Journal of personality and social psychology*, *76*(6), 893.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. & Miller, G. A. (1963). Introduction to the formal analysis of natural languages. In *Handbook of mathematical psychology* (Vol. 2, pp. 269–321). New York, NY: Wiley.

Cleland, A. A. & Pickering, M. J. (2003). The use of lexical and syntactic information in language production: evidence from the priming of noun-phrase structure. *Journal of Memory and Language*, *49*, 214–230.

Danescu-Niculescu-Mizil, C., Cheng, J., Kleinberg, J., & Lee, L. (2012). You had me at hello: how phrasing affects memorability. In *Proceedings of the 50th annual meeting of the Association for Computational Linguistics (ACL)* (pp. 892–901). Jeju Island, Korea: Association for Computational Linguistics. Retrieved from http://dl.acm.org/citation.cfm?id=2390524.2390647

Danescu-Niculescu-Mizil, C. & Lee, L. (2011). Chameleons in imagined conversations: a new approach to understanding coordination of linguistic style in dialogs. In *Proceedings of the 2nd workshop on cognitive modeling and computational linguistics* (pp. 76–87). Association for Computational Linguistics.

del Prado Martin, F. M. & Bois, J. W. D. (2015). Syntactic alignment is an index of affective alignment: an information-theoretical study of natural dialogue. In *Proceedings of the 37th annual meeting of the Cognitive Science Society*. Pasadena, CA: Cognitive Science Society.

Dell, G. S., Chang, F., & Griffin, Z. M. (1999). Connectionist models of language production: lexical access and grammatical encoding. *Cognitive Science*, *23*(4), 517–542.

Demberg, V. & Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition*, *109*(2), 193–210.

Elman, J. L. (1990). Finding structure in time. *Cognitive science*, *14*(2), 179–211.

Fusaroli, R., Bahrami, B., Olsen, K., Roepstorff, A., Rees, G., Frith, C., & Tylén, K. (2012). Coming to terms quantifying the benefits of linguistic coordination. *Psychological Science*, *23*(8), 931–939.

Garrod, S. & Anderson, A. (1987). Saying what you mean in dialogue: a study in conceptual and semantic coordination. *Cognition*, *27*, 181–218.

Genzel, D. & Charniak, E. (2003). Variation of entropy and parse trees of sentences as a function of the sentence number. In *Proceedings of the 2003 conference on empirical methods in natural language processing (EMNLP)* (pp. 65–72). Sapporo, Japan: Association for Computational Linguistics.

Gibson, E. (1998). Linguistic complexity: locality of syntactic dependencies. *Cognition*, *68*(1), 1–76. doi:10.1016/S0010-0277(98)00034-1

Gibson, E. A. F. (1991). *A computational theory of human linguistic processing: memory limitations and processing breakdown* (Doctoral dissertation, School of Computer Science, Carnegie Mellon University).

Gries, S. T. (2005). Syntactic priming: a corpus-based approach. *Journal of Psycholinguistic Research*, *34*(4), 365–399.

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

Hartsuiker, R. J., Bernolet, S., Schoonbaert, S., Speybroeck, S., & Vanderelst, D. (2008). Syntactic priming persists while the lexical boost decays: evidence from written and spoken dialogue. *Journal of Memory and Language*, *58*(2), 214–238.

Hartsuiker, R. J. & Westenberg, C. (2000). Persistence of word order in written and spoken sentence production. *Cognition*, *75*(2), B27–B39.

Healey, P. G., Purver, M., & Howes, C. (2014). Divergence in dialogue. *PLOS ONE*, *9*(6), e98598.

Henderson, J. & Ferreira, F. (2004). *Interface of language, vision and action: eye movements and the visual world*. New York, NY: Psychology Press.

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. (2006). *Redundancy and syntactic reduction in spontaneous speech* (Doctoral dissertation, Stanford University).

Jaeger, T. F. (2010). Redundancy and reduction: speakers manage syntactic information density. *Cognitive Psychology*, *61*(1), 23–62.

Jaeger, T. F. & Snider, N. (2007). Implicit learning and syntactic persistence: surprisal and cumulativity. *University of Rochester Working Papers in the Language Sciences*, *3*(1), 26–44.

Jaeger, T. F. & Snider, N. (2008). Implicit learning and syntactic persistence: surprisal and cumulativity. In *Proceedings of the 30th annual conference of the Cognitive Science Society* (pp. 1061–1066). Washington: Cognitive Science Society.

Jones, S., Cotterill, R., Dewdney, N., Muir, K., & Joinson, A. (2014). Finding zelig in text: a measure for normalising linguistic accommodation. In *Proceedings of the 25th international conference on computational linguistics (COLING)*. Dublin, Ireland.

Joshi, A. K. & Schabes, Y. (1997). Tree-adjoining grammars. In G. Rozenberg & A. Salomaa (Eds.), *Handbook of formal languages* (Vol. 3, pp. 69–124). Berlin, New York: Springer.

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.

Kaschak, M. P., Kutta, T. J., & Schatschneider, C. (2011). Long-term cumulative structural priming persists for (at least) one week. *Memory & Cognition*, *39*(3), 381–388.

Kaschak, M. P., Loney, R. A., & Borregine, K. L. (2006). Recent experience affects the strength of structural priming. *Cognition*, *99*, B73–B82.

Kennedy, A. & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research*, *45*(2), 153–168.

Kilgarriff, A. (2005). Language is never ever ever random. *Corpus Linguistics and Linguistic Theory*, *1*(2), 263–276.

Kootstra, G. J., van Hell, J. G., & Dijkstra, T. (2010). Syntactic alignment and shared word order in code-switched sentence production: evidence from bilingual monologue and dialogue. *Journal of Memory and Language*, *63*(2), 210–231.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*(3), 1126–1177.

Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, *10*(10), 447–454.

Manning, C. D. (2015). Computational linguistics and deep learning. *Computational Linguistics*, *41*(4), 701–707. doi:10.1162/COLIa00239

Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S. J., & McClosky, D. (2014). The Stanford CoreNLP natural language processing toolkit. In *Proceedings of 52nd annual meeting of the Association for Computational Linguistics: system demonstrations* (pp. 55–60). Baltimore, MD: Association for Computational Linguistics.

McKelvie, D. (1998). *SDP - spoken dialogue parser* (tech. rep. No. HCRC-TR/96). Human Communication Research Centre. Edinburgh, UK.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, & K. Q. Weinberger (Eds.), *Advances in neural information processing systems 26* (pp. 3111–3119). Curran Associates, Inc.

Mnih, A. & Hinton, G. E. (2009). A scalable hierarchical distributed language model. In D. Koller, D. Schuurmans, Y. Bengio, & L. Bottou (Eds.), *Advances in neural information processing systems 21* (pp. 1081–1088). Curran Associates, Inc.

Nenkova, A., Gravano, A., & Hirschberg, J. (2008). High frequency word entrainment in spoken dialogue. In *Proceedings of the 46th annual meeting of the Association for Computational Linguistics and Human Language Technologies: short papers* (pp. 169–172). Columbus, Ohio: Association for Computational Linguistics.

Niederhoffer, K. G. & Pennebaker, J. W. (2002). Linguistic style matching in social interaction. *Journal of Language and Social Psychology*, *21*(4), 337–360.

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)* (pp. 471–481). Lisbon, Portugal: Association for Computational Linguistics.

Ororbia II, A. G., Reitter, D., Wu, J., & Giles, C. L. (2015). Online learning of deep hybrid architectures for semi-supervised categorization. In *Machine learning and knowledge discovery in databases (ECML PKDD)* (Vol. 9284, pp. 516–532). Lecture Notes in Computer Science. Porto, Portugal: Springer.

Paxton, A., Abney, D., Kello, C. K., & Dale, R. (2014). Network analysis of multimodal, multiscale coordination in dyadic problem solving. In P. M. Bello, M. Guarini, M. Mc-

Shane, & B. Scassellati (Eds.), *Proceedings of the 36th annual meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Pickering, M. J. & Branigan, H. P. (1998). The representation of verbs: evidence from syntactic priming in language production. *Journal of Memory and Language*, *39*, 633–651.

Pickering, M. J. & Ferreira, V. S. (2008). Structural priming: a critical review. *Psychological Bulletin*, *134*(3), 427–459.

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

Pinker, S. (1991). Rules of language. *Science*, *253*(5019), 530–535.

Pollard, C. & Sag, I. A. (1994). *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.

Pullum, G. K. & Scholz, B. C. (2002). Empirical assessment of stimulus poverty arguments. *The Linguistic Review*, *18*(1-2), 9–50.

Qian, T. & Jaeger, T. F. (2011). Topic shift in efficient discourse production. In *Proceedings of the 33rd annual conference of the Cognitive Science Society* (pp. 3313–3318). Boston, MA.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The ez reader model of eye-movement control in reading: comparisons to other models. *Behavioral and Brain Sciences*, *26*(04), 445–476.

Reitter, D. (2008). *Context effects in language production: models of syntactic priming in dialogue corpora* (Doctoral dissertation, University of Edinburgh).

Reitter, D., Hockenmaier, J., & Keller, F. (2006a). Priming effects in combinatory categorial grammar. In *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 308–316). Sydney, Australia.

Reitter, D., Keller, F., & Moore, J. D. (2006b). Computational modeling of structural priming in dialogue. In *Proceedings of the Human Language Technology Conference/North American Chapter of the Association for Computational Linguistics (HLT/NAACL)* (pp. 121–124). New York, NY.

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

Reitter, D., Moore, J. D., & Keller, F. (2006). Priming of syntactic rules in task-oriented dialogue and spontaneous conversation. In *Proceedings of the 28th annual conference of the Cognitive Science Society* (pp. 685–690). Vancouver, Canada.

Scheepers, C. (2003). Syntactic priming of relative clause attachments: persistence of structural configuration in sentence production. *Cognition*, *89*, 179–205.

Schilling, H. E., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: word frequency effects and individual differences. *Memory & Cognition*, *26*(6), 1270–1281.

Smith, N. J. & Levy, R. (2013). The effect of word predictability on reading time is logarithmic. *Cognition*, *128*(3), 302–319.

Snider, N. & Jaeger, T. F. (2009). Syntax in flux: structural priming maintains probabilistic representations. Poster at the 15th Annual Conference on Architectures and Mechanisms of Language Processing. Barcelona, Spain.

Steedman, M. (2000). *The syntactic process*. Cambridge, MA: MIT Press.

Szmrecsanyi, B. (2005). Creatures of habit: a corpus-linguistic analysis of persistence in spoken english. *Corpus Linguistics and Linguistic Theory*, *1*(1), 113–149.

Szmrecsanyi, B. (2006). *Morphosyntactic persistence in spoken english. A corpus study at the intersection of variationist sociolinguistics, psycholinguistics, and discourse analysis*. Berlin, Germany: Mouton de Gruyter.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632–1634.

Travis, C. E. (2007). Genre effects on subject expression in spanish: priming in narrative and conversation. *Language Variation and Change*, *19*(02), 101–135.

Ferreira, V. S. & Bock, J. K. (2006). The functions of structural priming. *Language and Cognitive Processes*, *21*(7-8), 1011–1029.

Van Zaanen, M. (2000). ABL: alignment-based learning. In *Proceedings of the 18th conference on computational linguistics* (Vol. 2, pp. 961–967). COLING '00. Saarbrücken, Germany: Association for Computational Linguistics.

Wang, Y., Reitter, D., & Yen, J. (2014). Linguistic adaptation in online conversation threads: analyzing alignment in online health communities. In *Proceedings of the fifth workshop on cognitive modeling and computational linguistics (CMCL)* (pp. 55–62). Baltimore, MD.

Wang, Y., Yen, J., & Reitter, D. (2015). Pragmatic alignment on social support type in health forum conversations. In *Proceedings of the sixth workshop on cognitive modeling and computational linguistics (CMCL)* (pp. 9–18). Denver, CO: Association for Computational Linguistics.

Xu, Y. & Reitter, D. (2015). An evaluation and comparison of linguistic alignment measures. In *Proceedings of the sixth workshop on cognitive modeling and computational linguistics (CMCL)* (pp. 58–67). Denver, CO: Association for Computational Linguistics.

Xu, Y. & Reitter, D. (2016). Convergence of syntactic complexity in conversation. In *Proceedings of the 54th annual meeting of the Association for Computational Linguistics*. ACL '16. Berlin: Association for Computational Linguistics.

Zeldes, A. (2016). The GUM corpus: creating multilayer resources in the classroom. *Language Resources and Evaluation*, 1–32. doi:10.1007/s10579-016-9343-x